KANSAS

The Kansas Information Technology Architecture

Version 11.2
July 2009

The Kansas Information Technology Architecture
Version 11.2
July 2009

For further copies or questions regarding this document, please contact:

Kansas Information Technology Office
900 SW Jackson, Room 751-S
Topeka, Kansas 66612-1275
Phone: (785) 296-8055

Bill Roth
Kansas Chief Information Technology Architect
900 SW Jackson, Room 751-S
Topeka, Kansas 66612-1275
Phone: (785)296-2108

Email: William.Roth@da.ks.gov

This Architecture provided to the State of Kansas by the Kansas Technical Architecture Review Board (KTARB), presented to and approved by the Information Technology Executive Council (ITEC) on July 23, 2009.

© The State of Kansas 2009

Table of Contents

Executive Overview i

Part 1 Architecture scope, concepts, and objectives
Chapter 1 Introduction..... 1-1
Chapter 2 Kansas Enterprise Architecture overview 1-9
Chapter 3 Architecture Governance 1-14

Part 2 KITA Target summary
Chapter 4 KITA Targets 2-1

Part 3 Kansas Technical Reference Model
Chapter 5 Service Access & Delivery 3-1
Chapter 6 Service Platform & Infrastructure 3-25
Chapter 7 Component Framework..... 3-56
Chapter 8 Service Interface & Integration..... 3-70

Part 4 Kansas Service Component Reference Model
Chapter 9 Customer Services..... 4-1
Chapter 10 Process Automation 4-10
Chapter 11 Business Management Services..... 4-15
Chapter 12 Digital Asset Services 4-21
Chapter 13 Business Analytical Services 4-41
Chapter 14 Back Office Services 4-53
Chapter 15 Support Services..... 4-80

Appendices
Kansas Technical Architecture Review Board & Subcommittees A-1
KITA Version Change Control..... A-4
Technical Architecture Policies & Statutes A-6

Executive Overview

KITA today

The KITA V11.2 reflects 2 different views. It reflects the reality of what the state is currently operating and it reflects the direction that we have agreed to move toward. This direction is clarified in Chapter 4 TARGETS. We can not get away from our legacy technology, systems and investments. It will continue to grow and weigh us down. Classifying that legacy using TWILIGHT, CURRENT, EMERGING terms gives us focus on how to view that legacy in terms of support for our future.

The KITA V11.2 is an outcome of legislation (K.S.A. 75-7201 et seq) and is sponsored by the Information Technology Executive Council(ITEC) and the IT community of all branches of government. It provides the Technology aspect of the Kansas Enterprise Architecture (KEA).

The KITA and the Kansas Technical Architecture Review Board provide the mechanisms to exchange ideas on the State strategies (as reflected in latest Strategic Information Management Plan (SIM Plan)) and view of Targets to get to those goals versus agency strategies.

Kansas is working with other states and the federal agencies to move toward a consistent Architecture reference model that we can all refer. This Federal Enterprise Architecture (FEA) and our KEA provide those connections for Kansas. This will help agencies work with peer agencies in other states to share and compare like components and investments.

Future efforts to support our desired outcomes may include some of the following

- Collecting information about all of the State's IT assets in close to real-time to help identify actual investments carried today. Identify the actual numbers of licenses in use in Kansas and identified in the KITA. Identify technologies that are at risk due to obsolescence or isolation in vendor or personnel support.
- Identify what parts of the architecture will be utilized as part of the annual Three Year Plans, which describes the normal evolution of all agencies IT environment.
- Identify what parts of the architecture will be utilized, impacted and/or enhanced as part of major projects in the State of Kansas.
- Continue to develop an exchange relationship with NASCIO and Federal EA communities to exchange architecture models, standards, and solutions.
- Automate the change management of the KITA and improve collaboration for interested communities in parts of the architecture evolution.
- Automate agencies' yearly alignment with the KITA current and target architecture.

Architecture Usage

The KITA is only useful when it is actively being utilized to help the State of Kansas in every day technology investment decisions. The KITA is also used when IT projects are proposed for Kansas.

If the KITA is actively used the desired outcomes would include:

- Achieving a common view of how IT should be used to support Kansas' mission.
- Reduced barriers to cooperation and information sharing.
- Optimize use of the State's resources (IT and other).

Part 1

Architecture scope, concepts, and objectives

Chapter 1 Introduction

Vision

The vision of the KITA is to deliver an active technical architecture that is used to verify where Kansas currently stands and to focus our future information technology investments on supporting the business needs of State agencies and the citizens of Kansans.

The Kansas Strategic Information Management Plan documents a vision for the use of computers and networks in Kansas:

- **Services**: Support Kansas citizens and businesses so they can access needed Kansas government information and services electronically. KITA Service Reference model describes the Technology service delivery layer. The Targets describe how all government entities will deliver services.
- **Information Sharing**: Every branch and level of Kansas government can exchange and access information electronically, both internally within government and externally throughout the State, nation, and world. KITA supports Targets that describe how we will come together with all State entities to share information.
- **Collaboration**: Formal and informal groups work to break down barriers and provide consistent practices to share investments, efforts, and information in our goal to support our service delivery. KITA describes collaborative environments and content management that bring all thoughts and content together to support physical and virtual collaboration across Kansas.
- **Convergence**: Every tax dollar is maximized through Kansas government cooperation, coordination, and resource sharing, supported by cost-effective information technology. KITA describes the targets that we will be converging on. It also describes twilight and current standards that will need to be moved to the targets. And it describes emerging technologies that have been identified for research and monitoring only.

The Kansas Strategic Information Technology plan, when abstracted to its highest level, calls for the creation of an electronic ability to deliver services via the most effective means avalible.

Mission

Enterprise Architecture in Kansas will support the entirety of Kansas government functions, processes, performance considerations, information, technology and services. The State of Kansas is moving toward the Federal Enterprise Architecture Reference model with the support of National Association of State Chief Information Officers (NASCIO) and to support our relationships with other states, federal and local agencies.

The Kansas Information Technology Architecture (KITA) is a subset of the Kansas Enterprise Architecture and describes the information systems infrastructure that supports the business and applications used by the State. The purpose of the KITA is to guide the development and evolution of the information systems infrastructure. The KITA establishes consistency by helping:

- Provide managers and staff in the agencies and support services an understanding of the information systems infrastructure they are using and where the State is going.
- Provide a mechanism so State IT professionals have a consistent view of the information systems infrastructure and the methods they should employ to develop and deliver information systems services.
- Ensure the various development projects being managed in the State move us toward a target architecture that is supported by the consensus of agencies and management.

Structure

This document is in five parts:

- An introduction that provides background information and a description of the development of the architecture and principles that underlie it;
- A description of the target architecture for the future of IT in Kansas;
- An elaboration of the details of the Technical Reference Model (TRM);
- A focus on the details of the Service Reference Model (SRM);
- Appendices.

To Whom Does the Architecture Apply?

This Architecture applies to the Legislative, Executive, and Judicial Branch Agencies, Boards and Commissions of State Government and includes the Regents Institutions. Although the KITA is not controlling, counties and municipalities are encouraged to develop and implement their enterprise architectures in ways that are consistent with the KITA.

A Note to Vendors

In this document there are references to specific products as current or emerging standards, although our goal is to have a Standards Based Architecture throughout. In a number of cases we have chosen products as the standard where no widely accepted standards exist, or to achieve reduced total cost of ownership or improved service and support goals. There are several reasons for this including getting the IT community to recognize what they are using daily in the KITA, getting the state to come together on a smaller set of overall solutions for an area. For example, while we may not get to one target but we do not need 20 targets. In many instances, however, equivalent products, could comply with the architecture and meet the stringent requirements for lowest cost acquisitions that the State of Kansas uses in its purchasing decisions. Product equivalence determinations relative to the architecture are made by the Kansas Technology Architecture Review Board.

Scope & Business Rationale

The KITA will encompass all technology and technology support services that are essential to support an efficient and effective state government. Kansas agencies operate in a federated environment. This environment allows each agency to support their own Information Technology support staffs and budgeting. As a result, there are tremendous efficiencies to be gained when multiple agencies use similar technical and business approaches to accomplish their mission. The savings come in part from the initial investment. But also from the upkeep of each technical component working with all other components that become dependent on it. Savings also come from training on common approaches compared to many outdated or unique technical solutions that require a trained support staff long after the initial team and/or consultant has it working. Finally savings come from business solutions or information that can be shared and standardized across many agencies or branches of government. This allows business solutions or information to be created once and reused many times versus each unique group developing its own solution or collecting its own data.

This architecture describes:

- All of the information technology used and planned for immediate use in the State.
- The architectures of the agency and department applications, but not their application content.
- The ways in which business and information system services are delivered to users.
- The standards and architectural components that are to be used to govern and control the development and delivery of information systems services within the State of Kansas.

The intent is that the KITA will be applied to all agencies and departments in the State of Kansas.

.

Target Audiences

The KTARB reviewed the audiences that benefit from the KITA. The following tables, while not inclusive, do give a range of interests and perspectives that we are trying to support.

Primary Audiences		
Legislature	Clarify Convergence, Business support and service delivery.	
Governor		
ITEC		

Technology Audiences		
CITO's	ITAB	Identify strategies and similar efforts in other agencies that can be leveraged.
Federal EA council	NASCIO (EA Council)	
Agency CIO's	Agency IT Governance committees	
Technology community of interest		

Government Entities		
Cabinet Heads	Agency Exec IT Governance committees	Understand if we work together how it will affect LOB long term effectiveness
Business community of interest	Federal LOB agency	
Other States		

Other Audiences		
Vendors	Procurement	Understand why we are converging, how it impacts them, some strategies they can use to work with state easier
City/Counties	Citizens	

Principles

Late in 1998, an agency survey was completed by state IT directors to validate a set of architectural principles for driving key objectives and sub-architecture strategies in developing the statewide architecture. From the responses to the survey a list of architectural principles was developed. When the Kansas Information Technology Architecture Review Board was formed in June of 1999, the preliminary list of principles was modified to better reflect the views of the technical community charged with maintaining the architecture. The following modified list reflects the current effort.

1. Systems and technology infrastructure implemented by Kansas State government will be compliant with the architecture even though there may be some additional cost for architectural compliance on initial implementation (the cost to optimize for short term or agency benefits is a compounded cost when you combine total skill, evolution cost, etc… for long term).

2. We will identify opportunities for cross-functional systems and will implement systems in such a way that we can take advantage of standard components throughout Kansas State government.

- *Identical functions within different agencies should be performed with the same system components.*

- *Similar business functions within different agencies should be performed with common components, modified to fit the specific needs of each agency.*
- *System components developed by one agency should be available for reuse by any other agency.*
- *Responsibility for developing and supporting each common system component should be vested in a designated lead agency, with participation by all affected agencies.*

3. Applications and technology components (processors, network, etc.) should be implemented in such a manner that performance measurement and quality assurance data may be captured to support management and analysis of the IT environment.

4. Systems and components will use architecturally compliant commercial off-the-shelf (COTS) products where possible.

5. The architecture will support all forms of information (data, text, voice, video, image) in an integrated manner.

6. All Architecture areas should be considerate of the states need for dependability, stability,and recoverability and will support the necessary levels of privacy, and access to information.

7. Applications and technology components (processors, networks, etc.) should be architected in such a manner that the components are separable to be able to evolve independently. This allows part of the architecture and applications that use that part to evolve without wholesale replacement of applications and solutions.

8. Every architecture decision should take into consideration the FEA reference architecture, and the other States' architectures, if they have developed a similar architecture, so that technology investments can be minimized when multiple States come together on a project and information between state, federal and local government is optimized.

Goals

The overarching goal of the Strategic Information Management Plan is to manage IT from a consolidate approach, to provide citizen access to state services and information and to operate state government as effectively and efficiently as possible. The consolidated management model brings the IT community together to enhance communication and unifies state agencies and boards. This SIM Plan will be managed from a consolidated management philosophy. The SIM Plan Goals include the following:

1. Provide broad access to public information and services.
   - The KITA will support this goal with the clarity of security services and information delivery services to be able to deliver a consistent access model.
2. Use public and private resources effectively and efficiently.
   - By defining the correct combination of COTS components and customized software the state will deliver services and will utilize the IT industry solutions where appropriate. Projects will utilize the KITA for consistence and will support the Project Management Methodology (PMM) that allows use of consultants, or other agency personnel to deliver a solution to the State of Kansas.
   - The KITA evolution to a National Standard Architecture Reference model delivered by the Federal Enterprise Architecture CIO council and supported by the NASCIO Enterprise Architecture committee will allow Kansas to compare and build interstate and intergovernmental solutions with less confusion and overlap.
3. Manage government IT resources effectively and efficiently.
   - The KITA requires that performance measurement and quality assurance data may be captured to support management and analysis of the IT environment.
   - Automated collection of all IT and KITA Assets will clarify where our assets are, who uses them, and will help identify redundancy and underutilization.

- Defining the architecture and products that the state has agreed to support will allow for more efficient procurement practices.
- By clarifying the relationship between Project Management Methodologies in Kansas and the Architecture, we are helping determine the most effective and efficient implementation of infrastructure to support the business initiatives.

4. Promote economic development and citizen awareness in Kansas, and IT proficiency within Kansas state government.
   - By utilizing the Federal Enterprise Architecture as the basis for Kansas we will be able to promote IT proficiency as it supports multiple-state and Federal-State-Local initiatives. This proficiency will pay excellent long term benefits as interstate and intergovernmental sharing of initiatives and information become more common.
   - By clarifying the telecommunication and information sharing strategies of the State of Kansas we can help encourage competition and capabilities needed to support the State of Kansas needs and our business and citizen expectations.
   - By defining the current and target architecture of the State of Kansas we can help direct state and academic education programs to ensure that there is a steady supply of IT skill sets for current and future employees.

Architectures are not developed to bind organizations to inflexible rules; the goals of the architecture are coordination, simplification, improved performance and greater efficiency.

Strategies

This version of the KITA was developed and is targeted to evolve by utilizing the following strategies:

- Restructuring the KITA in alignment with the Federal Enterprise Architecture (FEA) reference models. Since this is a Technical Architecture the alignment includes the FEA's Technical Reference Model (TRM) and Service Reference Model (SRM). This strategy will allow state systems and projects to be matched and compared with other states and federal capabilities as we evolve together as a nation.

- Setting target architecture will bring visibility into where the state as a whole is going. This will support agency strategic and tactical planning efforts for both business initiatives and technical application and infrastructure initiatives.

- Future strategies:
  - Collecting the assets that make up the architecture with an automated asset inventorying system. The state will move toward a fully automated system to identify the redundancy in technology areas and to ease the burden of reporting assets used by agencies and universities. (Future)
  - Supporting efficiency in procurement and utilization practices. When we have an accurate inventory of the IT Assets in use and we have a clearly identified target architecture we should be able to support the procurement and utilization practices to ensure we have the appropriate contracts and license structures in place for current and future deployments. (Future)

Background

Legislative Overview

Kansas statutes organizing information technology coordination at the state level were amended in 1998. K.S.A. 75-7201 et. seq. replaced the Kansas Information Resource Council and the

Office of the Chief Information Architect with the Kansas Information Technology Executive Council (ITEC), Executive, Judicial, and Legislative Chief Information Technology Officers, and the Chief Information Technology Architect.

Adoption of a statewide information technology architecture was explicitly mandated. The portion of K.S.A 75-7201 et. seq. that relates to architecture is:

> (a) The information technology executive council is hereby authorized to adopt such policies and rules and regulations as necessary to implement, administer and enforce the provisions of this act.
>
> (b) The council shall:
>
> (1) Adopt:
>
> - Information technology resource policies and procedures and project management methodologies for all State agencies;
> - *An information technology architecture, including telecommunications systems, networks and equipment, that covers all State agencies;*
> - standards for data management for all State agencies; and
> - a strategic information technology management plan for the State;
>
> (2) provide direction and coordination for the application of the State's information technology resources;
>
> (3) designate the ownership of information resource processes and the lead agency for implementation of new technologies and networks shared by multiple agencies in different branches of State government; and
>
> (4) perform such other functions and duties as necessary to carry out the provisions of this act.

The development of the Kansas Information Technology Architecture (KITA) was in direct response to the statutory charge, and the resulting actions required of ITEC and the Information Technology Advisory Board (ITAB).

SIM Plan

In addition to the Information Technology Architecture, K.S.A 75-7201 et. seq. calls for a Kansas Strategic Information Management Plan (SIM Plan). The goal of the SIM Plan is to coordinate Information Technology (IT) development throughout Kansas State government, thereby promoting citizen access, information sharing, and improved government performance and communications at all levels across governmental boundaries. The Kansas Information Technology Architecture, with the SIM Plan, provides agencies the tools and framework to develop effective and efficient IT systems to meet individual agency needs and the needs of government in the digital age.

Architecture Project Summary

The development of the Kansas Information Technology Architecture in 1995 was not a small undertaking. Agencies needed a comprehensive technical framework for informed IT decisions, and the development process had to be done expeditiously. The Chief Information Technology Office (CITO) in the executive branch sanctioned the development of the first technical architecture, as well as its ensuing format and context.

The process for the first architecture prototype assembled commercial and governmental resources to provide a format and framework that met the needs of Kansas agencies and departments.

Since that original adoption, the Architecture has continued to develop, and is approaching a Maturity Level 4 as measured by the Enterprise Architecture Maturity Model (EAMM) of the National Association of State Chief Information Officers (NASCIO).

The architecture must continue to improve. With the KITA V11 release we modified it so it fits cleanly with the developing Federal Enterprise Architecture (see: http://www.feapmo.gov/fea.htm) and the GIS Architecture being developed by NSGIC. This was the start of being able to define a full Enterprise Architecture strategy for Kansas with the inclusion of a recognized Business Enterprise Architecture.

This update effort was initiated in Sept 2007.

Sep 2007	KTARB established and appointed by ITAB.
Oct-Dec 2007	KITA version 11.1 scope and approach established.
Jan 2009	KITA version 11.1 finalized, Published presented to ITAB and ITEC
Jan 2008	Subcommittee wrap-up meeting occurred, KITA version 11.1 process completed.

Architecture Project Team

The development of the KITA is supported by the Kansas Technical Architecture Review Board (KTARB). KTARB is responsible for keeping the architecture up to date. It determines architectural compliance by evaluating agency requests for waiver and analyzing projects referred to it. It also makes recommendations to the Information Technology Executive Council on changes and extensions to the KITA. It is chaired by the Chief Information Technology Architect in accordance with ITEC Policy 4000, 4010, and 4020.

The Subcommittees are staffed from the State's technical expert community who provide the basis of technical architecture decisions. The subcommittees have the option to invite members of private sector or vendor communities with the request and approval of the subcommittee lead. Please see Appendix I for subcommittee descriptions and members.

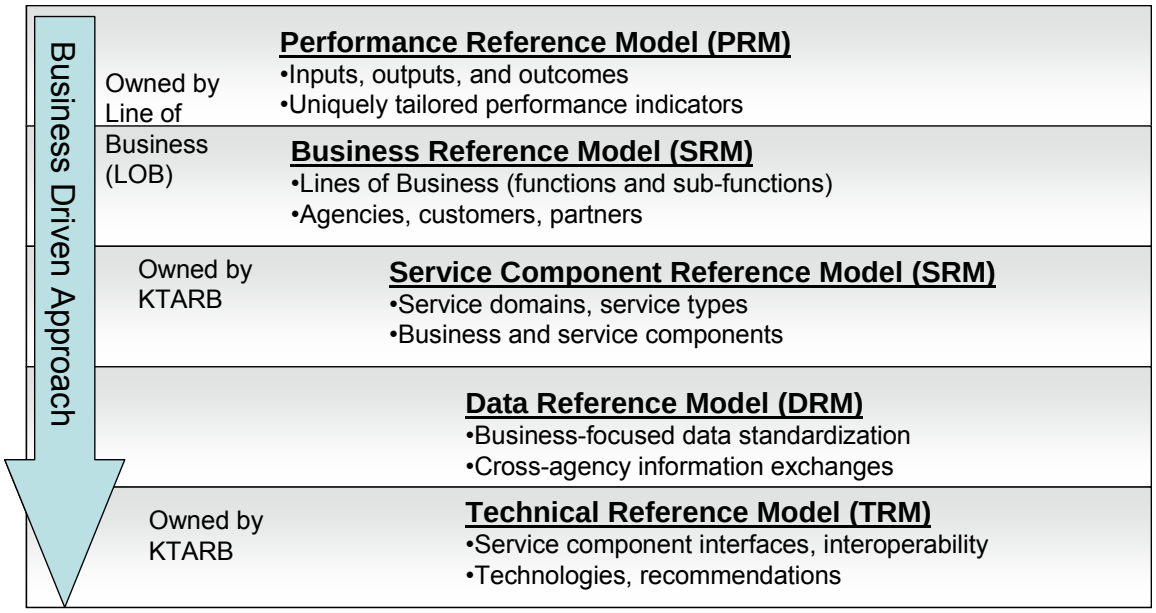
Chapter 2 Kansas Enterprise Architecture

Overview

The KITA is part of a set of interrelated "reference models" designed to facilitate cross-agency analysis and the identification of duplicative investments, gaps and opportunities for collaboration within and across agencies. Collectively, the reference models comprise a framework for describing important elements of the Kansas Enterprise Architecture KEA in a common and consistent way. The KEA is modeled after the Federal Enterprise Architecture (FEA) and follows directions established by the NASCIO Enterprise Architecture Working group as a means to align Federal-State-Local business and IT efforts.

Through the use of this common framework and vocabulary, IT portfolios can be better managed and leveraged across the state government. This chapter introduces the purposes and structures of the five KEA reference models:

- Performance Reference Model (PRM)
- Business Reference Model (BRM)
- Service Component Reference Model (SRM)
- Technical Reference Model (TRM)
- Data Reference Model (DRM)



Performance Reference Model (PRM)

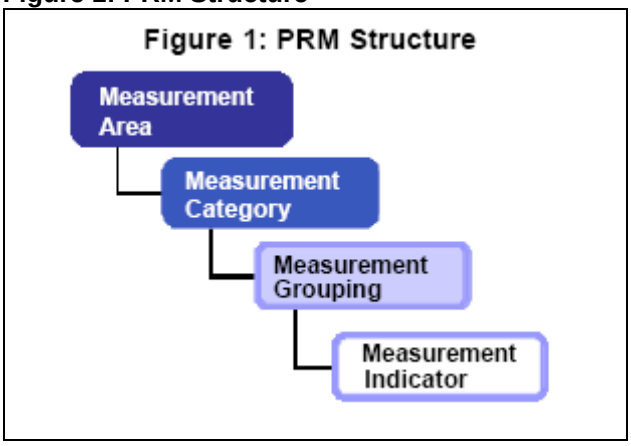
The PRM is a framework for performance measurement providing common output measurements throughout the state government. It allows agencies to better manage the business of government at a strategic level, by providing a means for using an agency's EA to measure the success of IT investments and their impact on strategic outcomes. The PRM accomplishes these goals by establishing a common language by which agency Enterprise Architects (EA) can describe the outputs and measures used to achieve program and business objectives. The model articulates the linkage between internal business components and the achievement of business and customer-centric outputs. Most importantly, it facilitates resource allocation decisions based on comparative determinations of which programs and organizations are more efficient and effective. The PRM focuses on three main objectives:

- Help produce enhanced performance information to improve strategic and daily decision making;
- Improve the alignment and better articulate the contribution of inputs to outputs, thereby creating a clear "line of sight" to desired results;

- Identify performance improvement opportunities that span traditional organizational structures and boundaries.

The PRM structure is designed to clearly express the cause–and-effect relationship between inputs and outputs. This "line of sight" is articulated through the use of the Measurement Area, Category, Grouping, and Indicator hierarchy. Refer to Figure 1 for the PRM structure.

Figure 1: PRM Structure



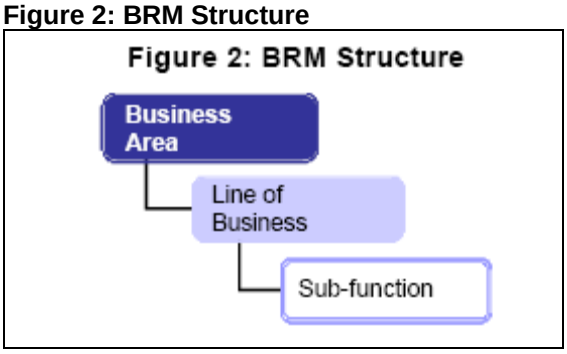
Business Reference Model (BRM)

The BRM provides a framework that facilitates a functional (rather than an organizational) view of the state government's lines of business (LOBs), including its internal operations and its services for citizens, independent of the agencies, bureaus and offices that perform them. The BRM maps the state government around common business areas instead of through a stove-piped, agency-by-agency view. It thus promotes agency collaboration and serves as the underlying foundation for the KEA and E-Gov strategies.

While the BRM does provide an improved way of thinking about government operations, its true utility as a model can only be realized when agencies effectively use it. The functional approach promoted by the BRM will do little to help accomplish the E-Gov strategic goals if it is not incorporated into business-focused enterprise architectures and the management processes of the State of Kansas. This is also known as the Enterprise Business Architecture (EBA) by the NASCIO EA community. This is one of the 2008 SIM Plan strategies to be completed by all agencies so that government services could be understood and mapped consistently across all branches and levels of government.

The BRM is structured into a tiered hierarchy representing the business functions of the state government. Refer to Figure 2 for the BRM tiered hierarchy.

Figure 2: BRM Structure



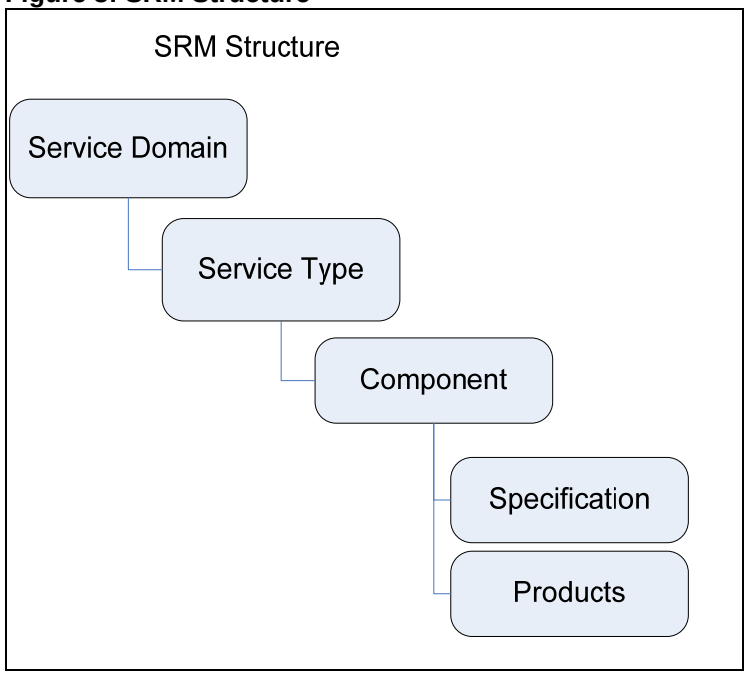
Service Component Reference Model (SRM)

The SRM is a business-driven, functional framework classifying Service Components according to how they support business and performance objectives. It serves to identify and classify horizontal and vertical

Service Components supporting federal agencies and their IT investments and assets. The model aids in recommending service capabilities to support the reuse of business components and services across the federal government.

The SRM is organized across horizontal service areas, independent of the business functions, providing a leverage-able foundation for reuse of applications, application capabilities, components, and business services. It is structured hierarchically as depicted in Figure 3.

Figure 3: SRM Structure

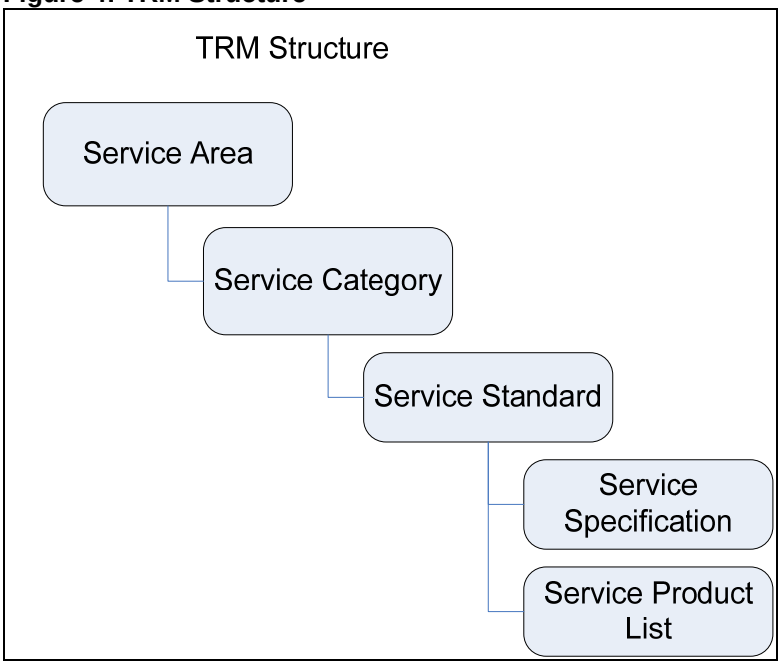


Technical Reference Model (TRM)

The TRM is a component-driven, technical framework that categorizes the standards and technologies to support and enable the delivery of Service Components and capabilities. It also unifies existing agency TRMs and E-Gov guidance by providing a foundation to advance the reuse and standardization of technology and Service Components from a government-wide perspective.

Aligning agency capital investments to the TRM leverages a common, standardized vocabulary, allowing interagency discovery, collaboration, and interoperability. Agencies and the state government will benefit from economies of scale by identifying and reusing the best solutions and technologies to support their business functions, mission, and target architecture. The TRM structure is depicted in Figure 4.

Figure 4: TRM Structure

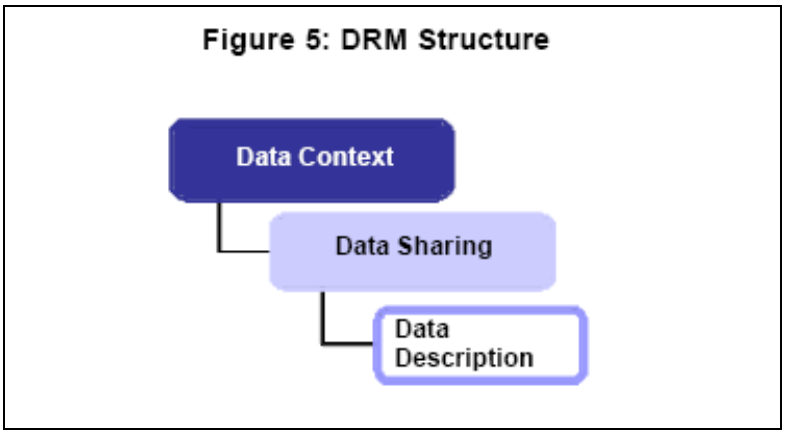


Data Reference Model (DRM)

The KEA Data Reference Model (DRM) is intended to promote the common identification, use, and appropriate sharing of data/information across the state government through its standardization of data in the following three areas: data context, data sharing, and data description (refer to Figure 5).

The current published version of the FEA DRM is undergoing revision. The FEA PMO is collaborating with members of the interagency DRM working group, chartered by the Architecture and Infrastructure Committee (AIC) of the Chief Information Officer (CIO) Council, to further enhance and improve this reference model. The DRM structure presented in Figure 5 is the updated description of the DRM based on the work being done by the FEA PMO and the interagency DRM working group. Because the new version of the DRM has not been completed, the latest published version is provided in this document for reference.

Figure 5: DRM Structure



KEA Profiles

A profile is a framework that cross-cuts the inter-related KEA reference models based upon a particular subject matter. The profiles describe how each reference model addresses a specific area and how agencies can utilize existing resources, standards, best practices, and use cases to implement or improve upon them. The following is a list of profiles that are under development by the FEA council

- **Security and Privacy Profile**: provides guidance on designing and deploying measures that ensure the protection of information resources.
  Security and Privacy Profile
- **Geospatial Profile (in development):** establishes a framework for more effective use and management of geospatial data and services as part of agencies' enterprise architectures. It describes how agencies can leverage geospatial data and technologies to enhance service delivery and mission accomplishment.
  FEA Geospatial Profile, version 1.1
- **Records Management Profile**: provides an overview of the FEA and explains how the reference models provide a context for applying effective records management practices.
  Federal Enterprise Architecture Records Management Profile

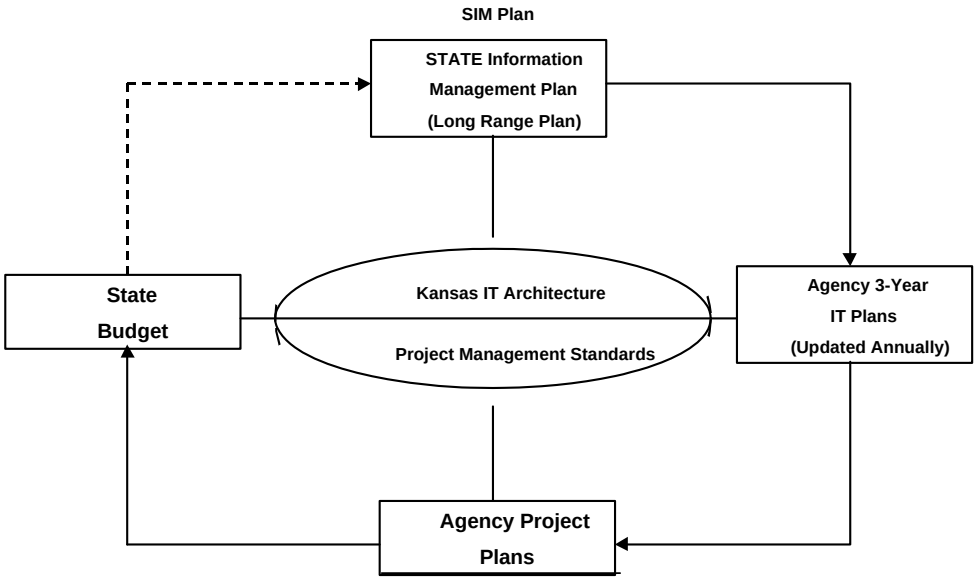
It is intended that Kansas will develop a series of profiles to support our cross cutting efforts. Due to the complexity and interrelations of our environment, Therefore gives us several advantages to separating the architecture details and aligning them with the KEA/FEA reference models as well as bringing together a complete picture for groups to use as a model for future development efforts. These profiles will be a first attempt with this version of the KITA and will be considered a work in progress as we develop them and refine their usage in context of projects and IT management.

Chapter 3 Architecture Governance

3.1 Relationships with other management concepts

This architecture is one of several architectures and sets of information systems documents. The relationships between the architectures and management concepts are shown in Figure 6.

Figure 6: Kansas IT Planning/Controls



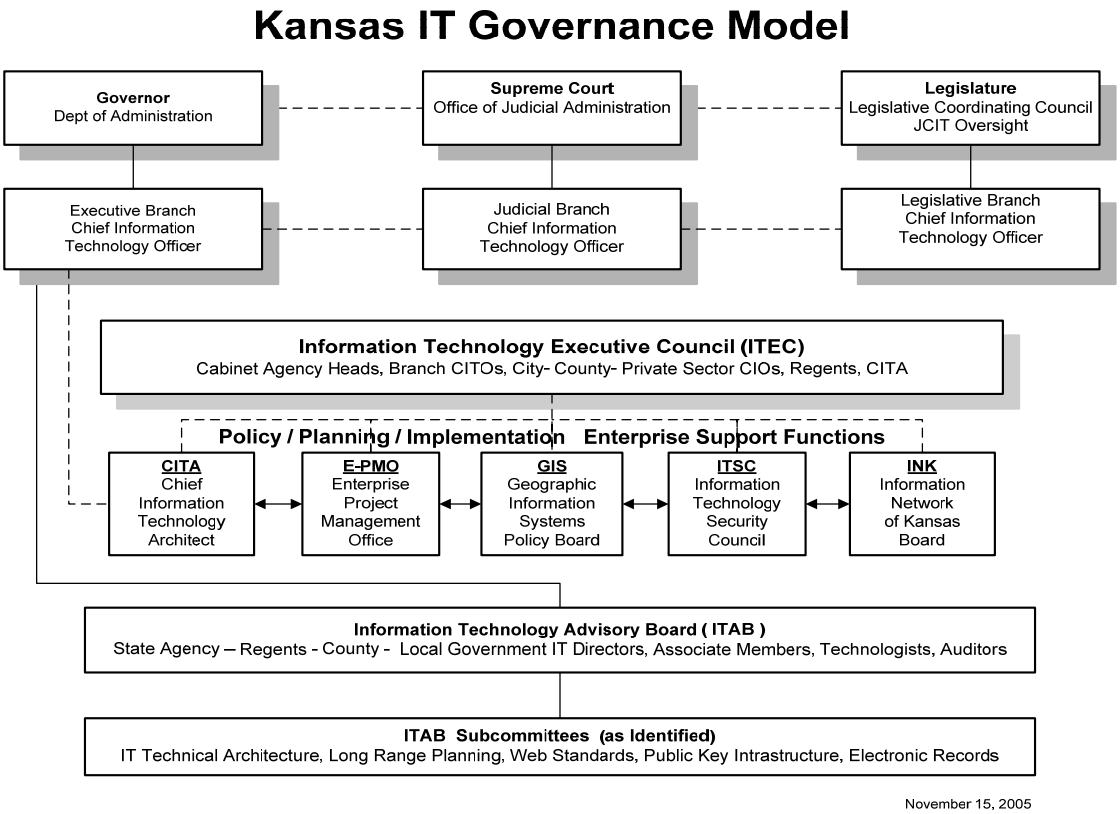
These concepts are described as:

- Strategic Information Management Plan (SIM Plan) updated every 2-3 years to provide updated strategic vision and objectives to enable and support government in Kansas.
- Kansas IT Architecture - The infrastructure that underpins the applications used within the State. This is the Technical and Service architecture and sub-architectures described in the rest of this document.
- Process and applications architectures – These elements are covered under K.S.A 75-7201 et seq., which calls for IT resource policies, procedures, and project management methodologies for all State agencies. Specific processes and policies that already exist will be identified and linked to sub-architecture elements within this document.
- Agency 3 year IT management and budget plans- Tactical Agency Plans (Engineering documents) – Development of an agency level technical architecture leads to a number of engineering documents. These documents will be created and maintained at the agency and department level.
- Project architectures – Developed for individual projects at a lower level to describe the changes being made by each project.

3.2 State IT Governance, Architecture Responsibility

A hierarchy of committees and organizations are involved with the long-term use and direction of the KITA. The relationships and organization of these groups is depicted in Figure 7.

Figure 7: KITA Governance Organization



Organization	Architecture Responsibility
ITEC	• Approve and oversee architecture directions through the CITA
CITO	• Project approval implies architecture compliance
CITA	• Architecture development responsibility under ITEC
ITAB	• Provides membership to the Technology Architecture Review Board
IT Technical Architecture Board (KTARB)	• Care and maintenance of KITA • Reports to ITAB on new directions, standards
Architecture Sub-committees	• Developed as needed to validate or update technical subsections of the KITA
Individual Agencies/ RITC / Regents	• Provide feedback on KITA and comment to CITA • Determine architecture compliance (checklists) for any new projects • Determine architecture compliance with any new purchase • Develop transition plans for KITA twilight components • Test new technologies and notify KTARB for Emerging

Organization	Architecture Responsibility
	technologies • Develop transition plans to KITA Targets
Procurement	• Removed from an IT policing role, now follows CITO approved requests for IT systems

3.3 Architecture Compliance, Enforcement

The architecture embodies a vision, a view of the future of information technology in the State. The architecture defined herein has taken as the definitive statement of that vision the work of the Kansas Strategic Information Management Plan. The commitment has been made to craft this architecture with the widest possible participation from members of the IT community within Kansas State government. Broad agreement with the standards and guidelines is sought as the architecture evolves.

The project approval process, as defined in K.S.A 75-7201 et. seq., should allow for the resolution of any architectural variances that may emerge over time. The architecture is intended to evolve as technology develops and as new initiatives emerge within the vast technology and organizational domains it spans.

In essence, the approval process coordinates many interests, including management, branch CITO, budget, and finally legislative committees.

- Non-compliance with the architecture is BAD
- Compliance with Current Standard is GOOD
- Compliance with Target is GREAT

Architectures are not developed to bind organizations to inflexible rules; the goals of the architecture are coordination, simplification, improved performance and greater efficiency.

3.4 Organization of this Document

The following chapters contain the sub-architecture details of the KITA. Each chapter is structured for easy reader access to information and technical content necessary in decision-making. Each chapter is organized with the following sections:

- Definition – A brief statement to set perspective and need for the sub-architecture.
- General Standards – Categorical tables that list the current standards, twilight standard, emerging standards and targets that apply to that element.
    - **Twilight Standards** (or retired technologies) are those that should be discouraged from future implementation. This is not to imply that existing resources should be rooted-out and replaced, but that the use of these products and services should not be extended in future planning and development.
    - **Current Standards** are the preferred directions and products that should impact project behavior when making implementation and design decisions. Products and technologies that fit in the phase 2, or current lifecycle as outlined in Section 3.4 above are in the current standard definitions. General State directions and emphasis will center around elements of this category.
    - **Emerging Standards** (or introductory technologies) include near-term directions and options that need continued monitoring to find applicability within the State IT infrastructure. Included are technologies that are not yet fully production-worthy, but are potential candidates for future implementations. Advanced users of technology,

such as academic researchers in the State Universities, will be the first implementers of these technologies. Continued market acceptance and adaptation will move these solutions toward the growth phase of the lifecycle, and inclusion in the statewide architecture may become possible.

- **Targets** are the direction that the state as a whole (or defined subset) has determined is the best long term approach. The target may be a combination of emerging and current standards. They may also be a reflection of the anticipated natural evolution of a technology area. Targets are the hardest to establish. While not every area needs to establish targets the clearer we can be on the enterprise targets the easier it is to align agency and State of Kansas initiatives. Targets allow Kansas to move outside of the tactical architecture issues and allow Kansas to look three to ten years into the future to forecast our evolution toward a more effective set of technology solutions.

- Context and Diagrams: Where applicable, high level views are outlined that describe the domain for each of the architecture sections.
- Related Policies and Procedures: Includes a listing of related (and published) Kansas standards, guidelines, and procedures that relate to the architecture category. This information is provided primarily for reference.
- Technical Product and Configuration Information: Provides reference to specific products with notes regarding their characteristics and use within Kansas government systems.

3.5 Bringing the State together via Architecture

Reading the previous 2 sections it is easy to identify that there will be many standards at any given time and that some standards can become targets. What we hope to achieve by using a single architecture for all state agencies is to bring all state agencies together to a smaller set of overall solutions.

- Instead of 20 email systems we get to 1 or 2
- Instead of 20 ways to write and deliver reports to people we get to 1 or 2
- Instead of 20 ways to share information with our peer agencies, states, federal partners we get to 1 or 2.
- Instead of 20 ways to capture and redistribute financial information we get to 1 or 2
- Instead of 20 ways for our business partners to exchange information with us we get to 1 or 2.

We will not get to this goal overnight, it will take a lot of continual dialog to keep investment choices visible to all parties who could eventually be able to leverage those investments. So instead of one agency doing research and buying a best product that fits their needs and then asks for that product to be put into the KITA as a standard we would need to get the subject matter experts to rationalize that investment and map a strategy to deploy that investment to all agencies. and to recognize that the decision, once made, on that product and standard discussion, is closed for foreseeable future.

Decision on how you get to a target should be based upon the best information available but the following criteria should be available to the decision makers.

- Market share of the product.
- Existing presence in Kansas State agencies.
- Cost of product and alternatives.
- Integration with other KITA environments (scale of effort to deploy).
- Skills available in state agencies (our employees move between agencies and if they get trained it is better to use that training than to start over).

- Stability of company, financial, leadership.
- Company focus on public sector and government (federal, state local).

Part 2

KITA Target Summary

Chapter 4 KITA Targets

Evolution Forecast & Implementation

Targets are the direction that the State as a whole has determined is the best long term technological approach. The target may be a combination of emerging and current standards. They may also be a reflection of the anticipated natural evolution of a technology area. Targets are the hardest to establish. While not every area needs to establish targets ,the clearer we can be on the enterprise targets the easier it is to align agency and State of Kansas initiatives. Targets allow Kansas to move outside of the tactical architecture issues to look two to ten years into the future to forecast our evolution toward a more effective set of technical solutions. The target may be an initiative in itself to establish a focus group, research, pilot, program definition, multi year rollout plan, etc.

The KTARB Subcommittees were established and populated by subject matter experts in state government. The KTARB members are also all long term career IT subject experts. The collection of thoughts from these teams allows Kansas to have the discussion about appropriate technological targets.

The KITA Targets are a series of compromises that balance:

- The need for common skills versus unique skill sets for each solution
- The need to leverage purchasing power of state versus the need to protect past product investment decisions;
- The need to come together as a state versus need to support unique agency requirements;
- The need to use COTS software versus the need to build solutions to support unique agency processes and practices;
- The need to look at cross agency common services versus a unique service suite for each agency;
- The need to support and evolve legacy systems versus best of breed next generation systems;
- The need to keep IT investment cost to absolute minimum versus the need to bring consistent functionality and solutions to the state and agencies that have a high degree of interaction.

This effort will help frame our discussions about the architectural evolution and will bring multiple agencies to the table to look for the right combination of compromises. While we know it is not wise to support twenty or more email systems, we are not sure what the right number is. Is the right number 1, 2 or 5 or 20? While we know it is not wise to have 30 different content management systems, is the right number 1 or 2 or more?

Kansas Technical Reference Model (TRM) Overview

The TRM is a component-driven, technical framework that categorizes the standards and technologies to support and enable the delivery of Service Components and capabilities. It also unifies existing agency TRMs and E-Gov guidance by providing a foundation to advance the reuse and standardization of technology and Service Components from a government-wide perspective.

Aligning agency capital investments to the TRM leverages a common, standardized vocabulary, allowing interagency discovery, collaboration, and interoperability. Agencies and the state government will benefit from economies of scale by identifying and reusing the best solutions and technologies to support their business functions, mission, and target architecture. The TRM is broken down into the following sections:

This chart shows the entire TRM details and their locations. Note that we will not have Targets established for every TRM section with this version of the KITA.

Chapter 5 - Service Access and Delivery

- 5.1 Access Channels
  Web Browser, Wireless / PDA, Collaboration / Communication, Other Electronic Channels
- 5.2 Delivery Channels
  Internet, Intranet, Extranet, Peer to Peer, Virtual Private Network
- 5.3 Service Requirements
  Legislative / Compliance, Authentication / Single Sign-On, Hosting
- 5.4 Service Transport
  Supporting Network Services, Service Transport

Chapter 6 - Service Platform and Infrastructure

- 6.1 Support Platforms
  Wireless / Mobile, Platform Independent. Platform Dependant
- 6.2 Delivery Servers
  Web Servers, Media Servers, Application Servers, Portal Servers
- 6.3 Software Engineering
  Integrated Development Environment, Software Configuration Management, Test Management, Modeling
- 6.4 Databases / Storage
  Databases, Storage
- 6.5 Hardware / Infrastructure
  Servers / Computers, Embedded Technology Devices, Periherals, Wide Area Network, Local Area Network, Network Devices / Standards, Video Conferencing

Chapter 7 - Component Framework

- 7.1 Security
  Certificates / Digital Signatures. Supporting Security Services
- 7.2 Presentation / Interface
  Static Display, Dynamic / Server Side Display, Content Rendering, Wireless / Mobile / Voice
- 7.3 Business Logic
  Platform Independent, Platform Dependent
- 7.4 Data Interchange
  Data Exchange
- 7.5 Data Management
  Database Connectivity, Reporting and Analysis

Chapter 8 - Service Interface and Integration

- 8.1 Integration
  Middleware. Enterprise Application Integration
- 8.2 Interoperability
  Data Formats / Classification, Data Types / Validation, Data Transformation
- 8.3 Interface
  Service Delivery, Service Description / Interface

Please refer to the appropriate chapter listed above as you look at the chapter references of the Target details.

Chapter 5. Service access & Delivery	TRM Component group	TARGET
5.1 Access Channels		
5.1.1 Web Browser		Latest 2 Versions
5.1.2 Wireless/PDA	Tablet PCs	Convertible
	Smartphone OS	Symbian(Open) Windows Mobile (proprietary)
5.1.3 Collaboration/communication	Email (Proprietary)	MS Exchange
5.2 Delivery Channels		
5.2.2 Intranet		Web technology with access controls Portals merging internal-facing and external-facing content and functions
5.2.3 Extranet		Web technology with access controls Portals merging internal-facing and external-facing content and functions
5.3 Service Requirements		
5.3.3 Hosting	Data Center Site Selection	Non-Flood Plane Non-Flight Paths Minimal Seismic Activity Lesser Populated Areas Non-Industrialized Areas
	Physical Structure Exterior	Concrete Construction No Exterior Windows Roof Drains Ground Drains Lightening Suppression Underground Utilities Electrical Ports on exterior walls Tiers nested or linearly segregated Exterior bermed above flood plane Modular Plan for scalable facility Resistant from nearby explosions or falling debris Free standing dedicated structures Underground locations based on datacenter needs(Berme or Salt Mine) Controlled perimeters such as barriers, bollard systems, tank traps as demanded by data center purpose. Access Ports built into external walls Tornado Resistant Roof Drains external to the building
	Remaining components	Targets for these areas can be found starting at page 3-14

Chapter 6. Service Platform & Infrastructure	TRM Component group	TARGET
6.1 Supporting Platforms		
6.1.2 Platform Independent	Operating Systems	Redhat Linux Red Hat Enterprise Linux
6.1.3 Platform dependent	Desktop Operating	Windows XP

Chapter 6. Service Platform & Infrastructure	TRM Component group	TARGET
	Systems	
	Midtier Server/OS	Windows 2008 Solaris 10
	Mainframe OS	Z/OS
6.2 Delivery Servers		
6.2.1 Web Servers		Apache (open) IIS (proprietary)
6.3 Software Engineering		
6.3.1 Integrated Development Environment	Common IDE components	Source code editor Compiler or interpreter Build-automation tools Debugger
	Programming Paradigm	Support of Service Oriented Support of Object Oriented
	Requirements	XML Capable Support development of Web-based environments Interoperability with multiple environments Code transparency Support Plugins
	Security	Single Signon with controlled permissions LDAP X.500 Support
6.3.2 Software Configuration management	Version Management	Ability to recover from previous versions Ability to distinguish different versions Ability to roll back to previous version Multiple checkouts with automated synchronization Auditable Release Control
	Issue Tracking	Task Assignment Task Ownership Metrics Reporting Capabilities
	Defect Tracking	Ability to identify, assign, and manage defects Ability to track defects Ability to capture bugs Ability to report Metrics
	Task Management	Ability to define tasks Ability to prioritize
	Deployment Management	Auto deployment and tracking (push or pull)
6.3.3 Test management	Test Management	Create reproducible test scripts Create reliable, reusable, reproducible test data sets Tests tracked to make sure all of the code is tested Easily generated load testing Documented scripts
	Usability Testing	Ability to test W3C compliance Ability to test for State requirements Ability to test for Federal requirements
	Unit Testing	Built-In unit testing
	Simulation Testing	Developer Test code
6.3.4 Modeling	Modeling	Tools that support UML

Chapter 6. Service Platform & Infrastructure	TRM Component group	TARGET
		Ability to link models Interoperable with other modeling programs Ability to create reusable models
6.4 Database/Storage		
6.4.1 Database		Relational Spatial XML CORBA
6.4.2 Storage	Tape	LTO 9840 Virtual Tape
	Disks	SAS Fibre Channel SATA
	Storage Systems	SAN
	SAN Switches	Shared storage between SANs
	Fiber connections	LC
6.5 Hardware Infrastructure		
6.5.4 Wide Area Networks	OSI Layer 3	Internet Protocol V6
6.5.5 Local Area Networks	OSI Layer 3	Internet Protocol V6

Chapter 7. Component Framework	TRM Component group	TARGET
7.5 Data management		
7.5.1 Database connectivity	Proprietary Protocols	Support Open Protocols
7.5.1 Database connectivity	Open Protocols	Java Database Connectivity (JDBC) Open Database Connectivity (ODBC)
7.5.2 Reporting & Analysis	System Architecture	Service Oriented Architecture XML SOAP WSDL WSRP
	Data Structure Support	Relational Database OLAP ROLAP MOLAP Java Online Analytical Processing (JOLAP)
	Data Integration and Export Format	XML XBRL XML for Analysis

Chapter 8. Service Interface & Integration	TRM Component group	TARGET
8.2 Interoperability		
8.2.1 Data format & Classification	Text	Plain text XML (includes XSD/XSL/XHTML, etc.; with included

		or accessible schema and character encoding explicitly specified) PDF/A Open Document Format
	Raster Image	TIFF (*.tif) (uncompressed) PNG (*.png) JPEG2000 (uncompressed)
	Vector Graphics	Scalable Vector Graphics (*.svg)
	Spreadsheet	Excel (.xls) OpenOffice (*.sxc)
	Database	Access (*.mdb) OpenOffice (*.sxc)
	Presentation	PowerPoint (*.ppt) OpenOffice (*.sxi)
	Audio	AIFF(uncompressed) (*.aif, *.aiff) WAVE (LPCM only) (*.wav)
	Video	Motion JPEG2000 (*.mj2) AVI (*.avi) (uncompressed) Motion JPEG (*.avi, *.mov)
8.2.2 Data Types/Validation	XML Data Types	Primitive Derived
	Database	Exact Numerics Approximate Numerics Date and Time Character Strings Unicode Character Strings Binary Strings Other Data Types
	Validation	XML Schemas applied against appropriate standardized profiles
8.2.3 Data Transformation		XSL (XSLT, XPath, XSL FO) OpenGIS©

Kansas Service Component Reference Model Overview

The SRM is a business-driven, functional framework classifying Service Components according to how they support business and performance objectives. It serves to identify and classify horizontal and vertical Service Components supporting federal agencies and their IT investments and assets. The model aids in recommending service capabilities to support the reuse of business components and services across the federal government.

The SRM is organized across horizontal service areas, independent of the business functions, providing a leverage-able foundation for reuse of applications, application capabilities, components, and business services.

This chart show the entire SRM details and their locations. Note that we will not have Targets established for every SRM section with this version of the KITA

Chapter 9 - Customer Service

- 9.1 Customer Relationship Management
  Call Center Management, Customer Analytics, Sales and Marketing, Product Management, Brand Management, Customer / Account Management, Customer Feedback, Surveys
- 9.2 Customer Preferences
  Personalization, Subscriptions, Alerts and Notifications
- 9.3 Customer Initiated Assistance
  Online Help, Online Tutorials, Self-Service, Reservations / Registration, Multi-Lingual Support, Assistance Request, Scheduling

Chapter 10 - Process Automation

- 10.1 Tracking and Workflow
  Process Tracking, Case Management, Conflict Resolution
- 10.2 Routing and Scheduling
  Inbound Correspondence Management, Outbound Correspondence Management

Chapter 11 - Business Management Services

- 11.1 Management of Process
  Change Management, Configuration Management, Requirements Management, Program / Project Management, Quality Management, Business Rule Management, Risk Management
- 11.2 Organization Management
  Workgroup / Groupware, Network Management
- 11.3 Investment Management
  Strategic Planning and Management, Portfolio Management, Performance Management

Chapter 12 - Digital Asset Management

- 12.1 Content Management
  Content Authoring, Content Review and Approval, Tagging and Aggregation, Content Publishing and Delivery, Syndication Management
- 12.2 Document Management
  Document Imaging and OCR, Document Referencing, Document Revisions, Library / Storage, document Review and Approval, Document Conversion, Indexing, Classification
- 12.3 Knowledge Management
  Information Retrieval, Information Mapping / Taxonomy, Information Sharing, Categorization, Knowledge Engineering, Knowledge Capture, Knowledge Distribution and Delivery, Smart Documents
- 12.4 Records Management
  Records Linking / Association, Document Classification, Document Retirement, Digital Rights Management

Chapter 13 - Business Analytical Services

- 13.1 Analysis and Statistics
  Mathematical, Structural / Thermal, Radiological, Forensics
- 13.2 Visualization
  Graphing / Charting, Imagery, Multimedia, Mapping / Geospatial / Elevation / GPS, CAD
- 13.3 Knowledge Discovery
  Data Mining, Modeling, Simulation
- 13.4 Business Intelligence
  Demand Forecasting / Management, Balanced Scorecard, Decision Support and Planning
- 13.5 Reporting
  Ad Hoc, Standardized / Canned, OLAP

Chapter 14 - Back Office Support

- 14.1 Data Management
  Data Exchange, Data Mart, Data Warehouse, Meta Data Management, Data Cleansing, Extraction and Transformation, Loading and Archiving, Data Recovery, Data Classification
- 14.2 Assets / Materials Management
  Computes / Automation Management
- 14.3 Development and Integration
  Legacy Integration, Enterprise Application Integration, Data Integration, Instrumentation and Testing, Software Development

Chapter 15 - Support Services

- 15.1 Security Management
  Identification and Authentication, Access Control, Encryption, Intrusion Detection, Verification, Digital Signatures, User Management, Role / Privilege Management, Audit Trail and Capture, Forensics, Incident Response, Risk Management
- 15.2 Collaboration
  Email, Threaded Discussions, Document Library, Shared Calendaring, Task Management
- 15.3 Search
  Query, Precision / Recall Ranking, classification, Pattern Matching
- 15.4 Communication
  Real Time Chat / Instant Messaging, Audio Conferencing, Video Conferencing, Event / News Management, Community Management, computer Telephony Integration, Voice Communications
- 15.5 Systems Management
  License Management, Remote Systems Control, System Resource Monitoring, Software Distribution, Issue Tracking
- 15.6 Forms Management
  Forms Creation, Forms Modification

Please refer to the appropriate chapter listed above as you look at the Chapter references of the Target details.

Chapter 9. Customer Services	**SRM Component group**	**TARGET**
9.1 Customer Relationship Management		
9.1.1 Call Center Management	Service Desk and Incident Management and Resolution	Integrated automated hardware, application, and network management tools to proactively alert service desk staff to potential problems

Chapter 10. Process Automation	**SRM Component group**	**TARGET**
10.1 Tracking and Workflow		
10.1.1 Process Tracking	Access	Supports both Internal and External Access

Chapter 10. Process Automation	SRM Component group	TARGET
	Security	Usage of electronic / digital signatures or certificates Auditing capabilities
	Processing	Parallel processing
	Workflow	Supports functions such as: • Process definition and mapping • Routing • Event triggers based on time, date, process completion, etc. • Group assignments • Individual assignments • Status reporting Can be tailored to domain-specific processes such as telecommunications management, service desk event management, document processing, personnel hiring, interlibrary loan requests, etc.
10.1.2 Case Management	Case Management	Data and communications protocol interoperability for inter-agency communications within the State, within the region, or nationally (i.e. communication with federal agencies)

10.1.3 Conflict Resolution	Business Rules	Functions include:
		• Rules composition
		• Rules deployment
		• Run-time engine
		• Component process and terminology definition
		  o Matching
		  o Conflict resolution
		  o Action
		• Rule set tracking
		• Policy testing tools
		• Authorization
		Rule Interchange Format

10.2 Routing and Scheduling		
10.2.1 Inbound Correspondence Management	Correspondence Management	Functions include: • Ability to capture and manage both electronic and paper inbound communications • Security for confidential correspondence • Workflow scheduling • Classification / categorization • Rules • Metadata creation and management • Search • Tracking • Data mining • Analysis • Response preparation • Disposition • Compliance management

Chapter 10. Process Automation	SRM Component group	TARGET
10.2.2 Outbound Correspondence Management	Correspondence Management	Functions include: • Ability to capture and manage both electronic and paper outbound communications • Security for confidential correspondence • Workflow scheduling • Classification / categorization • Rules • Metadata creation and management • Search • Tracking • Data mining • Analysis • Response preparation • One-to-one distribution • One-to-many distribution • Disposition • Compliance management

Chapter 11. Business Management Services	SRM Component group	TARGET
11.1 Management of Process		
11.1.4 Program / Project Management	Program / Project Management	State of Kansas PMM standards
11.1.7 Business Rule Management	Business Rule Management	Security Auditing Rules Engine standards such as JSR-94 Java Rule Engine API
	Risk Management	Kansas PMM standards Security auditing
11.2 Organizational Management		
11.2.1 Workgroup / Groupware	Portals	Shared document capability Channel support Role-based Contextual Seamless Guided Multimedia support Multi-platform access support (desktop, laptop, mobile device, etc.) Interoperability with multiple browsers
	E-mail	Shared Document capability Interoperability with multiple platforms, operating systems, and/or browsers
	Meeting / Conferencing	Shared Document capability Interoperability with multiple platforms, Operating systems, and/or browsers
	Instant Messaging	Group chat Individual chat Interoperability with multiple platforms, operating

Chapter 11. Business Management Services	SRM Component group	TARGET
		systems
11.3 Investment Management		
11.3.1 Strategic Planning and Management	Financial Modeling and Reporting	Interoperability between disparate Financial systems (XML based) Bloomberg Terminals
11.3.2 Portfolio Management	Financial Portfolio Mgmt	GAAP Compliant
	IT Project Portfolio Mgmt	Components should include: Portfolio and Investment Planning Application Portfolio Mgmt Project Portfolio Mgmt Project and Budget Mgmt Demand Management Resource Management Time Tracking Client Management Performance Management Knowledge Management Financial Management Workflow Reporting
11.3.3 Performance Management	Investment Tracking and Mgmt	GAAP Compliance

Chapter 12. Digital Asset Services	SRM Component group	TARGET
12.1 Content Management		
12.1.1 Content Authoring	Markup Languages	XML
12.1.2 Content Review an Approval	Version Control and Workflow	Embedded in Content / Document Management System
12.1.3 Tagging and Aggregation	Descriptive Metadata	Dublin Core Domain Specific Schemas / Formats Domain Specific Profiles`
	Semantic Markup	SGML XML Domain Specific Markup Profiles
12.1.4 Content Publishing and Delivery	XML Query Language	XPATH
	Presentation	CSS XSL
	XML Transformation	XSLT
	Content Syndication	RSS Services Atom Syndication Format
12.2 Document Management		
12.2.1 Document Imaging and OCR	Document Imaging	ANSI/AIIM MS53-1993 AIIM/TR 19-1993 AIIM/TR 29-1993

Chapter 12. Digital Asset Services	SRM Component group	TARGET
	OCR	ISO 1073/1-1976 ISO 1073/2-1976 FIPS PUB 32-1 ANSI X3.93M-1981 (R1989) ANSI X3.99-1983 9 (R1991)
12.2.2 Document Referencing	Item Identifiers	Archival Resource Key (ARK) Document Object Identifier (DOI) Handles Namespace Open URL Persistent URL
12.2.3 Document Revisions	Verson Control and Workflow	Embedded in Content / Document Management System
12.2.4 Library / Storage	Information Lifecycle Management	Long Term Archive and Compliance Storage Initiative (LTACSI) Standards
12.2.5 Document Review and Approval	Version Control and Workflow	Embedded in Content / Document Management System
12.2.6 Document Conversion	Document Conversion	Avoid Proprietary Formats and Convert to Open Formats

Chapter 13. Business Analytical Services	SRM Component group	TARGET
13.2 Visualization		
13.2.1 Graphing / Charting	Graphing / Charting	Standard Database Linkage (JDBC, ODBC) Integrated drawing tools Multiple graphing / charting types supported (i.e. bar, radar, etc.)
13.2.2 Imagery	Digital Scanning	Scan to PDF Scan to multiple file formats including TIFF, JPEG, JPG2000, ect.
	Microform (microfilm, microfiche)	NISO and ANSI standards including ANSI/NISO Z39.32 – 1996 (R2002) ANSI/NISO Z39.62 – 2000 ANSI/NISO Z39.74 – 1996 (R2002)
	Photographs	Multiple ANSI standards covering chemicals, film, processing, ect.
	Digital photographs	Multiple ANSI standards covering pixel-reporting, color-encoding, format profiles, ect.
13.2.4 Mapping / Geospatial / Elevation / GPS	Mapping / Geospatial / Elevation / GPS	Targets are available for this section starting at 4-41
13.4 Business Intelligence		
13.4.3 Decision Support and Planning		Visualization Dashboards
13.5 Reporting		
13.5.1 Ad hoc	Textual and Numeric	Interoperability with multiple data sources SQL query support XML support Ability to script

Chapter 13. Business Analytical Services	SRM Component group	TARGET
		Standard Database Connections (ODBC, JDBC) Web interface using forms Open, Standardized Metadata
13.5.2 Standardized / Canned	Textual and Numeric	Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms Open, Standardized Metadata
13.5.3 OLAP	Textual and Numeric	Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms Open, Standardized Metadata

Chapter 14. Back Office Support	SRM Component group	TARGET
14.1 Data Management		
14.1.3 Data Warehouse	All components	Same as current std
14.1.4 Meta Data Warehouse	All components	Same as current std
14.1.6 Extraction and Transformation	Extraction and Transformation	Integrate with XML. XML Data Interchange (XMI) Integrate with Message Brokers. Standards based metadata Leverage existing code as well as bring new functionality. Allow for the incorporation of existing working scripts and offer a complete tool set of ready to use transformations and functions as well as a comprehensive list of documented APIs and methods. Improve metadata management and administration as well as ensure data quality. Able to do upfront analysis and modeling and streamline the workflow process.
14.1.7 Loading and Archiving	Methodology	Customized applications and scripts Database import and export tools SQL ODBC, JDBC, OLE XML Automated capture and apply middleware Stored procedures
	Data Loading	Portable Support for o Oracle o Microsoft SQL Server o Sybase o IBM DB2

Chapter 14. Back Office Support	SRM Component group	TARGET
		o flat files o any JDBC-accessible data source.
	Application	Logging for: • Quality assurance • Audit trail • Error tracing and recovery
	Technology	Loading solutions built on an open architecture Re-usable custom extensions Open-standard Java-based server engine Platform independence –Minimally must support ODBC and OLE-DB connections. Database using single occurrence of DASD
14.1.8 Data Recovery	Device-based	End-to-end redundancy High fault tolerance SATA FireWire (IEEE-1394)
	Software-based	Automated recovery Point-in-time recovery System-managed storage
	Network-based	NDMP Version 5
14.1.9 Data Classification	General	Criteria: o Confidentiality o Integrity o Trustworthiness o Availability
	Information Lifecycle Management	Criteria based on access or availability and recovery requirements, and cost.
	Security	Same as current std
	Subject / Topic	Controlled Taxonomy within organizational units; Controlled Vocabulary Concept Mapping

Chapter 15. Support Services	SRM Component group	TARGET
15.1 Security Management		
15.1.13 Addition Security Areas	Network Security	IPSEC v6
15.2 Collaboration		
15.2.1 E-Mail	Server Side	Cross-platform support Supports IMAP and / or POP protocols Rules support GUI administration interface
	Client Side	*Current + Emerging*
15.2.3 Document Library	Repository	OAIS standard support for preservation repositories
15.2.4 Shared Calendaring	Access	Supports: • Group Access • Multiple User Accounts • Conference Rooms CalDAV
15.3 Search		

Chapter 15. Support Services	SRM Component group	TARGET
15.3.1 Query	Query	Intelligent Content Services Keyword with Boolean and proximity operators Fielded Search Phrase Search Fuzzy Match
	Query Language	SQL XPointer XPath OQL XQuery
15.3.2 Precision / Recall Ranking	Precision/ Recall Ranking	Boolean operators within query to better define search criteria Natural Language query using Linguistic Analysis Bayesian Probabilistic Model
15.3.3 Classification	Classification schemes	Dewey Decimal System (DDC) Universal Decimal Classification (UDC) Library of Congress Classification System (LCC) Subject-specific schemes such as: o NLM (National Library of Medicine) o Engineering Information (Ei) Classification Codes o Mathematics Subject Classification ACM Computing Classification System (CCS)
	Faceted Classification	Subject Domain specific standards
15.3.4 Pattern Matching	Sequence Patterns	Same as current std
	Tree Patterns	Same as current std
15.4. Communication		
15.4.4 Event / News Management	Management protocols	SNMP V2
15.5 Systems Management		
15.5.4 Software Distribution	Software and OS deployment	Elevated credential distribution Rollback and version control
15.5.5 Issue Tracking	Software Tools for IT Service Management	Integrated tools – platform independent

Part 3

Kansas Technical Reference Model (TRM)

Kansas Technical Reference Model (TRM) Overview

The TRM is a component-driven, technical framework that categorizes the standards and technologies to support and enable the delivery of Service Components and capabilities. It also unifies existing agency TRMs and E-Gov guidance by providing a foundation to advance the reuse and standardization of technology and Service Components from a government-wide perspective.

Aligning agency capital investments to the TRM leverages a common, standardized vocabulary, allowing interagency discovery, collaboration, and interoperability. Agencies and the state government will benefit from economies of scale by identifying and reusing the best solutions and technologies to support their business functions, mission, and target architecture. The TRM is broken down into the following sections:

Chapter 5 - Service Access and Delivery

- 5.1 Access Channels
  Web Browser, Wireless / PDA, Collaboration / Communication, Other Electronic Channels
- 5.2 Delivery Channels
  Internet, Intranet, Extranet, Peer to Peer, Virtual Private Network
- 5.3 Service Requirements
  Legislative / Compliance, Authentication / Single Sign-On, Hosting
- 5.4 Service Transport
  Supporting Network Services, Service Transport

Chapter 6 - Service Platform and Infrastructure

- 6.1 Support Platforms
  Wireless / Mobile, Platform Independent. Platform Dependant
- 6.2 Delivery Servers
  Web Servers, Media Servers, Application Servers, Portal Servers
- 6.3 Software Engineering
  Integrated Development Environment, Software Configuration Management, Test Management, Modeling
- 6.4 Databases / Storage
  Databases, Storage
- 6.5 Hardware / Infrastructure
  Servers / Computers, Embedded Technology Devices, Periherals, Wide Area Network, Local Area Network, Network Devices / Standards, Video Conferencing

Chapter 7 - Component Framework

- 7.1 Security
  Certificates / Digital Signatures. Supporting Security Services
- 7.2 Presentation / Interface
  Static Display, Dynamic / Server Side Display, Content Rendering, Wireless / Mobile / Voice
- 7.3 Business Logic
  Platform Independent, Platform Dependent
- 7.4 Data Interchange
  Data Exchange
- 7.5 Data Management
  Database Connectivity, Reporting and Analysis

Chapter 8 - Service Interface and Integration

- 8.1 Integration
  Middleware. Enterprise Application Integration
- 8.2 Interoperability
  Data Formats / Classification, Data Types / Validation, Data Transformation
- 8.3 Interface
  Service Delivery, Service Description / Interface

Chapter 5 Service Access and Delivery

The Service Access and Delivery Service Area defines the collection of Access and Delivery Channels that will be used to leverage the Service Component, and the legislative requirements that govern its use and interaction.

5.1 Access Channels

Access Channels define the interface between an application and its users, whether it is a browser, personal digital assistant or other medium.

5.1.1 Web Browser

Define the program that serves as your front end to the World Wide Web on the Internet. In order to view a site, you type its address (URL) into the browser's location field.

A **web browser** is a software application that enables a user to display and interact with text, images, and other information typically located on a web page at a website on the World Wide Web or a local area network. Text and images on a web page can contain hyperlinks to other web pages at the same or different websites. Web browsers allow a user to quickly and easily access information provided on many web pages at many websites by traversing these links.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Web Browser		W3C standards compliance Support for fundamental features including: • Bookmark manager • Content caching • Plug-in support for media-types Support for “annoyance removal” features including: • Pop-up blocker • Advertisement filtering • Controllable Downloads	Mobile Web Best Practices (W3C) Voice Browser Semantic Web support Support for W3C technology stack	Current + Emerging Standards

		Supports accessibility for persons with disabilities: • W3C User Agent Accessibility Guidelines (UAAG) (http://www.w3.org/WAI/intro/uaag.php) Security Controls SSL support Active X shutoff Automatic updates for security fixes and new releases Multiple language support for display		

Definitions of web browser technology include:

- **Web Browser Extensions** – Add-on software applications that give the web browser additional functionality. These are available for most browsers. Extensions produced by the browser developer are acceptable for use. Extensions developed by user communities should be used with some caution.

Examples of web browser technology include:

- Mozilla Firefox
- Microsoft Internet Explorer
- Safari
- Opera
- Netscape

Other Important information for Web Browser

- Web browsers should be updated regularly. The latest version of the browser is preferred but one version outside of the current version is also acceptable.

- W3C Consortium, http://www.w3.org (accessed 6/29/2006)

- "Web Browser Standards Support: Summary" http://www.webdevout.net/browser_support_summary.php (accessed 6/29/2006)

- Wikipedia, "Comparison of Web Browsers", http://en.wikipedia.org/wiki/Comparison_of_web_browsers. (accessed 6/29/2006)

- W3C Mobile Web Initiative, http://www.w3.org/Mobile/ (accessed 6/29/2006)

- W3C, Voice Browser Activity, http://www.w3.org/Voice/, (accessed 6/29/2006)

- Some of the more popular browsers include additional components. Protocols supported may include NNTP (network news transfer protocol), SMTP (simple mail transfer protocol), IMAP (Internet message access protocol), and POP (post office protocol). These browsers are often referred to as ***Internet suites*** or ***application suites*** rather than merely web browsers. An **Internet suite** is an Internet-related software suite that usually includes:
  - Web browser
  - Email client, news client, address book
  - Download manager
  - HTML editor
  - IRC client

  http://en.wikipedia.org/wiki/Web_browser and http://en.wikipedia.org/wiki/Internet_suite (accessed 6/29/2006)

5.1.2 Wireless / PDA

Define the technologies that use transmission via the airwaves. Personal Digital Assistant (PDA) is a handheld computer that serves as an organizer for personal information. It generally includes, at a minimum, a name and address database, to-do list and note taker.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
PDA				
Device		Blackberry		
OS		Windows CE Windows Mobile		
IPOD				
Device		IPOD		
OS		Windows CE		
Tablet PCs				
Device		Tablet PCs Slate Convertible		Convertible
OS		Windows XP Tablet		
Smartphones				
Device		Nokia J2ME Compliant		
OS		Symbian Palm OS Windows Mobile (Windows CE) BREW Linux based handheld		Symbian (open) Windows Mobile (proprietary)

Definitions of wireless / PDA technology include

- **Palm Operating System** – Palm is the leading Personal Digital Assistant (PDA). Version 5 of Palm OS provides multitasking and other capabilities that will provide an improved platform for E-Gov solutions.

- **Blackberry** – The leading email-enabled wireless device with wide use in several agencies.

- **Pocket PC Phone Edition** – Microsoft's environment for Internet-capable cellular phones.

- **Pocket PC 2000** – Microsoft's environment for PDA level devices.

- **Symbian Epoc** – A leading environment for web capable cellular phones.

- **Smartphones** - A smartphone is any electronic handheld device that integrates the functionality of a mobile phone, personal digital assistant (PDA) or other information appliance. This is often achieved by adding telephone functions to an existing PDA or putting "smart" capabilities, such as PDA functions, into a mobile phone. A key feature of a smartphone is that additional applications can be installed on the device. The applications can be developed by the manufacturer of the handheld device, by the operator or by any other third-party software developer.

5.1.3 Collaboration / Communication

Define the forms of electronic exchange of messages, documents, or other information. Electronic communication provides efficiency through expedited time of delivery.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Email (Proprietary)	Novell Groupwise Lotus Notes	MS Exchange		MS Exchange
Email (Open)			Open Source Email platforms	
Email (Hosted)			Hosted Email platfroms	
Workflow		Lotus Notes Silanis Captiva K2		
Document Management		ImageNow DocUShare FileNet Documentum Microsoft SharePoint		
Fax Server		RightFax™		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Kiosk				
Collaboration Environment		Microsoft SharePoint		

Definitions of Collaboration / Communication technology include

- **Electronic Mail (Email)** – Email (Electronic mail) is the exchange of computer generated and stored messages by telecommunication. An email can be created manually via messaging applications or dynamically/programmatically such as automated response systems.

- **Workflow** at its simplest is the movement of documents and/or tasks through a work process. More specifically, workflow is the operational aspect of a work procedure: how tasks are structured, who performs them, what their relative order is, how they are synchronized, how information flows to support the tasks and how tasks are being tracked. As the dimension of time is considered in Workflow, Workflow considers "throughput" as a distinct measure. Workflow problems can be modeled and analyzed using graph-based formalisms like Petri nets. While the concept of workflow is not specific to information technology, support for workflow is an integral part of document management and imaging software.

- **Facsimile (Fax)** – A fax is the digitized image of text and/or pictures, represented as a series of dots (bit map). Faxes are sent and received through telecommunication channels such as telephone or Internet.

- **Fax Server** - A fax server is a set of software running on a server computer which is equipped with one or more fax-capable modems attached to telephone lines. Its function is to accept documents from users, convert them into faxes, and transmit them, as well as to receive fax calls and either store the incoming documents or pass them on to users. Users may communicate with the server in several ways, through either a local network or the internet. In a big organization with heavy fax traffic, the computer hosting the fax server may be dedicated to that function, it which case the computer itself may also be known as a fax server.

  For outgoing faxes, several methods are available to the user:
  - An e-mail message (with optional attachments) can be sent to a special e-mail address; the fax server monitoring that address converts all such messages into fax format and transmits them.
  - The user can tell his computer to "print" a document using a "virtual printer" which, instead of producing a paper printout, sends the document to the fax server, which then transmits it.
  - A web interface can be used, allowing files to be uploaded, and transmitted to the fax server for faxing.
  - Special client software may be used.

  For incoming faxes, several user interfaces may be available:
  - The user may be sent an e-mail message for each fax received, with the pages included as attachments, typically in either TIFF or PDF format.
  - Incoming faxes may be stored in a dedicated file directory, which the user can monitor.
  - A website may allow users to login and check for received faxes.
  - Special client software may be used.

- **Kiosk** – A kiosk is a small physical structure (often including a computer and a display screen) that displays information for people walking by. Kiosks are common in public buildings. Kiosks are also used at trade shows and professional conferences.

5.1.4 Other Electronic Channels

Define the other various mediums of information exchange and interface between a user and an application.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Web Services		W3C standards including: • SOAP 1.2 or higher (Simple Object Access Protocol) • WSDL (Web Services Descriptive Language)		
Web Naming and Addressing		Uniform Resource Name (URN) Uniform Resource Identifier (URI) Uniform Resource Locator (URL)		

Definitions of other electronic technology include:

- **System to System** – System to System involves at least two computers that exchange data or interact with each other independent of human intervention or participation.

- **Web Service** – Web services (sometimes called application services) are services (usually including some combination of programming and data, but possibly including human resources as well) that are made available from a business's web server for web users or other web-connected programs.

- **Web Services** - Web services provide a standard means of interoperating between different software applications, running on a variety of platforms and/or frameworks. Web services are characterized by their great interoperability and extensibility, as well as their machine-processable descriptions thanks to the use of XML. They can be combined in a loosely coupled way in order to achieve complex operations. Programs providing simple services can interact with each other in order to deliver sophisticated added-value services.
  http://www.w3.org/2002/ws/Activity (accessed 6/29/2006)

- **Uniform Resource Locator (URL)** – URL is the global address of documents and other resources on the World Wide Web. The first part of the address indicates what protocol to use (i.e. “http://”), and the second part specifies the IP address or the domain name where the resource is located (i.e. “www.firstgov.gov”).

5.2 Delivery Channels

Delivery Channels define the level of access to applications and systems based upon the type of network used to deliver them.

5.2.1 Internet

The Internet is a worldwide system of computer networks in which users at any one computer can, if they have permission, get information from any other computer.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Internet		Internet I – commercial commodity internet Internet 2 – academic research internet IETF standards IANA registry ICANN registry		

Definitions of Internet technology include:

- The **Internet** is the worldwide, publicly accessible system of interconnected computer networks that transmit data by packet switching using the standard Internet Protocol (IP). It consists of millions of smaller domestic, academic, business, and government networks, which together carry various information and services, such as electronic mail, online chat, peer to peer networks like Bittorrent, and the interlinked Web pages and other documents of the World Wide Web.
- HTTP - **Hypertext Transfer Protocol** (**HTTP**) is the method used to transfer or convey information on the World Wide Web. It is a patented open internet protocol whose original purpose was to provide a way to publish and receive HTML pages.
- IP – a data-oriented protocol used for communicating data across a packet-switched internetwork
- TCP - **Transmission Control Protocol** (**TCP**) is one of the core protocols of the Internet protocol suite. Using TCP, applications on networked hosts can create *connections* to one another, over which they can exchange data or packets.

Other Important information for Internet

- Internet Engineering Task Force (IETF) http://www.ietf.org
- Internet Assigned Numbers Authority (IANA) http://www.iana.org
- Internet Corporation for Assigned Names and Numbers (ICANN) http://www.icann.org

5.2.2 Intranet

An Intranet is a private network that is contained within an enterprise. It may consist of many interlinked local area networks and is used to share company information and resources among employees.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Intranet		LAN technology Web technology with access controls Portals merging internal-facing and external-facing content and functions		Web technology with access controls Portals merging internal-facing and external-facing content and functions

Examples of Intranet / portal technology include:

- Microsoft Sharepoint
- uPortal
- Xythos

Other Important information for Intranet

"Five years ago, firms treated intranet and extranet sites differently because overlaps in user needs were few, and technology did not meet both needs neatly. Today, the user-controlled ease of intranets and the application integration of extranets have found a home in a single family of products: portal servers. Along with solving existing intranet and extranet requirements, modern portal servers provide advanced functionality — like user-driven customization, collaboration, and process automation — that other legacy technologies leave out. Firms with multiple, separate intranet and extranet sites that serve similar needs in distinct audiences should merge those efforts together and create enterprise portals that replace their aging infrastructure. […] Depending on business priorities, firms should take advantage of this technology's evolution and build enterprise portals that cross internal and external organizational boundaries, especially when users must share apps and information securely. Extended enterprise portals give firms:

- A single platform
- More security control
- Unified development standards
- Consistent, up-to-date information
- New features like customization, personalization, search, and collaboration
- Capabilities for composite applications and business intelligence."

Root, Nate L., Your Intranet And Extranet Are Obsolete: Upgrade To Portal Technology To Deliver Extra Value, Forrester Research, August 4, 2004
http://www.forrester.com/Research/Document/0,7211,35102,00.html

5.2.3 Extranet

An Extranet is a private network that uses the Internet protocol and the public telecommunication system to securely share part of a business's information or operations with suppliers, vendors, partners, customers, or other businesses. An extranet can be viewed as part of a company's intranet that is extended to users outside the company.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Extranet		LAN technology Web technology with access controls Portals merging internal-facing and external-facing content and functions		Web technology with access controls Portals merging internal-facing and external-facing content and functions

Other Important information for extranet

"Five years ago, firms treated intranet and extranet sites differently because overlaps in user needs were few, and technology did not meet both needs neatly. Today, the user-controlled ease of intranets and the application integration of extranets have found a home in a single family of products: portal servers. Along with solving existing intranet and extranet requirements, modern portal servers provide advanced functionality — like user-driven customization, collaboration, and process automation — that other legacy technologies leave out. Firms with multiple, separate intranet and extranet sites that serve similar needs in distinct audiences should merge those efforts together and create enterprise portals that replace their aging infrastructure. […] Depending on business priorities, firms should take advantage of this technology's evolution and build enterprise portals that cross internal and external organizational boundaries, especially when users must share apps and information securely. Extended enterprise portals give firms:

- A single platform
- More security control
- Unified development standards
- Consistent, up-to-date information
- New features like customization, personalization, search, and collaboration
- Capabilities for composite applications and business intelligence."

Root, Nate L., Your Intranet And Extranet Are Obsolete: Upgrade To Portal Technology To Deliver Extra Value, Forrester Research, August 4, 2004
http://www.forrester.com/Research/Document/0,7211,35102,00.html

5.2.4 Peer to Peer

Peer to Peer is a class of applications that operate outside the DNS system, have significant or total autonomy from central servers, and take advantage of resources available on the Internet.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Peer-to-Peer				
Distributed Computing				
Grid Computing		Grid architecture standards maintained by the Open Grid Forum (merged organization of the Enterprise Grid Alliance and the Global Grid Forum)		

Definitions of peer to peer technology include:

- A **peer-to-peer** (or **P2P**) computer network is a network that relies primarily on the computing power and bandwidth of the participants in the network rather than concentrating it in a relatively low number of servers. P2P networks are typically used for connecting nodes via largely *ad hoc* connections. Such networks are useful for many purposes. Sharing content files containing audio, video, data or anything in digital format is very common, and realtime data, such as telephony traffic, is also passed using P2P technology. A pure peer-to-peer network does not have the notion of clients or servers, but only equal *peer* nodes that simultaneously function as both "clients" and "servers" to the other nodes on the network. http://en.wikipedia.org/wiki/Peer-to-peer (accessed 6/29/2006)

- **Distributed computing** is decentralized and parallel computing, using two or more computers communicating over a network to accomplish a common objective or task. The types of hardware, programming languages, operating systems and other resources may vary drastically. It is similar to computer clustering with the main difference being a wide geographic dispersion of the resources. http://en.wikipedia.org/wiki/Distributed_computing (accessed 6/29/2006)

- **Grid computing** is an emerging computing model that provides the ability to perform higher throughput computing by taking advantage of many networked computers to model a virtual computer architecture that is able to distribute process execution across a parallel infrastructure. Grids use the resources of many separate computers connected by a network (usually the Internet) to solve large-scale computation problems. Grids provide the ability to perform computations on large data sets, by breaking them down into many smaller ones, or provide the ability to perform many more computations at once than would be possible on a single computer, by modeling a parallel division of labor between processes. Today resource allocation in a grid is done in accordance with SLAs (service level agreements). Grid computing reflects a conceptual framework rather than a physical resource. The Grid approach is utilized to provision a computational task with administratively-distant resources. The focus of Grid technology is associated with the issues and requirements of flexible computational provisioning beyond the local (home) administrative domain. http://en.wikipedia.org/wiki/Grid_computing (accessed 6/29/2006)

5.2.5 Virtual Private Network

A private data network that makes use of the public telecommunication infrastructure, maintaining privacy through the use of a tunneling protocol and security procedures.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Protocol	L2F,CHAP	IPSEC PPTP L2TP PPP IKE SSL	TLS	
Encryption	DES	3DES AES-128	AES-256	
Authentication Hashing	MD5	SHA-1	SHA-2	

Description	Example Products	Notes
IPSEC	Check Point, PIX, MS ISA	Site To Site and Client
SSL	Secure Access, FirePass, Access Gateway, WebVPN, OpenVPN	Client/Transaction

Definitions of Virtual Private Network technologies include:

- IP Security (IPSEC) – A set of protocols used to secure IP packet exchange. Tunnel and Transport are the two (2) modes supported by IPSEC. IPSEC uses certificates and Public Keys to authenticate and validate the sender and receiver.
- Point-to-Point Tunneling Protocol (PPTP) – a technology for creating Virtual Private Networks (VPNs) , developed jointly by Microsoft Corporation, U.S. Robotics, and several remote access vendor companies, known collectively as the PPTP Forum.

5.3 Service Requirements

Service Requirements define the necessary aspects of an application, system or service to include legislative, performance, and hosting.

5.3.1 Legislative / Compliance

Defines the prerequisites that an application, system or service must have mandated by congress or governing bodies.

Examples of Legislative / Compliance technology include:

- Section 508 – Federal Section 508 of the Rehabilitation Act of 1973, as amended (29 U.S.C. § 794d), applied to states by Title II of the Americans with Disabilities Act (28 CFR Part 35), requires that agencies' electronic and information technology is accessible to people with disabilities, including employees and members of the public.
- Web Content Accessibility – Refers to the ability of hardware and software that helps people who are physically or visually impaired to usefully interface with web content due to specific web content implementation characteristics, and to standards defining these characteristics. The Kansas standards for web content accessibility are established in IT Policy 1210: State of Kansas Web Accessibility Requirements.

5.3.2 Authentication / Single Sign-On

Refers a method that provides users with the ability to login one time, getting authenticated access to all their applications and resources.

5.3.3 Hosting

Refers to the service provider who manages and provides availability to a web site or application, often bound to a Service Level Agreement (SLA). The Hosting entity generally maintains a server farm with network support, power backup, fault tolerance, load balancing, and storage backup.

Data Center

A physical location used to house data processing equipment that requires a climate controlled and secure environment to operate. A data center can be used to serve an independent organization or can serve as a hosting center maintained by a service provider.

The following requirements represent key components that should be included when designing and constructing a Data Center or Hosting Center.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Site Selection	Office buildings Urban locations	Non-flood plane Non-flight paths Minimal Seismic Activity	Less populated areas provide benefit over dense populated areas Non-industrialized areas (especially due to toxicity)	Current + Emerging Standards
Physical Structure Exterior	Wooden const	Concrete const No Exterior Windows Roof Drains Ground Drains	Tiers nested or linearly segregated (lowest to highest reliability/security) Exterior bermed above flood plane	Current Standards + the following Resistant from nearby explosions or falling debris

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		Lightning suppression Underground Utilities Electrical Ports on exterior walls for access from outside electricity sources	for security & weather shelter Modular plan for scalable facility	Free standing dedicated structures Underground locations based on datacenter needs (Berme or Salt Mine) Controlled perimeters such as barriers, bollard systems, tank traps as demanded by data center purpose. Access Ports built into external walls Tornado Resistant Roof Drains external to the building
Physical Structure Interior	False Ceilings Floor tile subject to zinc whiskers	Sump drains Cable chases under raised floor Cable chases within/below ceiling void Open ceiling void or Data Center approved ceiling tiles as required Rust resistant post/beams to support raise floor Decontamination and encapsulation of sub floor Staging areas to unpack and prepare equipment	DC space free of non-IT equipment (no CRAC units, PDU's, etc.) Mechanical – electrical areas floor below or room beside DC Lower tier modules less costly / reliable / secure. Penetrations for direct liquid immersion cooling planned Penetrations for flexible air source& exhaust planned	Current + Emerging standards

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
			Penetrations for electrical outlets & expansion planned	
Mechanical Systems – A/C within DC space		Temp b/t 68-70 degrees RH b/t 48-52% Man Traps (Air Locks) Contaminant Controls 18" -36" raised floor High efficiency air filter system Humidity, temp monitoring and control.(with console alerts if possible) Dual loop cooling grid	Overhead supplemental cooling to address heat sensitive areas(waterless) Variable capacity floor-mount cooling Rack centric cooling modules Embedded and on-chip cooling Quick connect locations for cooling cabinets Upper tiers more reliable / secure / costly Flexible cooling air supply & return Direct liquid immersion cooling capable	Current + Emerging standards 24" -36" raised floor should replace 18" minimum
Mechanical Systems – A/C in public space		Pressurization less than adjacent data center	High efficiency air filter system Humidity, temp monitoring and control. (when equipment dictates) BACnet controls Individual climate control per cubical (applies to raised floor office environments)	Current + Emerging standards

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Mechanical Systems – Plumbing		Separate plumbing for DC and public areas Avoid overhead water sources Monitors for floor drains and sumps	Smart Monitors for floor drains and sumps (interfaced with multi-purpose building monitoring equipment) BACnet standard integrated building controls for HVAC	Current + Emerging standards
Mechanical Systems – Fire Suppression		Chemical Primary in DC (Dry power or Gas) Dry sprinkler backup in DC Wet sprinkler in public areas Monitoring Systems	In cabinet sensors. Gas Suppression Systems BACnet integrated building controls for fire	Current + Emerging standards
Electrical		Redundant Generators,UPS,and PDU's where possible Emergancy lighting in DC and access hallways Monitoring Systems Twistlock power cords under floor Isolated power feeds from public space	Dual Power Sources A/B power grid in machine space. DC power Modular Power Monitors for power providers Cabinet Monitors Load Banks or load testing agreements Direct Current only in Data Center with all backup power and conditioning in mechanical space	Current + Emerging standards
Security – Office Space		Key locks Card readers Access policy Video Monitoring	Biometric identification Integrated security systems BACnet integrated	Current + Emerging standards

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		Patrolled Security	building controls for security	
Security – Machine Space		Card key access Security clearances Access Policy Video monitoring Security Cabinets Security fencing	Biometric identification Integrated security systems Redundant BACnet integrated building controls for security	Current + Emerging standards
Security – Mechanical Space		Card key access to mechanical rooms, punch down locks, physical key locks Security policy Fenced security for mechanical yard Video Monitoring on interior and exterior space	Biometric identification Integrated security systems Redundant BACnet integrated building controls for security	Current + Emerging standards
Security – Building Exterior		Secured access to roof, air intake systems, and outside doors Security policy Video Monitoring of critical areas	Truck barriers and/or secured access points at a reasonable safe distance away from facility. Integrated security systems Video monitoring of exterior areas Simple berm with 2' raised walkway	Current + Emerging standards
Operational Issues		Reduction or elimination of employee workstations within the data center Remote access technologies to reduce human traffic within data center	Proactive monitoring systems with rules based actions (autonomics)	Current + Emerging standards

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Center Maintenance		Periodic professional sub floor and duct cleaning, air handler filters, and humidity control systems. Security checks on all maintenance support staff and vendors	Lifecycle replacement policies and budgeting Recovery systems and procedures testing	Current + Emerging standards Physical plant and facilities specialists dedicated to data center
Data Center Complacence		HIPPA (Physical security requirements) Professional audits (Ex, security, environmental) AADAG (ADA) access	Compliance with industry standards (ISO, IEEE, etc) Pick standards as ISO & IEEE are huge. NEC Article 645 for IT Equipment	Current + Emerging standards

Definitions of hosting technology include:

- **Internal (within agency)** – The hosting of a web site or application within an agency. The agency is responsible for the maintenance, support and availability of the web site or application.
- **External (ISP/ASP/FirstGov)** – The outsourcing of a web site or application with a managed service provider. An Internet Service Provider (ISP) provides telecommunications circuits , server collocation, and web site and application hosting. An Application Service Provider (ASP) offers software-based services for high end business applications and specific needs applications such as payroll, sales force automation, and human resources. FirstGov is the official managed service provider for the federal government.
- **Data Center** - The Physical environment that supports the hosting environment.
- **Man Trap (Air Lock) -** A short hallway leading into a Data Center with a door at each end effectively creating an air lock between the Data Center and the outside public area. Used to maintain the temperature and humidity within the data center upon entry by employees or service technicians. Also used as a security trap for holding individuals attempting unauthorized entry. Allows staff to view and review people requesting entry and contact security staff for support.

Other Important information for Hosting

Data Center Security policies and personnel security clearance policies are available for reference from DISC.

If just 10% of the State's Data Center applications need a very expensive Tier IV facility, the model should allow for 90% much cheaper lower Tier areas. If data spaces are modular, they allow for easy expansion and a gradual migration. The Uptime Institute guidelines below are just a subset of the available criteria. http://www.forsythe.com/Forsythe/advisory/dcservices/dcpd_design.jsp

- **Tier I Data Center** - Single path for power and cooling distribution with no redundant components - 99.671% availability. (**~$450** per square foot)

- **Tier II Data Center** - Single path for power and cooling distribution with redundant components - 99.741% availability. (**~$600** per square foot)

- **Tier III Data Center** - Multiple power and cooling distribution paths. These data centers only have one active path with redundant components. (**~$900** per square foot)

- **Tier IV Data Center** - Multiple active power and cooling distribution paths, redundant components and fault tolerance - 99.995% availability. (**~$1,100** per square foot)

The chances of the State funding top notch data centers improve if the requirements allow more expensive areas for high importance applications and avoid "gold plated" facilities for applications of lower importance (security / performance / time-sensitivity /criticality). Additionally, DFM strategic planning has identified data centers as inappropriate tenants for office buildings in the long run. It should be encouraged to design data centers with the DC being above the mechanical space and for levels to be planned for modular expansion. For State needs, this may suggest an "X" shape facility with agencies getting their own wing. Administrative areas can be the center of the "X" or maybe on a low cost wing.

5.4 Service Transport

Service Transport defines the end to end management of the communications session to include the access and delivery protocols.

5.4.1 Supporting Network Services

These consist of the protocols that define the format and structure of data and information that is either accessed from a directory or exchanged through communications.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
LAN Mail Access	X.400	MAPI SOAP		
Internet Mail Access		Thick client over IMAP4/S POP3/S	Clientless over HTTPS	
Mail Transport		SMTP ESMTP		
Directory Integration	NDS OID X.500	LDAP Active Directory		
Automatic Host Configuration	BOOTP	DHCP		
Domain Name Resolution	WINS NetBIOS	DNS		

Description	Example Products	Notes
LDAP server	iPlanet	

Definitions for Supporting Network Services include:

- Internet Message Access Protocol / Post Office Protocol (IMAP / POP3) – IMAP allows a client to access and manipulate electronic mail messages on a server. IMAP permits manipulation of remote message folders, called "mailboxes", in a way that is functionally equivalent to local mailboxes. IMAP also provides the capability for an offline client to resynchronize with the server. POP3 is the most commonly used protocol for retrieving email from a mail host.
- Simple Mail Transfer Protocol (SMTP) – SMTP facilitates transfer of electronic mail messages. It specifies how two systems are to interact, and the messages format used to control the transfer of electronic mail.
- Extended Simple Mail Transfer Protocol (ESMTP) – ESMTP allows new service extensions to SMTP to be defined and registered with Internet Assigned Numbers Authority (IANA).
- Lightweight Directory Access Protocol (LDAP) – LDAP is a subset of X.500 designed to run directly over the TCP/IP stack. LDAP is, like X.500, both an information model and a protocol for querying and manipulating it. LDAPv3 is an update developed in the IETF (Internet Engineering Task Force), which address the limitations found during deployment of the previous version of LDAP.
- Directory Services (X.500) – This is a network service that discovers and identifies resources on a network and makes them accessible to users and applications. The resources include users,

email addresses, computers, mapped drives, shared folders, and peripherals such as printers and PDA docking stations. Users and computers access these resources without the needing to know how or where the resources are connected.

- Dynamic Host Configuration Protocol (DHCP) – A protocol for assigning dynamic IP addresses to devices on a network. A device can receive a different IP address for every connection. Dynamic addressing provides reduced network administration over deploying and connecting user and peripheral devices.
- Domain Name System (DNS) – A protocol used for translating domain names (i.e. www.feapmo.gov) to their respective IP addresses. DNS is collectively a network of devices which store query results. As one DNS server or device cannot provide the translated IP address, it queries other DNS devices. This process is invisible to the user.
- X.400 – An ISO and ITU standard for email message addressing and transporting. X.400 supports Ethernet, X.25, TCP/IP and dialup transport methods.
- Messaging Application Programming Interface (MAPI) – A system built into Microsoft Windows that enables different e-mail applications to work together to distribute mail. As long as both applications are MAPI-enabled, they can share mail messages with each other.
- Oracle Internet Directory (OID) – An LDAP v3 directory which utilizes Oracle Database. Oracle Internet Directory serves as the central user repository for Oracle Identity Management for user administration in the Oracle environment and can provide a standards-based application directory for an enterprise.
- Windows Internet Naming Service (WINS) – A system that determines the IP address associated with a particular network computer. WINS uses a distributed database that is automatically updated with the names of computers currently available and the IP address assigned to each one.
- Network Basic Input Output System (NetBIOS) – An API that augments the DOS BIOS by adding special functions for local-area networks (LANs).
- Bootstrap Protocol (BOOTP) – An Internet protocol that enables a diskless workstation to discover its own IP address, the IP address of a BOOTP server on the network, and a file to be loaded into memory to boot the machine.
- Novell Directory Services (NDS) – The directory services for Novell Netware networks. NDS complies with the X.500 standard and provides a logical tree-structure view of all resources on the network so that users can access them without knowing where they're physically located.
- Simple Object Access Protocol – A lightweight XML-based messaging protocol used to encode the information in Web service request and response messages before sending them over a network. SOAP messages are independent of any operating system or protocol and may be transported using a variety of Internet protocols, including SMTP, MIME, and HTTP.

5.4.2 Service Transport

These consist of the protocols that define the format and structure of data and information that is either accessed from a directory or exchanged through communications.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Mail Encoding		MIME	S/MIME	
Internet Data Exchange		HTML EDI XML	XHTML	XML
Internet File Transfer	Point-To-Point	FTP SFTP		SFTP
VPN Communications		IPSEC PPTP		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Formatted File Transmittal	PostScript Image File (i.e. .jpg, .gif)	PDF		
Image Encoding	BMP GIF	JPEG PNG	SVG	
Mainframe Communication		TN3270 Secure TN3270		
Terminal Session	Telnet	SSH		
HTML Data Transmission		HTTP HTTPS		

Definitions of Service Transport technologies include:

- Hyper Text Transfer Protocol (HTTP) – The communications protocol used to connect to servers on the World Wide Web. Its primary function is to establish a connection with a web server and transmit HTML pages to the client browser.
- Hyper Text Transfer Protocol Secure (HTTPS) – The protocol for accessing a secure web server. Using HTTPS in the URL instead of HTTP directs the message to a secure port number rather than the default web port number of 80. The session is then managed by a security protocol.
- File Transfer Protocol (FTP) – A protocol used to transfer files over a TCP/IP network (Internet, UNIX, etc.). For example, after developing the HTML pages for a Web site on a local machine, they are typically uploaded to the Web server using FTP.
- Multipurpose Internet Mail Extensions (MIME) – MIME extends the format of Internet mail to allow non-U.S. American Standard Code for Information Interchange (ASCII) textual messages, non textual messages, multipart message bodies, and non USASCII information in message headers. MIME support allows compliant email clients and servers to accurately communicate embedded information to internal and external users.
- HyperText Markup Language (HTML) – The authoring language used to create documents on the World Wide Web.
- Extensible Hypertext Markup Language (XHTML) – A hybrid between HTML and XML specifically designed for Net device displays.
- Extensible Markup Language (XML) – A specification developed by the W3C. XML is a pared-down version of SGML, designed especially for Web documents. It allows designers to create their own customized tags, enabling the definition, transmission, validation, and interpretation of data between applications and between organizations.
- Portable Document Format (PDF) – A file format developed by Adobe Systems. PDF captures formatting information from a variety of desktop publishing applications, making it possible to send formatted documents and have them appear on the recipient's monitor or printer as they were intended.
- Scalable Vector Graphics (SVG) – A vector graphics file format that enables two-dimensional images to be displayed in XML pages on the Web. Vector images are created through text-based commands formatted to comply with XML specifications. In contrast to JPEG and GIF images on the Web, which are bitmapped and always remain a specified size, SVG images are scalable to the size of the viewing window and will adjust in size and resolution according to the window in which it is displayed.
- Secure Shell (SSH) – A program to log into another computer over a network, to execute commands in a remote machine, and to move files from one machine to another. It provides

strong authentication and secure communications over insecure channels. It is a replacement for rlogin, rsh, rcp, and rdist.

- Telnet – A terminal emulation program for TCP/IP networks such as the Internet.
- Telnet 3270 (TN3270) – A client program that includes the 3270 protocol for logging onto IBM mainframes. Using the TCP/IP networking protocol, it is widely used to connect a desktop computer to a mainframe and emulate a mainframe terminal (3270) session.

Chapter 6 Service Platform and Infrastructure

The Service Platform and Infrastructure Service Area define the collection of platforms, hardware and infrastructure standards that enable Component Based Architectures and Service Component reuse.

6.1 Supporting Platforms

Support platforms are hardware or software architectures. The term originally dealt with only hardware, and it is still used to refer to a CPU model or computer family.

6.1.1 Wireless / Mobile

Radio transmission via the airwaves. Various communications techniques are used to provide wireless transmission including infrared "line of sight," cellular, microwave, satellite, packet radio and spread spectrum.

Definitions of wireless / mobile technology include:

- **Java 2 Platform, Micro Edition (J2ME)** – Sun's Java environment for devices. It promises a relatively portable environment for those using Java for other tiers of the architecture.

- **Packet Radio -** Packet radio is a form of digital data transmission used in amateur radio to construct wireless computer networks. Its name is a reference to the use of packet switching between network nodes, which allows multiple virtual circuits to coexist on a single radio channel. Packet radio networks use the AX.25 data link layer protocol, derived from the X.25 protocol suite and designed for amateur radio use.

- **Satellite** – Communications satellites are artificial satellites stationed in space for the purposes of telecommunications using radio at microwave frequencies. Most communications satellites use geosynchronous orbits or near-geostationary orbits, although some recent systems use low Earth-orbiting satellites.

  Satellite Internet services are used in locations where terrestrial Internet access is not available and in locations which move frequently. Internet access via satellite is available worldwide, including vessels at sea. There are three types of satellite Internet service:

  - one-way multicast,
  - one-way with terrestrial return, and
  - two-way satellite access.

  A satellite telephone, satellite phone, or satphone is a mobile phone that communicates directly with orbiting communications satellites. Depending on the architecture of a particular system, coverage may include the entire Earth, or only specific regions.

- **Microwave -** Microwaves are used in broadcasting transmissions because microwaves pass easily through the earth's atmosphere with less interference than longer wavelengths. There is also much more bandwidth in the microwave spectrum than in the rest of the radio spectrum. Typically, microwaves are used in television news to transmit a signal from a remote location to a television station from a specially equipped van. Wireless LAN protocols, such as Bluetooth and the IEEE 802.11g and b specifications, also use microwaves in the 2.4 GHz ISM band, although 802.11a uses an ISM band in the 5 GHz range. Licensed long-range (up to about 25 km) Wireless Internet Access services can be found in many countries (but not the USA) in the 3.5–4.0 GHz range.
- **Infrared Communications** - IR data transmission is also employed in short-range communication among computer peripherals and personal digital assistants. These devices usually conform to standards published by IrDA, the Infrared Data Association. Remote controls and IrDA devices use infrared light-emitting diodes (LEDs) to emit infrared radiation which is

focused by a plastic lens into a narrow beam. The beam is modulated, i.e. switched on and off, to encode the data. The receiver uses a silicon photodiode to convert the infrared radiation to an electric current. It responds only to the rapidly pulsing signal created by the transmitter, and filters out slowly changing infrared radiation from ambient light. Infrared communications are useful for indoor use in areas of high population density. IR does not penetrate walls and so does not interfere with other devices in adjoining rooms. Infrared is the most common way for remote controls to command appliances.

Free space optical communication using infrared lasers can be a relatively inexpensive way to install a communications link in an urban area operating at up to 4 gigabit/s, compared to the cost of burying fiber optic cable.

Infrared lasers are used to provide the light for optical fiber communications systems. Infrared light with a wavelength around 1,330 nm (least dispersion) or 1,550 nm (best transmission) are the best choices for standard silica fibers.

- **Spread Spectrum** - Spread-spectrum techniques are methods in which energy generated at a single frequency is deliberately spread over a wide band of frequencies. This is done for a variety of reasons, including increasing resistance to natural interference or jamming and to prevent hostile detection.

6.1.2 Platform Independent

Defines the operating systems and programming languages that are able to execute and run on any platform or operating system. A platform is the underlying hardware and software comprising a system.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Architecture Framework		J2EE / J2EE5	AJAX W3C WAI-ARIA	
Operating System	Net BSD Open BSD Free BSD	Linux	Red Hat Enterprise Linux	Redhat Linux Red Hat Enterprise Linux

Definitions of Platform Independent technology include:

- **Java 2 Platform Enterprise Edition (J2EE)** – Sun's J2EE and Microsoft's .Net are the two dominant distributed computing architecture frameworks. J2EE provides portability of a single language (Java) over multiple operating systems and hardware platforms.

- **Linux** – Linux is an open source operating system that runs on multiple hardware platforms. With the ability to run on many platforms, including the PC and Macintosh, Linux has become an alternative to proprietary systems.

- **Red Hat Linux** was one of the most popular Linux distributions, assembled by Red Hat.

  It is one of the "middle-aged" Linux distributions; 1.0 was released in November 3, 1994. It is not as old as Slackware, but certainly older than many other distributions. It was the first Linux

distribution to use RPM as its packaging format, and over time has served as the starting point for several other distributions, such as the desktop-oriented Mandriva Linux (originally Red Hat Linux with KDE), Yellow Dog Linux (which started from Red Hat Linux with PowerPC support), and ASPLinux (Red Hat Linux with better non-Latin character support).

Since 2003, Red Hat has discontinued the Red Hat Linux line in favor of its new Red Hat Enterprise Linux. Red Hat Linux 9, the final release, hit its official end-of-life on April 30, 2004, although the Fedora Legacy project continues to publish updates.

- **AJAX - A**synchronous **J**avaScript **A**nd **X**ML, or its acronym, **Ajax** (Pronounced A-jacks), is a Web development technique for creating interactive web applications.

6.1.3 Platform Dependant

Defines the operating systems and programming languages that are able to execute and run on a specific platform or operating system. A platform is the underlying hardware and software comprising a system.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Architecture Framework		Microsoft .Net		
Desktop Operating Systems	Windows 2000 Windows 95 Windows ME Windows 98 Windows NT4 Solaris 7 and below Mac OS AIX 5.1 and below OS/2	Windows XP Windows Vista Solaris 8 and above AIX 5.2 and above HPUX Mac OSX	Windows 7	Windows XP
Mid Tier SERVER /OS	Windows 2000 Windows NT4 Novell Netware Solaris 8 and below DEC Alpha	Windows 2008 Windows 2003 Solaris 9 and above AIX 5.2 and above HPUX OS/400		Windows 2008 Solaris 10
MAINFRAME / OS	OS/390	Z/OS		Z/OS

Definitions of Platform Dependent technology include:

- **Windows 2000** – Also known as "Win2K" and "W2K," it is a major upgrade to Windows NT 4. Launched in February 2000, Windows 2000 comes in one client and three server versions. Windows 2000 looks like Windows 95/98, but adds considerably more features, dialogs and options.

- **Windows.Net** – Microsoft's .Net and Sun's J2EE are the two dominant distributed computing architecture frameworks. .Net supports a wide range of languages but is primarily tied to the Microsoft Windows operating system and Intel hardware.

- **Mac OS X** – Mac OS X is Apple's UNIX based operating system based on industry standards. Launched in March 2001, OS X has advanced built-in security functions and complete interoperability with both Internet standards and Microsoft products.

- **Windows Vista** – Windows Vista is Microsoft's upcoming operating system, which was formerly code-named "Longhorn". At present, Longhorn is reserved for Server Version since the official launch and rename to Vista.

- **Windows Server Longhorn -** Windows Server "Longhorn" is the former working name of an upcoming server operating system from Microsoft. Windows Vista, the replacement for the "Longhorn" codename.

6.2 Delivery Servers

Delivery Servers are front-end platforms that provide information to a requesting application. It includes the hardware, operating system, server software, and networking protocols.

Note: Hardware and Operating Systems are included in the Platform Independent (6.1.2) and Platform Dependant (6.1.3) sections. Network protocols will be addressed in the Network sections (5.4 and 6.5) of the architecture.

6.2.1 Web Servers

A *web server* is a computer that provides World Wide Web services on the Internet. It includes the hardware, operating system, web server software, TCP/IP protocols and the web site content (web pages). If a web server is used internally and not by the public it may be known as an "intranet server."

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Web Servers	Other proprietary Web servers	Apache IIS Websphere Application Server		Apache (open) IIS (proprietary)

Definitions of Web Servers technology include:

- **Apache** – A widely used public domain, Platform independent based web server from the Apache Group (www.apache.org). It is based on, and is a plug-in replacement for, NCSA's HTTP server Version 1.3. The name came from a body of existing code and many "patch files."

- **Internet Information Server (IIS)** – Web server software from Microsoft that runs on Windows platforms and Microsoft.Net.

6.2.2 Media Servers

Provide optimized management of media based files such as audio and video streams and digital images. Comprises the content delivery service environment for streaming live or on demand video and audio content.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Media Servers		Apple Quick Time Server MacroMedia Flash media server Windows media server Helix media delivery systems		
Media Players		Apple Quick Time player Windows media player MacroMedia Flash player		

6.2.3 Application Servers

In a three tier environment, a separate computer (application server) performs the business logic, although some part may still be handled by the user's machine. After the web exploded in the mid 1990s, application servers became web-based.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Proprietary		Websphere Application Server Oracle application server Windows .Net framework		
Open		Tomcat		

Definitions of Application Servers technology include:

- **WebSphere -** WebSphere refers to a brand of IBM software products, although the term also popularly refers to one specific product.

6.2.4 Portal Servers

Portals represent focus points for interaction, providing integration and single source corporate information.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Portal Servers	Pearl	Sharepoint Metadot Jetspeed Pluto		

Definitions of Portal Servers technology include:

- **Portals** – Portals are a very common application server mechanism by which organizations can manage information. They provide a single point of entry for all users, they can access Web services transparently from any device, and they are highly flexible. Portals can work inside or outside of the organization, and they can attach themselves to any part of it.

- **uPortal** – uPortal is a Java-based framework for creating web portals for educational institutions. uPortal is open source under a BSD-style license. uPortal has integrated Apache Software Foundation's Pluto software to become JSR 168 compliant allowing it to host Portlets.

6.3 Software Engineering

Software engineering covers the technology associated with building software systems as well as technical solutions supporting management issues, such as testing, modeling and versioning. The TRM is concerned with component technical architecture, not engineering processes.

6.3.1 Integrated Development Environment

This consists of the hardware, software and technology that facilitate the development of software applications and systems.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Common IDE Components		Source code editor Compiler or interpreter Build-automation tools Debugger	Integrated version control GUI construction tools Class browser Object inspector Class hierarchy diagram	Source code editor Compiler or interpreter Build-automation tools Debugger
Programming Paradigm	Structured	Support of Object Oriented	Support of Service Oriented	Support of Service Oriented Support of Object Oriented
Requirements	Text Based	XML Capable Support development of Web-based environments Interoperability with multiple environments Code transparency Support Plugins	Smart Client Development	XML Capable Support development of Web-based environments Interoperability with multiple environments Code transparency Support Plugins
Security		LDAP X.500 Support	Single Signon with controlled permissions	Single Signon with controlled permissions LDAP X.500 Support

Description	Example Products	Notes
Java-based	• Java • JBoss AS • Sonic MQ (JMS only) • Eclipse IDE • MyEclipse Plug-in • Eclipse RCP (Rich Client Platform) • Rational Application Developer • (RAD) 6.1 • NetBeans • JBuilder • Web Sphere Studio • Apache • Struts • Tomcat	• (current or current -1 version) • (J2EE compliant application server)
Windows-based	Visual Studio.Net Visual Studio 2008 Visual Studio Team System Cold Fusion	
Linux-based	• LAMP • Apache • Cold Fusion	
Other	PeopleTools (PeopleSoft) TSO / Roscoe (Cobol) Advantage:GEN (Cobol) Software AG NSC Software AG Natural Editor (Natural) Cold Fusion	 Twilight technology with current application in legacy applications Natural Security Control

Definitions of Integrated Development Environments technology include:

- **Code Transparency –** The ability to see the source code generated by the IDE.
- **XML Capable –** Ability to create, read, and transfer XML data.
- **Web Sphere Studio** – Integrated Java (J2EE) environment for programmers building Java, web, and web services applications – and the successor to IBM Visual Age.
- **Visual Studio** – A complete development system providing the tools for analyzing and modeling all aspects of an application before a single component is built so that developers can design efficient architectures and reduce time to market. Developers can choose the programming language they know best and the language that is best suited to the solution, including Microsoft Visual Basic, Visual C++, Visual J++, and Visual FoxPro. Visual Studio is used to build scalable, data driven Web sites and applications.

- **Visual Studio.Net** – A comprehensive tool set for rapidly building and integrating XML web services, Microsoft Windows–based applications, and web solutions. This is the successor to Visual Studio.
- **Rational Application Developer** - based on Eclipse and provides extensive support for Java and J2EE artifacts, as well as other languages such as COBOL and PL/I.

6.3.2 Software Configuration Management

Technology applicable to all aspects of software development from design to delivery specifically focused on the control of all work products and artifacts generated during the development process. Several technical solutions on the market provide the integration of the software configuration management functions.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Version Management	Manual Process	Ability to recover from previous versions Ability to distinguish different versions Ability to roll back to previous version Multiple checkouts with automated synchronization Auditable Release Control		Ability to recover from previous versions Ability to distinguish different versions Ability to roll back to previous version Multiple checkouts with automated synchronization Auditable Release Control
Issue Tracking		Task Assignment Task Tracking Assign Ownership Metrics Reporting capabilities		Task Assignment Task Tracking Assign Ownership Metrics Reporting capabilities
Defect Tracking		Ability to identify, assign and manage defects Ability to track defects Ability to capture bugs Ability to report Metrics		Ability to identify, assign and manage defects Ability to track defects Ability to capture bugs Ability to report Metrics

Task Management		Ability to define tasks Ability to prioritize		Ability to define tasks Ability to prioritize
Change Management	Manual process or No process In-house developed	ITIL process compliant		
Deployment Management		Auto deployment and tracking (Push or Pull)	Virtual Auto Deployment	Auto deployment and tracking (Push or Pull)
Requirements Management and Traceability		No process or Manual process	Automation with manual processes Tie requirements to functionality and to test cases	

Description	Example Products	Notes
Version Management	• CA Librarian (MF) • PVCS (server side) • Subversion • Serena Dimensions • Team Systems • In-house Developed • Source Safe • Visual Studio Team • ClearCase	• Various clients available • Twilight technology with current application in legacy applications • Twilight technology with current application in legacy applications
Issue Tracking	• Team Systems • HEAT • Bugtrack • ISIS • Microsoft Project • OpenWorkbench • Clarity • BMC Remedy • Bugzilla • BMC Remedy • IBM Infoman • Microsoft Office (Project, Excell, Word) • ClearQuest	Twilight technology with current application in legacy applications

Description	Example Products	Notes
Deployment Management	• Team Systems • Maven • Cruise Control • Altiris • Ant • CA Unicenter TNG • Build Forge	Also refer to the System's Management Subcommittee area
Defect Tracking	See Issue Tracking above	
Task Management	• MS Project • OpenWorkbench • Clarity	
Change Management	• BMC Remedy • Serena Dimensions • ClearCase	
Requirements Management and Traceability	• Doors • Requisite Pro • MS Office (Project, Excel, Word)	

Definitions of Software Configuration Management technology include:

- **Version Management** – Refers to tracking and controlling versions of files. Version Management includes capabilities such as labeling, branching, merging, version content comparisons, and security and permission management across version controlled projects.
- **Defect Tracking** – Refers to the identification, assignment, and management of discovered defects within an application, product or solution. Defect tracking tools provide searchable defect data to identify urgent and related defects or bugs. The architecture should be built to facilitate the pushing of software patches across the enterprise.
- **Issue Management** – Refers to the management of business, technical, and infrastructure issues throughout the entire lifecycle of a project.
- **Task Management** – Requirements, testing, and issues assignments are transformed into prioritized tasks. Task Management tools provide automation features for managing, delivering, assigning, reminding, and collaborating task management and execution.
- **Change Management** – Refers to the management of application code and content changes across the software development lifecycles.
- **Deployment Management** – Refers to the capability of software delivery to remote networked desktops, servers, and mobile devices across an enterprise. Deployment automation tools provide centralized and accelerated delivery of applications to users via push technologies, eliminating the need for manual installation and configuration.
- **Requirements Management and Traceability** – Consists of information discovery, capture, storage and dissemination. Requirements management reduces software development costs and associated risks through documenting, measuring, and analyzing deviations to project requirements. Traceability refers to tracking requirements artifacts to their source, and changes in requirements to include the impact analysis of the change. Requirements traceability is an integral component in quality software implementation and the management of document succession.
- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

6.3.3 Test Management

Technology which supports the consolidation of all testing activities and results. Test Management activities include test planning, designing (test cases), execution, reporting, code coverage, and heuristic and harness development.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Test Management	Manual Process	Create reproducible test scripts. Create reliable, reusable, reproducible test data sets Tests tracked to make sure all of the code is tested Easily generated load testing Documented scripts		Create reproducible test scripts. Create reliable, reusable, reproducible test data sets Tests tracked to make sure all of the code is tested Easily generated load testing Documented scripts
Functional Testing	Manual Process	Scripting Data Sets		
Business Cycle Testing				
Usability Testing		Ability to test W3C compliance Ability to test for State requirements Ability to test for Federal requirements (Section 508)		Ability to test W3C compliance Ability to test for State requirements Ability to test for Federal requirements (Section 508)
Performance Profiling	Manual Process	Scripting Data Sets		
Load / Stress / Volume Testing	Manual Process	Scripting Data Sets		
Security and Access control Testing				
Reliability Testing	Manual Process	Scripting Data Sets		
Configuration Testing				

Installation Testing				
Unit Testing		Built-in Unit Testing		Built-in Unit Testing
Simulation Testing		Developer test code	Automated simulation testing	Developer test code

Description	Example Products	Notes
Test Creation	Microsoft .NET Team Systems	
Test tracking	Microsoft .NET Team Systems	
Functional Testing		
Business Cycle Testing		
Usability Testing	• Bobby • JAWS • Morae (Techsmith)	
Performance Profiling	CompuWare ClientVantage (web) CompuWare STROBE (mainframe)	
Load/Stress/Volume Testing	CompuWare QA Load (web) Jmeter BadBoy	
Security and Access control Testing		
Reliability Testing		
Configuration Testing		
Installation Testing		
Unit Testing	JUnit Cactus	Includes recursion testing
Simulation Testing	Parallels VM Canoe Webtags	Windows and Linux

Definitions of Test Management technology include:

- **Functional Testing** – This type of test focuses on any requirements that can be traced directly to use cases (or business functions), business rules, and design.
- **Business Cycle Testing** – Refers to the emulation of activities performed over a period of time that is relevant to the application under test.
- **Usability Testing (508 Testing)** – Refers to a test to ensure that the application navigation, functionality, and GUI allow a user to effectively and efficiently do their work in a way that they are satisfied with the application.
- **Performance Profiling** – Refers to a performance test that measures and evaluates response times and transaction rates.
- **Load/Stress/Volume Testing** – Refers to tests that measure and evaluate how a system performs and functions under varying workloads, large amounts of data and/or resource utilization.

- **Security and Access Control Testing** – Focuses on the technical, administrative and physical security controls that have been designed into the system architecture in order to provide confidentiality, integrity and availability.
- **Reliability Testing** – Refers to the verification that failover methods are invoked properly and the system recovers properly.
- **Configuration Testing** – Refers to a test to ensure that the application or system can handle all hardware and software variables and requirements that have been defined.
- **Installation Testing** – Refers to the verification that the software installation process works properly in different environments and among varying conditions

6.3.4 Modeling

Technology that supports the process of representing entities, data, business logic, and capabilities for aiding in software engineering.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Modeling	Manual Process	Tools that support UML Ability to link models Interoperable with other modeling programs Ability to create reusable models	BPEL	Tools that support UML Ability to link models Interoperable with other modeling programs Ability to create reusable models

Description	Example Products	Notes
Modeling	• MS Visio • Provision • MS PowerPoint • System Architect • Metis • MyEclipse • Rational Rose	• UML portion only (Rational Rose)

Definitions of Modeling technology include:

- **Unified Modeling Language (UML)** – A general-purpose notational language for specifying and visualizing complex software, especially large, object-oriented projects.
- **Case Management** – Computer Aided Software Engineering (CASE) software that provides a development environment for programming teams. CASE systems offer tools to automate, manage and simplify the development process.

6.4 Database / Storage

Database / Storage refers to a collection of programs that enables storage, modification, and extraction of information from a database, and various techniques and devices for storing large amounts of data.

6.4.1 Database

Refers to a collection of information organized in such a way that a computer program can quickly select desired pieces of data. A database management system (DBMS) is a software application providing management, administration, performance, and analysis tools for databases.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Databases	Flat File	Relational Spatial XML	Object-oriented databases CORBA	Relational Spatial XML CORBA

Description	Example Products	Notes
	Database 2 (DB2)	DB2 is a family of relational database products offered by IBM. DB2 provides an open database environment that runs on a wide variety of computing platforms.
	Oracle	Relational database product.
	SQL Server	Data management server product developed by Microsoft.
	Sybase	Data management and synchronization server products developed by Sybase.
	PostgreSQL	Open Source relational database.
	MySQL	Open Source relational database.
	SAS	
	Adabas	Non-relational database software from SoftwareAG.
	MS Access	

Preferences for Databases:

- Relational databases should support SQL.
- Support XML data formats.
- Support Spatial Data formats.
- Support large objects (BLOBS).
- Support federated/heterogeneous services for data interchange.
- Support various storage options/partitioning.
- Support data mirroring.
- Ease of recoverability.
- Support 24x7availability.

Definitions of Database technologies include:

- **CORBA** - Common Object Request Broker Architecture is an architecture and specification for creating, distributing, and managing distributed program objects in a network. It allows programs at different locations and developed by different vendors to communicate in a network through an "interface broker." CORBA was developed by a consortium of vendors through the Object Management Group (OMG), which currently includes over 500 member companies. Both

International Organization for Standardization (ISO) and X/Open have sanctioned CORBA as the standard architecture for distributed objects. CORBA 3 is the latest level.

- **Object-Oriented Database** - a database structure that organizes, manipulates, and retrieves classes of objects, such as sound, video, text, and graphic files. CORBA is emerging as the industry standard for object-oriented technologies.
- **Relational Database** - A relational database is a collection of data items organized as a set of formally-described tables from which data can be accessed or reassembled in many different ways without having to reorganize the database tables. The relational database was invented by E. F. Codd at IBM in 1970. The standard user and application program interface to a relational database is the *structured query language* (SQL). SQL statements are used both for interactive queries for information from a relational database and for gathering data for reports. Source: http://searchsqlserver.techtarget.com/sDefinition/0,,sid87_gci214260,00.html
- **SQL** – Structured Query Language. - A standard interactive and programming language for getting information from and updating a database. Although SQL is both an ANSI (American National Standards Institute) and an ISO (International Standards Organization) standard, many database products support SQL with proprietary extensions to the standard language. Queries take the form of a command language that lets you select, insert, update, find out the location of data, and so forth. Source http://searchsqlserver.techtarget.com/sDefinition/0,,sid87_gci214230,00.html
- **Spatial database system** - a database system that offers spatial data types in its data model and query language, and supports spatial data types in its implementation, providing at least spatial indexing and spatial join methods. Spatial database systems offer the underlying database technology for geographic information systems and other applications. Source: acm.org at http://portal.acm.org/citation.cfm?coll=GUIDE&dl=GUIDE&id=615206
- **XML** - XML (Extensible Markup Language) is a flexible way to create common information formats and share both the format and the data n the World Wide Web, intranets, and elsewhere. Source: http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci213404,00.html

Other Important information for Database Technologies

- Source for other XML-capable database products: http://www.rpbourret.com/xml/XMLDatabaseProds.htm.
- Standards for Object-Oriented technology (Object Management Group): http://www.omg.org/

6.4.2 Storage

Storage devices are designed to provide information to direct attached servers or provide non-volatile digital storage media to support information processing in a local and a network environment. These devices provide extended storage capabilities to the network with reduced costs compared to traditional file servers.

Note: The State of Kansas has chosen to limit this area to only Storage. Databases are covered within the Database section (6.4.1) of the architecture.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Tape	AIT DLT/SDLT 3480, 3490	LTO Virtual Tape 9840/3590		LTO 9840 Virtual Tape

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Disks	IDE MFM ESDI	SCSI EIDE Solid State 3390 SDRam Flash Based Drives Fibre Channel ATA/PATA/SATA		SAS Fibre Channel SATA
Optical		Re-writeable CDs and DVDs–R WORM Drive		
Storage Systems	NAS	Local RAID SAN	iSCSI	SAN
SAN Switches	McData	Brocade Cisco Inter-VSAN routing		Shared storage between SANs
Fiber Connections	SC	LC ST		LC
SAN Management		TekTools		

Definitions of storage technologies include:

- **Network Attached Storage (NAS)** – A NAS device is a server that is dedicated to nothing more than file sharing.
- **Network-attached storage** (commonly abbreviated to **NAS**) systems are generally computing-storage devices that can be accessed over a computer network (usually TCP/IP), rather than directly being connected to the computer (via a computer bus such as SCSI). This enables multiple computers to share the same storage space at once, which minimizes overhead by centrally managing hard disks. NAS systems usually contain one or more hard disks, often arranged into logical, redundant storage containers or RAID arrays.

  The protocol used with NAS is a file based protocol such as NFS or Microsoft's Common Internet File System (CIFS). In reality, there is a miniature operating system on the device such as DART on EMC's Celerra devices or Data ONTAP on NetApp NAS devices.

  A storage area network (SAN) is very similar, except it uses a block-based protocol and generally runs over an independent, specialized storage network.

NAS devices become logical file system storage for a local area network. Thus the performance of NAS devices depends heavily on cached memory (the equivalent of RAM) and network interface overhead (the speed of the router and network cards). The benefit is that the device can become a giant neighborhood hard drive for a whole building. The disadvantage is that any constrictions in the local network will slow down the resulting access time.

NAS was developed to address problems with direct attached storage, which included the effort required to administer and maintain "server farms", and the lack of scalability, reliability, availability, and performance. They can deliver significant ease of use, provide heterogeneous data sharing and enable organizations to automate and simplify their data management.

In addition, the price of NAS appliances has plummeted in recent years, offering flexible network based storage to the home consumer for little more than the cost of a regular USB or Firewire external hard disk.

- **Storage Area Network (SAN)** – A SAN is a high-speed sub network of shared storage devices. A storage device is a machine that contains nothing but a disk or disks for storing data.

  In computing, a **storage area network** (**SAN**) is a network designed to attach computer storage devices such as disk array controllers and tape libraries to servers. As of 2006, SANs are common in enterprise storage.

  There are two variations of SANs:

  1. A network whose primary purpose is the transfer of data between computer systems and storage elements. A SAN consists of a communication infrastructure, which provides physical connections, and a management layer, which organizes the connections, storage elements, and computer systems so that data transfer is secure and robust. The term SAN is usually (but not necessarily) identified with block I/O services rather than file access services.

  2. A storage system consisting of storage elements, storage devices, computer systems, and/or appliances, plus all control software, communicating over a network.

- **iPod** - The **iPod** is a brand of portable media players designed and marketed by Apple Computer. Devices in the iPod family provide a simple user interface designed around a central scroll wheel (with the exception of the iPod shuffle). The standard iPod model stores media on a built-in hard drive, while the smaller iPod shuffle and iPod nano use flash memory. Like most digital audio players, an iPod can serve as an external data storage device when connected to a computer.

- **Portable Media Device** - A **portable multimedia player**, or **PMP** is a hard disk or flash memory based electronic device, such as an "MP3 player", which is capable of storing and playing files in one or more media formats.

- A **USB flash drive** is essentially NAND-type flash memory integrated with a USB 1.1 or 2.0 interface. It is a small, lightweight, removable and rewritable data storage device of up to 64 GB, the most popular retail sizes being somewhere between 512 MB and 2 GB (as of 2006).[1]

- **Serial Attached SCSI** (SAS) is a new generation serial communication protocol for devices designed to allow for much higher speed data transfers and is compatible with SATA. SAS uses serial communication instead of the parallel method found in traditional SCSI devices but still uses SCSI commands for interacting with SAS devices.

  The physical SAS connector is available in 3 different variants:

  - SFF 8482 — which is form factor compatible with SATA.
  - SFF 8484 — hi-density internal connector for connecting up to 4 devices.
  - SFF 8470 — hi-density external connector (aka Infiniband connector) for connecting up to 4 devices.

  Serial Attached SCSI supports three transport protocols:

  - Serial SCSI Protocol (SSP) — Supporting SAS disk drives.
  - Serial ATA Tunneling Protocol (STP) — Supporting SATA disks.
  - Serial Management Protocol (SMP) — for managing SAS Expanders.

  A SAS domain is a set of SAS ports communicating with each other. A SAS domain contains one or more SAS devices and a service delivery subsystem. A SAS domain may be a SCSI domain. Each SAS device is assigned a World Wide Name (aka SAS address) assigned by IEEE for the particular vendor. The WWN uniquely identifies the device in an SAS domain just as a SCSI ID identifies a device in a parallel SCSI bus.

- **Linear Tape-Open** (or **LTO**) is a computer storage magnetic tape format developed as an open alternative to the proprietary Digital Linear Tape (DLT). The format was developed and initiated by Certance, Hewlett-Packard, and IBM. An implementation of the LTO tape format goes by the name "Ultrium".

  Current manufacturers of LTO tape drives are IBM, Hewlett-Packard, Certance (acquired by Quantum in 2005), NEC, Seagate, Qualstar, Exabyte and Tandberg Data.

  LTO Ultrium manufacturers: IBM StorageSmart, Seagate Viper, HP SureStore, Dell

  LTO is considered an entry point type of type backup system from an economic perspective as compared to a 9840 tape backup system which is more expensive and is usually used on mainframe systems.

- **Digital Linear Tape (DLT)** is considered a *de facto* standard for magnetic tape technology used for computer data storage. It was invented by Digital Equipment Corporation in 1984, and was purchased by Quantum Corporation in 1994, who currently manufactures drives and licenses the technology and trademark.

  A variant with higher capacity is called **Super DLT** (SDLT). The lower cost "value line" was initially manufactured by Benchmark Storage Innovations. Quantum acquired Benchmark in 2002.

  DLT uses linear serpentine recording with multiple tracks on half-inch (12.6 mm) wide tape. SDLT adds an optical servo system that reads servo patterns on the back of the tape. DLT7000 and 8000 tilt the head forward and backward to reduce crosstalk between adjacent tracks through azimuth; this is called Symmetric Phase Recording.

The cartridges contain a single reel and the tape is pulled out of the cartridge by means of a leader tape attached to the take-up reel inside the drive. The drive leader tape is buckled to the cartridge leader during the load process. Tape speed and tension are controlled electronically via the reel motors; there is no capstan. The tape is guided by 4 to 6 rollers that touch only the back side of the tape. Tape material is metal particle tape (MP/AMP.)

Media are guaranteed for 30 years of data retention under specified environmental conditions; however, they are easily damaged by mishandling (dropping or improper packaging during shipment.)

Current manufacturers of cartridges for the DLT/SDLT market are Fujifilm, Hitachi/Maxell and Imation. VStape is made by Sony. All other companies/brands (even Quantum) are contractors and/or resellers of these companies.

All (S)DLT drives support hardware data compression. The often-used compression factor of 2:1 is optimistic and generally only achievable for text data; a more realistic factor across a file system is 1.3:1 to 1.5:1.

A new naming convention took effect in 2005, calling the performance line DLT-S and the value line DLT-V.

- **Fibre Channel - Fibre Channel** is a gigabit speed network technology primarily used for Storage Networking. Fibre Channel is standardized in the T11 Technical Committee of the InterNational Committee for Information Technology Standards (INCITS), an American National Standard Institute (ANSI) accredited standards committee. It started for use primarily in the supercomputer field, but has become the standard connection type for storage area networks in enterprise storage. Despite its name, Fibre Channel signaling can run on both twisted-pair copper wire and fiber optic cables.

- **Tiered Storage –** Storage can be classified in four tiers. These tiers are defined as follows:

  Tier 1: **Enterprise Storage**-High end mainframe and midrange storage; Highest performing FiberChannel drives, high end data caching.

  Tier 2: **Department Storage**-High performing on-line storage; often without same high-end caching or the fastest FiberChannel drives.

  Tier 3: **Near-line and Content storage**-economy storage architecture; often SATA drive architecture.

  Tier 4: **Virtual Tape and Tape storage-**economy disk storage with migration to tape for seldom used data.

  As the cost for disk storage comes down as the technology improves, the lines between these tiers continues to blur and overlap.

- **Solid state disk** (**SSD**, also called **solid state drive**) is a data storage device that uses memory chips, such as SDRAMs, to store data, instead of the spinning platters found in conventional hard disk drives. While not technically "disks" in any sense, these devices are so named because they are typically used as replacements for disk drives in situations where space, power supply, or ruggedness concerns would make conventional drives impractical.
  SDRAM-based
  The term "SSD" is used for two different kinds of products. The first, SSDs based on fast, volatile memory such as ***SDRAM***, are categorized by extremely fast data access, less than 0.01 milliseconds (over 250 times faster than the fastest hard drives in ***2004***) and are used

primarily to accelerate applications that would otherwise be held back by the latency of disk drives.

Since these SSDs use volatile memory, they typically incorporate internal ***battery*** and backup disk systems to ensure data persistence. If power is lost for whatever reason, the battery would keep the unit powered long enough to copy all data from ***random access memory*** (RAM) to backup disk. Upon the restoration of power, data is copied back from backup disk to RAM and the SSD resumes normal operation.

These devices are especially useful on a computer which already has the maximum amount of RAM. For example, some ***x86*** architectures have a 4 ***GB*** limit, but this can effectively be extended by putting the ***swap file*** on a SSD. These SSD do not provide as fast storage as main RAM because of the ***bandwidth bottleneck*** of the ***bus*** they connect to.

FLASH-based
The second type of SSD uses flash memory to store data. These products, usually the same size as conventional storage are typically used as low power, rugged replacements for hard drives, especially in installations exposed to extreme conditions. To avoid confusion with the first type, these disks are generally referred to as "flash disks." Occasionally, the term is used to include keydrives, since they are technically solid-state and RAM-based. The flash memory cells tend to fail after around a million writes which made early devices unsuitable for storage which is often updated in place, such as swap files. To address this problem, some SSD vendors have introduced wear-levelling techniques that track usage and transparently relocate the data in highly utilized extents of storage to extents that have been less utilized.

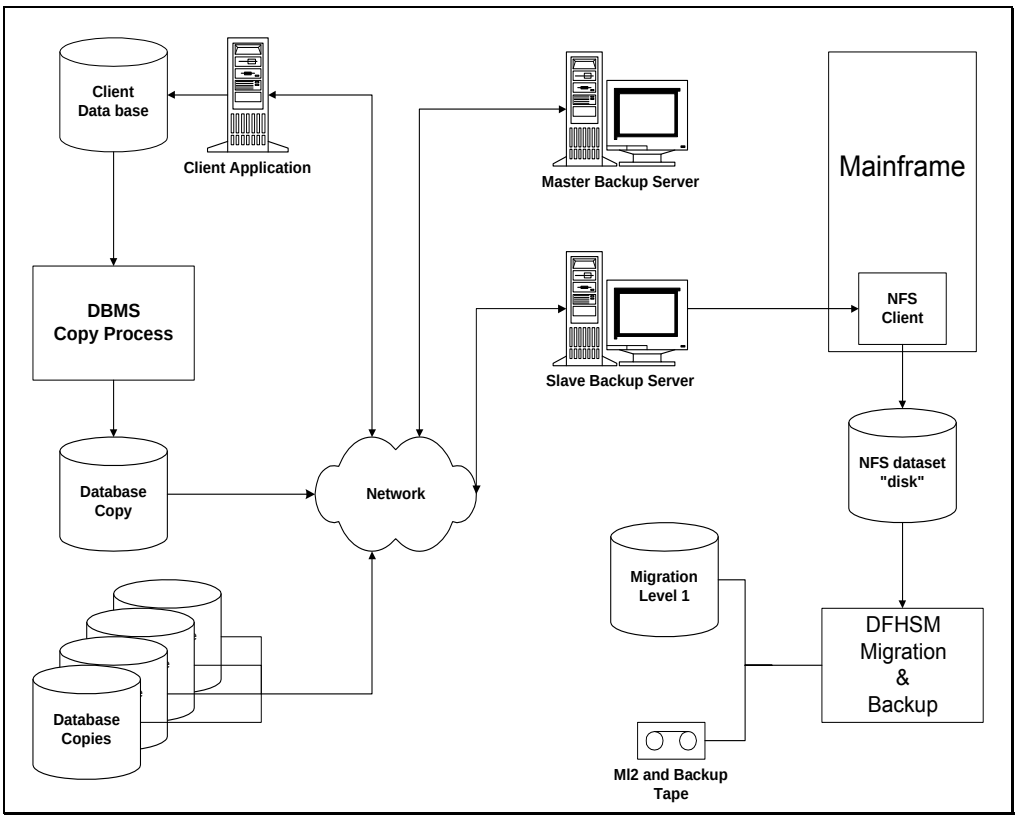
Another use for flash drives is running lightweight operating systems designed specifically for turning general-purpose PCs into network appliances comparable to more expensive routers and firewalls. In this situation a write protected flash drive containing the whole operating system is booted from. A similar system could boot from CD, floppy disk or a traditional hard drive but flash memory is a good choice because of very low power consumption and failure rate.

Both types of solid state disk are relatively expensive and low-capacity compared to hard drives. Their unique ability to solve specific problems of data storage, however, has created a small but consistent demand for these products since the 1970s.

- **RAM-Disk –** A RAM-Disk, Ramdisk or Ramdrive is a virtual solid state disk that uses a segment of active computer memory, RAM, as secondary storage, a role typically filled by hard drives. Access times are greatly improved, because RAM is approximately a thousand times faster than hard drives. However, the volatility of RAM means that data will be lost if power is lost, e.g. when the computer is turned off. RAM disks can be used to store temporary data or hold uncompressed programs for short periods.

Other Important information for Storage:

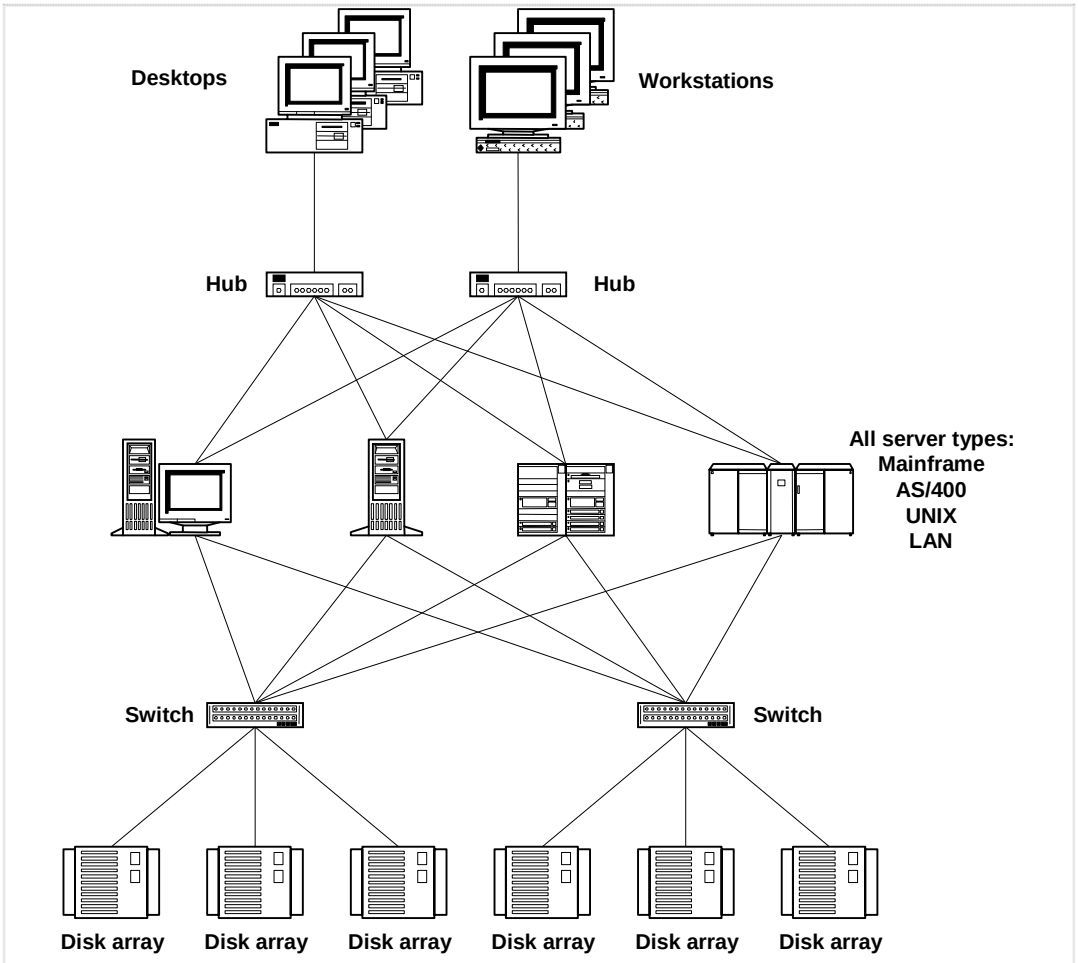
Move toward Hierarchical Storage Management
Another strong storage solution is Hierarchical Storage Management (HSM), which is designed to minimize storage costs. HSM uses “intelligent” software to move files between a hierarchy of storage devices that are ranked in terms of cost per megabyte of storage, speed of storage and retrieval, and overall capacity limits. Files are migrated along the hierarchy to less expensive forms of storage based on rules tied to the frequency of data access (illustrated below).



Agency & Enterprise Storage Area Networks (SAN)

Application availability is a top priority in the State's distributed networks. Server failover via clustering is occasionally deployed as part of a high availability architecture for continuous access, data protection, and disaster tolerance. This is where external storage, especially centralized, consolidated shared storage attached via a high performance fault-tolerant Storage Area Networks (SAN), should be employed.

The SAN method of attaching storage provides improvements in availability and performance. As shown in the figure below, SANs can be used to connect shared storage arrays, cluster servers for failover, interconnect mainframe disk or tape resources to distributed network servers and clients, and create parallel or alternate data paths for high performance computing environments. A SAN is simply an additional network constructed from storage interfaces and generally passing block-level protocols. The processes of moving large amounts of data across the storage network through data vaulting, data interchange, and disaster recovery are typically daily events. Application storage is better served by external shared repositories than by bus-attached storage. Clustered servers operate with improved availability and data accessibility if the data is shared and external. SANs also enable high availability data protection processes by moving data in the background, and placing no load on the server or the primary application network.


Desktops
Workstations
Hub
Hub
All server types:
Mainframe
AS/400
UNIX
LAN
Switch
Switch
Disk array
Disk array
Disk array
Disk array
Disk array
Disk array

6.5 Hardware / Infrastructure

Defines the physical devices, facilities and standards that provide the computing and networking within and between enterprises.

6.5.1 Servers / Computers

This refers to the various types of programmable machines which are capable of responding to sets of instructions and executing programs.

Note: Detail information for programmable machines is addressed within the Platform Independent/Dependant (6.1.2 and 6.1.3) sections.

Definitions of server / computer technologies include:

- **Enterprise Server** – A computer or device on a network that manages network resources and shared applications for multiple users.
- **Mainframe** – A very large computer capable of supporting hundreds, or even thousands, of users simultaneously. Mainframes support simultaneous programs.
- **Other Definitions –** If Needed

Other Important information for Storage:

Special Purpose System Servers

Additional categories of special-purpose (non-compliant) computing exist for a small segment of State use. Most of these requirements reside in the academic and research domains of the Kansas Regents/Universities. Because of the special nature and exclusive use of these computers, these systems are not to be included in State wide architectures, but are listed here for referential content.

- **Fax Server** - A fax server is a set of software running on a server computer which is equipped with one or more fax-capable modems attached to telephone lines. Its function is to accept documents from users, convert them into faxes, and transmit them, as well as to receive fax calls and either store the incoming documents or pass them on to users. Users may communicate with the server in several ways, through either a local network or the internet. In a big organization with heavy fax traffic, the computer hosting the fax server may be dedicated to that function, it which case the computer itself may also be known as a fax server.

- **VoiceMail Server -** Voicemail (or voice mail, vmail or VMS) is a centralized system of managing telephone messages for a large group of people. In its simplest form it mimics the functions of an answering machine, uses a standard telephone handset for the user interface, and uses a centralized, computerized system rather than equipment at the individual telephone. Voice mail systems are much more sophisticated than answering machines in that they can:

    - answer many phones at the same time
    - store incoming voice messages in personalized mailboxes associated with the user's phone number
    - enable users to forward received messages to another voice mailbox
    - send messages to one or more other user voice mailboxes
    - add a voice introduction to a forwarded message
    - store voice messages for future delivery
    - make calls to a telephone or paging service to notify the user a message has arrived in his/her mailbox
    - transfer callers to another phone number for personal assistance
    - play different message greetings to different callers.

Voicemail messages are stored on hard disk drives, media generally used by computers to store other forms of data. Messages are recorded in digitized natural human voice similar to how music is stored on a CD. To retrieve messages, a user calls the system from any phone, logs on using Touch-tones (clearing security), and his/her messages can be retrieved immediately. Many users can retrieve or store messages at the same time on the same voice mail system.

Many voicemail systems also offer an automated attendant facility. Automated attendants enable callers to a "main" business number to access directory service or self-route the call to various places such as a specific department, an extension number, or to an informational recording in a voice mailbox, etc.

- **Automated Call Distribution Server -** In telephony, an automated attendant (also auto attendant or auto-attendant, or sometimes autoattendant or AA) system allows callers to be automatically transferred to a user's extension without the intervention of a receptionist. A receptionist, who acts as the telephone operator, can be reached by pressing 0 on most systems. The automated attendant is a feature on most modern PBX and key phone systems.

  There is often a dial-by-name directory to find user on a system. The dial-by-name directory is usually set up last name followed by first name. Once the user name is announced the caller can press # and it will automatically ring the extension, sometimes after announcing the correct extension number for the caller's future reference.

- **Interactive Voice Response Server -** In telephony, interactive voice response, or IVR, is a computerized system that allows a person, typically a telephone caller, to select an option from a voice menu and otherwise interface with a computer system. Generally the system plays pre-recorded voice prompts to which the person presses a number on a telephone keypad to select the option chosen, or speaks simple answers such as "yes", "no", or numbers in answer to the voice prompts.

  The latest systems use natural language speech recognition to interpret the questions that the person wants answered. The newest trend is Guided Speech IVR which integrates live human agents into the design and workflow of the application to help the speech recognition with human context.

  Other innovations include the ability to speak complex and dynamic information such as an e-mail, news report or weather information using Text-To-Speech (TTS). TTS is computer generated synthesised speech and is no longer the robotic voice people associate with computers. Real voices are used to create the speech in tiny fragments that are glued together before being played to the caller.

- **Predictive Dialer Server -** A predictive dialer is a computerized system that automatically dials batches of telephone numbers for connection to agents assigned to sales or other campaigns. Predictive dialers are widely used in call centers.

- **High Performance Computing (HPC) -** refers to the use of (parallel) supercomputers and computer clusters, that is, computing systems comprised of multiple (usually mass-produced) processors linked together in a single system with commercially available interconnects. This is in contrast to mainframe computers, which are generally monolithic in nature. While a high level of technical skill is undeniably needed to assemble and use such systems, they can be created from off-the-shelf components. Because of their flexibility, power, and relatively low cost, HPC systems increasingly dominate the world of supercomputing. Usually, computer systems in or above the teraflop-region are counted as HPC-computers.

  The term is most commonly associated with computing used for scientific research. A related term, High-performance technical computing (HPTC), generally refers to the engineering applications of cluster-based computing (such as computational fluid dynamics and the building and testing of virtual

prototypes). Recently, HPC has come to be applied to business uses of cluster-based supercomputers, such as data warehouses, line-of-business (LOB) applications and transaction processing.

Many ideas for the new wave of grid computing were originally borrowed from HPC.

- **Veritas Cluster Server (VCS) -** is a High-availability cluster software, for Unix, Linux and Microsoft Windows computer systems, created by Veritas Software (now part of Symantec). It provides application cluster capabilities to systems running Databases, file sharing on a network, electronic commerce websites or other applications.

  Most Veritas cluster server implementations, attempt to build availability into a cluster, eliminating single points of failure by making use of redundant components like multiple network cards, storage area networks in addition to the use of VCS.

- **High Availability Clusters (HAC) -** High availability clusters (HAC) improve *availability* of applications by *failing* them over or *switching* them over in a group of systems as opposed to High Performance Clusters which improve *performance* of applications by allowing them to run on multiple systems simultaneously.

6.5.2 Embedded Technology Devices

This refers to the various devices and parts that make up a Server or Computer as well as devices that perform specific functionality outside of a Server or Computer.

Note: Detail information for Embedded Technology Devices is addressed within the Support Platform (6.1) and Database/Storage (6.4) sections.

6.5.3 Peripherals

A **peripheral** is a type of computer hardware that is added to a host computer in order to expand its abilities. More specifically the term is used to describe those devices that are optional in nature, as opposed to hardware that is either demanded, or always required in principle.

The term also tends to be applied to devices that are hooked up externally, typically through some form of computer bus like USB. Typical examples include joysticks, printers and scanners. Devices such as monitors and disk drives are not considered peripherals when they are not truly optional, and video capture cards are typically not referred to as peripheral because they are internal devices.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Printers	Impact Printers	Laser Printers InkJet Technologies	Multi-function office equipment that incorporate Print, Copier, Scanner, and Fax capabilities	
Scanners				

Definitions of Peripherals technologies include:

- **Printer** – A computer printer, or more commonly just a printer, is a device that produces hard copy (permanent human-readable text and/or graphics of documents stored in electronic form, usually on physical print media such as paper or transparencies). Many printers are primarily used as computer peripherals, and are permanently attached to a computer which serves as a document source. Other printers, commonly known as **network printer**, have

built-in network interfaces (typically wireless or Ethernet), and can serve as a hardcopy device for any user on the network. In addition, many modern printers can directly interface to electronic media such as memory sticks or memory cards, or to image capture devices such as digital cameras, scanners; some printers are combined with a scanners and/or fax machines in a single unit. A printer which is combined with a scanner can essentially function as a photocopier.

- **Scanner** – Device that can read text or illustrations printed on paper and translates the information into a form the computer can use. A scanner works by digitizing an image dividing it into a grid of boxes and representing each box with either a zero or a one, depending on whether the box is filled in.

6.5.4 Wide Area Networks

A data network typically extending a LAN outside a building or beyond a campus is known as a WAN. This is typically created by using circuits and routers to connect geographically separated LANs. The Internet is a WAN consisting of many interconnected private WANs and LANs.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Media – fiber		Multi/single mode		
Wireless point-to-point	Microwave	Laser Infrared Cellular RF		
Multiplexing	Time/frequency division Coarse wave division	Dense wave division	Optical Solutions	Dense wave division
Carrier services	Fractional T1 DS0 T1 IMA ISDN	DS3 T3 OC3 DSL cable modem	> OC3	
OSI Layer 2	FDDI Frame relay ATM	SONET ethernet		
OSI Layer 3	HDLC/SDLC Bisynch SNA X.25	Internet Protocol v4 MPLS multicast	Internet Protocol v6	Internet Protocol v6
Routing protocol	RIP	BGPv4 MPLS	MBGP IS-IS	

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		OSPF		
Quality of Service	Time/Frequency division multiplexing	DiffServ, over provisioning		

Description	Example Products	Notes
IP routers	Various models from Cisco, Juniper, Foundry, etc.	

Definitions of WAN technologies include:

- Frame Relay – A packet switching protocol for connecting devices on a Wide Area Network (WAN). Frame Relay networks in the U.S. support data transfer rates at T1 (1.544 Mbps) and T3 (45 Mbps) speeds.
- Asynchronous Transfer Mode (ATM) – A high bandwidth, high speed, controlled delay, fixed size packet switching and transmission system integrating multiple data types (voice, video, and data). Uses fixed size packets also known as "cells" (ATM is often referred to as "cell relay").
- Wave division multiplexing – the simultaneous transmission of multiple information flows over different wavelengths of light on fiber optic media. Coarse usually involves 16 or fewer wavelengths, dense is more.
- SONET – Synchronous Optical NETwork. A time division multiplexed technology used primarily by large carriers, but also suitable for statewide or large corporate networks.
- Internet Protocol (IP) – the protocol upon which the internet is based. Most applications use version 4, but version 6 (especially because of its larger address size) is becoming more common.
- Multi-Protocol Label Switching (MPLS) - A network routing protocol that is based on switching through the use of tag labels. MPLS is often used for performance or QOS reasons.
- Multicast – Sending a single sourced information flow to multiple recipients without having to create duplicate flows. This is often used with video transmission, group conferencing, and similar applications.
- Routing protocols – The “rules” used by IP routers to exchange route information, thereby enabling them to determine how to get a packet to its intended destination.
- Quality of Service (QOS) – Some type of guarantee (or at least bounds) or probability that specific transmission characteristics (e.g. latency, jitter, bandwidth) will be provided to an information flow. QOS is often needed for certain real-time applications, e.g. video transmissions over IP.

6.5.5 Local Area Networks

A network that interconnects devices over a geographically small area, typically in one building or a part of a building.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Media installation		EAI/TIA 568-B		
Media – copper	Coax Twinax Category 3	Category 5, 5E	Category 6E	
Media – fiber		Multi/single mode		
Wireless	Vendor proprietary	IEEE 802.11 a,b,g	802.11N	
Topology	Bus	Star		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
	ring	Mesh point-to-point		
OSI Layer 1	Repeaters Concentrators hubs 10Base2/10BaseT ethernet	100BaseT Ethernet 1000BaseT Ethernet	10GB BaseT	
OSI Layer 2	Bridges token ring NetBEUI Netbios	IEEE 802.1-3 ethernet ethernet switches VLANs (IEEE 802.1q)		
OSI Layer 3	DECnet Appletalk, IPX	Internet Protocol v4	Internet Protocol v6 SIP	Internet Protocol v6
Quality of Service	Time/Frequency division multiplexing	IEEE 802.3p, over provisioning		

Description	Example Products	Notes
Ethernet switches	Many products from vendors such as Cisco, Extreme, 3Com, Foundry, etc.	KanWIN moving to CISCO as a single vendor for both routing and switching.

Definitions of Local Area Network technologies include:

- Ethernet – A Local Area Network (LAN) standard that uses a bus or star topology and supports data transfer rates of 10 Mbps, 100 Mbps (Fast Ethernet) or 1 Gbps (gigabit Ethernet). The Ethernet specification served as the basis for the IEEE 802.3 standard, which specifies the physical and lower software layers. Ethernet uses the CSMA/CD access method to handle simultaneous demands. It is the most widely implemented LAN standard.
- Token Ring – A type of computer network in which all the computers are arranged (schematically) in a circle. A token, which is a special bit pattern, travels around the circle. To send a message, a computer catches the token, attaches a message to it, and then lets it continue to travel around the network. This technology is rarely used anymore.
- Virtual LAN (VLAN) – A network of computers that behave as if they are connected to the same wire even though they may actually be physically located on different segments of a LAN. VLANs are configured through software rather than hardware, which make them extremely flexible. VLANS are commonly used to logically group a set of systems/users with a need for the same type of “network profile” together.

6.5.6 Network Devices / Standards

A group of stations (computers, telephones, or other devices) connected by communications facilities for exchanging information. Connection can be permanent, via cable, or temporary, through telephone or other communications links. The transmission medium can be physical (i.e. fiber optic cable) or wireless (i.e. satellite).

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
OSI layer 1 ethernet	Repeaters Concentrators hubs			
OSI Layer 2 ethernet	Bridges	Ethernet switches		
OSI layer 3 IP		Routers Ethernet switches	CISCO Systems	

Definitions of Network Devices/Standards include:

- Hub – A common connection point for devices in a network. Hubs are commonly used to connect segments of a LAN. A hub contains multiple ports. When a packet arrives at one port, it is copied to the other ports so that all segments of the LAN can see all packets.
- Switch – In networks, a device that filters and forwards packets between LAN segments. Switches operate at the data link layer (layer 2) and sometimes the network layer (layer 3) of the OSI Reference Model and therefore support any packet protocol. LANs that use switches to join segments are called switched LANs or, in the case of Ethernet networks, switched Ethernet LANs.
- Router – A device or setup that finds the best route between any two networks, even if there are several networks to traverse. Like bridges, remote sites can be connected using routers over dedicated or switched lines to create WANs.
- Network Interface Card (NIC) – Often abbreviated as NIC, an expansion board you insert into a computer so the computer can be connected to a network. Most NICs are designed for a particular type of network, protocol, and media, although some can serve multiple networks.
- Transceivers – Short for transmitter receiver, a device that both transmits and receives analog or digital signals. The term is used most frequently to describe the Component in local area networks (LANs) that actually applies signals onto the network wire and detects signals passing through the wire. For many LANs, the transceiver is built into the network interface card (NIC). Some types of networks, however, require an external transceiver.
- Gateway – Gateways are points of entrance to and exit from a communications network. Viewed as a physical entity, a gateway is that node that translates between two otherwise incompatible networks or network segments.
- Integrated Services Digital Network (ISDN) – ISDN is a system of digital phone connections which has been available for over a decade. This system allows data to be transmitted simultaneously across the world using end to end digital connectivity.
- T1/T3 – T1 service delivers 1.544 Mbps. Typically channel into 24 DS0s, each capable of carrying a single voice conversation or data stream. The European T1 or E1 transmission rate is 2.048 Mbps. A T3 circuit communicates at 45 Mbps, or 28 T1 lines.
- Digital Subscriber Line (DSL) – Refers collectively to all types of digital subscriber lines, the two main categories being ADSL and SDSL. Two other types of DSL technologies are High data rate DSL (HDSL) and Very high DSL (VDSL).
- Firewall – This refers to the network device that is designed to prevent unauthorized access to or from a private network. Firewalls can be implemented in both hardware and software, or a combination of both. Firewalls are frequently used to prevent unauthorized Internet users from accessing private networks connected to the Internet, especially intranets. There are several types of firewall techniques and firewalls may implement one or more simultaneously. Packet filtering inspects inbound and outbound packets, validating against defined business rules.

Application gateways apply security rules against applications. Circuit level gateways apply security rules against physical connection attempts to and from the network. Proxy servers mask the internal requestor by inspecting and augmenting the packet header. Four common architectures of firewalls include the packet filtering router, the screened host firewall system, the dual homed host firewall, and the screened subnet firewall (with a DMZ), which is one of the most secure implementations.

6.5.7 Video Conferencing

Communication across long distances with video and audio contact that may also include graphics and data exchange. Digital video transmission systems typically consist of camera, codec (coder decoder), network access equipment, network, and audio system.

Chapter 7 Component Framework

The Component Framework Service Area defines the underlying foundation and technical elements by which Service Components are built, integrated and deployed across Component-Based and Distributed Architectures. The Component Framework consists of the design of application or system software that incorporates interfaces for interacting with other programs and for future flexibility and expandability. This includes, but is not limited to, modules that are designed to interoperate with each other at runtime. Components can be large or small, written by different programmers using different development environments and may be platform independent. Components can be executed on standalone machines, a LAN, Intranet or the Internet.

7.1 Security

Security defines the methods of protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide integrity, confidentiality and availability.

7.1.1 Certificates / Digital Signatures

Software used by a certification authority (CA) to issue digital certificates and secure access to information. The evolution of Public Key Infrastructure (PKI) is based on the verification and authentication of the parties involved in information exchange.

7.1.2 Supporting Security Devices

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Firewalls		Vendor specific Checkpoint NGR	Checkpoint NGX	
Virtual Private Networks (VPN)		Vendor specific Checkpoint VPN		
Intrusion Detection Systems (IDS)		Vendor specific ISS Real Secure	Anomaly detection	
Intrusion Protection Systems (IPS)		Vendor specific ISS Proventia ISS Real Secure Server Sensor		
Wireless		Vendor Specific Alcatel		
Virus Management		Vendor Specific McAfee EPO		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Authentication and Authorization		Win 2000 Radius Server		
Encryption	DES SSL	3DES TLS	AES	
Hard Authentication		RSA Tokens		
Vulnerability Scanners		Vendor Specific		
Patch Management	Microsoft SMS	Vendor specific	Microsoft System Center Operations Manager	
Monitoring Auditing Log Analysis		Vendor specific		
Workstation Policy Assurance		Vendor specific		
Spam Management		Vendor specific		
Web Filtering		Vendor Specific		
Secure Email		Tumbleweed IronMail		
System health/status		Nagios Whats Up		
System health/status		Nagios		

Description	Example Products	Notes
Firewalls	Checkpoint Cisco PIX Cisco IOS FW	
Virtual Private Networks (VPN)	Checkpoint VPN VPNet VPN Cisco VPN	
Intrusion Detection Systems (IDS)	SNORT Tripwire ISS	
Intrusion Protection Systems (IPS)	ISS Proventia ISS RealSecure Server Sensor ISS RealSecure Desktop sensor Symantec Norton Desktop Cisco CS-MARS	
Wireless	Alcatel Cisco Aironet	
Virus Management	Norton Symantec TrendMicro McAfee Sophos F-Secure	
Encryption	PGP	
Vulnerability Scanners	Nessus Eeye/Sourcefire ISS Internet Scanner NetRecon Computer Associates Qualys	
Patch Management	Microsoft SMS Patchlink Computer Associates Symantec Ipatch	
Monitoring Auditing Log Analysis	Sourcefire ARC Sight HP OpenView Patchlink	
Workstation Policy Assurance	Symantec Client Security Patchlink Tivoli SCM Pedestal	
Spam Management	Symantec Brightmail Sophos Anti-spam ISS Proventia M TrendMicro McAfee Webshield Computer Associates FrontBridge VamSoft Open Relay Filter	

Description	Example Products	Notes
Web Filtering	BlueCoat CISCO SurfControl Spam Assassin Iron Mail St. Bernard iPrism ISS Proventia M	

7.2 Presentation / Interface

This defines the connection between the user and the software, consisting of the presentation that is physically represented on the screen.

7.2.1 Static Display

This consists of the software protocols that are used to create a predefined, unchanging graphical interface between the user and the software.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Static Display		HTML W3C Standards ITEC Policy 1210 SGML		

Definitions of Static Display technology include:

- **Hyper Text Markup Language (HTML)** – The language used to create web documents and a subset of Standard Generalized Markup Language (SGML)

7.2.2 Dynamic / Server Side Display

This consists of the software that is used to create graphical user interfaces with the ability to change while the program is running.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Dynamic / Server Side Display		JSP ASP ASP.NET PHP		

Definitions of Dynamic / Server Side Display technology include:

- **Java Server Pages (JSP)** – JSP is part of Sun's J2EE architecture and provide template capabilities for presenting dynamically generated web content. ASPs are text files written in a combination of standard HTML tags, JSP tags, and Java code.
- **Active Server Pages (ASP)** – A web server technology from Microsoft that allows for the creation of dynamic, interactive sessions with the user.
- **Active Server Pages .Net (ASP.Net)** – ASP.NET is a set of technologies in the Microsoft .NET Framework for building web applications and XML web services. ASP.NET pages execute on the server and generate markup such as HTML, WML or XML that is sent to a desktop or mobile browser.

7.2.3 Content Rendering

This defines the software and protocols used for transforming data for presentation in a graphical user interface.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Content Rendering		DHTML XHTML CSS XSLT Javascript W3C Standards ITEC Policy 1210		

Definitions of Content Rendering technology include:

- **Dynamic HTML (DHTML)** – A collective term for a combination of new Hypertext Markup Language (HTML) tags and options, style sheets, and programming that will allow web pages that are more animated and more responsive to user interaction than previous versions of HTML.
- **Extensible HTML (XHTML)** – The W3C's recommendation for the next generation of HTML.
- **Cascading Style Sheets (CSS)** – A style sheet format for HTML documents endorsed by the World Wide Web Consortium. CSS1 (Version 1.0) provides hundreds of layout settings that can be applied to all the subsequent HTML pages that are downloaded.

7.2.4 Wireless / Mobile / Voice

This consists of the software and protocols used for wireless and voice enabled presentation devices.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Mobile Device presentation		WML	XHTMLMP VXML	

Definitions of Wireless / Mobile / Voice technology include:

- **Wireless Markup Language (WML)** – An XML-based protocol designed for Wireless devices.
- **XHTML Mobile Profile (XHTMLMP)** – XHTMLMP is designed for resource-constrained web clients that do not support the full set of XHTML features, such as mobile phones, PDAs, pagers and set top boxes. It extends XHTML Basic with modules, elements and attributes to provide a richer authoring language. XHTML replaces the Wireless Markup Language (WML).
- **Voice XML (VXML)** – VXML is an XML vocabulary for specifying IVR (Integrated Voice Response) Systems.

Other Important information for Wireless / Mobile / Voice

- W3C Mobile Web Initiative, http://www.w3.org/Mobile/ (accessed 6/29/2006)
- W3C, Voice Browser Activity, http://www.w3.org/Voice/, (accessed 6/29/2006)

7.3 Business Logic

Defines the software, protocol or method in which business rules are enforced within applications.

7.3.1 Platform Independent

Consists of all software languages that are able to execute and run on any type of operating system or platform

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Platform Independent		Java Enterprise Edition (J2EE) Java Standard Edition (J2SE) Java Servlet (JSR 53) Java Portlet API (JSR 168) JavaScript Enterprise Java Beans (EJB) Java XML-RPC SAS AS Perl PHP Python C, C++ COBOL	WSRP	

Definitions of Platform Independent technology include:

- **Enterprise Java Beans (EJB)** – a software component in Sun's J2EE platform, which provides a pure Java environment for developing and running distributed applications.
- **C, C++** – C is a procedure programming language. C++ is an object-oriented version of C that has been widely used to develop enterprise and commercial applications.
- **JavaScript** – A scripting language that runs within a web browser.
- **Java Servlet (JSR 53)** – Java Servlets provide reusable web components that can be incorporated into portals.
- **Java Portlet API (JSR 168)** – Java Portlet API enables interoperability between Portlets and Portals by defining APIs that address the areas of aggregation, personalization, presentation and security.
- **Web Services for Remote Portals (WSRP)** – WSRP defines an XML and web services standard that will allow the plug and play of visual, user facing web services with portals or other intermediary web applications.

- **Apache Ant** (or simply Ant) - an XML-based build scripting language used heavily by the Open Source community. Ant is a software tool for automating software build processes.

7.3.2 Platform Dependant

Consists of the programming languages and methods for developing software on a specific operating system or platform.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Mainframe	FORTRAN Assembler	Natural RPG		
Windows		Visual Basic Visual Basic .NET VB Script C# ASP ASP .NET PowerBuilder		
Other		PeopleTools		

Definitions of Platform Dependent include:

- **Visual Basic** – A version of the BASIC programming language from Microsoft specialized for developing Windows applications.
- **Visual Basic .Net (VB.Net)** – A version of the BASIC programming language from Microsoft specialized for developing Windows applications that is used within Microsoft's .NET environment.
- **C-Sharp (C#)** – An object-oriented programming language from Microsoft that is based on C++ with elements from Visual Basic and Java.
- **VB Script** – A scripting language from Microsoft. A subset of Visual Basic, VBScript is widely used on the web for both client processing within a web page and server side processing in Active Server Pages (ASPs).

7.4 Data Interchange

Define the methods in which data is transferred and represented in and between software applications.

7.4.1 Data Exchange

Data Exchange is concerned with the sending of data over a communications network and the definition of data communicated from one application to another. Data Exchange provides the communications common denominator between disparate systems.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Open Protocols for Databases		Java Database Connectivity (JDBC) Open Database Connectivity (ODBC)		Java Database Connectivity (JDBC) Open Database Connectivity (ODBC)
Transport Method		SMTP (Email) FTP (Batch Xfer) HTTP (Web Transport) Embedded Services Web Services (ex. WebDAV) Portal Services		SMTP (Email) FTP (Batch Xfer) HTTP (Web Transport) Embedded Services Web Services (ex. WebDAV) Portal Services
Definition of Data		XML Schemas EDI	XML Schemas	XML Schemas EDI
Messaging Template		Data Exchange Templates (DETs)		Data Exchange Templates (DETs)
Data Exchange Architecture		Data Synchronization Exchanges Data Publishing Exchanges		Data Synchronization Exchanges Data Publishing Exchanges
Error Handling and Failure Recovery				Implement error handling and recovery when constructing data exchanges
Authentication				
Verification				

Definitions of Data Exchange technologies include:

- **XMI** – Enables easy interchange of metadata between modeling tools (based on the OMG UML) and metadata repositories (OMG MOF based) in distributed heterogeneous environments. XMI integrates three key industry standards: XML, UML, and MOF. The integration of these three standards into XMI marries the best of OMG and metadata and modeling technologies, allowing developers of distributed systems to share object models and other metadata over the Internet.
- **XQuery** – A language used for processing and evaluating XML data. The XQuery language provides results of expressions allowing the use of evaluations to the implementation of XQuery.

- **Simple Object Access Protocol (SOAP)** – SOAP provides HTTP/XML based remote procedure call capabilities for XML Web Services.
- **Electronic Business using XML (be-XML)** – A modular suite of specifications that enables enterprises to conduct business over the Internet: exchanging business messages, conducting trading relationships, communicating data in common terms and defining and registering business processes.
- **Resource Description Framework (RDF)** – RDF provides a lightweight ontology system to support the exchange of knowledge on the web. It integrates a variety of web-based metadata activities including sitemaps, content ratings, stream channel definitions, search engine data collection (web crawling), digital library collections, and distributed authoring, using XML as interchange syntax. RDF is the foundation for the Semantic Web envisioned by Tim Berners-Lee – an extension of the current web in which information is given well-defined meaning, to better enable computers and people to work in cooperation.
- **Web Services User Interface (WSUI)** – WSUI uses a simple schema for describing a WSUI "component" that can be used in a portal to call backend SOAP and XML services. WSUI uses XSLT style sheets to construct user facing views to enable users to interact with the services.
- **XML Schema** – The formal definition of the structure and format of the data.
- **Data Exchange Templates** – A spreadsheet outlining each data element within the Schema along with definitions, validation rules, and example content. This is a more human readable version of XML Schema. The DET also defines minimal data content and data quality considerations.
- **Data Synchronization exchanges** – traditional regulatory exchanges where one or more data providers regularly submit a bulk set of data to a data consumer.
- **Data Publishing** – make data available to data consumers via the Internet.

Other Important information for Data Exchange

- http://www.soaprpc.com
- Environmental Information Exchange Network (www.exchangenetwork.net/).
- Links to documents and tools on data exchange schema are available on the Exchange Network website (http://www.exchangenetwork.net/dev_schema/).

7.5 Data Management

Data management is the management of all data/information in an organization. It includes data administration, the standards for defining data and the way in which people perceive and use it.

7.5.1 Database Connectivity

Defines the protocol or method in which an application connects to a data store or database.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Proprietary Protocols		Active Data Objects (ADO) Active Data Objects .Net (ADO.Net) Object Linking and Embedding/Database (OLE/DB) Data Access Objects (DAO) DB2 Connect		Support Open Protocols
Open Protocols		Java Database Connectivity (JDBC) Open Database Connectivity (ODBC)		Java Database Connectivity (JDBC) Open Database Connectivity (ODBC)

Preferences for Database Connectivity:

- Support Open Protocols
- Supports multiple clients for querying and reporting.
- Supports Two-Phase Commit.

Definitions of Database Connectivity technologies include:

- **Java Database Connectivity (JDBC)** – JDBC provides access to virtually any tabular data source from the Java programming language. It provides cross-DBMS connectivity to a wide range of SQL databases, and other tabular data sources, such as spreadsheets or flat files.
- **Open Database Connectivity (ODBC)** – A database programming interface from Microsoft that provides a common language for Windows applications to access databases on a network. ODBC is made up of the function calls programmers write into their applications and the ODBC drivers themselves.
- **Active Data Objects (ADO)** – A programming interface from Microsoft that is designed as "the" Microsoft standard for data access. First used with Internet Information Server, ADO is a set of COM objects that provides an interface to OLE DB. The three primary objects are Connection, Command and Record set.
- **Active Data Objects .Net (ADO.Net)** – ADO.Net is the data access component of the Microsoft's .NET Framework. It provides an extensive set of classes that facilitate efficient access to data from a large variety of sources, enable sophisticated manipulation and sorting of data.
- **Object Linking and Embedding/Database (OLE/DB)** – A Microsoft low-level API designed to provide connections to different data sources. OLE/DB allowed connectivity to ODBC-based SQL providers/sources as well as other formats such as text and comma delimited.

- **Data Access Objects (DAO)** – DAO is the Microsoft library for accessing Microsoft Jet engine data sources such as Microsoft Office based applications. DAO is replaced by ADO and ADO.Net.
- **DB2 Connect** – An IBM connectivity API to access DB2 sources.
- **Two-Phase Commit** – A computer technique used to ensure a transaction that updates multiple databases and successfully updates all the appropriate files throughout the distributed database environment. The two-phase commit guarantees that the work performed preserves the integrity of the data by verifying that each database involved in the transaction receives the modifications, applies them to the database, and that they are recoverable. If each database involved in the transaction reports these tasks were successfully accomplished, then each database is told to commit the transaction. The commit makes the changes permanent. The process is not considered finished until all the involved databases confirm that their commit completed successfully.

7.5.2 Reporting and Analysis

Consist of the tools, languages and protocols used to extract data from a data store and process it into useful information.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
System Architecture	Thick Client (Win32 Application)	Service Oriented Architecture o XML o SOAP o WSDL o WSRP		Service Oriented Architecture o XML o SOAP o WSDL o WSRP
Data Structure Support	Hierarchical Database	Relational Database Online Analytical Processing (OLAP) ROLAP MOLAP Java Online Analytical Processing (JOLAP)		Relational Database OLAP ROLAP MOLAP Java Online Analytical Processing (JOLAP)
Data Integration and Export Format	Flat Files (CSV, Fixed Width) Microsoft Excel	XML Microsoft Excel Adobe PDF	Extensible Business Reporting Language (XBRL) XML for Analysis	XML XBRL XML for Analysis

Description	Example Products	Notes
Enterprise Business Intelligence	Cognos 8 BI	• Simple and intuitive single query access to complex, multi-query, reports across heterogeneous data sources.
	Business Objects XI	• Four primary components: graphical designer, Web-based administrator, data server, metadata repository • Graphical interface for performing all tasks involved with building, testing, and managing an ETL job • End-to-end data lineage and impact analysis • Supports creation of semantic layers • Can create data marts from historical Crystal Reports instances • Graphical dataflow and workflow modeler • Extensive library of reusable transformations and functions • Interactive debugging and error-trapping • Distributed processing • Grid computing support • Comprehensive changed data capture (CDC) support • Batch and real-time data movement • Deep integration with ERP and CRM applications

Preferences for Reporting and Analysis:

- Support Open Database Connectivity Protocols.
- Support Graphical Interfaces for summarizing data, e.g. desktop dashboards.
- Support for ad-hoc and "canned" queries.
- Provide guided report creation as well as programmatic control of report creation.

Definitions for Reporting and Analysis Technologies Include:

- **Simple Object Access Protocol (SOAP) -** A lightweight XML-based messaging protocol used to encode the information in Web service request and response messages before sending them over a network.
- **Web Service Description Language (WSDL) -** an XML-formatted language used to describe a Web service's capabilities as collections of communication endpoints capable of exchanging messages.
- **Web Services for Remote Portlets (WSRP) -** WSRP defines how to plug remote web services into the pages of online portals and other user-facing applications.

- **Extensible Business Reporting Language (XBRL)** – Extensible Business Reporting Language (XBRL is an open specification which uses XML-based data tags to describe financial statements for both public and private companies.
- **Online Analytical Processing (OLAP)** – Decision support software that allows the user to quickly analyze information that has been summarized into multidimensional views and hierarchies.
- **Relational Online Analytical Processing (ROLAP)** - A product that provides multidimensional analysis of data, aggregates and metadata stored in an RDBMS. The multidimensional processing may be done within the RDBMS, a mid-tier server or the client. A 'merchant' ROLAP is one from an independent vendor which can work with any standard RDBMS.
- **Multidimensional Online Analytical Processing (MOLAP) -** a multidimensional optimization of OLAP. It is often applied to data warehousing information, which is stored in fact tables. In a star schema, these tables contain the basic information and dimensional links to supporting attribute tables.
- **XML for Analysis** – XML for Analysis uses the Simple Object Access Protocol (SOAP) to let Web browser based programs access backend data sources for data analysis. The specification allows companies to build online analytical processing (OLAP) and data mining applications that work over the web.

Chapter 8 Service Interface and Integration

The Service Interface and Integration Service Area defines the discovery, interaction and communication technologies joining disparate systems and information providers. SOAs leverage and incorporate Service Interface and Integration standards to provide interoperability and scalability.

8.1 Integration

Integration defines the software services enabling elements of distributed business applications to interoperate. These elements can share function, content, and communications across heterogeneous computing environments. In particular, service integration offers a set of architecture services such as platform and service location transparency, transaction management, basic messaging between two points, and guaranteed message delivery.

8.1.1 Middleware

Middleware increases the flexibility, interoperability, and portability of existing infrastructure by linking or "gluing" two otherwise separate applications.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Security		Shibboleth InCommon (identity management)		
Web Based		XML SOAP Web Services SOA		
Message Oriented		MOM: • IBM WebSphere MQ • MS Message Queue (MSMQ) JMS (Java Message Service) Other example products: • ActiveMQ • iPush Server • Oracle Advanced Queuing • SonicMQ • Tuxedo • Advanced Message Queuing Protocol (AMQP)		
Remote Procedure Calls		RPC Jave RMI • JRMP • RMI-IIOP		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Database Access		SQL: • PL/SQL • ISQL/w • Open ANSI SQL/92 • SPARQL Net8		
Transaction Processing Monitors		TMON		
Object Request Broker		Interface Description Language (IDL) CORBA COM DCOM COM+		

Definitions of middleware technology include:

- **Remote Procedure Call (RPC)** – RPC is a protocol allowing a program on a client computer to invoke a program on a server computer.
- **Message Oriented Middleware (MOM): IBM WebSphere MQ** – Software solution providing APIs, queue management, message routing, automatic failover, and workload balancing. Message Oriented Middleware (MOM) is software residing in both sides of the client/server architecture providing support for asynchronous calls, or messages, between applications. Message queues are used to track and store requests waiting for execution by the source application. Messaging allows otherwise complex programming and networking details to be abstracted from the developer.
- **Message Oriented Middleware (MOM): Microsoft Message Queue (MSMQ)** –Software technology providing synchronous and asynchronous message queuing, routing, and security. Message Oriented Middleware (MOM) is software residing in both sides of the client/server architecture providing support for asynchronous calls, or messages, between applications. Message queues are used to track and store requests waiting for execution by the source application. Messaging allows otherwise complex programming and networking details to be abstracted from the developer.
- **Database Access: PL/SQL** – Oracle's procedural extension to industry standard SQL. Database Access provides access to and across multiple database technologies in a distributed environment. Database Access is provided through the use of native database Application Programming Interfaces (APIs), client side APIs, or server side database gateways.
- **Database Access: ISQL/w** – Microsoft's implementation of ANSI SQL. Database Access provides access to and across multiple database technologies in a distributed environment. Database Access is provided through the use of native database Application Programming Interfaces (APIs), client side APIs, or server side database gateways.
- **Database Access: OPEN ANSI SQL/92)** – SQL is the information processing industry standard language of relational database management systems (RDMS). ANSI X3.1351992 (also referred to as SQL92 and ANSI SQL) is the industry standard for Database Language SQL. This standard promotes the portability and interoperability of database application programs and facilitates maintenance of database systems across heterogeneous data processing environments. SQL92 provides a standardized way for embedding SQL statements into application development languages. Database Access provides access to and across multiple database technologies in a

distributed environment. Database Access is provided through the use of native database Application Programming Interfaces (APIs), client side APIs, or server side database gateways.

- **Database Access: NET8** – NET8 (called SQL*NET prior to Oracle8) is Oracle's client/server middleware product that offers transparent connection from client tools to the database, or from one database to another. SQL*Net/ Net8 works across multiple network protocols and operating systems. Previous versions referred to as SQL*Net. Database Access provides access to and across multiple database technologies in a distributed environment. Database Access is provided through the use of native database APIs, client side APIs, or server side database gateways.
- **Transaction Processing Monitor** – Software providing synchronous messaging and queuing along with other transaction management services designed to support the efficient processing of high volumes of transactions. Core services include load balancing, rollback/commit, and recovery. Transaction Processing provides cost-effective scalability to applications and database systems by managing and throttling transactions on behalf of the database system.
- **Object Request Broker (ORB): Common Object Request Broker Architecture (CORBA)** – An architecture that enables objects to communicate with one another regardless of what programming language they were written in or what operating system they're running on. Object Request Broker (ORB) is a technology enabling distributed objects to communicate and exchange data with remote objects. ORB encapsulates the locality and implementation of the objects, allowing users to develop applications that leverage components by accessing the Components interface.
- **Object Request Broker (ORB): Component Object Model (COM)** – A software architecture created by Microsoft to design and build component-based applications. COM object capabilities are accessible from exposed interfaces. Object Request Broker (ORB) is a technology enabling distributed objects to communicate and exchange data with remote objects. ORB encapsulates the locality and implementation of the objects, allowing users to develop applications that leverage Components by accessing the Components interface.
- **Object Request Broker (ORB): Distributed Component Object Model (DCOM)** – An extension of the Component Object Model (COM) that allows COM components to communicate across network boundaries. Traditional COM components can only perform inter-process communication across process boundaries on the same machine. Object Request Broker (ORB) is a technology enabling distributed objects to communicate and exchange data with remote objects. ORB encapsulates the locality and implementation of the objects, allowing users to develop applications that leverage components by accessing the components interface.
- **Object Request Broker (ORB): Component Object Model + (COM+)** – COM+ is an extension of the COM that provides a runtime and services that are readily used from any programming language or tool, and enables extensive interoperability between Components regardless of how they were implemented. Object Request Broker (ORB) is a technology enabling distributed objects to communicate and exchange data with remote objects. ORB encapsulates the locality and implementation of the objects, allowing users to develop applications that leverage components by accessing the components interface.

Other Important information for Middleware

Refer to *Software Engineering (6.3) and Business Logic (7.3)* for additional information.

8.1.2 Enterprise Application Integration

Refers to the processes and tools specializing in updating and consolidating applications and data within an enterprise. EAI focuses on leveraging existing legacy applications and data sources so that enterprises can add and migrate to current technologies.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Querying		Z39.50 SRW\SRU		

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Harvesting		OAI-PMH		
Transforming and formatting		MPEG 21 XSLT ETL		
Business Process Management		BPM BPEL		

Definitions of Enterprise Application Integration technology include:

- **Business Process Management** – This process is responsible for the definition and management of cross application business processes across the enterprise and/or between enterprises.
- **Application Connectivity** – This process provides reusable, noninvasive connectivity with packaged software. This connectivity is provided by uni- or bidirectional adapters.
- **Transformation and Formatting** – This process is responsible for the conversion of data, message content, information structure, and syntax to reconcile differences in data amongst multiple systems and data sources. For other issues on Transformation and Formatting look at the ETL processes in the data management section of the document.

Other Important information for Enterprise Application Integration

See Enterprise Application Integration **service** section (14.3.2).

8.2 Interoperability

Interoperability defines the capabilities of discovering and sharing data and services across disparate systems and vendors.

The service areas below are defined in terms of **data** while the above definition of Interoperability includes discovery and sharing of **services**. One comprehensive method of discovery and sharing of services is Universal Description, Discovery and Integration (UDDI). This protocol is one of the major building blocks required for successful Web services. UDDI creates a standard interoperable platform that enables companies and applications to quickly, easily, and dynamically find and use Web services over the Internet. UDDI also allows operational registries to be maintained for different purposes in different contexts. UDDI is a cross-industry effort driven by major platform and software providers, as well as marketplace operators and e-business leaders within the OASIS standards consortium. http://www.uddi.org/.

The UDDI project takes advantage of World Wide Web Consortium (W3C) and Internet Engineering Task Force (IETF) standards such as Extensible Markup Language (XML), and HTTP and Domain Name System (DNS) protocols. UDDI uses Web Services Description Language (WSDL) to describe the services. WSDL is an XML-based language used to describe the services a business offers and to provide a way for individuals and other businesses to access those services electronically. WSDL is the cornerstone of the Universal Description, Discovery, and Integration (UDDI) initiative spearheaded by Microsoft, IBM, and Ariba. UDDI is an XML-based registry for businesses worldwide, which enables businesses to list themselves and their services on the Internet. WSDL is the language used to do this. WSDL is derived from Microsoft's Simple Object Access Protocol (SOAP) and IBM's Network Accessible Service Specification Language (NASSL). WSDL replaces both NASSL and SOAP as the means of expressing business.

Additionally, cross platform programming features are addressed by adopting early versions of the proposed Simple Object Access Protocol (SOAP) known as XML Protocol messaging specifications found at the W3C Web site. The UDDI protocol is the building block that will enable businesses to quickly, easily and dynamically find and transact with one another using their preferred applications.

UDDI.org has released the UDDI Version 3 specification. UDDI Version 3 builds on the vision of UDDI: a "meta service" for locating web services by enabling robust queries against rich metadata. Learn more about UDDI Version 3 by reading the UDDI Version 3 Features List.

Specifications and Standards for UDDI can be found at *http://www.soaprpc.com/specifications.html*; and for WSDL *http://searchwebservices.techtarget.com/sDefinition/0,,sid26_gci521683,00.html*.

8.2.1 Data Format / Classification

Defines the structure of a file. There are hundreds of formats, and every application has many different variations (database, word processing, graphics, executable program, etc.). Each format defines its own layout of the data. The file format for text is the simplest.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Text	PDF with encryption or without embedded fonts HTML versions	Plain text XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema and character encoding	PDF/A Open Document Format OpenOffice (.sxw)	Plain text XML (includes XSD/XSL/XHTML, etc.; with included or accessible schema and character encoding

	prior to 4.x	explicitly specified) PDF (*.pdf) (embedded fonts, no encryption) Rich Text Format (*.rtf) Cascading Style Sheets (*.css) HTML 4.x (*.htm) Microsoft Word (*.doc) Word Perfect (*.wpd) Postscript (*.ps) Microsoft Project (.mpp) Assessable to individuals with disabilities		explicitly specified) PDF/A Open Document Format Assessable to individuals with disabilities
Raster Image		Bitmap (*.bmp) JPEG/JFIF (*.jpg) JPEG2000 (uncompressed) GIF (*.gif) PhotoShop (*.psd)		TIFF (*.tif) (uncompressed) PNG (*.png) JPEG2000 (uncompressed)
Vector Graphics		Computer Graphics Metafile (*.cgm) Encapsulated Postscript (*.eps) Macromedia Flash (*.swf)		Scalable Vector Graphics (*.svg)
Spreadsheet	Quattro Pro (*.wb_)	Excel (*.xls)	OpenOffice (*.sxc)	Excel (.xls) OpenOffice (*.sxc)

Database	Paradox Filemaker Pro	DBF (*.dbf) Access (*.mdb) SQL DDL	OpenOffice (*sxc)	Access (*.mdb) OpenOffice (*.sxc)
Presentation	Corel Presentations	PowerPoint (*.ppt)	OpenOffice (*.sxi)	PowerPoint (*.ppt) OpenOffice (*.sxi)
Audio		AIFF(uncompressed) (*.aif, *.aiff) WAVE (LPCM only) (*.wav) RealNetworks 'Real Audio' (*.ra, *.rm, *.ram) Windows Media Audio (*.wma) MP3 (MPEG 1/2, Layer 3) (*.mp3)		AIFF(uncompressed) (*.aif, *.aiff) WAVE (LPCM only) (*.wav)
Video		MPEG1, MPEG2 (*.mpg, *.mpeg) AVI (compressed) (*.avi) QuickTime Movie (*.mov) MPEG 4 (*.mp4) RM (RealNetworks; 'Real Video') (*.rv) Windows Media Video (*.wmv) Motion JPEG2000 (*.mj2) AVI (*.avi) (uncompressed) Motion JPEG (*.avi, *.mov)		Motion JPEG2000 (*.mj2) AVI (*.avi) (uncompressed) Motion JPEG (*.avi, *.mov)

Definitions of Data Format / Classification technologies include:

- **Extensible Markup Language (XML)** – XML has emerged as the standard format for web data, and is beginning to be used as a common data format at all levels of the architecture. Many specialized vocabularies of XML are being developed to support specific Government and Industry functions.

- **XML Linking Language (XLINK)** – A language used to modify XML documents to include links, similar to hyperlinks, between resources. XLINK provides richer XML content through advanced linking integration with information resources.
- **Namespaces** – Namespaces are qualified references to URI (Uniform Resource Identifier) resources within XML documents.
- **Electronic Data Interchange (EDI)** – Defines the structure for transferring data between enterprises. EDI is used mainly used for purchase related information. ANSI X.12 refers to the approved EDI standards.
- **Definitions** – If Needed.

Other Important information for Data Format / Classification

- **FILExt –** http://filext.com/
- **PRONOM** – an on-line information system about data file formats and their supporting software products. Originally developed to support the accession and long-term preservation of electronic records. PRONOM holds information about software products, and the file formats which each product can read and write. http://www.nationalarchives.gov.uk/PRONOM/default.htm
- **DROID** (Digital Record Object Identification) – a platform-independent Java tool developed by The National Archives to perform automated batch identification of file formats. It is the first in a planned series of tools developed by The National Archives under the umbrella of its PRONOM technical registry service. http://droid.sourceforge.net
- **JHOVE** – This Java tool has been developed by Harvard University to allow the automatic identification, validation and characterization of a range of digital object types. http://hul.harvard.edu/jhove/jhove.html
- **Typed Object Model** – TOM is an object-based infrastructure for describing diverse data formats. TOM describes formats using type descriptions and also implements "type brokers", which receive and maintain descriptions of data formats, describe them to clients, and contact servers that interpret and translate data in those formats. http://tom.library.upenn.edu
- **Digital Formats of Library of Congress Collections** – This LC website is collecting technical information about file formats relevant to the Library's digital collections, in order to inform preservation decisions. It also includes an overview of factors which may affect the sustainability of formats over the long-term. http://www.digitalpreservation.gov/formats/index.shtml
- ***KU Digital Initiatives – Recommended Standards and Best Practices for Digital Projects,* University of Kansas, January 2003 (includes format standards and references)** http://kudiglib.ku.edu/planning/dli_standards.pdf
- **Florida Center for Library Automation Digital Archives Recommended File Formats** – (Recommended data formats for Preservation Purposes in the FCLA Digital Archives) Guidelines for preparing and submitting files to the FCLA Digital Archives. http://www.fcla.edu/digitalArchive/pdfs/recFormats.pdf

8.2.2 Data Types / Validation

Refers to standards used in identifying and affirming common structures and processing rules. This technique is referenced and abstracted from the content document or source data.

Category	Twilight Standard	Current Standard	Emerging Standard	Target
XML Data Types		Primitive Derived		Primitive Derived
Database		Exact Numerics Approximate Numerics Date and Time		Exact Numerics Approximate Numerics Date and Time

Category	Twilight Standard	Current Standard	Emerging Standard	Target
		Character Strings Unicode Character Strings Binary Strings Other Data Types		Character Strings Unicode Character Strings Binary Strings Other Data Types
Validation		XML Schemas applied against appropriate standardized profiles		XML Schemas applied against appropriate standardized profiles

Definitions of Data Types/Validation technologies include:

- **Document Type Definition (DTD)** – DTD is used to restrict and maintain the conformance of an XML, HTML, or SGML document. The DTD provides definitions for all tags and attributes within the document and the rules for their usage. Alterations to the document are validated with the referenced DTD.
- **XML Schema** – XML Schemas define the structure, content, rules and vocabulary of an XML document. XML Schemas are useful in automation through embedding processing rules.
- **Profiles** –
- **Database Validation** – Can be handled at user input, within the application code or with database constraints. Types of validation include data type validation, range checking, code checking and complex validation (verifying data against process-based business rules).

Examples of Data Types:

- XML Primitive – string, Boolean, decimal, float, double, duration, dateTime, time, date, gYearMonth, gYear, gMonthDay, gDay, gMonth, hexBinary, base64Binary, anyURI, QName, NOTATION
- XML Derived – normalized String, token, language, NMTOKEN, NMTOKENS, Name, NCName, ID, IDREF, IDREFS, ENTITY, ENTITIES, integer, nonPositiveInteger, negativeInteger, long, int, short, byte, nonNegativeInteger, unsignedLont, unsignedInt, unsignedShort, unsignedByte, positiveInteger
- Database Data Type Categories –
  Exact Numerics – bigint, int, decimal, numeric, smallint, tinyint, bit, money, smallmoney
  Approximate Numerics – float, real
  Date and Time – datetime, smalldatetime
  Character Strings – char, varchar, text
  Unicode Character Strings – nchar, nvarchar, ntext
  Binary Strings – binary, varbinary, image
- Other Data Types – cursor, sql_variant, table, timestamp, uniqueidentifier, xml, BLOB, CLOB

8.2.3 Data Transformation

Data Transformation consists of the protocols and languages that change the presentation of data within a graphical user interface or application.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data		XSL (XSLT, XPath,	OpenGIS©	XSL (XSLT, XPath,

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Transformation		XSL FO)		XSL FO) OpenGIS©

Description	Example Products	Notes
Data Transformation	See link below.	

Definitions of Data Transformation technologies include:

- **CSS** (Cascading Style Sheets) – A style sheet format for HTML documents endorsed by the World Wide Web Consortium. CSS1 (Version 1.0) provides hundreds of layout settings that can be applied to all the subsequent HTML pages that are downloaded. CSS2 (Version 2.0) adds support for XML, oral presentations for the visually impaired, downloadable fonts and other enhancements. CSS3 has been in ongoing development for several years and extends Version 2 with vertical text, elaborate borders and backgrounds, user interaction and speech, among the highlights. For updates, visit www.w3.org/Style/CSS/.
- **Style Sheet** - A master page layout used in document creation systems such as word processing, desktop publishing and the Web. The style sheet is a file that is used to store margins, tabs, fonts, headers, footers and other layout settings for a particular category of document. When a style sheet is selected, its format settings are applied to all the documents created under it, saving the page designer or programmer from redefining the same settings over and over again for each page.
- **XPath** - A sublanguage in an XSL style sheet that is used to identify XML elements for processing. It is also used to calculate numbers and manipulate strings. XPath and XSLT expressions are intertwined. For example, although XPath can reference a variable, the variable must be created and given a value in XSLT. XPath syntax is somewhat like the directory addressing in Unix, which uses a slash for the root directory as well as the separator between hierarchies.
- **XSL** (e**X**tensible **S**tylesheet **L**anguage) A standard from the W3C for describing a style sheet for XML documents. It is the XML counterpart to the Cascading Style Sheets (CSS) in HTML and is compatible with CSS2. XSL is made up of three components: (1) XSL Transformations (XSLT) is the processing language for XSL. It is used to convert XML documents into HTML or other document types and may be used independently of XSL. (2) XML Path Language (Xpath) is used to identify and select tagged elements within an XML document, and (3) XSL Formatting Objects (XSL FO) provides the format vocabulary.
- **XSLT (Extensible Style sheet Language Transform)** – Transforms XML document from one schema into another. Used for data transformation between systems using different XML schema, or mapping XML to different output devices.

Other Important information for Data Transformation

- Useful links:
  - List of XML conversion tools: http://www.perfectxml.com/soft.asp?cat=2
  - W3C (World Wide Web Consortium) Style Sheet page: http://www.w3.org/Style/.
  - OpenGIS© Standards: http://www.opengeospatial.org/standards
  - Kansas GIS Standards: http://www.kansasgis.org/docs/browse_docs.cfm

8.3 Interface

Interface defines the capabilities of communicating, transporting and exchanging information through a common dialog or method. Delivery Channels provide the information to reach the intended destination, whereas Interfaces allow the interaction to occur based on a predetermined framework.

8.3.1 Service Discovery

Service Discovery

Defines the method in which applications, systems or web services are registered and discovered.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Service Discovery		UDDI Z39.50 SRW / SRU		

Definitions of Service Discovery technology include:

- **Universal Description Discovery and Integration (UDDI)** – UDDI provides a searchable registry of XML Web Services and their associated URLs and WSDL pages.

8.3.2 Service Description / Interface

Defines the method for publishing the way in which web services or applications can be used.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Service Description / Interface		WSDL API – service specific XML gateways – service specific		

Definitions of Service Description / Interface technology include:

- **Web Services Description Language (WSDL)** – WSDL is an XML based Interface Description Language for describing XML Web Services and how to use them
- **Application Program Interface (API) / Protocol** – A language and message format used by an application program to communicate with the operating system or some other control program such as a database management system (DBMS) or communications protocol. APIs are implemented by writing function calls in the program, which provide the linkage to the required subroutine for execution. Thus, an API implies that some program module is available in the computer to perform the operation or that it must be linked into the existing program to perform the tasks.

Part 4

Kansas Service Component Reference Model (SRM)

Kansas Service Component Reference Model Overview

The SRM is a business-driven, functional framework classifying Service Components according to how they support business and performance objectives. It serves to identify and classify horizontal and vertical Service Components supporting federal agencies and their IT investments and assets. The model aids in recommending service capabilities to support the reuse of business components and services across the federal government.

The SRM is organized across horizontal service areas, independent of the business functions, providing a leverage-able foundation for reuse of applications, application capabilities, components, and business services.

Chapter 9 - Customer Service

- 9.1 Customer Relationship Management
  Call Center Management, Customer Analytics, Sales and Marketing, Product Management, Brand Management, Customer / Account Management, Customer Feedback, Surveys
- 9.2 Customer Preferences
  Personalization, Subscriptions, Alerts and Notifications
- 9.3 Customer Initiated Assistance
  Online Help, Online Tutorials, Self-Service, Reservations / Registration, Multi-Lingual Support, Assistance Request, Scheduling

Chapter 10 - Process Automation

- 10.1 Tracking and Workflow
  Process Tracking, Case Management, Conflict Resolution
- 10.2 Routing and Scheduling
  Inbound Correspondence Management, Outbound Correspondence Management

Chapter 11 - Business Management Services

- 11.1 Management of Process
  Change Management, Configuration Management, Requirements Management, Program / Project Management, Quality Management, Business Rule Management, Risk Management
- 11.2 Organization Management
  Workgroup / Groupware, Network Management
- 11.3 Investment Management
  Strategic Planning and Management, Portfolio Management, Performance Management
- Supply Chain Management
  Procurement, Sourcing Management, Inventory Management, Catalog Management, Ordering / Purchasing, Invoice / Requisition Tracking and Approval, Storefront / Shopping Cart, Warehouse Management, Returns Management, Logistics and Transportation

Chapter 12 - Digital Asset Services

- 12.1 Content Management
  Content Authoring, Content Review and Approval, Tagging and Aggregation, Content Publishing and Delivery, Syndication Management
- 12.2 Document Management
  Document Imaging and OCR, Document Referencing, Document Revisions, Library / Storage, document Review and Approval, Document Conversion, Indexing, Classification
- 12.3 Knowledge Management
  Information Retrieval, Information Mapping / Taxonomy, Information Sharing, Categorization, Knowledge Engineering, Knowledge Capture, Knowledge Distribution and Delivery, Smart Documents
- 12.4 Records Management
  Records Linking / Association, Document Classification, Document Retirement, Digital Rights Management

Chapter 13 - Business Analytical Services

- 13.1 Analysis and Statistics
  Mathematical, Structural / Thermal, Radiological, Forensics
- 13.2 Visualization
  Graphing / Charting, Imagery, Multimedia, Mapping / Geospatial / Elevation / GPS, CAD
- 13.3 Knowledge Discovery
  Data Mining, Modeling, Simulation
- 13.4 Business Intelligence
  Demand Forecasting / Management, Balanced Scorecard, Decision Support and Planning
- 13.5 Reporting
  Ad Hoc, Standardized / Canned, OLAP

Chapter 14 - Back Office Support

- 14.1 Data Management
  Data Exchange, Data Mart, Data Warehouse, Meta Data Management, Data Cleansing, Extraction and Transformation, Loading and Archiving, Data Recovery, Data Classification
- 14.2 Assets / Materials Management
  Computes / Automation Management
- 14.3 Development and Integration
  Legacy Integration, Enterprise Application Integration, Data Integration, Instrumentation and Testing, Software Development
- Human Resources
  Recruiting, Resume Management, Career Development and Retention, Time Reporting, Awards Management, Benefit Management, Retirement Management, Personnel Administration, Education / Training, Health and Safety, Travel Management
- Financial Management
  Billing and Accounting, Credit / Charge, Expense Management, Payroll, Payment / Settlement, Debt Collection, Revenue Management, Internal Controls, Auditing, Activity Based Management, Currency Translation
- Human Capital / Workforce Management
  Resource Planning and Allocation, Skills Management, Workforce Directory / Locator, Team / Org Management, Contingent Workforce Management, Workforce Acquisition / Optimization

Chapter 15 - Support Services

- 15.1 Security Management
  Identification and Authentication, Access Control, Encryption, Intrusion Detection, Verification, Digital Signatures, User Management, Role / Privilege Management, Audit Trail and Capture, Forensics, Incident Response, Risk Management
- 15.2 Collaboration
  Email, Threaded Discussions, Document Library, Shared Calendaring, Task Management
- 15.3 Search
  Query, Precision / Recall Ranking, classification, Pattern Matching
- 15.4 Communication
  Real Time Chat / Instant Messaging, Audio Conferencing, Video Conferencing, Event / News Management, Community Management, computer Telephony Integration, Voice Communications
- 15.5 Systems Management
  License Management, Remote Systems Control, System Resource Monitoring, Software Distribution, Issue Tracking
- 15.6 Forms Management
  Forms Creation, Forms Modification

Chapter 9 Customer Services

The Customer Services Domain defines the set of capabilities that are directly related to an internal or external customer, the business's interaction with the customer, and the customer driven activities or functions. The Customer Services Domain represents those capabilities and services that are at the front end of a business and interface at varying levels with the customer.

9.1 Customer Relationship Management

Capabilities within this Service Type are used to plan, schedule, and control the activities between the customer and the enterprise, both before and after a product or service is offered.

9.1.1 Call Center Management

Handle telephone sales and/or service to the end customer

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Service Desk and Incident Management and Resolution	Bendata HEAT Support Magic Home grown tools Peregrine Service Center	BMC Remedy BMC Magic HP OpenView HP Service Center FRS HEAT CA Unicenter Track-It (Small organizations)	Axios (ITIL based) Service Now Newmera Footprints	Integrated automated hardware, application, and network management tools to proactively alert service desk staff to potential problems
Change Management	Manual Processes Peregrine Service Center CA- Endeavor	BMC Magic		
Problem Management	Manual Processes	Manual Processes		

Definitions for Call Center Management include:

- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

9.1.2 Customer Analytics

Allow for the analysis of an organization's customers, as well as the scoring of third-party information as it relates to an organization's customers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Customer Analytics		Crystal Reports 9.0		

9.1.3 Sales and Marketing

Facilitate the promotion of a product or service and capture of new business

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Sales and Marketing				

9.1.4 Product Management

Facilitate the creation and maintenance of products and services

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Service Catalog		Web-based list		

9.1.5 Brand Management

Support the application of a trade name to a product or service as well as developing an awareness for the name

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Brand Management				

9.1.6 Customer / Account Management

Support the retention and delivery of a service or product to an organization's clients

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Customer Relationship Management (CRM)				

9.1.7 Contact and Profile Management

Provide a comprehensive view of all customer interactions, including calls, email, correspondence and meetings; also provides for the maintenance of a customer's account, business and personal information

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Customer Relationship Management (CRM)				

9.1.8 Partner Relationship Management

Provide a framework to promote the effective collaboration between an organization and its business partners, particularly members of the distribution chain (e.g., channel and alliance partners, resellers, agents, brokers, and dealers) and other third parties that support operations and service delivery to an organization's customers; includes performance evaluation of partners, if necessary

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Partner Relationship Management				

9.1.9 Customer Feedback

Is used to collect, analyze and handle comments and feedback from an organization's customers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Customer Feedback				

9.1.10 Surveys

Are used to collect useful information from an organization's customers.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Assessment tools		Survey Monkey	TechQual	

9.2 Customer Preferences

Capabilities within this Service Type allow an organization's customers to change a user interface and the way that data is displayed.

9.2.1 Personalization

Change a user interface and how data is displayed

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Portal – personalization		uPortal Kansas.gov portal Kansas.gov content management system	Microsoft SharePoint MySites	

9.2.2 Subscriptions

Allow a customer to join a forum, listserv, or mailing list

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Listserv		Listproc Majordomo		
Subscription feed services		RSS		

9.2.3 Alerts and Notifications

Allow a customer to be contacted in relation to a subscription or service of interest

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Alerts and Notifications		Emergency Notification Systems		

9.3 Customer Initiated Assistance

Capabilities within this Service Type allow customers to proactively seek assistance and service from an organization.

9.3.1 Online Help

Provide an electronic interface to customer assistance

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Online Help		Web ITEC Policy 1210		

9.3.2 Online Tutorials

Provide an electronic interface to educate and assist customers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Online Tutorials		Web ITEC Policy 1210		

9.3.3 Self-Service

Allow an organization's customers to sign up for a particular service at their own initiative

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Self-Service		Web forms ITEC Policy 1210		

9.3.4 Reservations / Registration

Allow electronic enrollment and confirmations for services

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Reservations / Registration		In-house registration system		

9.3.5 Multi-Lingual Support

Allow access to data and information in multiple languages

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Multi-Lingual Support				

9.3.6 Assistance Request

Support the solicitation of support from a customer

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Assistance Request				

9.3.7 Scheduling

Define the set of capabilities that support the plan for performing work or service to meet the needs of an organization's customers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Project management		MS Project Manager		

Chapter 10 Process Automation

The Process Automation Services Domain defines the set of capabilities that support the automation of process and management activities that assist in effectively managing the business. The Process Automation Services domain represents those services and capabilities that serve to automate and facilitate the processes associated with tracking, monitoring, and maintaining liaison throughout the business cycle of an organization.

10.1 Tracking and Workflow

Capabilities within this Service Type provide automatic monitoring and routing of documents to the users responsible for working on them to support each step of the business cycle.

10.1.1 Process Tracking

Allow the monitoring of activities within the business cycle

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Access		Supports both Internal and External Access		Supports both Internal and External Access
Security		Usage of X.509 electronic / digital signatures or certificates Auditing capabilities		Usage of X.509 electronic / digital signatures or certificates Auditing capabilities
Processing		Parallel processing		Parallel processing
Workflow		Process definition and mapping Routing Event triggers based on time, date, process completion, etc. Group assignments Individual assignments Status reporting Can be tailored to domain-specific processes such as telecommunications management, service desk event management, document		Process definition and mapping Routing Event triggers based on time, date, process completion, etc. Group assignments Individual assignments Status reporting Can be tailored to domain-specific processes such as telecommunications management, service desk event management, document processing, personnel hiring, interlibrary loan

		processing, personnel hiring, interlibrary loan requests, etc.		requests, etc.

Description	Example Products	Notes
Process Tracking	• Ad-hoc Workflow • FormSoft Forms Warehouse • ImageNow • PeopleAdmin • BMC Remedy • ACE*Comm • Webtrieve • OCLC Interlibrary Loan system • RAPID ILL system	Twilight Technology

Other Important information for Process Tracking

- Please also refer to the *Document Management / Content Management* (13.2) and S*ervice Desk / Call Center (9.1)* sections.

10.1.2 Case Management

Manage the life cycle of a particular claim or investigation within an organization to include creating, routing, tracing, assignment and closing of a case as well as collaboration among case handlers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Case Management		Data and communications protocol interoperability for inter-agency communications within the State, within the region, or nationally (i.e. communication with federal agencies)		Data and communications protocol interoperability for inter-agency communications within the State, within the region, or nationally (i.e. communication with federal agencies)

Description	Example Products	Notes
Case Management	Specific Line-of-Business applications	Examples include law enforcement, corrections, social services, hospitals / medical centers, immigration, revenue, etc.

10.1.3 Conflict Resolution

Support the conclusion of contention or differences within the business cycle

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		Manual Accounting and Budget Processes		
Business Rules		Rules composition Rules deployment Run-time engine Component process and terminology definition o Matching o Conflict resolution o Action Rule set tracking Policy testing tools Authorization Domain-specific business rules languages	Rule Interchange Format	Rules composition Rules deployment Run-time engine Component process and terminology definition o Matching o Conflict resolution o Action Rule set tracking Policy testing tools Authorization Rule Interchange Format

Definitions of Conflict Resolution technology include:

- **Conflict Resolution:** The conflict resolution stage compares all candidate rules on an agenda to determine if any execution sequencing conflicts exist in the logic of the rules. If any are determined, the conflicts are resolved according to the execution sequence defined by the user. The engine then sorts the agenda based on priority (high to low). The rule with the highest priority will execute first.

  BizTalk Server Business Rules Framework
  http://www.microsoft.com/technet/prodtechnol/biztalk/biztalk2004/planning/business-rules-framework-overview.mspx

- **Rule Interchange Format:** a core rule language plus extensions which together allow rules to be translated between rule languages and thus transferred between rule systems.

  W3C Rule Interchange Format Working Group
  http://www.w3.org/2005/rules/wg

Other Important information for Conflict Resolution

- References for (business) rule languages: http://xml.coverpages.org/ni2005-02-21-a.html#ruleLanguages

10.2 Routing and Scheduling

Capabilities within this Service Type provide automatic directing, assignment, or allocation of time for a particular action or event.

10.2.1 Inbound Correspondence Management

Manage externally initiated communication between an organization and its stakeholders.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Correspondence Management		Ability to capture and manage both electronic and paper inbound communications Security for confidential correspondence Workflow scheduling Classification / categorization Rules Metadata creation and management Search Tracking Data mining Analysis Response preparation Disposition Compliance management		Ability to capture and manage both electronic and paper inbound communications Security for confidential correspondence Workflow scheduling Classification / categorization Rules Metadata creation and management Search Tracking Data mining Analysis Response preparation Disposition Compliance management

Definitions of Inbound Correspondence Management technology include:

- **Correspondence Management:** The application of records management principles to written communications.

Pearce-Moses, Richard, *A Glossary of Archival and Records Terminology*, Society of American Archivists, http://www.archivists.org/glossary/term_details.asp?DefinitionKey=640

Other Important information for Inbound Correspondence Management

- Rosi, Jan, "Correspondence Management: Taking the Fast Lane from the Inbox to the Outbox", KMWorld, June 1 2006, http://www.kmworld.com/Articles/ReadArticle.aspx?ArticleID=15847

- Refer to the Records Management section (12.4)

10.2.2 Outbound Correspondence Management

Manage internally initiated communication between an organization and its stakeholders

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Correspondence Management		Ability to capture and manage both electronic and paper outbound communications Security for confidential correspondence Workflow scheduling Classification / categorization Rules Metadata creation and management Search Tracking Data mining Analysis Response preparation One-to-one distribution One-to-many distribution Disposition Compliance management		Ability to capture and manage both electronic and paper outbound communications Security for confidential correspondence Workflow scheduling Classification / categorization Rules Metadata creation and management Search Tracking Data mining Analysis Response preparation One-to-one distribution One-to-many distribution Disposition Compliance management

Definitions of Outbound Correspondence Management technology include:

- **Correspondence Management:** The application of records management principles to written communications.

Pearce-Moses, Richard, *A Glossary of Archival and Records Terminology*, Society of American Archivists, http://www.archivists.org/glossary/term_details.asp?DefinitionKey=640

Other Important information for Outbound Correspondence Management

- Rosi, Jan, "Correspondence Management: Taking the Fast Lane from the Inbox to the Outbox", KMWorld, June 1 2006, http://www.kmworld.com/Articles/ReadArticle.aspx?ArticleID=15847
- Refer to the Records Management section (12.4)

Chapter 11 Business Management Services

The Business Management Services Domain defines the set of capabilities that support the management of business functions and organizational activities that maintain continuity across the business and value-chain participants. The Business Management Services Domain represents those capabilities and services that are necessary for projects, programs and planning within a business operation to successfully be managed.

11.1 Management of Process

Capabilities within this Service Type regulate the activities surrounding the business cycle of an organization.

11.1.1 Change Management

Control the process for updates or modifications to the existing documents, software or business processes of an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Document Management		Microsoft SharePoint	ITIL CMDB BMC Remedy	

11.1.2 Configuration Management

Control the hardware and software environments, as well as documents of an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
			ITIL Configuration Management	

11.1.3 Requirements Management

Gather, analyze and fulfill the needs and prerequisites of an organization's efforts

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Requirements Management		Manual Process	ITIL Process Management	

Definitions for Requirements Management include:

- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

11.1.4 Program / Project Management

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Program / Project Management		Manual Process State of Kansas	ITIL Process Management Version Three	State of Kansas PMM standards

		PMM standards	Project Portfolio Management	

Definitions for Program / Project Management include:

- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

11.1.5 Governance / Policy Management

Influence and determine decisions, actions, business rules and other matters within an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Governance / Policy Management		Manual Process	ITIL Process Management Workflow Systems (such as those in BMC Remedy or AceComm)	
Rules		Manual Process		
Environment		Dynamic Policy Microsoft SharePoint Document Management Systems		

Definitions for Governance / Policy Management include:

- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

11.1.6 Quality Management

Help determine the level that a product or service satisfies certain requirements

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Quality Management		Manual Process Assessment Measurements LIBqual	ITIL Processes RLG/NARA Audit Checklist for Certifying Digital Repositories	

			TECHqual	

Other Important information for Quality Management

- RLG/NARA Task Force on Digital Repository Certification, 2005, *Audit Checklist for Certifying Digital Repositories*, <http://www.rlg.org/en/pdfs/rlgnara-repositoriesschecklist.pdf>
- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

11.1.7 Business Rule Management

Manage the enterprise processes that support an organization and its policies

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Business Rule Management		ISO 27000 Security Standards Security Auditing Rules Engine standards such as JSR-94 Java Rule Engine API		Security Auditing Rules Engine standards such as JSR-94 Java Rule Engine API

11.1.8 Risk Management

Support the identification and probabilities or chances of hazards as they relate to a task, decision or long-term goal

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Risk Management		Kansas PMM standards Security auditing ISO 27000 security standard		Kansas PMM standards Security auditing

11.2 Organizational Management

Capabilities within this Service Type support both collaboration and communication within an organization.

11.2.1 Workgroup / Groupware

Support multiple users working on related tasks

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Portals		Shared document capability Channel support Role-based Contextual Seamless Guided Multimedia supportMulti-platform access support (desktop, laptop, mobile device, etc.) ITEC Policy 1210 Compliance Interoperability with multiple browsers		Shared document capability Channel support Role-based Contextual Seamless Guided Multimedia support Multi-platform access support (desktop, laptop, mobile device, etc.) ITEC Policy 1210 Compliance Interoperability with multiple browsers
E-mail		Shared Document capability Interoperability with multiple platforms, operating systems, and/or browsers		Shared Document capability Interoperability with multiple platforms, operating systems, and/or browsers
Meeting / Conferencing		Shared Document capability Interoperability with multiple platforms, Operating systems, and/or browsers		Shared Document capability Interoperability with multiple platforms, Operating systems, and/or browsers
Instant Messaging		Group chat Individual chat Interoperability with multiple platforms, operating systems		Group chat Individual chat Interoperability with multiple platforms, operating systems

Blog (Weblog) platforms		Authoring Editing Linking Web syndication Ability to publish to the web Multi-language support Ability for other users to post comments ITEC Policy 1210 Compliance	Photo-blogging Video-blogging	
Wiki platforms		Authoring Editing Linking Web syndication Ability to publish to the web Multi-language support Ability for group editing and commenting ITEC Policy 1210 Compliance		

Description	Example Products	Notes
Portals	• uPortal • Sharepoint • Xythos	
Blog (Weblog) Platforms	• Drupal • Roller • Six Apart – Movable Type • Six Apart – TypePad • Telligent Systems • Traction Software • UserLand Software • WordPress	

Description	Example Products	Notes
	• iUpload	
Instant Messaging	• Jabber • AOL • Yahoo! • Microsoft • Integrated domain-specific products such as OCLC's Questionpoint (library application)	
Wiki platforms	• Twiki • Minerva • TikiWiki • Zwiki	
Email	See Collaboration section	

Definitions of Workgroup / Groupware technology include:

- **Weblog software** (also called **blog software** or **blogware**) is a category of software which consists of a specialized form of Content Management Systems specifically designed for creating and maintaining weblogs.
- **Wiki software** is a type of collaborative software that runs a Wiki system. This typically allows web pages to be created and edited using a common web browser. It is usually implemented as a server-side script that runs on one or more web servers, with the content generally stored in a relational database management system, although some implementations use the server's file system instead.

Other Important information for Workgroup / Groupware

- Trends in Collaboration:
  - Collaboration capabilities continue to move into the infrastructure; tension develops between application and infrastructure approaches
  - Microsoft emerges as an early collaboration platform leader; IBM and Oracle expend great energy trying to keep up
  - Application vendors enrich their collaboration features
  - Organizations swap out point products for infrastructure and application features
  - Market consolidation continues
  - Formal collaboration and interoperability standards are needed

  Driver, Erica, *Trends 2005: Collaboration*, December 10, 2004, Forrester Research, Inc.
  http://www.forrester.com/Research/Document/0,7211,36024,00.html

- Li, Charlene, *The Forrester Wave™: Blogging Platforms, Q2 2006,* June 27, 2006, Forrester Research Inc., http://www.forrester.com/Research/Document/0,7211,37147,00.html

- See section on *Collaboration* (16.2) for additional information.

11.2.2 Network Management

Monitor and maintain a communications network in order to diagnose problems, gather statistics and provide general usage.

11.3 Investment Management

Capabilities within this Service Type manage the financial assets and capital of an organization.

11.3.1 Strategic Planning and Management

Support the determination of long-term goals and the identification of the best approach for achieving those goals

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Financial Modeling and Reporting		Excel Access SAS Spreadsheets Financial programs (Peachtree, QuickBooks) SAP Financials In-house developed Bloomberg PeopleSoft		Interoperability between disparate Financial systems (XBML) Bloomberg Terminals PeopleSoft

Definitions for Strategic Planning and Management include:

- Financial Modeling and Reporting – The determination of long term goals and best practices to support financial systems.
- Bloomberg – Proprietary software providing a portal into the Investment world tracking bonds,

11.3.2 Portfolio Management

Support the administration of a group of investments held by an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Financial Portfolio Mgmt		In-house developed Portfolio and Investment Planning DESKTOP Excel Quicken Money		GAAP Compliant

		SAS		
Investment Tracking		The Vault – In house developed Pool Manager – Proprietary		
IT Project Portfolio Mgmt		Components could include: Application Portfolio Mgmt Project Portfolio Mgmt Project and Budget Mgmt Demand Management Resource Management Time Tracking Client Management Performance Management Knowledge Management Financial Management Workflow Reporting		Components could include: Portfolio and Investment Planning Application Portfolio Mgmt Project Portfolio Mgmt Project and Budget Mgmt Demand Management Resource Management Time Tracking Client Management Performance Management Knowledge Management Financial Management Workflow Reporting

Definitions for Portfolio Management include:

- IT Project Portfolio Mgmt (PPM). – PPM represents a shift away from one-off, ad hoc approaches to Project Management. PPM establishes a set of values, techniques and technologies that enable visibility, standardization, measurement and process improvement. PPM enables organizations to manage the continuous flow of projects from concept to completion.

11.3.3 Performance Management

Measure the effectiveness of an organization's financial assets and capital

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Investment Tracking and Mgmt		The Vault Pool Manager		GAAP Compliance

Chapter 12 Digital Asset Services

The Digital Asset Services Domain defines the set of capabilities that support the generation, management, and distribution of intellectual capital and electronic media across the business and extended enterprise. A *digital asset* is an electronic object that has value for some purpose. It may have been created digitally or it may have been digitized from a non-digital original source. Examples of digital assets include word processing documents, databases, websites, organizational records, digital recordings of musical performances, etc. ("Preservation Planning for Digital Information: Final Report of the HVC[2] Digital Preservation Task Force," KU, November 2004)

Trends Toward Integrating Services Into One Technology

Digital Asset Services-related technologies have traditionally fallen into four functional categories (paralleling the four service types outlined below): content management, document management, knowledge management, and records management. Over the past five years or so, vendors in this technology space have been acquired by larger companies who have focused on building products that integrate the categories as modules within their product line, commonly referred to as "Enterprise Content Management" (ECM). This trend will continue as large vendors attempt to meet the demand by their customers requiring access to digital assets across the organization, and needing to manage those assets from the time of creation onward. There is also a trend with some organizations to build systems based on open source products, bringing together the various functionalities to manage digital information across the entire enterprise.

Digital Asset Life-Cycle Management

The **ISO Archiving Standard** defines the International Organization for Standardization (ISO) Reference Model for an **Open Archival Information System** (OAIS Reference Model http://ssdoo.gsfc.nasa.gov/nost/isoas/ref_model.html). An OAIS is an archive consisting of an organization of people and systems that has accepted the responsibility to preserve information and make it available for a **Designated Community**. It meets a set of such responsibilities as defined in the standard, and this allows an OAIS archive to be distinguished from other uses of the term 'archive'. The term 'Open' in OAIS is used to imply that this Recommendation, as well as future related Recommendations and standards, are developed in open forums, and it does not imply that access to the archive is unrestricted.

"Digital Asset Lifecycle Management" is the process by which access to usable and accurate digital information is managed and maintained over time (Based on "Preservation Planning for Digital Information: Final Report of the HVC[2] Digital Preservation Task Force," KU, November 2004). Effective lifecycle management requires that we understand and attend to preservation from the time of creation throughout the full lifecycle of digital objects. This lifecycle has two key dimensions:

- A chronological dimension (see figure 1), which highlights the various stages of an object's life from the point of creation forward. Examples of these stages include being filed awaiting further work; being copied by a new user into a new version; being consulted by a user authorized to read it but not change it; being copied by a systems administrator as part of a nightly backup; being copied into slow-access storage; being transformed so as to be usable with new versions of software.

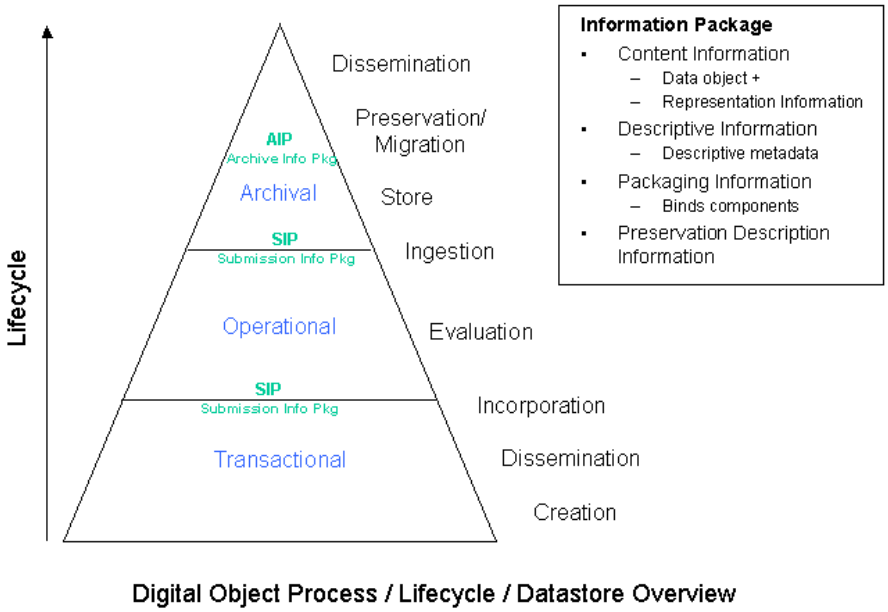


Figure 1: The Chronological Dimension of the Information Lifecycle

Archival Information Package (AIP): An Information Package, consisting of the Content Information and the associated Preservation Description Information (PDI), which is preserved within an OAIS.

Archival Information Unit (AIU): An Archival Information Package whose Content Information is not further broken down into other Content Information components, each of which has its own complete Preservation Description Information. It can be viewed as an 'atomic' AIP. An example of an AIU would be a table of numbers representing temperatures in a certain region with all the associated documentation describing how and where the temperatures were measured, what instruments were used to make the measurements, who made the measurements, why they were made, what processing has been performed on the measurements and who has had custody of these measurements since they were first created, how the measurements relate to other information, how the measurements can be uniquely referenced by others, etc.

Information Package: The Content Information and associated Preservation Description Information which is needed to aid in the preservation of the Content Information. The Information Package has associated Packaging Information used to delimit and identify the Content Information and Preservation Description Information.

An Information Package is a conceptual container of two types of information called **Content Information** and **Preservation Description Information (PDI)**. The Content Information and PDI are viewed as being encapsulated and identifiable by the **Packaging Information**. The resulting package is viewed as being discoverable by virtue of the **Descriptive Information**.

Submission Information Package (SIP): An Information Package that is delivered by the Producer to the OAIS for use in the construction of one or more AIPs.

- A functional dimension (see figure 2), which highlights the different roles that are occupied through an object's interaction with creators, editors, users, stewards, policymakers, and others over the course of its lifetime.

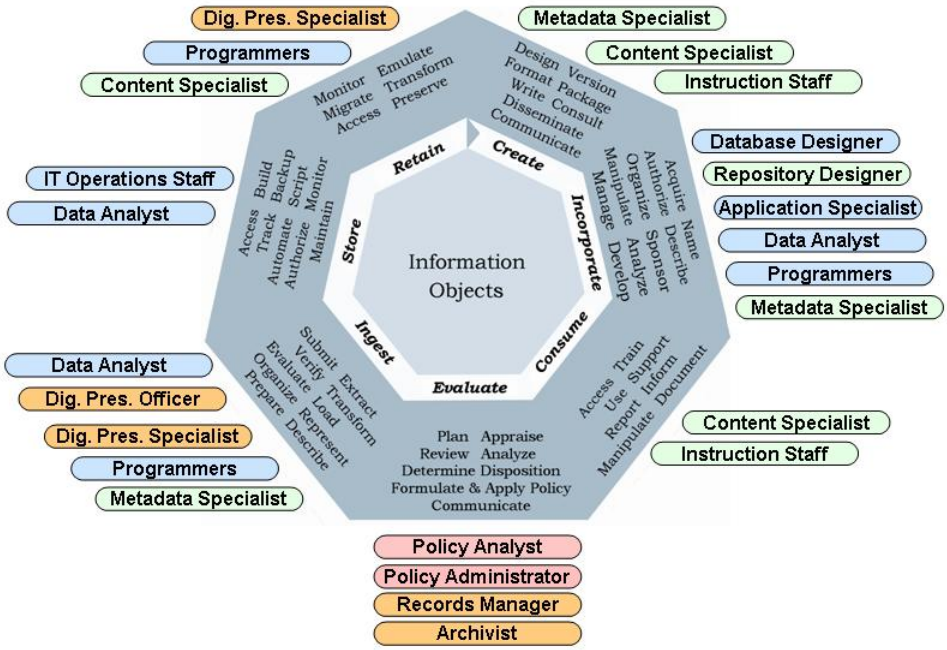


Figure 2: The Functional Dimension of the Information Lifecycle

Access: The OAIS entity that contains the services and functions which make the archival information holdings and related services visible to Consumers.

Ingest: The OAIS entity that contains the services and functions that accept Submission Information Packages from Producers, prepares Archival Information Packages for storage, and ensures that Archival Information Packages and their supporting Descriptive Information become established within the OAIS.

Create: Creation brings ideas into existence as information objects representing data, geography, literature, art, or other forms of usable expression. In the lifecycle of information, the role of creator is shared by individuals, organizations, publishers and information systems.

Another role important in the creation phase if objects must persist is the role of digital consultant. Consultants give creators early awareness of the impact that various choices for format, file standards, descriptive metadata, and dissemination hold for the long term viability of the their objects. Moving digital preservation efforts upstream through early consultation is essential for the organization.

Incorporate: Incorporation brings information objects into the organization's business environment through formal or informal acquisition and organization of information. While incorporation by itself does not guarantee digital preservation, it is an essential early step.

Roles and processes pertinent to incorporation may overlap with those required to retain objects indefinitely as part of a digital preservation program. Precise roles depend on the nature of the information and the processes required. Managers oversee aspects of incorporation at generally a high level and may be individuals, organizations, or systems. Specialists have knowledge and skills needed for a particular process or set of processes to succeed. Specialists are also involved with data services in areas such as metadata modeling and development, data administration and stewardship, repository architecture and design, and quality assurance throughout the lifecycle of information. Because information objects are already part of an organization's business environment, these roles are established, but may require enrichment to support digital preservation.

Consume: Consumption of information includes access, use, manipulation, and reuse of information. Persons, organizations, and systems making information consumable and those accompanying services that facilitate its use and reuse sustain the role of consumer. There must also be staff with the role of developing the specialized architecture of both information and systems to insure that information can be preserved.

Evaluate: Evaluation considers the institutional framework of policy and standards supporting digital preservation program as decisions are made.

The roles related to evaluation include some already present, legal counsel, data steward, and some that would need further definition including preservation officer, records manager, and digital data policy development specialist.

Ingest: Information objects and their accompanying support systems are brought into the retention phase of the information lifecycle through ingest following evaluation. Ingest may require further identification and organization of information objects sharing some similarity to the earlier incorporation phase. Ingest may also require transformation to represent the information in a new manner.

While some of the roles are similar to those previously discussed, a degree of specialization particularly related to data services may be new to this phase.

Store: Information objects destined for permanent retention will require appropriate storage for the level of ongoing retrieval expected as well as ongoing backup and monitoring. Roles are similar to those currently required for storage of all organizational data and emphasize good management practice in conjunction with carefully considered storage architectural requirements for long term retention.

Retain: As information objects age, the challenge increases in making them persistently available to consumers. Objects permanently retained will require ongoing monitoring to ensure that they remain available. Not only will the physical environment require ongoing evaluation and adjustment, but also systems and services must exist to monitor the future availability and usability of preserved objects.

Specialists knowledgeable about the organization's digital assets, their information formats, and the requirements for making formats available to consumers as well as the processes required to migrate or emulate older data formats and systems are critical to the future digital preservation. These specialists may have much in common and work closely with the digital consultants present at the outset of the creation phase of the information lifecycle.

Digital Migration*:* The transfer of digital information, while intending to preserve it, within the OAIS. It is distinguished from transfers in general by three attributes:

– a focus on the preservation of the full information content;

– a perspective that the new archival implementation of the information is a replacement for the old; and
– an understanding that full control and responsibility over all aspects of the transfer resides with the OAIS.

Digital Asset Lifecycle Management needs to be taken into consideration regardless of the service type or function of the system (Content Management, Document Management, Knowledge Management or Records Management) in which the digital object resides.

12.1 Content Management

Capabilities within this Service Type manage the storage, maintenance and retrieval of documents and information of a system or web site.

The term "Content Management" is often used to describe the management of websites, but in other instances refers to the management of all information across the whole of an enterprise. It also includes any business rules and editorial processes associated with content.
(See "CM vs DM vs KM vs DAM vs SCM vs DRM -- Which One is Right for You?" by Tony Byrne (2001-11-17) at http://web.archive.org/web/20041113063511/www.cmswatch.com/Features/OpinionWatch/FeaturedOpinion/?feature_id=53 for more information.)

12.1.1 Content Authoring

Allow for the creation of tutorials, CBT courseware, web sites, CD-ROMs and other interactive programs

Content authoring includes the capture of discrete digital assets, concern for formatting dependencies, rendering schemas, and authenticities, and version control of digital assets. The trend is from creating content and presentation at the same time for one-time use to separating content creation from presentation. This allows for the reuse of content for different purposes and various presentation methods (e.g., web site, e-mail, mobile devices, print).

Capabilities:

- Creating reusable digital assets, such as text, data, images, video, etc.
- Storage of digital assets in a system that captures and maintains the various contexts in which the individual asset has been used.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Markup Languages		SGML/HTML XML		XML
Authoring Tools (Examples)		XML: XML Spy Xmetal XML Cookbook Dreamweaver Camtasia Studio	OpenOffice	
Structural Description		IMS		

Other Important information for Content Authoring

- Instructional Management Systems (IMS) Specifications: http://www.imsglobal.org/specifications.html

12.1.2 Content Review and Approval

Allow for the approval of interactive programs

This includes the implementation of an organization's policies, procedures, and processes for review and approval of digital assets.

Capabilities:

- Version Control
- Workflow

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Version Control		Embedded in Content/Document Management System		Embedded in Content/Document Management System
Workflow		Embedded in Content/Document Management System		Embedded in Content/Document Management System

12.1.3 Tagging and Aggregation

Support the identification of specific content within a larger set of content for collection and summarization

Capabilities:

- Tagging to facilitate searching and navigation.
- Semantic markup.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Descriptive Metadata		Dublin Core Domain Specific Schemas/Formats Domain Specific Profiles		Dublin Core Domain Specific Schemas/Formats Domain Specific Profiles
Semantic Markup		SGML XML Domain Specific Markup Profiles	Web 2.0 Tagging Resource Description Framework (RDF)	SGML XML Domain Specific Markup Profiles

12.1.4 Content Publishing and Delivery

Allow for the propagation of digital assets

Capabilities:

- Creation of stylesheets for various delivery methods.
- Information syndication (RSS)

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
XML Query Language		XPATH		XPATH

Presentation		CSS XSL		CSS XSL
XML Transformation		XSLT		XSLT
Content Syndication		RSS Services Atom Syndication Format		RSS Services Atom Syndication Format

Definitions

- **Atom Syndication Format – an XML language used for web feeds.**
  - The name **Atom** applies to a pair of related standards. The Atom *Syndication Format* is an XML language used for web feeds, while the Atom *Publishing Protocol* (*APP* for short) is a simple HTTP-based protocol for creating and updating Web resources. www.wikepedia.com
  - Atom is an XML-based document format that describes lists of related information known as "feeds". Feeds are composed of a number of items, known as "entries", each with an extensible set of attached metadata. For example, each entry has a title. The primary use case that Atom addresses is the syndication of Web content such as weblogs and news headlines to Web sites as well as directly to user agents. See IETF RFC 4278, http://tools.ietf.org/html/rfc4287.

Other Important information for Content Publishing and Delivery

- **RSS** – Really Simple Syndication – delivers its information as an XML file called an "RSS feed", "webfeed", "RSS stream", or "RSS channel". These RSS feeds provide a way for users to passively receive newly released content (such as text, web pages, sound files, or other media); this might be the full content itself or just a link to it, possibly with a summary or other metadata (data describing the content).
- RSS Standards – RSS 2.0 (http://www.rssboard.org/rss-specification)
- Atom syndication format - IETF proposed standard, specified in RFC 4287 (http://tools.ietf.org/html/rfc4287)

12.1.5 Syndication Management

Control and regulate an organization's brand

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Digital Rights Management – see Records Management topic below				

Definitions – See Records Management topic below.

Other Important information for Syndication Management

- Coyle, Karen, *Rights Expression Language: A Report for the Library of Congress*, Feb. 2004, (http://www.loc.gov/standards/relreport.pdf)

12.2 Document Management

Capabilities within this Service Type control the capture and maintenance of an organization's documents and files.

Electronic Document Management Systems (EDMS) is becoming an all-encompassing term, referring to the integration of the underlying technologies including:

- Document imaging
- Document/Library services
- Workflow
- Enterprise Report Management (ERM)
- Forms Management
- Optical Character Recognition (OCR)/Intelligent Character Recognition (ICR) Technologies

Electronic document management systems provide users with access to more applications within a common user interface, through the utilization of industry standard Internet browser technology. One of the primary reasons users prefer this level of technology is the distributed functionality that becomes available almost immediately after implementation. (AIIM, "Recommended Practice: Analysis, Selection, and Implementation Guidelines Associated with Electronic Document Management Systems (EDMS)," April 2006)

12.2.1 Document Imaging and OCR

Support the scanning of documents

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Document Imaging		Standard Recommended Practice, File Format for Storage and Exchange of Images; Bi-Level Image File Format: Part I (ANSI/AIIM MS53-1993) Electronic Imaging Output/Display Devices (AIIM/TR19-1993) Electronic Imaging Output/Printers (AIIM/TR29-1993)		Standard Recommended Practice, File Format for Storage and Exchange of Images; Bi-Level Image File Format: Part I (ANSI/AIIM MS53-1993) Electronic Imaging Output/Display Devices (AIIM/TR19-1993) Electronic Imaging Output/Printers (AIIM/TR29-1993)

Optical Character Recognition (OCR)		Alphanumeric Character Sets for Optical Recognition, Part I: "Character Set of OCR-A- Shapes and Dimensions of the Printed Image" (ISO 1073/1-1976) Alphanumeric Character Sets for Optical Recognition, Part II: "Character Set OCR-B- Shapes and Dimensions of the Printed Image" (ISO 1073/2-1976) Character Sets for Optical Character Recognition, Adopted 13 September, 1989 (FIPS PUB 32-1; Incorporates OCR-A and OCR-B. Adopts ANSI X3.2-1970 (R1976), ANSI X3.49-1975 (R1982)) Optical Character Recognition Positioning (ANSI X3.93M-1981 (R1989)) Guideline for Optical Character Recognition Print Quality (ANSI X3.99-1983 (R1991))		Alphanumeric Character Sets for Optical Recognition, Part I: "Character Set of OCR-A- Shapes and Dimensions of the Printed Image" (ISO 1073/1-1976) Alphanumeric Character Sets for Optical Recognition, Part II: "Character Set OCR-B- Shapes and Dimensions of the Printed Image" (ISO 1073/2-1976) Character Sets for Optical Character Recognition, Adopted 13 September, 1989 (FIPS PUB 32-1; Incorporates OCR-A and OCR-B. Adopts ANSI X3.2-1970 (R1976), ANSI X3.49-1975 (R1982)) Optical Character Recognition Positioning (ANSI X3.93M-1981 (R1989)) Guideline for Optical Character Recognition Print Quality (ANSI X3.99-1983 (R1991))

Other Important information for Document Imaging and OCR

- Information on document or digital imaging and OCR can be found through the Association for Information and Image Management (AIIM) at www.aiim.org/.

12.2.2 Document Referencing

Support the redirection to other documents and information for related content

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Item Identifiers		Archival Resource Key (ARK) Document Object Identifier (DOI) Handles Namespace Open URL Persistent URL	XML Pointer (XPointer)	Archival Resource Key (ARK) Document Object Identifier (DOI) Handles Namespace Open URL Persistent URL

Definitions for Document Referencing include:

- **Archival Resource Key** – A specially constructed, actionable, and persistent URL encapsulating a globally unique identity that is independent of the current service provider.
- **Document Object Identifier** – The DOI system provides a framework for persistent identification, managing intellectual content, managing metadata, linking customers with content suppliers, facilitating electronic commerce, and enabling automated management of media. DOI names can be used for any form of management of any data, whether commercial or non-commercial (www.doi.org).
- **Namespace** – an abstract container providing context for the items (names, or technical terms, or words) it holds and allows disambiguation of items having the same name (residing in different namespaces - compare: URL).
- **Open URL** – Provides a standardized format for transporting bibliographic data about objects between information services.
- **Handles** – Unique persistent identifiers for Internet resources (www.cnri.net).

12.2.3 Document Revisions

Support the versioning and editing of content and documents

This includes the implementation of an organization's policies, procedures, and processes for review and approval of digital assets.

Capabilities:

- Version Control
- Workflow

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Version Control		Embedded in Content/Document Management System		Embedded in Content/Document Management System
Workflow		Embedded in Content/Document Management System		Embedded in Content/Document Management System

12.2.4 Library / Storage

Support document and data warehousing and archiving

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Warehousing		See Section 14.1.3, Data Management, Data Warehouse		
Information Lifecycle Management – Physical Storage Considerations			Long Term Archive and Compliance Storage Initiative (LTACSI) Standards	Long Term Archive and Compliance Storage Initiative (LTACSI)

Other Important information for Library / Storage

- Information on LTACSI can be found at www.snia.org/tech_activities/dmf/ltacsi/.

12.2.5 Document Review and Approval

Support the editing and commendation of documents before releasing them

This includes the implementation of an organization's policies, procedures, and processes for review and approval of digital assets.

Capabilities:

- Version Control
- Workflow

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Version Control		Embedded in Content/Document Management System		Embedded in Content/Document Management System
Workflow		Embedded in Content/Document Management System		Embedded in Content/Document Management System

12.2.6 Document Conversion

Support the changing of files from one type of format to another

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		In General Avoid Proprietary Formats Where Possible and Convert to Open Formats See Section 8.2.1		In General Avoid Proprietary Formats Where Possible and Convert to Open Formats See Section 8.2.1

12.2.7 Indexing Support

The rapid retrieval of documents through a structured numbering construct

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Structured Classification Schemes		See Data Interoperability section (14.1.9)		

12.2.8 Classification

Support the categorization of documents

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Classification Schemes		See Data Interoperability section (14.1.9)		

12.3 Knowledge Management

Capabilities within this Service Type identify, gather and transform documents, reports and other sources into meaningful information. Knowledge Management is a business process for managing intellectual assets in such a way that improves an employee's comprehension in a specific area of interest.

1) A business process for managing intellectual assets. It is a discipline that promotes an integrated approach to creating, capture, organization, access, and use of an organization's knowledge and information assets. Examples are structured databases, textual data, etc., and the tacit knowledge and expertise of individual employees.
2) The systematic process of finding, selecting, organizing, distilling and presenting information in a way that improves an employee's comprehension in a specific area of interest. (Source: Educause, http://www.educause.edu/Browse/645?PARENT_ID=229, accessed 5/26/06)

12.3.1 Information Retrieval

Allow access to data and information for use by an organization and its stakeholders

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Search		See Section 15.3		

12.3.2 Information Mapping / Taxonomy

Support the creation and maintenance of relationships between data entities, naming standards and categorization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Search, Classification		See Section 15.3		

12.3.3 Information Sharing

Support the use of documents and data in a multi-user environment for use by an organization and its stakeholders

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		See Section 11.2.1, Workgroup / Groupware		

12.3.4 Categorization

Allow classification of data and information into specific layers or types to support an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Classification		See Section 14.1.9		

12.3.5 Knowledge Engineering

Support the translation of knowledge from an expert into the knowledge base of an expert system

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.3.6 Knowledge Capture

Facilitate collection of data and information

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.3.7 Knowledge Distribution and Delivery

Support the transfer of knowledge to the end customer.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.3.8 Smart Documents

Support the interaction of information and process (business logic) rules between users of the document. (i.e. the logic and use of the document is embedded within the document itself and is managed within the document parameters)

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.4 Records Management

Capabilities within this Service Type store, protect, archive, classify and retire documents and information.

Capabilities within this Service Type capture, categorize, assign disposition, associate, ensure authenticity, and execute disposition of an organization's records in accordance with state and federal statutes and regulations.

Other electronic records management analyses and standards are concerned with identical and in some cases broader than just the records life cycle, but they address different topics and/or viewpoints. For example, ANSI/ARMA/AIIM TR48-2004, Framework for Integration of Electronic Document Management Systems and Electronic Records Management Systems, is concerned with the integration of two types of standalone applications where records management functionality is provided well after records are created and managed in other applications. Similarly, DoD 5015.2-STD, Design Criteria Standard for Electronic Records Management Applications, addresses design for one type of records management implementation – not for records management services which can be implemented at the beginning of the record life cycle in any electronic environment to include a services-oriented architecture. Finally, ISO 15489-1: Information and Documentation - Records Management - Part 1: General; ISO/TR 15489-2 Information and Documentation - Records Management - - Part 2: Guidelines, discuss guidelines and best practices for setting up an organizational records management program. (*Functional Requirements and Attributes for Records Management Services, December 7, 2005*; Interagency Project Team and the Records Management Service Components Program Office of the National Archives and Records Administration)

Records Management components break down as (from *Records Management Services (RMS) Overview*, National Archives and Records Administration, 2005, http://www.archives.gov/era/pdf/rmsc-overview.pdf):

Records Management Component	Definition
Capture Record	Capture Information with associated attributes in an electronic system.
Assign Disposition	Using an established disposition authority, assign the disposition schedule, item number, and disposition instructions to the record.
Categorize Record	Utilizing agency business rules, assign an appropriate descriptive label to the records to facilitate management in an electronic system.
Search Repository	Query all or selected system repositories of records (transitory, temporary, and permanent) across the enterprise for content and/or attributes, in order to determine the existence and location of matching records.
Retrieve Record	Using the search results, allow for the selective display of the full record and/or associated attributes for an authorized purpose.
Ensure Authenticity	Ensure the acceptability of a record as genuine, based on its characteristics such as structure, content, and context.
Associate Record	Provide the capability to associate a record to one or more other records through a Record Association attribute.
Execute Disposition	Implement destruction, transfer, or continued retention of a record in accordance with the established disposition authority. After validation that the disposition action is valid, execute the disposition action, and record the transaction.

Other Important information for Records Management

- *Statutes Related to Records Management*
  - Government Records Preservation Act (KSA 45-401 through KSA 45-413) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
  - Kansas Acts Against Discrimination (KSA 44-1001 et seq)

- Open Records Act (KSA 45-215 through 45-223) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
- Public Records Act (KSA 75-3501 through 75-3518) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
- Records made on Electronically-accessed Media; Authorization, Conditions and Procedures, Application, Notice to State Records Board (KSA 45-501) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
- Tampering with a Public Record (KSA 21-3821) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
- Telecommunications services of certain state agencies; extension to certain private, nonprofit agencies or governmental entities; records of services (KSA 75-4709) (online at http://www.kshs.org/government/records/stategovt/recordslaw.htm)
- Federal statutes that may apply to specific agencies

- ***Standards Related to Records Management***
  - Department of Defense 5015.2-STD, "Design Criteria Standard for Electronic Records Management Applications" (Online at http://jitc.fhu.disa.mil/recmgt/#standard)
  - ISO 23081-1:2006: "Information and documentation - Records management processes - Metadata for records - Part 1: Principles" (Available for purchase through ANSI at http://webstore.ansi.org/ansidocstore/product.asp?sku=ISO+23081%2D1%3A2006)
  - ISO/TR 15489-2:2001: "Information and documentation - Records management - Part 2: Guidelines" (Available for purchase through ANSI at http://webstore.ansi.org/ansidocstore/product.asp?sku=ISO%2FTR+15489%2D2%3A2001)
  - Federal Section 508 Electronic and Information Technology Accessibility Standards (36 CFR § 1194) (Online at http://www.section508.gov/)
  - • Information Technology Policy 1210: State of Kansas Web Accessibility Requirements (Online at http://www.da.ks.gov/itec/documents/itecitpolicy1210.htm)
  -

12.4.1 Record Linking / Association

Support the correlation between logical data and information sets

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.4.2 Document Classification

Support the categorization of documents and artifacts, both electronic and physical

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.4.3 Document Retirement

Support the termination or cancellation of documents and artifacts used by an organization and its Stakeholders

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

12.4.4 Digital Rights Management

Support the claim and ownership of intellectual capital and artifacts belonging to an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Digital Rights Management		ISO MPEG REL Adobe Content Manager Creative Commons MPEG-21/5 PRISM	ERMI FDRM METSR ODRL XACML	

Definitions for Digital Rights Management include:

- Adobe Content Manager – Used extensively for e-books and enterprise documents (http://www.adobe.com/products/acrobatpro/main.html)
- Creative Commons - (http://www.creativecommons.org)
- ISO MPEG REL – an XML-based Rights Expression Language used to specify terms and conditions for the authorized distribution and use of any digital content. (http://www.chiariglione.org/mpeg)
- ERMI – Electronic Resource Management Initiative – (http://www.library.cornell.edu/cts/elicensestudy/home.html)
- FDRM – Federated Digital Rights Management (http://www.dlib.org/dlib/july02/martin/07martin.html)
- MPEG-21/5 – Information Technology – Multimedia Framework – Part 5: Rights Expression Language. ISO/IEC FDIS 21000-5:2003(E) (http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm)
- METSR – METSRights (http://www.loc.gov/standards/rights/METSRights.xsd)
- ODRL – Open Digital Rights Language (http://www.odrl.net)
- PRISM – Publishing Requirements for Industry Standard Metadata (http://www.prismstandard.com)
- XACML – eXtensible Access Control Markup Language (http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml)

Chapter 13 Business Analytical Services

The Business Analytical Services Domain defines the set of capabilities supporting the extraction, aggregation, and presentation of information to facilitate decision analysis and business evaluation

13.1 Analysis and Statistics

Capabilities within this Service Type examine business issues, problems and their solutions.

Other General Information for Analysis and Statistics

Currently, data analysis is fragmented into separate too-related processes that produce independent insights and, sometimes, predictive models. There is a need for an integrated analysis platform that allows a coherent action strategy including automatic alert triggers and critical data routing. Advantages of an integrated approach:

- It supports the full suite of analytical, planning, and reporting functions for comprehensive data-driven solutions.
- It offers modules for data mining, analysis, outcome prediction and simulation, programmable action strategies and reporting.
- It provides an integrated development environment for customizing, testing and publishing analytical applications for operational use.
- It supports sharing and reuse of analytical applications as components of new applications.

Definition: an analytical application is a self-contained run-time software executable that has been designed to perform a data analysis an/or action function. It may be composed of single or multiple components. Examples of components include:

- Mixed data extractors and processors
- Prediction models
- Classifiers
- Event detectors
- Action strategies
- Report definitions
- Connectors to other systems for data routing and alerting

From: Lantrip, David B, *Analytics Applications Provide Actionable Results,* KMWorld, June 2006

13.1.1 Mathematical

Support the formulation and mathematical analysis of probabilistic models for random phenomena and the development and investigation of methods and principles for statistical inference

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Description	Example Products	Notes
General Quantitative and Statistical Analysis	SAS SPSS	
Domain-specific tools		This area has too many specific products to catalog

13.1.2 Structural / Thermal

Support the use of data flow and data modeling diagrams for applying systematic analysis of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Description	Example Products	Notes
Domain-specific tools		This area has too many specific products to catalog

13.1.3 Radiological

Support the use of radiation and x-ray technologies for analysis and scientific examination

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Description	Example Products	Notes
Domain-specific tools		This area has too many specific products to catalog

13.1.4 Forensics

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Description	Example Products	Notes
Domain-specific tools		This area has too many specific products to catalog

13.2 Visualization

Capabilities within this Service Type convert data into graphical or picture form.

13.2.1 Graphing / Charting

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Graphing / Charting		Standard Database Linkage (JDBC, ODBC) Integrated drawing tools Multiple graphing / charting types supported (i.e. bar, radar, etc.)		Standard Database Linkage (JDBC, ODBC) Integrated drawing tools Multiple graphing / charting types supported (i.e. bar, radar, etc.)

Description	Example Products	Notes
Graphing / Charting	• Microsoft Visio • SAS • SPSS • Crystal Reports • Microsoft Project • Microsoft Excel • Erwin • SQL Reports • PhotoShop • Jfree Chart • ILog Domain-specific tools	

13.2.2 Imagery

Support the creation of film or electronic images from pictures or paper forms

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Digital Scanning		Scan to PDF Scan to multiple file formats including TIFF, JPEG, JPG2000, etc.		Scan to PDF Scan to multiple file formats including TIFF, JPEG, JPG2000, etc.

Microform (microfilm, microfiche)		NISO and ANSI standards including: ANSI/NISO Z39.32 - 1996 (R2002) Information on Microfiche Headers ANSI/NISO Z39.62 - 2000 Eye-legible Information on Microfilm Leaders and Trailers and on Containers of Processed Microfilm on Open Reels ANSI/NISO Z39.74 - 1996 (R2002) Guides to Accompany Microform Sets		NISO and ANSI standards including: ANSI/NISO Z39.32 - 1996 (R2002) Information on Microfiche Headers ANSI/NISO Z39.62 - 2000 Eye-legible Information on Microfilm Leaders and Trailers and on Containers of Processed Microfilm on Open Reels ANSI/NISO Z39.74 - 1996 (R2002) Guides to Accompany Microform Sets
Photographs		Multiple ANSI standards covering chemicals, film, processing, etc. http://www.nssn.org/search/IntelSearch.aspx		Multiple ANSI standards covering chemicals, film, processing, etc. http://www.nssn.org/search/IntelSearch.aspx
Digital photographs		Multiple ANSI standards covering pixel-reporting, color-encoding, format profiles, etc. http://www.nssn.org/search/IntelSearch.aspx		Multiple ANSI standards covering pixel-reporting, color-encoding, format profiles, etc. http://www.nssn.org/search/IntelSearch.aspx

Other Important information for Imagery

- **ISO 10196:2003 Document imaging applications -- Recommendations for the creation of original documents** [Scope: ISO 10196:2003 provides guidance on the creation of printed documents so that they may be easily reproduced as microforms or scanned images. ISO 10196:2003 does not apply to technical drawings for which requirements are given in ISO 5457 and ISO 6428. It also does not apply to special micrographics or scanning-related applications (scanning of bank cheques or bar codes).] http://www.nssn.org/search/DetailResults.aspx?docid=24828&selnode=
- **ISO 6200:1999 Micrographics -- First generation silver-gelatin microforms of source documents -- Density specifications and method of measurement**, http://www.nssn.org/search/DetailResults.aspx?docid=33149&selnode=
- **ISO 8514-1:2000 Micrographics -- Alphanumeric computer output microforms -- Quality control -- Part 1: Characteristics of the test slide and test data** http://www.nssn.org/search/DetailResults.aspx?docid=35459&selnode=

- **ISO 8514-2:2000 Micrographics -- Alphanumeric computer output microforms -- Quality control -- Part 2: Method**
  http://www.nssn.org/search/DetailResults.aspx?docid=35460&selnode=

13.2.3 Multimedia

Support the representation of information in more than one form to include text, audio, graphics, animated graphics and full motion video

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Multimedia				

Description	Example Products	Notes
Multimedia	• Microsoft Powerpoint • OpenOffice presentation software • Apple Macintosh Software	

13.2.4 Mapping / Geospatial / Elevation / GPS

Provide for the representation of position information through the use of attributes such as elevation, latitude, and longitude coordinates

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Mapping / Geospatial / Elevation / GPS		Data input o Manual digitizing o Scanning o Keyed bulk data entry o Automatic checking and corrections for digitizing errors o Acceptance of existing raster and vector data Data manipulation o Data revisions o Thinning and weeding of digital line work o Sliver polygon removal o Transformation between map projections o Edge matching of adjoining map files o Transformation of data to fit specified control points (i.e. rubber sheeting) o Raster to vector conversion o Merging polygons with common		Data input o Manual digitizing o Scanning o Keyed bulk data entry o Automatic checking and corrections for digitizing errors o Acceptance of existing raster and vector data Data manipulation o Data revisions o Thinning and weeding of digital line work o Sliver polygon removal o Transformation between map projections o Edge matching of adjoining map files o Transformation of data to fit specified control points (i.e. rubber sheeting) o Raster to vector conversion o Merging polygons with common

		attributes o Computing distance buffers o Aggregating data within specified parameters Data analysis o Point, line, and polygon overlay analysis o Geometric measurements and calculations o Analysis of proximity and contiguity o Spatial data queries o Attribute data queries o Coordinate geometry calculations o Digital terrain modeling and analysis o Network analysis Data presentation o Display and plot of raster and/or vector data o Display and plot of data at user-defined scales o Display and plot of digital terrain models o Automatic plot of attribute data as map text o Automatic generation of map symbols based on attribute data o Automatic dimensioning o Specific printer and plotter capabilities o Specific report and map output formats		attributes o Computing distance buffers o Aggregating data within specified parameters Data analysis o Point, line, and polygon overlay analysis o Geometric measurements and calculations o Analysis of proximity and contiguity o Spatial data queries o Attribute data queries o Coordinate geometry calculations o Digital terrain modeling and analysis o Network analysis Data presentation o Display and plot of raster and/or vector data o Display and plot of data at user-defined scales o Display and plot of digital terrain models o Automatic plot of attribute data as map text o Automatic generation of map symbols based on attribute data o Automatic dimensioning o Specific printer and plotter capabilities o Specific report and map output formats

Description	Example Products	Notes
Mapping / Geospatial / Elevation / GPS	• ESRI ArcGIS • OpenARC • OpenGIS standard products	

13.2.5 CAD

Support the design of products with computers

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
2D Design				
3D Design				

Description	Example Products	Notes
CAD	AutoCAD Microstation Domain-specific tools	

Other Important information for CAD

- ISO CAD standard is divided in three parts:
  - ISO 13567-1:1998
    - Technical product documentation -- Organization and naming of layers for CAD -- Part 1: Overview and principles
  - ISO 13567-2:1998
    - Technical product documentation -- Organization and naming of layers for CAD -- Part 2: Concepts, format and codes used in construction documentation
  - ISO/TR 13567-3:1999
    - Technical product documentation -- Organization and naming of layers for CAD -- Part 3: Application of ISO 13567-1 and ISO 13567-2

  AIA Cad Layer Guidelines, 2nd edition (1997)

- The capabilities of modern CAD systems include:
  - Wireframe geometry creation
  - 3D parametric feature based modeling, Solid modeling
  - Freeform surface modeling
  - Automated design of assemblies, which are collections of parts and/or other assemblies
  - create Engineering drawings from the solid models
  - Reuse of design components
  - Ease of modification of design of model and the production of multiple versions
  - Automatic generation of standard components of the design
  - Validation/verification of designs against specifications and design rules
  - Simulation of designs without building a physical prototype
  - Output of engineering documentation, such as manufacturing drawings, and Bills of Materials to reflect the BOM required to build the product
  - Import/Export routines to exchange data with other software packages
  - Output of design data directly to manufacturing facilities
  - Output directly to a Rapid Prototyping or Rapid Manufacture Machine for industrial prototypes
  - maintain libraries of parts and assemblies
  - calculate mass properties of parts and assemblies
  - aid visualization with shading, rotating, hidden line removal, etc...
  - Bi-directional parametric associatively (modification of any feature is reflected in all information relying on that feature; drawings, mass properties, assemblies, etc... and counter wise)
  - kinematics, interference and clearance checking of assemblies
  - sheet metal
  - hose/cable routing

- electrical component packaging
- inclusion of programming code in a model to control and relate desired attributes of the model
- Programmable design studies and optimization
- Sophisticated visual analysis routines, for draft, curvature, curvature continuity.

13.3 Knowledge Discovery

Capabilities within this Service Type facilitate the identification of useful information from data.

13.3.1 Data Mining

Provide for the efficient discovery of non-obvious, valuable patterns and relationships within a large collection of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Mining		Domain-specific tools		

13.3.2 Modeling

Develop descriptions to adequately explain relevant data for the purpose of prediction, pattern detection, exploration or general organization of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Modeling		Domain-specific tools		

13.3.3 Simulation

Utilize models to mimic real-world processes

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Simulation		Domain-specific tools		

13.4 Business Intelligence

13.4.1 Demand Forecasting / Management

Facilitate the prediction of sufficient production to meet an organization's sales of a product or service

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Demand Forecasting and Management		Manual Process Domain-specific tools		

13.4.2 Balanced Scorecard

Support the listing and analyzing of both positive and negative impacts associated with a decision

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Balanced Scorecard		Manual Process COBIT SQL Reporting		

13.4.3 Decision Support and Planning

Support the analysis of information and predict the impact of decisions before they are made

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Decision Support and Planning		Manual Process	Visualization Dashboards	Visualization Dashboards

13.5 Reporting

Capabilities within this Service Type organize data into useful information

13.5.1 Ad hoc

Support the use of dynamic reports on an as needed basis

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Textual **Numeric**		Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms	Open, Standardized Metadata	Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms ITEC Policy 1210 Compliant Open, Standardized Metadata

Description	Example Products	Notes
	• Crystal Reports • SAS • Microsoft Excel • Microsoft Access • JReport • Alphablox (IBM) • Business Objects • Oracle/PeopleSoft	
Specialized Programs		

13.5.2 Standardized / Canned

Support the use of pre-conceived or pre-written reports

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Textual **Numeric**		Interoperability with multiple data sources SQL query support XML support	Open, Standardized Metadata	Interoperability with multiple data sources SQL query support XML support

		Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms		Ability to script Standard Database Connections (ODBC, JDBC) Web interface using forms ITEC Policy 1210 Compliant Open, Standardized Metadata

Description	Example Products	Notes
	Crystal Reports SAS	

13.5.3 OLAP

Support the analysis of information that has been summarized into multidimensional views and hierarchies

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Textual **Numeric**		Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC)	Open, Standardized Metadata	Interoperability with multiple data sources SQL query support XML support Ability to script Standard Database Connections (ODBC, JDBC) Open, Standardized Metadata

Description	Example Products	Notes
	SAS	
	Data warehouse system integrated reporting	

Other Important information for OLAP

- See *Data Management* sections (7.5 and 16.3)

Chapter 14 Back Office Services

The Back Office Services Domain defines the set of capabilities that support the management of enterprise planning and transactional-based functions.

14.1 Data Management

Capabilities within this Service Type provide for the usage, processing and general administration of unstructured information.

14.1.1 Data Exchange

Support the interchange of information between multiple systems or applications; includes verification that transmitted data was received unaltered

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Transport Method	Diskettes Tapes Paper	Email Batch File Transfer Internet Embedded application distribution services		Email Batch File Transfer Internet Embedded application distribution services
Data Format and Content Identification	Use of specific proprietary or legacy system formats agreed upon on a case-by-case basis.	Use of specific proprietary or legacy system formats agreed upon on a case-by-case basis.	Adhere to XML, EDI, or other similar conventions to publish or exchange content in a technology-neutral, standard way Adopt SOAP philosophies.	Adhere to XML, EDI, or other similar conventions to publish or exchange content in a technology-neutral, standard way. Adopt SOAP philosophies.
Security, Integrity, and Quality Assurance		Use of Encryption Technologies, Security Tokens, Checksum algorithms, and Digital Signatures Adherence to FIPS 186-2 (DSS). Establish verification procedures Use of SSL.	Adherence to FIPS 186-3 (DSS). Use of TLS.	Use of Encryption Technologies, Security Tokens, Checksum algorithms, and Digital Signatures Adherence to FIPS 186-3 (DSS). Establish verification procedures Use of TLS.
Platform and Language Independence		ODBC	Use of SOAP (for data exchange), ODBC (for query-type “exchange”),	Use of SOAP (for data exchange), ODBC (for query-type “exchange”),

			or other similar philosophies.	or other similar philosophies.
Data Compression		Adaptive Coding with Embedded Dictionary (DCLZ Algorithm) for Information Interchange (ANSI INCITS 223-1995 (R2001) Adaptive Coding with Sliding Window for Information Interchange (ANSI INCITS 241-1994 (R1999)) Adaptive Lossless Data Compression (ALDC) Algorithm for Information Interchange (ANSI INCITS 280-1996 (R2001)) Adaptive Lossless Data Compression algorithm (ALDC) (ISO/IEC 15200:1996) Binary Arithmetic Coding Algorithm (ISO/IEC 12042:1993) DCLZ Algorithm (ISO/IEC 11558:1992) FBI Wavelet/Scalar Quantization Specification (compression of digitized fingerprint images) ISO/IEC 14495-2:2003 (Lossless and near-lossless compression of continuous-tone still images)		Adaptive Coding with Embedded Dictionary (DCLZ Algorithm) for Information Interchange (ANSI INCITS 223-1995 (R2001) Adaptive Coding with Sliding Window for Information Interchange (ANSI INCITS 241-1994 (R1999)) Adaptive Lossless Data Compression (ALDC) Algorithm for Information Interchange (ANSI INCITS 280-1996 (R2001)) Adaptive Lossless Data Compression algorithm (ALDC) (ISO/IEC 15200:1996) Binary Arithmetic Coding Algorithm (ISO/IEC 12042:1993) DCLZ Algorithm (ISO/IEC 11558:1992) FBI Wavelet/Scalar Quantization Specification (compression of digitized fingerprint images) ISO/IEC 14495-2:2003 (Lossless and near-lossless compression of continuous-tone still images)

		JPEG 2000 image coding system - Part 1: Core coding system (INCITS/ISO/IEC 15444-1-2000) Streaming Lossless Data Compression algorithm (SLDC) (ISO/IEC 22091:2002)		JPEG 2000 image coding system - Part 1: Core coding system (INCITS/ISO/IEC 15444-1-2000) Streaming Lossless Data Compression algorithm (SLDC) (ISO/IEC 22091:2002)

Definitions for Data Exchange include:

- Checksum - A method for checking the integrity of transmitted data. A checksum is an integer value computed from a sequence of octets by treating them as integers and computing the sum. The value is recomputed at the receiving end and compared for verification.
- Digital Signature - A digital signature is an electronic rather than a written signature that can be used by someone to authenticate the identity of the sender of a message or of the signer of a document. It can also be used to ensure that the original content of the message or document that has been conveyed is unchanged.
- DRDA - (Distributed Relational Database Architecture) An IBM architecture for distributing data across multiple heterogeneous platforms. It also serves as a protocol for access to these databases from IBM and non-IBM platforms.
- DSS – Digital Signature Standard, currently FIPS 186-2 (Federal Information Processing Standard).
- EDI – (Electronic Data Interchange) is the computer-to-computer exchange of business data in standard formats.
- SOAP - (Simple Object Access Protocol) is a way for a program running in one kind of operating system (such as Windows XP) to communicate with a program in the same or another kind of an operating system (such as Linux) by using the World Wide Web's Hypertext Transfer Protocol (HTTP)and its XML as the mechanisms for information exchange.
- ODBC – (Open Database Connectivity) A Microsoft interface that allows applications to access data in various database management systems using SQL as a standard for accessing the data.
- Encryption - Encryption is a secure process for keeping your sensitive and confidential information private. It is a process by which bits of data are mathematically jumbled with a password-key. The Encryption process makes the data unreadable unless or until decrypted.
- SSL – Secure Sockets Layer is a commonly-used protocol for managing the security of a message transmission on the Internet.
- TLS – Transport Layer Security is an emerging protocol based upon SSL, also used for managing the security of a message transmission on the Internet.
- XML - Extensible Markup Language is a flexible way to create common information formats and share both the format and the data on the World Wide Web, intranets, and elsewhere.

Other Important information for Data Exchange

- www.niem.gov – National Information Exchange Model. An organization and effort formed to establish data exchange standards between federal and other government entities.
- www.w3.org – World Wide Web Consortium, a key resource for work done within the NIEM, promoting the health and growth of the World Wide Web by developing uniform protocols and guidelines.
- www.openarchives.org - The Open Archives Initiative (OAI) develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.

- www.disa.org – Data Interchange Standards Association, endorsing and providing a base for both XML and EDI standards.
- www.x12.org – the ASCX12 or Accredited Standards Committee, an organization charted by ANSI (the American National Standards Institute) for the purpose of developing and maintaining uniform standards for electronic exchange of business transactions, or what became known as EDI. The EDI standards are known as the X12 standards.
- www.oasis-open.org – Organization for the Advancement of Structured Information Standards, a global consortium that drives the development, convergence, and adoption of e-business standards.
- www.ws-i.org – An open industry organization chartered to promote Web services interoperability across platforms, operating systems, and programming languages.
- http://csrc.nist.gov/publications/fips/fips186-2/fips186-2-change1.pdf - The current DSS (Digital Signature Standard).
- http://csrc.nist.gov/publications/drafts.html - Page containing draft of FIPS 186-3, the proposed update to FIP 186-2, the DSS (Digital Signature Standard).
- ftp://wwwc3.lanl.gov/pub/misc/WSQ/FBI_WSQ_FAQ -- Frequently Asked Questions on the FBI's Wavelet/Scalar Quantization Specification for compression of digitized gray-scale fingerprint images.

14.1.2 Data Mart

Support a subset of a data warehouse for a single department or function within an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Definitions for Data Mart include:

- **Data mart** – a subset of highly summarized data from the data warehouse optimized for end-user access to facilitate data analysis and designed to provide answers to a distinct set of questions.

Other Important information for Data Mart:

- *See the Data Warehouse section (14.1.3) for Data Mart Standards.*

14.1.3 Data Warehouse

Support the archiving and storage of large volumes of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Staging		Code Library Management and Version Control Data Staging Job Control	Metadata-Driven Data Staging	Code Library Management and Version Control Data Staging Job Control
Metadata Repository Development and Maintenance		Information catalog integration/merge (e.g., from the data model to the database to the front end reporting tool) Manage and display graphical or tabular representation of metadata repository contents (metadata browser)		Information catalog integration/merge (e.g., from the data model to the database to the front end reporting tool) Manage and display graphical or tabular representation of metadata repository contents (metadata browser)
Data Access (Reporting)		Operational Data Store (Isolated Reporting Environment) Parameter or variable driven capabilities Time and event based scheduling of report execution Iterative execution Support multiple delivery methods and formats (email, web, network directory) Report linking Mass report distribution support via email, web or network directory User interface controls for conceptual representations such as gauges and sophisticated charting. Alerts or exception controls that monitor specific values, ranges or differences that notify the user when targeted levels		Operational Data Store (Isolated Reporting Environment) Parameter or variable driven capabilities Time and event based scheduling of report execution Support multiple delivery methods and formats (email, web, network directory) Report linking Mass report distribution support via email, web or network directory User interface controls for conceptual representations such as gauges and sophisticated charting. Alerts or exception controls that monitor specific values, ranges or differences that notify the user when targeted levels are met or exceeded. (Performance

		are met or exceeded. (Performance management)		management)
Data Access (Ad Hoc Query)		Multi-pass SQL Highlighting and conditional formatting Successive constraints Direct SQL Entry Basic calculations on the result set Dataset pivoting Column calculations on pivot results Sorting Charting and Graphs Metadata access Export to multiple formats (XML, EXCEL, PDF, CSV)		Multi-pass SQL Highlighting and conditional formatting Successive constraints Direct SQL Entry Basic calculations on the result set Dataset pivoting Column calculations on pivot results Sorting Charting and Graphs Metadata access Export to multiple formats (XML, EXCEL, PDF, CSV)
Data Modeling and Mining		Clustering Support Classifying Estimating and Predictive Analysis Affinity Grouping		Clustering Support Classifying Estimating and Predictive Analysis Affinity Grouping

Guiding Principles for Data Warehouse:

- Improve stakeholder satisfaction by increasing the effectiveness and efficiency of agency decision support processes.
- Align decision support and data warehouse design and development efforts with agency business processes and functional requirements.
- Operational and Informational/Analytical data stores are housed in separate isolated database instances.
  - Separate data sources isolate operational systems, which perform critical business processing, from ad hoc queries, reporting and online analytical data processing (OLAP). The separation of operational and informational/analytical systems decreases the risk of adversely impacting operational system performance.
- Data stored and managed within a Data Warehouse is to be structured in a way that provides high quality, secure, accessible and reliable data to the end business user community.

Definitions for Data Warehouse include:

- **Operational Data** – Data that is used in day-to-day transactions.
- **Informational/Analytical Data** – Data that is compiled from operational data for use as the basis for business decision support.
- **Data Staging** – The storage of data prior to being loaded into a data warehouse or data mart.
- **Data Staging Job Control** – Staging job definition, job scheduling (time and event), monitoring, logging, exception handling, error handling, and notification.
- **Metadata –** Metadata describes the meaning and structure of data, as well as how it is created, accessed and used.
- **Operational Data Store –** The operational data store is a subject-oriented, integrated, updateable collection of data designed and structured to allow better integration among operational data used for immediate inquiry and reporting. The operational data store reflects, at any point, the current operational state of its subject matter and enhances performance by removing processing load from the transactional system.
- **DBMS –** Database Management System
- **Clustering -** Clustering (or "Cluster Analysis" or "Typology") consists of partitioning the set of observations in a data base into a small number or groups, or clusters, so that:
  - Observations in one cluster are as similar as possible.
  - Observations belonging to different clusters are as different from one another as possible.

Other Important information for Data Warehouse

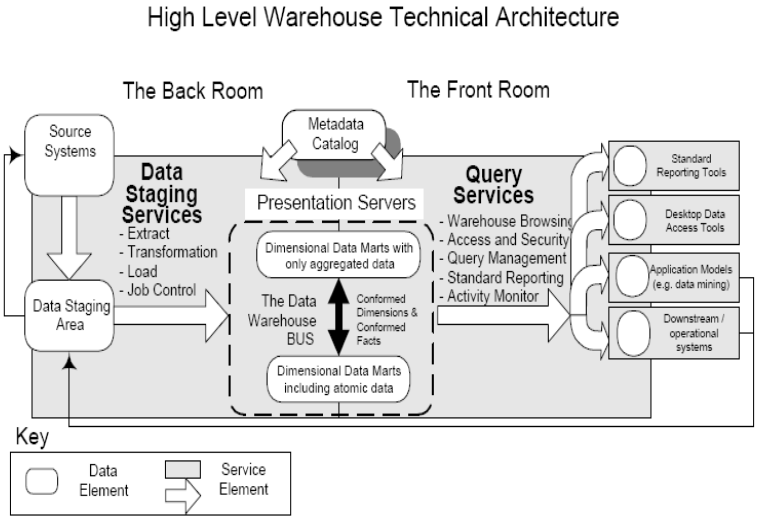


Figure 1 – High Level Data Warehouse Technical Architecture

14.1.4 Meta Data Warehouse

Support the maintenance and administration of data that describes data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Format	Meta Data Coalition Open Information Model 1.0	Common Warehouse Metamodel (CWM)		Common Warehouse Metamodel (CWM)
Metadata Modeling		Unified Modeling Language (UML) 2.0		Unified Modeling Language (UML) 2.0

Guiding Principles for Data Warehouse:

- Through gathering technical, informational, and knowledge data from the transactional systems and its own data stores, the metadata repository helps in the maintenance, coordination, and publication of the data warehouse systems.
  - Provides the context for the data
  - Some metadata, such as definitions are readily seen. Yet, there is other metadata behind the scenes at a system level, which is necessary for the warehouse operation.
  - Data elements that are most used throughout an organization's systems need to be prioritized and identified. Definitions for these elements should be the first collected.
  - Metadata for core elements must match across systems to support integration and linkage of the organization's systems. This effort requires coordination and support from the various transactional system administrators and the warehouse staff.
- Metadata is no different from other business data; it too must be integrated across different tools, different systems, application upgrades and replacements.

Definitions for Meta Data Management include:

- **Common Warehouse Metamodel (CWM™):** Standard interfaces that can be used to enable easy interchange of warehouse and business intelligence metadata between warehouse tools, warehouse platforms and warehouse metadata repositories in distributed heterogeneous environments.
- **Unified Modeling Language™ (UML®):** A specification defining a graphical language for visualizing, specifying, constructing, and documenting the artifacts of distributed object systems.

Other Important information for Meta Data Management

- Object Management Group (OMG; http://www.omg.org)
- Meta Data Coalition (MDC; http://www.mdcinfo.org)
- "Nearly every data warehousing expert will agree that **metadata management** is the key to data warehousing success. A metadata repository contains the organization's business models and rules, data view definitions, data usage model, report dictionary, user profiles, conceptual, logical, and physical data models, source file data dictionaries, data element descriptions, and data conversion rules. In effect, the metadata repository provides the context for the data that resides in the data warehouse. Without this context, the data warehouse becomes a pit of meaningless scraps of data. Without metadata, without the context in which to view data, we run the risk of misapplying it." (Corporate Information Designs, Inc.)
- **Catalog of OMG specifications: (**http://www.omg.org/technology/documents/spec_catalog.htm)
- **Catalog of OMG Modeling and Metadata Specifications:** http://www.omg.org/technology/documents/modeling_spec_catalog.htm
  - **Common Warehouse Metamodel (CWM™):** Standard interfaces that can be used to enable easy interchange of warehouse and business intelligence metadata between warehouse tools, warehouse platforms and warehouse metadata repositories in distributed heterogeneous environments.
  - **Common Warehouse Metamodel (CWM™) Metadata Interchange Patterns (MIPS):** The purpose of this specification is to add a semantic context to the interchange of metadata in terms of recognized sets of objects or object patterns. It introduces the term "Unit of Interchange" (UOI) to define a valid, recognizable CWM interchange. From this information, a user of CWM, working in conjunction with CWM MIP, should be able to produce truly interoperable tools.
  - **Meta-Object Facility (MOF™):** MOF is an extensible model driven integration framework for defining, manipulating and integrating metadata and data in a platform independent manner. MOF-based standards are in use for integrating tools, applications and data.
  - **MOF™ Query / Views / Transformations:** This specification is one of a series related to developing the 2.0 revision of the OMG Meta Object Facility specification, referred to as MOF 2.0. This specification addresses a technology neutral part of MOF and pertains to: 1.) queries on models; 2.) views on metamodels; and 3.) transformations of models.

- **Reusable Asset Specification (RAS):** This Specification provides a set of guidelines and recommendations about the structure, content, and descriptions of reusable software assets. Recognizing that there are different categories of reusable software assets, the specification identifies some categories, or rather types or profiles, and provides general guidelines on these profiles. RAS addresses the engineering elements of reuse. It attempts to reduce the friction associated with reuse transactions through consistent, standard packaging.
- **Unified Modeling Language™ (UML®):** ***:*** A specification defining a graphical language for visualizing, specifying, constructing, and documenting the artifacts of distributed object systems.
- **UML Diagram Interchange:** This specification enables a smooth and seamless exchange of documents compliant to the UML standard (referred to as UML models) between different software tools. While this certainly includes tools for developing UML models, it also includes tools such as whiteboard tools, code generators, word processing tools, and desktop publishing tools. Special attention is given to the Internet as a medium for exchanging and presenting UML models.
- **UML Human-Usable Textual Notation (HUTN):** A specification for a Human-Usable Textual Notation (HUTN) for expressing other specifications in terms of the UML Profile for Enterprise Distributed Computing (EDOC) and its companion UML Profile for Common Object Request Broker Architecture (CORBA). HUTN offers three main benefits. (1.) It is a generic specification that can provide a concrete HUTN language for any MOF model; (2.) The HUTN languages can be fully automated for both production and parsing; and (3.) The HUTN languages are designed to conform to human-usability criteria.
- **XML Metadata Interchange (XMI®):** XMI is a model driven XML Integration framework for defining, interchanging, manipulating and integrating XML data and objects. XMI-based standards are in use for integrating tools, repositories, applications and data warehouses. XMI provides rules by which a schema can be generated for any valid XMI-transmissible MOF-based metamodel. XMI provides a mapping from MOF to XML. As MOF and XML technology evolved, the XMI mapping is being updated to comply with the latest versions of these specifications. Updates to the XMI mapping have tracked these version changes in a manner consistent with the existing XMI Production of XML Schema specification (XMI Version 2).

14.1.5 Data Cleansing

Support the removal of incorrect or unnecessary characters and data from a data source

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Cleansing				

Definitions for Data Cleansing include:

- **Data cleansing** – the act of detecting and correcting (or removing) corrupt or inaccurate records from a data set. After cleansing, a data set will be consistent with other similar data sets in the system. The inconsistencies detected or removed may have been originally caused by different data dictionary definitions of similar entities in different stores, may have been caused by user entry errors, or may have been corrupted in transmission or storage. Preprocessing the data will also guarantee that it is unambiguous, correct, and complete. The actual process of data cleansing may involve removing typos or validating and correcting values against a known list of entities. The validation may be strict (such as rejecting any address that does not have a valid ZIP code) or fuzzy (such as correcting records that partially match existing, known records). Data cleansing is synonymous with the less frequently-used term *data scrubbing*. Data cleansing differs from *data validation* in that validation almost invariably means data is rejected from the system at entry and is performed at entry time, rather than on batches of data.

Other Important information for Data Cleansing

- Han, J., Kamber, M. *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2001.
- Kimball, R., Caserta, J. *The Data Warehouse ETL Toolkit*, Wiley and Sons, 2004.

14.1.6 Extraction and Transformation

Support the manipulation and change of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Extraction and Transformation	Hand-written scripts	Hand-written scripts and tools bundled with the underlying database that have limited out-of-the box functionality. XML Data Interchange (XMI) Integrate with XML. Standards based metadata	Leverage existing code as well as bring new functionality. Allow for the incorporation of existing working scripts and offer a complete tool set of ready to use transformations and functions as well as a comprehensive list of documented APIs and methods. Improve metadata management and administration as well as ensure data quality. Able to do upfront analysis and modeling and streamline the workflow process. Integrate with Message Brokers.	Integrate with XML. XML Data Interchange (XMI) Integrate with Message Brokers. Standards based metadata Leverage existing code as well as bring new functionality. Allow for the incorporation of existing working scripts and offer a complete tool set of ready to use transformations and functions as well as a comprehensive list of documented APIs and methods. Improve metadata management and administration as well as ensure data quality. Able to do upfront analysis and modeling and streamline the workflow process.

Preferences and Requirements for Extraction and Transformation:

- Integrate data from multiple sources. Should easily connect to varied data sources and targets.
- Cleanse and migrate data
- Convert data types
- Reconcile data inconsistencies
- Eliminate duplicate records
- Manage complex transformations
- Provide real time, near real time, and scheduled updates
- Offload production systems
- Support ad-hoc reporting

- Facilitate analytical processing
- Build and maintain data marts, and data warehouses – an effective solution should be able to quickly capture information from multiple sources to create an "instant data mart" that answers a specific business question.
- Ease of use – (easy to install and configure)
- Cost effective
- Handle large volumes of data (Parallel processing)

Other Important information for Extraction and Transformation

- Market Overview Update: ETL – Giga Information Group Inc. (2002)
- ETL – Based Integration for All of Us – Embarcadero Technologies (2003)
- Liberate Data for use by all Business Managers – Embarcadero Technologies (no date)

14.1.7 Loading and Archiving

Support the population of a data source with external data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Methodology	• Copy to floppy disk	Customized applications and scripts Database import and export tools SQL ODBC, JDBC, OLE XML Vendor supplied applications Copy to CDs and DVDs Print to Paper Reformat to Microfilm	Automated capture and apply middleware Stored procedures Virtual warehouses Converting archive data into XML format RMA Information Lifecycle Management issues (*see discussion under Digital Asset Services domain*)	Customized applications and scripts Database import and export tools SQL ODBC, JDBC, OLE XML Automated capture and apply middleware Stored procedures
Data Loading	• Import of data from static batch extractions	Conceptual data model Logical data model Physical data model Similarity of source and target structures Simplified data inter-dependency	Metadata repositories Load utilities that access the metadata of the source and target to automate the transform and load processes	Portable Support for o Oracle o Microsoft SQL Server o Sybase o IBM DB2 o flat files o any JDBC-accessible data source.
Application		Logging for: • Quality assurance • Audit trail		Logging for: • Quality assurance

		• Error tracing and recovery		• Audit trail • Error tracing and recovery
Technology		Platform independence – Minimally must support ODBC and OLE-DB connections. Database using single occurrence of DASD	Data Capture and Apply – allows automated loading of changed data from the source data store to the target data store. Disk mirroring for speed and convenience during the load process	Loading solutions built on an open architecture Re-usable custom extensions Open-standard Java-based server engine Platform independence – Minimally must support ODBC and OLE-DB connections. Database using single occurrence of DASD
Preservation		Data migration Disk upgrades	Preservation strategy built into lifecycle management processes	*See section on Digital Asset Management for additional information*

Guiding Principles for Loading and Archiving:

- Models allow for the correct level of summarization required to support the business.
- Loading of unnecessary data elements increases loading time, environmental requirements and data retrieval times.
- The more diverse target data structures are from the source, the more complex the algorithms and processing required to accurately load the data. When the source and the target differ, the load processes typically must, for example, parse the records, normalize repeating groups, transform values, validate values, substitute code values and generate keys. Greater complexity often requires custom written solutions to translate requirements from the source data structures to the target data structure
- Data inter-dependency determines the order that data must be loaded, and if parallel loading operations are possible. Complex dependencies complicate load strategies and encourage bottlenecks.
- Metadata repositories provide a single, secure, and standardized method for providing metadata
- Metadata must be visible and accessible to both programs and people
- An effective load solution should be portable so that it can interoperate with all enterprise key databases, file types, operating systems, and hardware platforms.
- Minimize network traffic by moving data from sources directly to targets; basing an ETL solution on Java-based components ensures that it will run on a wide variety of operating systems and hardware platforms.
- The load process should log crucial information collected during the load process. This information is critical for quality assurance and serves as an audit trail. Minimally, this log should include date processed, number of records read and written, errors encountered and the

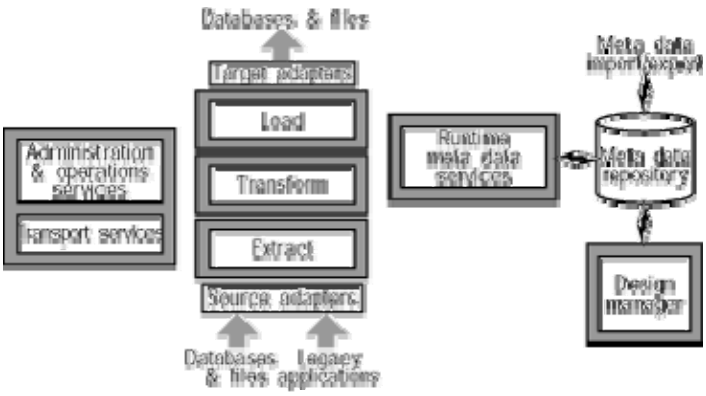
business rules applied. This log should be available while the processes are running to allow early determination of problems and in establishing estimated time of completion. The load process must provide information about its current state in case of failure. Recovering from a system failure can be complicated, error prone, and time consuming. All necessary information to correct the problem must be available. Ideally, the load process includes logic to allow restarting the load from the point of the last known successfully loaded records.

- Methods for moving the data from one source to another must have compatible architectures. Source data that cannot be easily accessed by the target will be difficult to maintain. Minimally must support ODBC and OLE-DB connections. Choose methods where in-house or vendor expertise is readily available. Avoid relying on mechanisms where prompt support is difficult to obtain.
- The next generation of loading solutions should be built on an open architecture, providing these key benefits:
  - Portable
  - Rich in functions
  - Extensible
  - Scalable
  - Easy to use
  - Cost-effective

  Because no ETL solution can anticipate the needs of every customer, an open architecture will support the creation of custom routines. JavaScript provides this capability. The custom extensions should also be re-usable so that work can be shared among team members and across projects. An open standard, Java-based server engine can scale to keep pace with growing data volumes, processing many gigabytes per day within the time available for batch processing.
- Policies must be in place to assure the security of the archived data. Access to archived files requires the same scrutiny as live data. Every attempt must be made to ensure document integrity or fixity. Every attempt must be made to secure archived content from modification in order to guarantee integrity. Policies in place to regularly upgrade the archive's cryptographic algorithms should be in place to avoid exposure.
- Archived data should be evaluated to ascertain if, and when, any or all of the archived data meets sunset criteria. A strategy should be in place to appropriately and securely remove data that is no longer needed.
- Long-term data preservation strategies and processes must be in place for archived data.
- *Long-term* implies the data must be available for a period long enough to generate concern about the impact of changing technologies, including support for new media and data formats, or with a changing user community. This problem may be further complicated by the proprietary nature of some of the software. Encrypted or compressed data would be rendered useless without retaining suitable decryption software, original decryption keys, and compression algorithms.

Definitions for Loading and Archiving include:

- **Disk mirroring** - Creating on two or more physical disk drives exact duplicates of a disk volume to make files accessible in case of failure of one drive of the mirror set. In the case of "load", this provides flexibility during the load by allowing temporary discontinuance of the mirroring process, then, while one copy is loaded the other copy remains available.
- **ETL** - Extract/Transform/Load. This is a 3 stage process delivers data from source systems into a data warehouse. First, the extract function reads data from a specified source database and extracts a desired subset of data. Next, the transform function works with the acquired data - using rules or lookup tables, or creating combinations with other data - to convert it to the desired state. Finally, the load function is used to write the resulting data (either all of the subset or just the changes) to a target database.



The core components of an ETL product.
Courtesy of Intelligent Business Strategies.

The diagram above depicts the major components involved in ETL processing. The bullets below describe each component in more detail:

- **Design manager:** Provides a graphical mapping environment that lets developers define source-to-target mappings, transformations, process flows, and jobs. The designs are stored in a meta data repository.
- **Meta data management:** Provides a repository to define, document, and manage information (i.e., meta data) about the ETL design and runtime processes. The repository makes meta data available to the ETL engine at run time and other applications via an import/export mechanism.
- **Extract:** Extracts source data using adapters, such as ODBC, native SQL formats, or flat file extractors. These adapters consult meta data to determine which data to extract and how.
- **Transform:** ETL tools provide a library of transformation objects that let developers transform source data into target data structures and create summary tables to improve performance.
- **Load:** The ETL load element's primary responsibility involves inserting transformed data into a data warehouse database through the use of target data adapters, such as SQL or native bulk loaders, to insert or modify data in data warehouse databases or files. Generally, data warehouses are updated periodically rather than continuously, and large numbers of records are often loaded to multiple tables in a single data load. The data warehouse may be taken offline during update operations so that data can be loaded faster. The design of the loading element should focus on efficiency and performance to minimize the data warehouse offline time. Although ETL processes primarily support data warehouses, nearly twenty percent of ETL applications are in non-data warehouse applications. Examples of this include moving data between applications, customer data integration, and database consolidation. ETL tools
- **Transport services:** ETL tools use network and file protocols (e.g., FTP) to move data between source and target systems and in-memory protocols (e.g., data caches) to move data between ETL run-time components.
- **Administration and operation:** ETL utilities let administrators schedule, run, and monitor ETL jobs as well as log all events, manage errors, recover from failures, and reconcile outputs with source systems.

These components are used in most vendor-supplied ETL tools, but data warehouse developers can also build them. A majority of companies have a mix of packaged and homegrown ETL applications. In some cases, they use different tools in different projects, or they use custom code to augment the functionality of an ETL product.

- **Open Architecture** - A term used to describe any computer or peripheral design that has published specifications. A published specification lets third parties develop add-on hardware for an open-architecture computer or device.

- **Virtual Data Warehouse -** a simple data warehousing tool that provides end users with direct access to operational data on legacy databases. In the 'virtual' scenario, data is not moved from source databases to a target data warehouse as in a traditional data warehouse scenario. In fact, a target database and data warehouse does not exist within a virtual environment. Instead, users employ intuitive end-user tools to access data directly from operational files, using common business terms.
- **Archive** - Once a data warehouse or other data store is constructed and functioning, at some future point, warehouse management will most likely face challenges in managing query performance, ETL performance, and storage. To ensure that a balance is maintained between providing customers with the warehouse information they need, and ensuring that less critical historical information is accessible in some way, warehouse managers must establish an archive strategy. Archiving entails the aging out of data that is no longer required for immediate access by the data warehouse user population. The purpose of archiving is to help ensure that relevant data is accessible with as little wait time as possible. An archive strategy takes into account the entire storage requirements for the data in the data warehouse and balances this with the needs of the end users. Warehouse managers must take into account the totality of the types of trending and analytics required by their customers and what those needs translate into as storage requirements. Based on the analysis of these needs, the warehouse team can begin devising an archive strategy.
- **RMA** – Records Management Application, a software program that files, stores, retrieves, and manages the retention of electronic records. These products work with a variety of common desktop applications such as e-mail software, digital imaging tools, electronic spreadsheet and presentation products. RMAs have a variety of useful features, such as: Centralized filing and storage of electronic records, version control, robust search and retrieval, and automated disposition that is compliant with legal requirements.

References:

- Ekerson, W. (2003). *The evolution of ETL*. http://www.tdwi.org/

- ISO Archive Standards – Reference Model Papers. http://nssdc.gsfc.nasa.gov/nost/isoas/ref_model.html

14.1.8 Data Recovery

Support the restoration and stabilization of data sets to a consistent, desired state

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Device-based	IDE SCSI PATA.	Fiber-Channel based. Image-based or snapshot technology. RAID 0 – 5 EIDE	Image-based or snapshot technology RAID 6 SATA FireWire (IEEE-1394) LTO-3 AIT-6 Sarbanes-Oxley	End-to-end redundancy High fault tolerance SATA FireWire (IEEE-1394)
Software-based		Automated recovery.		Automated recovery

		Point-in-time recovery. System-managed storage		Point-in-time recovery System-managed storage
Network-based		Adherence to NDMP	NDMP Version 5 SANs based upon IP	NDMP Version 5

Recommendation for Data Recovery:

- Establishment, testing, and maintenance of a disaster recovery or business contingency plan.
- Ability to take backups while data is online to users.
- Ability to perform up-to-the-minute (point-in-time) recovery.
- Ability to take complete or incremental (changes only) backups.
- Ability to backup data logically (specific files/datasets) as well as physically (device-based).
- Use of storage technology (hardware) with built-in redundancy for inherent, transparent recoverability (i.e. data striping techniques used with RAID (Redundant Array of Independent Disks), data mirroring, etc.)
- Establishment and maintenance of backup metadata.
- Use of compression algorithms to maximize backup storage media.
- Flexibility of recovery technology to recover from backups taken by older software versions.
- Use of system-managed storage systems with built-in recovery mechanisms such as automated migration and recall facilities.
- Ability to operate within the network capacity of the enterprise.

Definitions for Data Recovery include:

- **ATA** – Advanced Technology Attachment. A mass storage device interface, also called IDE (Integrated Drive Electronics).
- **Data Mirroring** – The act of copying data from one location to a storage device in real time. Because the data is copied in real time, the information stored from the original location is always an exact copy of the data from the production device. Data mirroring is useful in the speedy recovery of critical data after a disaster. Data mirroring can be implemented locally or offsite at a completely different location.
- **Data Recovery** – The process of salvaging data from damaged, failed, wrecked or inaccessible primary storage media when it cannot be accessed normally. Often the data is being salvaged from storage media formats such as hard disk drive, storage tapes, CDs, DVDs, RAID, and other electronics. This can be due to physical damage to the storage device or logical damage to the file system that prevents it from being mounted by the host operating system. Although there is some confusion as to the term, data recovery can also be the process of recovering delete information from a storage media in for example forensic purposes. http://en.wikipedia.org/wiki/Data_recovery.
- **Data Striping** – The writing of data across multiple drives and adding of redundancy to permit recovery of the data if one of the drives fails. It is a RAID technique.
- **IEEE 1394** – an international high-performance serial-bus standard which TI has released in a chipset that offers the real-time data transfer of video, audio and peripheral applications through a universal I/O interface. With this technology, digital cameras, CD-ROMs, printers, hard-disk drives and audio/stereo equipment can move data at high speeds to desktops and portable computers through a single cable. Also known as FireWire.
- **Multipathing** – Multipathing software lets you define and control redundant physical paths to I/O devices, such as storage devices and network interfaces. If the active path to a device becomes unavailable, the software can automatically switch to an alternate path to maintain availability. This capability is known as *automatic failover.*

- **NDMP** – Network Data Management Protocol. The NDMP initiative was launched to create an open standard protocol for network-based backup for network-attached storage.
- **PATA** – Parallel Advanced Technology Attachment. A device interface based upon parallel signaling technology.
- **RAID** – Redundant Array of Independent Disks. A RAID array is a series of drives which together act as a single storage system. In most configurations this storage system can tolerate the failure of a drive without losing data and depending on how it is configured, can rebuild itself once the failed component is replaced. See this link for various levels and definitions of RAID: http://www.integratedsolutions.org/raid_ov.htm
- **Sarbanes-Oxley** – Pub. L. No. 107-204, 116 Stat. 745, also known as the *Public Company Accounting Reform and Investor Protection Act of 2002* and commonly called SOX or SarbOx
- **SAN** – Storage Area Network. A storage area network (SAN) is a high-speed special-purpose network (or subnetwork) that interconnects different kinds of data storage devices with associated data servers on behalf of a larger network of users. Typically, a storage area network is part of the overall network of computing resources for an enterprise.
- **SATA** – Serial Advanced Technology Attachment. A device interface based upon serial signaling technology.

Other Important information for Data Recovery

- Tanenbaum, A. & Woodhull, A.S. (1997). *Operating Systems: Design And Implementation*, 2nd ed. New York: Prentice Hall.
- Network-based backup standard (NDMP): www.ndmp.org

- Backup products: http://www.ndmp.org/products/index.shtml
- "Enterprise Design for Backup and Recovery Services". http://www.microsoft.com/technet/itsolutions/wssra/raguide/BackupandRecoveryServices/igbrbp_2.mspx?mfr=true
- Tape media summary: http://www.pcmag.com/encyclopedia_term/0,2542,t=magnetic+tape&i=46499,00.asp
- Compliance laws/acts affecting storage and other requirements for recoverability: http://www.archivas.com:8080/product_info/z3_pdfs_gh23/Compliance.pdf

14.1.9 Data Classification

Allow the classification of data

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
General		Criteria: o Confidentiality Integrity o Trustworthiness o Availability	Business requirements-based Multi-faceted taxonomy serving: ILM enterprise content management compliance data mining and decision support security	Criteria: o Confidentiality o Integrity o Trustworthiness o Availability

			Basis for enterprise information asset management	
Information Lifecycle Management		By-product of hierarchical storage management (HSM) implementations where the pervasive classification criterion was the age of the data. Criteria based on access or availability and recovery requirements, and cost.	Data classification assigns a level of business importance, availability, sensitivity, security, and regulation requirements to data. This process matches classifications of data with their proper tier of the storage infrastructure, as well as the appropriate security, compliance, data protection, migration and disaster recovery levels. Establishing a solid classification methodology is fundamental to organizational adoption and the ILM strategy. Example categories: o Mission-critical online data o Business-critical online data o Accessible online data o Nearline data o Offline data	Criteria based on access or availability and recovery requirements, and cost. Data classification assigns a level of business importance, availability, sensitivity, security, and regulation requirements to data. This process matches classifications of data with their proper tier of the storage infrastructure, as well as the appropriate security, compliance, data protection, migration and disaster recovery levels. Establishing a solid classification methodology is fundamental to organizational adoption and the ILM strategy. Example categories: o Mission-critical online data o Business-critical online data o Accessible online data o Nearline data o Offline data
Security		Classification of data determines the extent to which the data needs to be controlled / secured Classification of data		Classification of data determines the extent to which the data needs to be controlled / secured • Classification of

		is indicative of value in terms of *Business Assets*. Essential to distinguish what is of little value and what is highly sensitive and confidential Example classification scheme: o *Confidential Data* – Do Not Share o *Restricted Data* – Business Need To Know o *Sensitive Data* – Share with Care o *Public Data*		data is indicative of value in terms of *Business Assets*. Essential to distinguish what is of little value and what is highly sensitive and confidential Example classification scheme: o *Confidential Data* – Do Not Share o *Restricted Data* – Business Need To Know o *Sensitive Data* – Share with Care o *Public Data*
Subject / Topic	Uncontrolled	Controlled Taxonomy within organizational units; Controlled Vocabulary	Controlled Taxonomy and Vocabulary for the enterprise; Concept Mapping	Controlled Taxonomy within organizational units; Controlled Vocabulary Concept Mapping

Definitions for Data Classification include:

- **Confidentiality** – the privacy of an asset
- **Integrity** – If data is damaged or incorrectly altered, how important is it to restore it to a trustworthy state with minimum loss
- **Availability** – requirement that an asset be available to an authorized person, entity, service or device; general rule: the more critical the data is, the higher its availability ranking will be
- **Data classification** – the conscious decision to assign a level of sensitivity to data as it is being created, amended, enhanced, stored, or transmitted. The classification of the data should then determine the extent to which the data needs to be controlled / secured and is also indicative of its value in terms of *Business Assets*. The classification of data and documents is essential to distinguish what is of little value and what is highly sensitive and confidential.

Other Important information for Data Classification

Data can be classified from a variety of perspectives. Three primary perspectives are:

- Information Lifecycle Management
- Security
- Subject / Topic (Refer to the *Support Services Domain, Interoperability section* for further information.)

"The future state of data classification will involve a much broader perspective, motivated by business requirements. The taxonomy will serve several needs (including ILM, enterprise content management,

compliance, data mining and decision support, and security) and will emerge as a basis of enterprise information asset management."

Example Security categories:

- "*Top Secret*: Highly sensitive internal documents e.g. pending mergers or acquisitions; investment strategies; plans or designs; that could seriously damage the organization if such information were lost or made public. Information classified as Top Secret has very restricted distribution and must be protected at all times. Security at this level is the highest possible.
- *Highly Confidential*: Information that, if made public or even shared around the organization, could seriously impede the organization's operations and is considered critical to its ongoing operations. Information would include accounting information, business plans, sensitive customer information of bank's, solicitors and accountants etc., patient's medical records and similar highly sensitive data. Such information should not be copied or removed from the organization's operational control without specific authority. Security at this level should be very high.
- *Proprietary*: Information of a proprietary nature; procedures, operational work routines, project plans, designs and specifications that define the way in which the organization operates. Such information is normally for proprietary use to authorized personnel only. Security at this level is high.
- *Internal Use Only*: Information not approved for general circulation outside the organization where its loss would inconvenience the organization or management but where disclosure is unlikely to result in financial loss or serious damage to credibility. Examples would include, internal memos, minutes of meetings, internal project reports. Security at this level is controlled but normal.
- *Public Documents*: Information in the public domain; annual reports, press statements etc.; which has been approved for public use. Security at this level is minimal."

Legal Considerations:

Sarbanes-Oxley – Pub. L. No. 107-204, 116 Stat. 745, also known as the *Public Company Accounting Reform and Investor Protection Act of 2002* and commonly called SOX or SarbOx

Gramm-Leach-Bliley – *Public Law 106-102, Title V, Privacy;* Financial Modernization Act*; regulates the sharing of personal information about individuals who obtain financial products or services from financial institutions.*

HIPAA – *Health Insurance Portability and Accountability Act (HIPAA)*; Pub. L. 104-191, 110 Stat. 1936 enacted by the U.S. Congress in 1996.

FERPA – *The Family Educational Rights and Privacy Act of 1974* (FERPA or the Buckley Amendment) is a United States federal law codified at 20 U.S.C. § 1232g, with implementing regulations in title 34, part 99 of the Code of Federal Regulations.

References:

UT-Austin Data Classification Guidelines,
www.utexas.edu/its/policies/opsmanual/dataclassification.html
Adapted from "Classification of Data"
(http://www.stanford.edu/group/security/classification/classification_of_data.html), with permission from Stanford University, Stanford, California 94305-4102.]

Butterfield, Rich, Data Classification: a Prerequisite to ILM, *Storage Networking World Online*, May 30, 2005, www.snwonline.com/implement/data_class_05-30-05.asp

Storage Networking Industry Association, Information Lifecycle Management Initiative,
(http://www.snia-dmf.org/ilmi/index.shtml)
Various whitepapers on Data Classification.
Group is working to set industry standards for ILM (storage perspective)

The Information Security Glossary – Data Classification, www.yourwindow.to/information-security/gl_dataclassification.htm

StorageTek, *White Paper: Best practices in data classification for information lifecycle management*, January 2005

Example Vendors (ILM data classification):

- Abrevity
- Arkivio
- BridegeHead Software
- Index Engines
- Kazeon Systems
- Njini
- Scentric
- StoredIQ
- Trusted Edge

14.2 Human Resources

Capabilities within this Service Type provide for the recruitment and management of personnel.

14.3 Financial Management

Capabilities within this Service Type provide the accounting practices and procedures that allow or the handling of revenues, funding and expenditures.

14.4 Assets / Materials Management

Capabilities within this Service Type support the acquisition, oversight and tracking of an organization's assets.

14.4.1 Computers Automation Management

Support the identification, upgrade, allocation and replacement of physical devices, including servers and desktops, used to facilitate production and process driven activities.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Device Management		Standard Replacement cycles (implies inventory of assests) MS Report Manager 2006 Compaq ZeroAdmin Microsoft MMC		
Event Management	Unicenter TNG Novell Manage Wise	CA – OPS SNMP Protocol MIB Jobtrack CA Unicenter TNG	PeopleSoft Asset Manager	

14.5 Development and Integration

Capabilities within this Service Type provide communication between hardware/software applications and the activities associated with deployment of software applications.

14.5.1 Legacy Integration

Support the determination of long-term goals and the identification of the best approach for achieving those goals

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Legacy Integration Planning		Migrate from mainframe to client-server "Rip-and-Replace" – replacement of large, installed bases of mature software with new products	Service Oriented Architecture Service Oriented Programming Expose legacy transactions as services Application rationalization Application Portfolio Management (APM) Decision metrics: • application size (in lines of code or function points) • Complexity • maintenance activity • calendar year cost Compliance exposure (HIPAA, Sarbanes Oxley, FERPA, etc.)	
Legacy Integration Strategies	Componentize	Web-to-Host "Green-screen in a browser" Interface re-engineering: converts the 3270 interface to a graphical UI and permits navigational changes to the		

		screen flow Screen components: wrap screen functions as Enterprise JavaBeans (EJB), XML, and component object model (COM) components Web-Service-to-Host Migrate from Platform Migrate from DBMS Migrate from Language Outsource Replacement • Rewrite application • Use packaged application • Use hosted application • Acquire open-source version Application retirement		

Definitions for Legacy Integration include:

- **Legacy systems:** Applications that are more than three years old. Applications written three or more years ago using HTML, XML, Perl, C#, and Java qualify as legacy applications because they share so many attributes with their older siblings — the original authors are gone, they were poorly documented, and as a result, they are now poorly understood, so nobody wants to work on them. Applications like this aren't really legacy, but they do share legacy characteristics — they share knowledge loss, and knowledge loss increases the cost of keeping an application operational by a large margin.
- **Web-to-host:** focus at the presentation or user interface (UI) layer and expose 3270 screens in a number of ways
- **Web-service-to-host:** functions similarly to Web-to-host, except that the end target is a service. The tools wrap the 3270 transactions in WSDL and expose it to calling programs via SOAP. Some of these tools and the newer releases of CICS work with distributed program link (DPL) technology that expose the CICS COMMAREA as part of the service interface.
- **Application Portfolio Management (APM):** process and technology that inventories IT applications and develops application metrics to permit IT managers to streamline inventories and reduce overall costs

Other Important information for Legacy Integration

- Murphy, Phil, *Got Legacy? Four Fates Await Your Applications*, Forrester Research, Inc. January 10, 2006 http://www.forrester.com/Research/Document/0,7211,35370,00.html

14.5.2 Enterprise Application Integration

Support the redesigning of disparate information systems into one system that uses a common set of data structures and rules

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Integration Design		Manual Process	Application Portfolio Management (APM)	
Application Integration Methods		Enterprise Service Bus (ESB) Rich clients Composite Applications Collaboration Platforms (proprietary) Platform vendor technology Independent integration vendor technology; features such as: • Embedded process modeling • XML-based messaging • Vertical industry templates • Standards-based (J2EE, Web Services, BPM, BAM, .NET, etc.)	SOA Multichannel integration Business Process Management (BPM) using BPEL Collaboration Platforms (open standards-based) Integration Platforms with functions including: • Message bus • Integrated process management • Presentation and user interaction features • Life-cycle management	

Definitions for Enterprise Application Integration include:

- **Composite Applications:** facilitate business processes across multiple functions and systems. They are specific to user roles, accomplish business goals (instead of needing multiple applications), foster collaborative work, respond to business events, and adapt to new requirements.

14.5.3 Data Integration

Support the organization of data from separate data sources into a single source using middleware or application integration as well as the modification of system data models to capture new information within a single system

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Data Integration Methods		OAI-PMH standard Enterprise Information Integration platforms for distributed query and aggregation, data grids, ETL Federated search technology Enterprise content management	Universal Data Integration Platforms (consolidate selected data integration functions to satisfy requirements for real-time information): • Data access control • Service registries • • Transformation, query, aggregation, metadata mapping • Data process control • Business intelligence components • Quality remediation • Quality of Service • Life-cycle support Data Integration Services (wrap often separate data integration functions in Web services and/or other widely used interfaces to enable reuse and interconnection of integration silos. These data integration services don't actually perform the data integration functions; they	

			merely provide standard access points to them.): Similar functions as above	

Other Important information for Data Integration

- Rymer, John R., Mike Gilpin, and Ken Vollmer, *Integration Landscape 2005*, Forrester Research, Inc., December 22, 2004. http://www.forrester.com/Research/Document/0,7211,35370,00.html

14.5.4 Instrumentation and Testing

Support the validation of application or system capabilities and requirements

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Functional Testing		Manual Automatic		
Performance Testing		Manual Automatic		

Definitions for Instrumentation and Testing include:

- **Functional Testing:** Testing to verify that application functionality conforms to predefined specifications.
- **Automated Functional Testing Tools:** Tools to build and automatically execute suites of functional tests.

Other Important information for Instrumentation and Testing

- Schwaber, Carey, *The Forrester Wave™: Functional Testing Solutions, Q2 2006*, Forrester Research, Inc., May 31, 2006, http://www.forrester.com/Research/Document/0,7211,37587,00.html
- Schwaber, Carey, *Performance-Driven Software Development*, Forrester Research, Inc., February 28, 2006, http://www.forrester.com/Research/Document/0,7211,37519,00.html

14.5.5 Software Development

Support the creation of both graphical and process application or system software.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Chapter 15 Support Services

The Support Services Domain defines the set of cross-functional capabilities that can be leveraged independent of Service Domain objective and/or mission.

15.1 Security Management

15.1.1 Identification and Authentication

Support obtaining information about those parties attempting to log on to a system or application for security purposes and the validation of those users

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
User Security	Address based	Token based 2 factor Certificates (x.509) Passwords Radius TACACS	Smart Card Kerberos Biometrics	

15.1.2 Access Control

Support the management of permissions for logging onto a computer or network

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Authorization / Access Control	Access control lists X.500 Password Protected directories OS based systems	Directory based services **L**ightweight **D**irectory **A**ccess **P**rotocol (LDAP) Active Directory		

Other Important Information for Access Control:

- The management of admission to system and network resources. The first part of access control is authenticating the user, which proves the identity of the user or client machine attempting to log on. The second part is granting the authenticated user access to specific resources based on company policies and the permission level assigned to the user or user group.

15.1.3 Encryption

Supports the encoding of data for security purposes.

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Encryption	Data Encryption Standard (DES)	3 DES/RSA	Advanced Encryption Standard (AES)	

Definitions for Encryption include:

- Encryption is the reversible transformation of data from the original (the plaintext) to a difficult-to-interpret format (the cipher text) as a mechanism for protecting its confidentiality, integrity and sometimes its authenticity. Encryption uses an encryption algorithm and one or more encryption keys.

15.1.4 Intrusion Detection

Support the detection of illegal entrance into a computer system

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
IDS		Signature based	Heuristic / anomaly based	
IPS		Signature based	Heuristic / anomaly based	

Definitions for Intrusion Detection include:

- Intrusion Detection System (IDS) Software that detects an attack on a network or computer system. A Network IDS (NIDS) is designed to support multiple hosts, whereas a Host IDS (HIDS) is set up to detect illegal actions within the host. Most IDS programs typically use signatures of known cracker attempts to signal an alert. Others look for deviations of the normal routine as indications of an attack.
- Intrusion Prevention System (IPS) is a preemptive approach to network security used to identify potential threats and respond to them swiftly. Like an intrusion detection system (IDS), an intrusion prevention system (IPS) monitors network traffic. However, because an exploit may be carried out very quickly after the attacker gains access, intrusion prevention systems also have the ability to take immediate action, based on a set of rules established by the network administrator.

15.1.5 Verification

Supports the confirmation of authority to enter a computer system, application or network

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Definitions for Verification include

- **Verification -** The process of affirming that a claimed identity is correct by comparing the offered claims of identity with previously proven information stored in the identity card or PIV system**.**

15.1.6 Digital Signature

Guarantee the unaltered state of a file

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Definitions for Digital Signature include

- **Digital Signature -** A digital signature is an electronic rather than a written signature that can be used by someone to authenticate the identity of the sender of a message or of the signer of a document. It can also be used to ensure that the original content of the message or document that has been conveyed is unchanged.

15.1.7 User Management

Support the administration of computer, application and network accounts within an organization

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
User Management	Product specific	Directory based services		
Security Administration		Product specific		

15.1.8 Role / Privilege Management

Support the granting of abilities to users or groups of users of a computer, application or network

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

15.1.9 Audit Trail Capture and Analysis

Support the identification and monitoring of activities within an application or system

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
	SYSLOG	Vendor specific Network analysis tools		
		Vendor specific Forensic analysis tools		
		Vendor specific DB administration tools		
		Vendor specific Host based application tools		

Definitions for Audit Trail Capture and Analysis include:

- An audit trail is a record showing who has accessed an Information Technology (IT) system and what operations the user has performed during a given period. This can relate to an individual system, data base or network traffic.

Related Policies & Procedures

Confidential Information

State law and regulations require agencies to keep some information confidential, or give the agency the option of keeping it confidential. Following are examples of such statutes.

- Kansas Open Records Act (K.S.A. 45-215 et seq.) It is the general policy of this Act that public records be open for inspection by anybody. Additionally, it specifies that computer data are considered a record. However, the act contains a long list of information that <u>may</u> be kept confidential by agencies. For example, security information, medical treatment records, personnel performance records, or information that could reveal the identity of any undercover law enforcement officer can all be kept confidential at the discretion of the agency holding the records.

Other laws <u>mandate</u> closing records or parts of records:

- Child abuse records and reports (K.S.A. 38-1507)

- Juvenile offense records (K.S.A. 38-1607; 38-1608(c))
- Individually identifiable drug abuse treatment records (K.S.A. 65-4608; K.S.A. 45-221(a)(3))
- Criminal history record information (K.S.A. 22-4701 et seq.; K.S.A. 22-4707; K.S.A. 22-4907; K.A.R. 10-12-1; 10-12-2.)
- Election ballots (K.S.A. 25-2422; K.S.A. 25-2708; K.S.A. 25-3008)
- Judicial Pre-sentence reports (K.S.A. 21-4605)
- Grand jury proceeding records (K.S.A. 22-3012)
- Parole Board pre-parole report and supervision history (K.S.A. 22-3711)
- Mental patients commitment and treatment records (K.S.A. 59-212; K.S.A. 59-2931; K.S.A. 59-2941)
- Long-term care facility residents' information (K.S.A. 75-5921)
- Health-Care peer review records (K.S.A. 65-4915(b))
- Adoption records (K.S.A. 59-212; K.S.A. 59-214; K.S.A. 59-2122; K.S.A. 65-2423)
- Income tax reports and returns (K.S.A. 79-3234(b))
- Census data (K.S.A. 11-204(b))
- Portions of corporate annual report data held by the Secretary of State (K.S.A. 17-7514; K.S.A. 17-7656)
- Information acquired by the State Child Death Review Board (K.S.A. 22a-243(j))
- Complaints filed with the Election Commission (K.S.A. 25-4161(b))
- Racing Commission background checks (K.S.A. 74-8804(o))
- Marriage, birth and death certificates (K.S.A. 65-2422)
- Information maintained by the Crime Victim Compensation Board (K.S.A. 74-7308(c))
- Diversion agreements in district court (K.S.A. 22-2911)
- Legislative Post Audit has the same duty of confidentiality as the audited agency (K.S.A. 46-1106(g))

WLAN Security Mechanisms

To meet the stated security goals of confidentiality, integrity, and availability, the State of Kansas will implement the current state of the art security mechanisms for wireless local area networks, including 802.11i authentication and encryption.. Security mechanisms will be deployed in the wireless local area network in the following areas:

- Security of the Control Channel and User Channel between Access Points and the Access Controller
- Secure authentication of the wireless user to the network with 802.11i authentication and encryption when accessing State agency networks
- Captive Portal authentication for Guest Access
- Segregation of wireless network traffic in the wired network through Virtual LANs
- Intrusion Protection – preventing unauthorized users and access points from connecting to the network or intercepting wireless transmission of authorized users.
- Single sign-on capability leveraging existing authentication databases

Securing the Control Channel and User Channel between the Access Point and the Access Controller

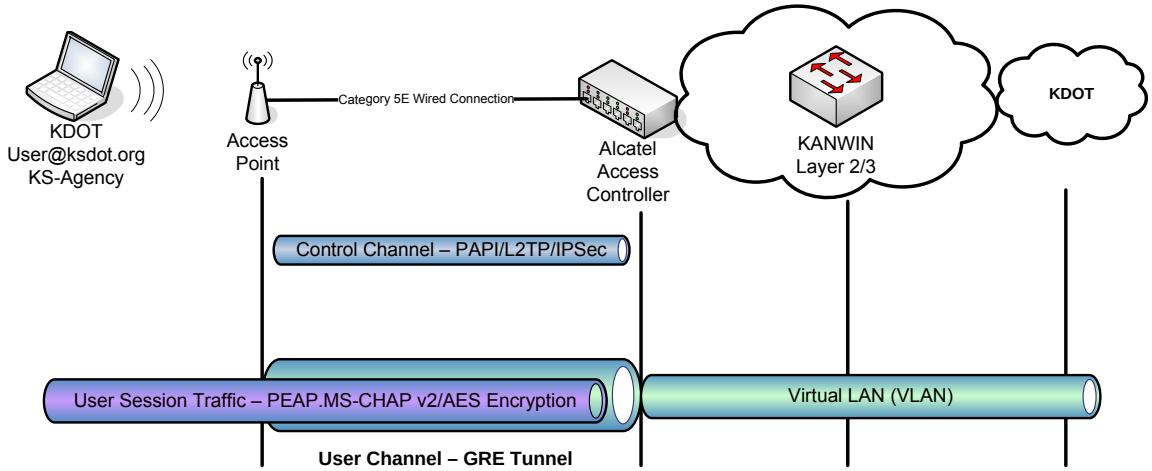


Figure 1 - Control Channel and User Channel Security

15.1.10 Forensics

15.1.11 Incident Response

15.1.12 Risk Management

15.1.13 Addition Security Areas

Physical Security

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Physical access	Key locks	Cypher lock Key card	Smart Card Biometrics	

Application Security

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Application Security		S/MIME PGP SSL Middle-ware Signed Java Anti-virus	Transport layer Security (TLS)	

Hardware/System Security

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target

Hardware / System Security		Windows Domains TOPSECRET / RACF TCACS NDS/Novell Anti-virus control Intrusion Detection/prevention systems		

Data Security

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
	Embedded passwords	CORBA Anti-virus control PGP	AES (encryption)	

Network Security

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
		Firewalls / router ACL IPSEC v4 Encryption (3DES/RSA) Encrypted VPN Intrusion Detection/Prevention	AES (encryption) IPSEC v6	IPSEC v6

15.2 Collaboration

Capabilities within this Service Type allow for the concurrent, simultaneous communication and sharing of content, schedules, messages and ideas within an organization.

15.2.1 E-Mail

Support the transmission of memos and messages over a network

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Server Side		Cross-platform support Supports IMAP and / or POP protocols Rules support	Exchange server roles such as bridgehead services (an internal routing server that ties into Active Directory and applies policy), unified messaging services, client access services (for mobile and remote Exchange access and Web services), and mailbox services (for free/busy lookups and public folder administration) Support for Web Services GUI administration interface Unified messaging Enterprise Collaboration platform	Cross-platform support Supports IMAP and / or POP protocols Rules support GUI administration interface
Client Side		Ability to tie to a user account Filtering Junk mail filters Folders Support Address-book Directory Support	Personal Information Management Systems Anti-Phishing support	*Current + Emerging*

		Attachment Support Supports IMAP and / or POP protocols Customizable Supports certificates and encryption		
Mobile Devices				

Description	Example Products	Notes
Server Side	• Exchange • Groupwise • Open-Xchange • Mirapoint • Chandler / Westwood • Novell Hula Project • Oracle Collaboration Suite • Sendmail • Live Meeting	
Client Side	• Outlook • Outlook Web Access • Eudora • Mozilla Thunderbird • Mail (Macintosh)	
Mobile Devices		

15.2.2 Threaded Discussion

Support the running log of remarks and opinions about a given topic or subject

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Forums				
Wikis		[see section above on *Workgroup / Groupware*]		
Blogs		[see section above on *Workgroup / Groupware]*		

Other Important information for Threaded Discussion

- Threaded discussion technologies are still in an infant stage in state government. Like other organizations, they are beginning to gain more traction in our technology landscape. Nonetheless deployment of these types of applications is something that is emerging, and ubiquitous deployment is still many years out. As a result, there are no set standards that the State is willing to currently adopt, but it is an area that the State is following closely.

15.2.3 Document Library

Support the grouping and archiving of files and records on a server

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Repository			OAIS standard support for preservation repositories	OAIS standard support for preservation repositories

Description	Example Products	Notes
Document Library / Repository	Novell Server / LAN Stores DSpace Document management systems Content management systems	

Definitions of Document Library technology include:

- *Reference Model for an Open Archival Information System (OAIS)* – ISO 14721, 2002. <http://public.ccsds.org/publications/archive/650x0b1.pdf>

Other Important information for Document Library

- Refer to the Document Management / Digital Asset group for technical specifications.

15.2.4 Shared Calendaring

Allow an entire team as well as individuals to view, add and modify each other's schedules, meetings and activities

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Security		User Level Security		
Access		Group Access Multiple User AccountsConferenc e Rooms	CalDAV	Group Access Multiple User Accounts Conference Rooms CalDAV

Description	Example Products	Notes
	Exchange Groupwise e-Groupwise Google Calendar Chandler / Westwood Lotus Notes	

Description	Example Products	Notes
	Open-Xchange Oracle Collaboration Suite	

15.2.5 Task Management

Support a specific undertaking or function assigned to an employee

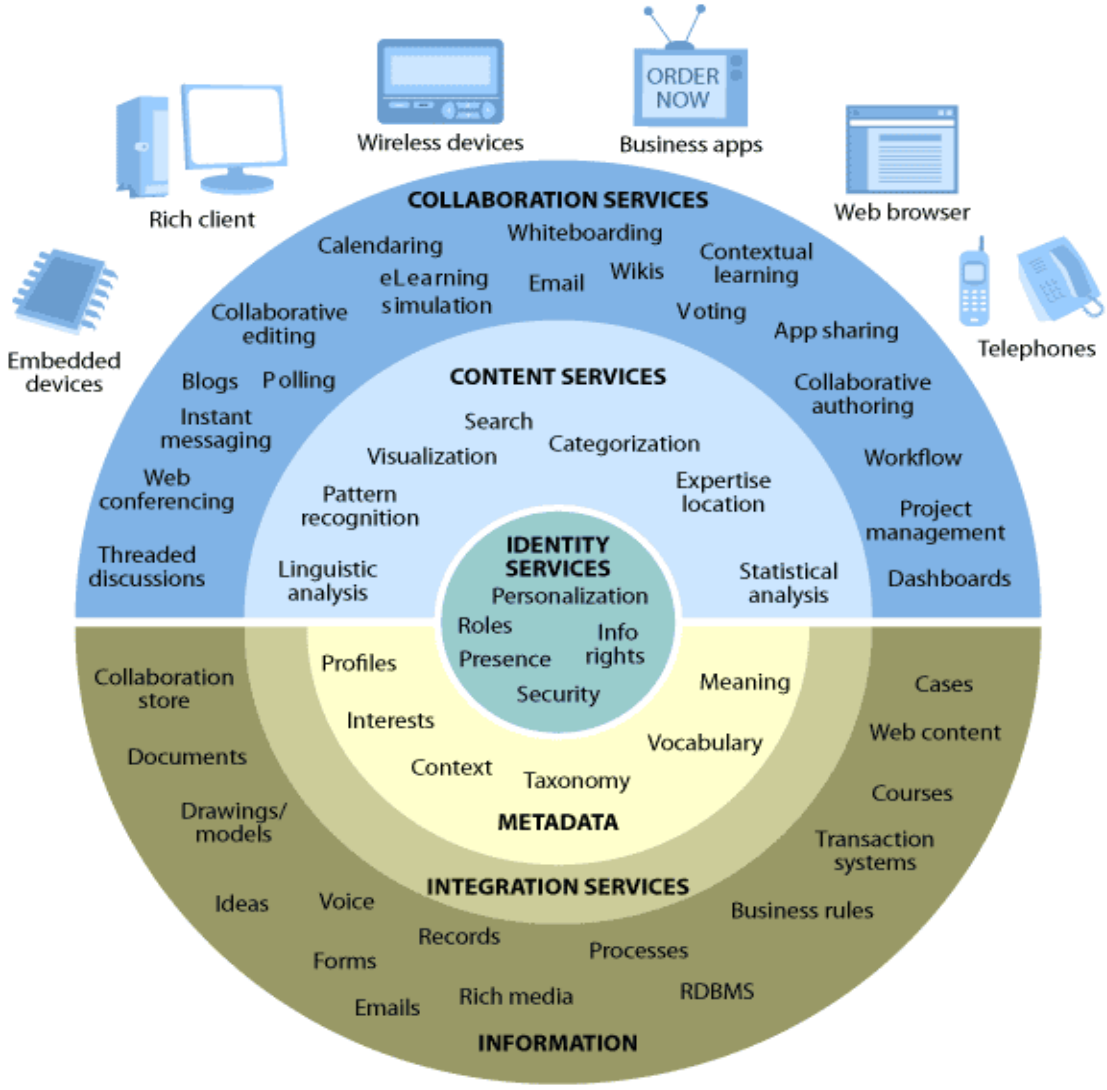
Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Task Management		Add Resources Progress Tracking Ability to Baseline Event Triggers		

Description	Example Products	Notes
	Exchange Groupwise Microsoft Project Open Workbench Clarity Lotus Notes	

Other Important information for Task Management

Additional Collaboration Services include:

- Web conferencing
- Polling
- Voting
- Collaborative editing
- Collaborative authoring
- Workflow
- eLearning simulation
- Whiteboarding
- Contextual learning
- Application sharing
- Project management
- Dashboards



Source: Forrester Research, Inc.

Moore, Connie, and Erica Driver, *The Information Workplace Will Redefine The World Of Work At Last*, June 1, 2005, Forrester Research Inc., http://www.forrester.com/Research/Document/0,7211,36217,00.html

15.3 Search

Capabilities within this Service Type provide for the probing and lookup of specific data from a data source.

“It is taken for granted that government information will always be distributed among many resources that are separately maintained. The searching challenge is not to define standards to be implemented at any one site, but to define searching standards that are common to many sites.

“No search service standard should or could fully supplant the many other search mechanisms optimized for particular technologies or communities of practice. Rather, the search service standard for government data and information resources should be implemented typically as a supplement to other search mechanisms that may be required for reasons other than broad scale interoperability. Otherwise, rather than enhancing public access to diverse information sources across government, the ultimate affect might be to collapse that diversity and trivialize information search techniques generally.

“Major Requirements:

- **Direct access to data:** the search service standard should be of a general-purpose design that is independent of specific access mechanisms. Designing on an abstract model of the search function allows the search service standard to accommodate a wide range of catalogs, directories and databases. The search service standard is required to support interoperable searching of such model registries as these become well-defined. The search service standard should support different levels of access control, such as restrictions by service, session, distributed resource, database, record, or data element.
- **Library catalog search:** Any search service standard must address the huge installed base of public mechanisms for search of government data and information supported among the nation's libraries, and a significant portion of that information is in paper and other tangible media. Public policy interests require that access to information in the public domain must not be encumbered by constraints other than those explicitly provided under law.
- **Internet search engines:** As a practical matter, any search interface standard advocated for use by governments must be able to work with commonly available Internet search engines.
- **Government locaters:** In addition to the public need for cross-agency coherence in descriptions of information holdings, the search service standard should be useful for agencies in organizing information for its internal needs. Government organizations require a search service standard addressing all types of government information held over the long term electronically or in hard copy
- **Search Request / Response:** As a standard specification, a search service is a formal definition of what request messages are handled at the service interface and what response messages may be returned. To maximize interoperability, the search service should be at least implementable over the Internet using TCP/IP, HTTP/HTTPS, HTTP GET and HTTP POST. The search service should be specified in a widely recognized interface definition language such as WSDL (Web Services Definition Language), and that definition should include a query language, a query syntax, and standardization of a result set schema. The query language of the search service standard must precisely define how searches are expressed and communicated between a client component and a server component. The response from a search service will be a result set, that typically lists matching documents sorted by some criteria and each item being available in a selection of original formats or generalized formats such as XML.
- **Structured Information:** To meet requirements expressed elsewhere in this document, the search service standard must support searching of structured information.
- **Semantic Mapping:** The query evaluation function of the search service must be able to handle "abstract concepts" such as name, category, and date according to what they mean semantically rather than merely how they may be labeled syntactically. Handling search concepts at the semantic level is crucial for interoperability, as it is impossible for searchers to know the particular structural elements of every searchable collection.
- **File attributes and object properties:** The search service standard should adopt readily to the underlying data model of named properties and property sets defined for objects.

- **Digital document metadata:** The search service standard should have a history of production use for searching metadata variants.
- **Collection organization:** In addition to organizing content within a searchable collection, designers of search facilities are often faced with compiling information as a logical collection. The search service standard itself should not standardize on any particular approach, but should be compatible with many and diverse approaches to compiling collections of information
- **Service registries:** A gateway approach would allow the search service standard to interoperate with business and services registries using either the ebXML model or the Universal Description, Discovery, and Integration (UDDI) model. This approach would likely parallel the approach for interoperable searching of any registry supporting the ISO 11179 "meta-model" guidelines. These are of particular interest from an infrastructure architecture perspective in that such registries can also hold the key semantic concepts that are essential to further evolution of the search service standard itself. As new communities of practice converge on interoperable searching using their key semantic concepts, these can be placed into a searchable semantic registry.
- **Scalability:** Optimization of any particular implementation should not be constrained by the service itself. A separate issue is the degree to which the standard search service may be operated in parallel against many resources. The standard search service should also be scalable in terms of supporting arbitrarily complex searches.
- **Extensibility:** The search service standard should be adaptable to virtually all search tasks. In addition to the usual sets of data structures and relations, other structure and relation semantics should be definable through profiles.
- **Internationalization:** The search service standard must have be in use worldwide in many languages. It should support negotiation between client and server as to each other's language capabilities for the session. Character set negotiation should also be supported, with support for Latin-1 as a minimum for U.S. Federal Government applications."

From: *Categorization of Government Information (CGI) Working Group, U.S. Federal Interagency Committee on Government Information, December 2004*
http://www.search.gov/interop/requirements.html

Additional Requirements include:

- Supports different levels of access control, such as restrictions by service, session, distributed resource, database, record, or data element
- Supports authentication of user identity through an ancillary service (e.g., e-Authentication)
- Supports verification of the integrity of delivered data, metadata, or other information
- Supports the search service standard for library catalogs accessible over network technologies (ISO 23950, identical to ANSI/NISO Z39.50)
- Supports the library standard for catalog records, Machine-Readable Cataloging
- Supports access to data without mandating proprietary technologies, nor proprietary vocabularies or thesauri
- Can be readily accommodated by leading search products, including Internet search engines
- Supports search of information that may be unstructured (often called "full-text"), semi-structured (typically represented with inline "markup"), or structured (sometimes known as "fielded")
- Supports search of HTML meta element contents and other varieties of metadata embedded within particular types of files (e.g., PDF, e-mail, etc)
- Supports customizable search of other varieties of structured metadata through common mechanisms such as SQL and LDAP
- Provides for interoperable search across locators for information and collections of information
- Interoperable with the international standard search service supporting the U.S. National Spatial Data Infrastructure Clearinghouse of geospatial data

- Implementable over the Internet using TCP/IP, HTTP/HTTPS, *data query structures such as* HTTP GET and HTTP POST, *and XQUERY*
- Precisely defined as to how searches are expressed and communicated between a client component and a server component, including a query language, a query syntax, and standardization of a result set schema
- Specified in an interface definition language such as Web Services Definition Language (WSDL)
- Supports searching of structured information using a nested Boolean query, e.g., (date > '20040101') AND ((subject = 'earthquake') OR (subject = 'temblor'))
- Supports the usual sets of data structures (word, phrase, date, URL.) and relations (equal, greater than, less than)
- Includes a query evaluation function to handle "abstract concepts" (e.g., name, category, date) according to what they mean semantically rather than merely how they may be labeled syntactically
- Supports abstract concepts that are produced by semantic mapping without requiring any particular semantic mapping technique
- Supports gateway to Internet Anonymous FTP Archive (IAFA) file system catalogs and Distributed Authoring and Versioning for the Web (WebDAV)
- Adopts readily to the underlying data model of named properties and property sets that is defined for objects addressable by software
- Already in production use for searching metadata variants such as Dublin Core Metadata Initiative, ISO 15836 Encoded Archival Description, and ISO 8879Standard Generalized Markup Language (SGML)
- Compatible with many and diverse approaches to compiling collections of information, without mandating any particular approach
- Supports interoperable search of business and services registries, modeled on ISO 11179 Metadata Registries, ebXML, or the Universal Description, Discovery, and Integration (UDDI) model
- Scalable in terms of supporting arbitrarily complex searches
- Scalable in not foreclosing concurrent searches on multiple servers
- Extensible to search tasks with unusual data structures and relations, definable through profiles or equivalent
- Provides extension mechanisms to nurture innovation in areas not yet ready for the broadest level of standardization
- Has been in use worldwide in many languages
- Supports negotiation between client and server as to each other's language capabilities for the session
- Supports character set negotiation, with Latin-1 as a minimum for U.S. Federal Government applications

From: ***Recommendations On The Categorization Of Government Information***
December 16, 2004, http://www.search.gov/interop/Recommendations-CGI-final.doc

15.3.1 Query

Support retrieval of records that satisfy specific query selection criteria

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Query	Single Word Keyword	Keyword with Boolean and proximity operators Fielded Search Phrase search String Search Fuzzy Match	Intelligent Content Services	Intelligent Content Services Keyword with Boolean and proximity operators Fielded Search Phrase Search Fuzzy Match
Query Language		SQL	XPointer XPath OQL XQuery	SQL XPointer XPath OQL XQuery

Definitions for Query include:

- Query: To interrogate a collection of data such as records in a database. The term may also be used to search a single file or collection of files such as HTML files on the Web. However, in addition to obtaining lists of records that match the search criteria, queries to a database allow for counting items and summing amounts. A query on the Web yields only a list of matching pages and is more often called a "search." [*Computer Desktop Encyclopedia. Computer Language Company Inc., 2006. Answers.com 19 Jul. 2006. http://www.answers.com/topic/query*]
- A statement of information needs, typically keywords combined with Boolean operators and other modifiers, in the field of information retrieval.
- A database query, which is a specification of a result to be calculated from a database. Queries are often specified using SQL. Other examples are the CODASYL database language, "network" databases, relational algebra, relational calculus, Datalog, SQL3, QUEL, XPointer, XPath and OQL.
- A request for information from a database. There are three general methods for posing queries:
  - Choosing parameters from a menu: In this method, the database system presents a list of parameters from which you can choose. This is perhaps the easiest way to pose a query because the menus guide you, but it is also the least flexible.
  - Query by example (QBE): In this method, the system presents a blank record and lets you specify the fields and values that define the query.
  - Query language: Many database systems require you to make requests for information in the form of a stylized query that must be written in a special *query language*. This is the most complex method because it forces you to learn a specialized language, but it is also the most powerful.
    [http://www.webopedia.com/TERM/Q/query.html]
- Intelligent Content Services (ICS): uses linguistic analysis, fact extraction, automated categorization, and taxonomies to help organize and make sense out of large volumes of disorganized, unstructured content.

- Intelligent Content Services (ICS): a semantically smart content-centric set of software services that enhance the relationship between information workers and computing systems by making sense of content, recognizing context, and understanding the end user's requests for information.

Other Important information for Query

- Resource: Brown, Matt and Laura Ramos "Searching For A Better Search," Forrester Research Inc., August 29, 2005, 7 pgs.

- Basic search focuses on word count and proximity instead of meaning, and errs on the side of recall over precision.

- ICS accomplishes four main goals:
  - Imposes structure on a body of content
  - Discovers the inherent structure with documents or text
  - Exposes meaning to aid discovery and enable more relevant retrieval
  - Monitors content as it arrives and alerts users to interesting matches.

15.3.2 Precision / Recall Ranking

Support selection and retrieval of records ranked to optimize precision against recall

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Precision/ Recall Ranking		Boolean operators within query to better define search criteria	Natural Language query using Linguistic Analysis Bayesian Probabilistic Model	Boolean operators within query to better define search criteria Natural Language query using Linguistic Analysis Bayesian Probabilistic Model

Criteria for Precision / Recall Ranking

- Relevancy should consider placement of words when applying weighting - in the document, filename or metadata. A word appearing for the first time at the end of a document should not receive the same weight as a word in the filename or metadata.
- Relevancy should consider file source, i.e. e-mail, file system, website
- Relevancy should be accurate across file formats, i.e. Text, image, audio, video. Non-text formats may be "mapped" or use metadata.
- All results of a query should be relevant
- All relevant results of a query should be returned so no important material is left out.

Definitions for Precision / Recall Ranking include:

- **Precision Ranking** - # relevant hits in hit-list / # hits in hit-list.
  Precision is a measure of how well the engine performs in not returning non-relevent documents.

- **Recall Ranking** - # of relevant hits / # of relevant items in the collection
  Recall is a measure of how well the engine performs in finding relevant documents. Recall is 100% when every relevant document is retrieved. Ranking is the art of sorting results in relevance to the submitted query.

- **Linguistic Analysis** – Break down text to determine accurate concepts and meanings. Allows search engine to overcome ambiguities in the user's search terms to increase search precision and recall rankings.
  - Parts of Speech Ambiguity – Words can be used with different meanings in a sentence. "How to set a table." "How to set the sets of a table."
  - Topical Ambiguity – A word can have multiple meanings. i.e. Report. Can be a document, sound of a rifle being fired, or as a verb "to report on the topic".
- **Bayesian Probabilistic Model** – The likelihood that something will happen can be plausibly estimated by how often it occurred in the past. An example is Google anticipating your search criteria "Did you mean..." by caching your previous searches and using that data to predict what you're looking for now to provide a more relevant results list.

- For any given retrieved set, ***Recall*** is the number of retrieved relevant items as a proportion of all Relevant items. Recall is, therefore, a measure of effectiveness in retrieving (or selecting) performance and can be viewed as a measure of effectiveness in including relevant items in the retrieved set.
- For any given retrieved set, ***Precision*** is the number of retrieved Relevant items as a proportion of the number of retrieved items. Precision is, therefore, a measure of purity in retrieval performance, a measure of effectiveness in excluding non-relevant items from the retrieved set

Buckland, Michael and Fredric Gey, "The relationship between Recall and Precision",
Journal of the American Society for Information Science, 45(1):12-19, 1994.
DOI: 10.1002/(SICI)1097-4571(199401)45:1<12::AID-ASI2>3.0.CO;2-LUS:
http://dx.doi.org/10.1002/(SICI)1097-4571(199401)45:1<12::AID-ASI2>3.0.CO;2-L

Other Important information for Precision / Recall Ranking

- Organizations should determine the acceptable parameters for Precision and Recall Rankings within their system to determine the system requirements. These parameters are measurable and testable using the formulas in the definitions

15.3.3 Classification

Support selection and retrieval of records organized by shared characteristics in content or context

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Classification schemes		Examples of "traditional" schemes include: o Dewey Decimal System (DDC) o Universal Decimal Classification (UDC) o Library of Congress Classification System (LCC) Subject-specific schemes such as: o NLM (National Library of Medicine) o Engineering Information (Ei) Classification Codes o Mathematics	Semantic web	Dewey Decimal System (DDC) Universal Decimal Classification (UDC) Library of Congress Classification System (LCC) Subject-specific schemes such as: o NLM (National Library of Medicine) o Engineering Information (Ei)

		Subject Classification o ACM Computing Classification System (CCS)		Classification Codes o Mathematics Subject Classification o ACM Computing Classification System (CCS)
Faceted Classification		Subject Domain specific standards		Subject Domain specific standards

Definitions for Classification include:

- Data can be classified from a variety of perspectives. Three primary perspectives are:
  - Information Lifecycle Management (Refer to the *Back Office Services Domain, Data Management* section for further information.)
  - Security (Refer to the *Back Office Services Domain, Data Management* section for further information.)
  - Subject / Topic
- Classification systems provide ways to separate concepts into relatively broad topics. Many classification systems also include specific terminology to be used (for example, object names). Each item has a "correct" (or, at least, agreed upon) place somewhere in a single, large, hierarchically organized classification system.
- **Faceted Classification:** "A faceted classification differs from a traditional one in that it does not assign fixed slots to subjects in sequence, but uses clearly defined, mutually exclusive, and collectively exhaustive aspects, properties, or characteristics of a class or specific subject. Such aspects, properties, or characteristics are called facets of a class or subject […]".
- "[I]ndividual facets can be accessed and retrieved either alone or in any desired combination. This feature is especially important for computerized retrieval, which has been successfully applied to faceted classification, and in online retrieval as a complement to verbal retrieval by subject headings or keywords."
- "Since the 1960s, all major classification schemes (with the exception of LCC) either have been partially restructured on a faceted basis or display a fully faceted structure."

Wynar, Bohdan S. *Introduction to cataloging and classification*. 8th edition. 1999. pgs. 320, 322

- The facet classification is an analytic-synthetic scheme. It is analytic because it subdivides broader elements into single concepts that are clearly defined through facet analysis. It is synthetic in that new elements can be developed.

Other Important information for Classification

- In order for the categorization of Government information to add value for the information user, it should meet several general major requirements:
  - Enhance public access to Government information resources.
  - Render a predictable level of granularity among the search returns from decentralized data sources.
  - Be a realistic mandate for Government entities, many of which operate with less than optimal levels of funding or IT support, to carry out.
  - Be compatible with existing information characterization and retrieval mechanisms.
  - Be flexible enough to allow for technological advances in information management, publishing, or discovery and retrieval.

From: ***Recommendations On The Categorization Of Government Information***

December 16, 2004, http://www.search.gov/interop/Recommendations-CGI-final.doc

- A faceted classification system has the following important characteristics:
  - A faceted system focuses on the important, essential or persistent characteristics of content objects, helping it to be useful for fine-grained rapidly changing repositories.
  - You don't have to know the name of the category (or categories) into which a document is placed. In a business world in which terminology changes faster than you can blink, this is a big asset.
  - The absence of polyhierarchy is implied, at least, by having mutually orthogonal facet hierarchies. The ordering principle in a facet is not necessarily hierarchical (general-specific, whole-part, etc.), although that will be true in most cases. It might even be alphabetical.
  - It's easy to add a new facet at any time.
  - Flexibility in general. Makes few assumptions about the scope and organization of the domain. So it's hard to "break" a faceted classification schema.
  - Should be easier to construct the facet hierarchies.
  - Combining elements from separate facets using a defined syntax -- for example, to express the functions of a product, make assertions, or frame questions in a structured way -- is an extremely powerful method of precise retrieval.
  - Adding persistent (even "typed") relationships between elements in different facets -- for example, John Smith (in the Person facet) <relationship: *is an employee of*> Generic Company (Organization facet) provides substantial useful representation of knowledge in the faceted classification schema itself.

The Knowledge Management Connection
http://www.kmconnection.com/DOC100100.htm

15.3.4 Pattern Matching

Support retrieval of records generated from a data source by imputing characteristics based on patterns in the content or context

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Sequence Patterns		Exact string match Substitutions Wildcards		Exact string match Substitutions Wildcards
Tree Patterns		Structured Programming Neural Networks		Structured Programming Neural Networks

Definitions for Pattern Matching include:

- **Pattern matching** is the act of checking for the presence of the constituents of a given pattern. In contrast to pattern recognition, the pattern is rigidly specified. Such a pattern concerns conventionally either sequences or tree structures. Pattern matching is used to check that things have the desired structure, to find relevant structure, to retrieve the aligning parts, and to substitute the matching part with something else. Pattern matching includes testing whether an input sentence or structure matches a pattern that may contain variables, such as the use of wildcards * and ? as part of a filename specification, to select multiple files with similar names.. Pattern matching is followed by substitution into an output pattern to transform expressions. Substitution is the replacement of variables in an expression or pattern by values of the variables.

Other Important information for Pattern Matching

- **Sequence (or specifically text string) patterns** are often described using regular expressions and matched using respective algorithms. Sequences can also be seen as trees branching for

each element into the respective element and the rest of the sequence, or as trees that immediately branch into all elements.

- **Tree patterns** can be used in programming languages as a general tool to process data based on its structure. Some functional programming languages such as Haskell, ML and the symbolic mathematics language Mathematica have a special syntax for expressing tree patterns and a language construct for conditional execution and value retrieval based on it. Because of simplicity and efficiency reasons these tree patterns lack some features that are available in regular expressions. Depending on the languages, pattern matching can be used for function arguments, in case expressions, whenever new variables are bound, or in very limited situations such as only for sequences in assignment in Python. Often it is possible to give alternative patterns that are tried one by one. Pattern matching can benefit from guards.
- **Pattern recognition** is a field within the area of machine learning. Alternatively, it can be defined as "the act of taking in raw data and taking an action based on the category of the data" (CITATION). As such, it is a collection of methods for supervised learning. Pattern recognition aims to classify data (patterns) based on either *a priori* knowledge or on statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations, defining points in an appropriate multidimensional space. A complete pattern recognition system consists of a sensor that gathers the observations to be classified or described; a feature extraction mechanism that computes numeric or symbolic information from the observations; and a classification or description scheme that does the actual job of classifying or describing observations, relying on the extracted features.

15.4. Communication

Capabilities within this Service Type transmit data, messages and information in multiple formats and protocols.

15.4.1 Real Time Chat / Instant Messaging

Support the conferencing capability between two or more users on a local area network or the internet

15.4.2 Audio Conferencing

Support audio communications sessions among people who are geographically dispersed

15.4.3 Video Conferencing

Support video communications sessions among people who are geographically dispersed

15.4.4 Event / News Management

Monitor servers, workstations and network devices for routine and non-routine events

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Management protocols	CMIP Novell SNMP 1.x	SNMP V1/V2	SNMP V3	SNMP V2
Network Monitoring	Home grown & proprietary products	HP OpenView What's Up Gold NetSaint, Orion, Nagios, Servers Alive	IFIX	
Network Capacity	RMON Probes	Lanalyzer eHealth Solar Winds		
Element Managers		CiscoWorks Spectrum Optivity		
Tools		NG Sniffer EtherPeak TCPDump App. Vantage	Alterpoint Opsware	

Definitions for Event / News Management include:

- SNMP – The Simple Network Management Protocol facilitates the exchange of management information between network devices and management systems. It is part of the Transmission Control Protocol/Internet Protocol (TCP/IP) protocol suite. SNMP enables network administrators to manage network performance, find and solve network problems, and plan for network growth.
- Element Managers - Element managers are designed to provide management services for a specific set of products as opposed to all SNMP devices in a network. This is typically vendor specific.

15.4.5 Community Management

Support the administration of online groups that share common interests

15.4.6 Computer / Telephony Integration
Support the connectivity between server hardware, software and telecommunications equipment into a single logical system

15.4.7 Voice Communication
Provide telephony or other voice communications

15.5 Systems Management

Capabilities within this Service Type support the administration and upkeep of an organization's technology assets, including the hardware, software, infrastructure, licenses, and components that comprise those assets.

15.5.1 License Management

Support the purchase, upgrade and tracking of legal usage contracts for system software and applications

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
License Tracking		Software/Hardware inventories Manual notifications for renewals	ITIL processes for process compliance	

Description	Example Products	Notes
License Tracking	JobTrack CA Unicenter TNG	

Other Important information for License Management

- Information in this category applies to hardware, software, and content databases.
- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

15.5.2 Remote Systems Control

Support the monitoring, administration and usage of applications and enterprise systems from locations outside of the immediate system environment

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Applications Usage		RDP		
Monitoring				
Administration		Change tracking/ auditing SMS ManageWise CA Unicenter TNG		

15.5.3 System Recourse Monitoring

Support the balance and allocation of memory, usage, disk space and performance on computers and their applications

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Real-time Monitoring		Resource usage Alerts Automated response engines		
System Availability Monitoring		Short term capacity adjustments Tivoli NetView	Capacity sharing Load Balancing	
Performance/ Capacity Planning (long term monitoring)		Trends based on historical use. Tuning efforts Forecasting		

15.5.4 Software Distribution

Support the propagation, installation and upgrade of written computer programs, applications and components

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Desktop				
Server				
Software and OS deployment		Remote installation and maintenance Monitoring and reporting Base lining for fallback ability Novell Zenworks MS SMS/MOM	Unattended, timed and group installations	Elevated credential distribution rollback and version control

15.5.5 Issue Tracking

Receive and track user-reported issues and problems in using IT systems, including help desk calls

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Software Tools for IT Service Management		Unique tools selected to meet needs of the service being offered. (Help Desk, support of a single IT function or application)	Multi-functional service desk tool to support the business as well as IT with reporting and analytical capabilities ITIL process compliant Open Data Standards for information integration / sharing Workflow engines which allow for issue routing.	Integrated tools – platform independent

Definitions for Issue Tracking include:

- **Issue Tracking** –
- **Open Data Standards** – Refer to Data Interoperability Section of the KITA.
- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

15.6 Forms Management

Capabilities within this Service Type support the creation, modification, and usage of physical or electronic documents used to capture information within the business cycle.

15.6.1 Forms Creation

Support the design and generation of electronic or physical forms and templates for use within the business cycle by an organization and its stakeholders

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Form Format		Print and Fill Fill and Print Workflow Enabled (Ad hoc)		
File Format		PDF fill-able Forms Spreadsheet Forms Word Processor Forms		
Access		Access Database on Demand XML Ability to update databases Ability to integrate with databases Ability to integrate with workflow		
Security		Electronic Signature Support		
Web		Postscript Support Macro / POST support	Web Service Support	

Description	Example Products	Notes
Forms Creation	InfoPath MS Office Suite OpenOffice suite	

15.6.2 Forms Modification

Support the maintenance of electronic or physical forms, templates and their respective elements and fields

Architecture Component	Twilight Standard	Current Standard	Emerging Standard	Target
Support		Maintainable Versioning	Processes conform to ITIL Change Management	
Form Format		Print and Fill Fill and Print Workflow Enabled (Ad hoc)		
File Format		PDF Fill-able Forms Spreadsheet Forms Word Processor Forms		
Access		Access Database on Demand XML Ability to update databases Ability to integrate with databases Ability to integrate with workflow		
Security		Electronic Signature Support		
Web		Postscript Support Macro / POST support	Web Service Support	

Definitions for Forms Modification include:

- **Information Technology Infrastructure Library (ITIL) -** A framework of best practice approaches intended to facilitate the delivery of high quality information technology (IT) services. ITIL outlines an extensive set of management procedures that are intended to support businesses in achieving both quality and value for money in IT operations. These procedures are supplier independent and have been developed to provide guidance across the breadth of IT infrastructure, development, and operations. (http://en.wikipedia.org/wiki/ITIL accessed 2006.10.17)

Appendices

Appendix I: Technical Architecture Review Board

Structure of the Technical Architecture Review Board

The Technical Architecture Review Board has committees responsible for the sub architectures summarized in the current release of the Kansas Information Technical Architecture. Each sub-committee will have at least three members and it reports to an Executive Council of the branch CITOs and the Chief Information Technology Architect (CITA). The CITA calls the Council, arranges the agenda, and chairs the Architecture Review Board.

Subcommittee Chairs

Subcommittee Chairs are the individuals that belong to the KTARB and also are responsible for specific areas of the architecture document. Listed below are the members of the KTARB and their specialty area regarding the KITA.

KTARB Specialty area	Name	Organization
Platform	Allan Haverkamp	KDOT
Network	Steve Selaya	KUMC
System Management	Jay Coverdale	DISC
Disaster Recovery & Customer Support	Jerry Merryman	DISC
Application	Bryan Dreiling	DISC
Information Management	Scott Leonard	KSHS
Security	Larry Kettlewell	DISC

These seven individuals led a total of fourteen subcommittees to complete the subcommittee drafts of the KITA. Listed below are the fourteen subcommittees and their corresponding subcommittee chairs.

Subcommittee Name	**Contact**
Systems Management	Jay Coverdale
Data Center Management	Jay Coverdale
Data Management	Scott Leonard
Data Interoperability	Scott Leonard
Digital Assets Management	Scott Leonard
Architecture Security	Larry Kettlewell
Platform	Alan Haverkamp
Console Event Management	Jerry Merriman
Help Desk and Problem Management	Jerry Merriman
Business Continuity	Jerry Merriman
Network	Steve Selaya
Development Environments	Bryan Dreiling
Solutions to Support Systems	Bryan Dreiling
Interface and Integration	Bryan Dreiling

The following members staffed the fourteen architecture subcommittees:

Systems Management	
Jay Coverdale - Chair	DISC
Darin Bodenhamer	SRS
Kathie Harris	SRS
Jeff Lampe	KDOT
Doug Quinn	DofA
Jim Schick	JJA
Brain Schmelzle	KHP
Beth Warner	KU
Randy Robert	Shawnee County
Data Center Management	
Jay Coverdale Chair	*DofA*
Bill Kritikos	KDOT
Norma Jean Schaefer	KDHE
Tom Stevens	KDHE
Jim Brook	DofA
Marsha Rogers	DofA
John Lowrey	DofA
Richard Wadleigh	KDOR
Ron Dugger	KDOL
David Best	SRS
Bill Pesek	KU
Gloria Robertson	KSU
Data Management	
Scott Leonard - Chair	*KSHS*
Cathy Nicoletti	DofA
Janet Litfin	KDOR
Carol Baldry	KDOT
Jim Fox / Frank Golos	SRS
Beth Zlotky	DofA
Tanya Keys	JJA
Stan Pfeifer	KBI
Beth Warner	KU
Thomas "TJ" Rains	ESU
Terri Clark	LAS
Data Interoperability	
Scott Leonard - Chair	*KSHS*
Raf Lorie	KDOR
Rick Miller	KDOT
Kathie Harris	SRS
Beth Zlotky	DofA
Shawn Brown	KBI
Beth Warner	KU
Terri Clark	LCS
Digital Assets Management	
Scott Leonard - Chair	*KSHS*
Brad Sedlacek	KDOT
Matt Veatch	KSHS
Stan Pfeifer	KBI
Beth Warner	KU
Lynn Carlin	KSU
Terri Clark	LCS
Doug Cruce	Judicial
Architecture Security	
Larry Kettlewell - Chair	*DofA*
Norm Bacon	KDOC
Stan Wiechert	KDOR
Ben Woodward	KDOT
Shawn Finley	KDOL
Jake Wall	KDOT
Shayne Casey	JJA
Robert Louis	KBI
Beth Warner	KU
Sherry Callahan	KUMC
Royce Gilbert	KSU
Allan Foster	LPA
Platform	
Allan Haverkamp - Chair	*KDOT*
Jay Coverdale	DofA
Glen Yancey	KDOR
Ron Holthaus	KDOT
Hank Sipple	DofA
Anthony Fadale	DofA
Kevin Cronister	DofA
Shayne Casey	JJA
Derek Welsh	KBI
Harvard Townsend	KSU
Console Event Management	
Jerry Merryman - Chair	*DofA*
Richard Wadleigh	KDOR
Linda Schramek	DofA
Helen Gibbon	DofA
Beth Warner	KU
Help Desk and Problem Management	
Jerry Merryman - Chair	*DofA*
Kevin Cronister	DofA
Danny Kuckelman	KDOL
Richard Wadleigh	KDOR
Leslie Fisher	SRS
Rita Hodges	KDOT
Kelli Bulter	ESU
Beth Warner	KU
Business Continuity	
Jerry Merryman - Chair	*DofA*
Cheryl Shaughnessy	DofA
Vickie Rogers	DofA
Stan Wiechert	KDOR
Bill Kritikos	KDOT

Ben Woodward	KDOT
Paul Houser	KDOL
Royce Gilbert	KSU
Beth Warner	KU
Network	
Steve Salaya - Chair	*KUMC*
Sean Buffum	KDOR
Alan Chapas	KDOT
Ron Dugger	KDOL
Alex Hoover	KDOL
Dave Timpany	DofA
Jim Schick	JJA
Joe Wallingford	KHP
Robert Louis	KBI
Mike Erickson	WSU
Bill Klein	KU
Development Environments	
Bryan Dreiling - Chair	*DISC*
Raf Lorie	KDOR
Cathy Jones	DofA
Dyann Berglund	KDOT
Cecilia Burke	SRS
Jameson Watkins	KUMC
Beth Warner	KU
Solutions to Support Systems	
Bryan Dreiling - Chair	*DISC*
Julie Tooley	KDOT
Jeremy Anschutz	KDHE
Shaun Miller	KDOE
Beth Warner	KU
Interface and Integration	
Bryan Dreiling - Chair	*DISC*
Libby Peters	SRS
Jeanette Branam	KDOT
George Hubka	KDOL
Shawn Brown	KBI
Beth Warner	KU

Appendix II: KITA Change Control

Change Control

Draft versions 1.x included work in progress during the fourth quarter, calendar year 1998.

Draft Version 2.0 released on December 18, 1998 for limited review and format approval by ITAB and ITEC.

Draft Version 2.1 released in January 1999 included the working version of platform Architecture and completed versions of the Network Architecture.

Draft Version 2.18 released in February 1999 included the completed version of platform Architecture.

Draft Version 3.10 included synthetic versions of sub-architectures in chapters 6-20. These preliminary chapters were gathered from empirical data and interviews with State Agencies, and should be used for reference only (see appendix V for the discussion on the synthetic chapters).

Draft Version 3.11 was a minor revision of Version 3.10. New sections related to the structure of the Kansas Technical Architecture Review Board and its membership.

Draft Version 3.12 was a minor revision of Version 3.11. New sections included Appendix IX, on the Kansas Technical Architecture Review Board, and a revised Chapter 11, Middleware.

Draft Version 4.0 was a major revision of the document structure bringing it into conformance with the architectural groupings of the Kansas Technical Architecture Review Board.

KSTA Version 5.0 was released October 1, 1999, and included several sub-architectures moved from 'synthetic' to 'organic' through KTARB committee activity. Chapters moved to organic (or State directed content) included: Applications Development and Management, Data Management, Platform, and Internet architectures. The Network section had been previously moved to organic status. The remaining sub-architectures continue in their synthetic or as-is status until the release of future versions.

KSTA Version 6.0 was released in January 2000, included updates to all Chapters in Section 1, Chapter 4 (Network Architecture; Voice Network) Chapter 7 (Platform Architecture) and Chapter 18 (Data Management). New, or ‘organic’ content, for Chapters 9 (Asset Management), 10 (Console Event Management), 11 (Help Desk and Problem Management), 15 (Application Structure), and portions of 19 (Information Management) is also included in release Version 6.0. Roughly 60% of the content of the KITA has been moved from ‘synthetic’ to ‘organic’ as of the release of Version 6.0.

KSTA Version 7.0, released in April, 2000, included updates and new "organic" content for the following chapters: Chapter 7 (Platform Architecture); Chapter 13, Change & Configuration Management; Chapter 19, Information Management; and Chapter 20, Security. Chapters remaining to be converted from "synthetic" to "organic" content include: Chapters 8, 12, 16 and 17. Approximately 75% of the content of the KITA has been converted from synthetic to organic as of release of Version 7.0.

KSTA Version 8.0, released in July, 2000, brought the document up to full "organic" content and included new material for the following chapters: 8 Storage Architecture, 12 Change and Configuration Management, 16 Application Infrastructure Services, 17 Workgroup Services, and a new chapter 20, Electronic Records Management. Chapter 20 from Version 7.0, Security Architecture became a new Chapter 21 in V.8.0. Additionally, updated material was included for Chapters 7 Platform Architecture, 9 Asset Management, 14 Application Development and Management, and 15 Application Structure.

KSTA Version 9.0 was released in October 2001, included updates to a majority of the chapters. The “Sub-Assembly” data Appendices in chapters 14 through 17 were relocated to the Appendices at the back

of the document. Document structural changes included the designation of "PART" and associating "CHAPTER" with the applicable chapter numbers. Formatting within and between chapters was enhanced and improved.

KSTA Version 9.5 released in October 2002, included updates to some of the tables within the chapters and minor amounts of verbiage changes.

With Version 10.0, released in April 2003, the document's title was changed to be the Kansas Information Technology Architecture. Related naming conventions (references to the Kansas Information Technology Architecture Review Board, for example) were also updated. A major addition was included within chapter 4 (Network Architecture) to treat emerging wireless activities. Significant modifications were also made within chapter 21 (Security Architecture).

With Version 11.0, released in September 2006, Kansas has restructured the KITA to align with the Federal Enterprise Architecture and the NASCIO community. This restructuring will allow us to be able to share architecture standards, directions project statements and targets will any other state or federal agencies using an agreed upon reference structure that supports the business model and services delivered as well as supporting technical architecture standards.

Appendix IV : Technical Architecture Policies

Information Technology Policy #4000 Revision #0

Information Technology Policy #4010 Revision #0

Information Technology Policy #4020 Revision #0

Appendix V :Product Lifecycles

The definitions and directions included in this architecture will include product and technology standards. It is important to remember that the computer hardware lifecycle is currently 18-24 months. Hardware, software, products, applications and languages tend to follow a predictable lifecycle pattern. This pattern involves three general phases that can be described graphically as shown in Figure 8.

The x-axis represents time, with the left-most part of the graph representing new product introduction. The y-axis represents market acceptance potential for the specific technology element.

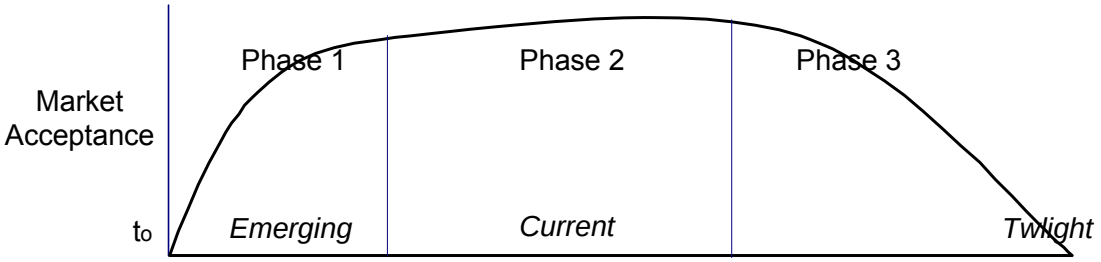


Figure 8: Technology Lifecycle

Phase 1 Emerging. This is the beginning of the product lifecycle, when the technology is introduced. General market introduction occurs here. Only early adopters will implement these technologies into production environments. For Kansas agencies, procurement of systems and technology from this sector should be approved only under controlled or guarded situations. New technology risks and the lack of operational support usually outweigh any technical features that may exist. Competition will be little or none, and prices will be high. Staff skills to implement and support technologies in this area will be scarce and expensive.

Products, standards and technologies are listed in this category to identify future candidates for standard acceptance and support. As new items progress through the emerging stage of market acceptance, Kansas organizations can request these to be placed into the appropriate emerging sections of the KITA so that future trends and movements can be tracked and considered. Eventually, some elements of the emerging categories will be promoted to the current technical standard. Others in this category will simply drop-off from consideration.

Phase 2 Current. With further product integration and development, standards, technology and products achieve market acceptance, both inside and outside the Kansas Information Technology Architecture. Improved market acceptance and need to utilize proven components in state systems will allow for current standards to be defined and outlined.

Agencies wishing to select proven, stable and supported technologies will look to items in this category for near-term implementation. Prices for current standards will have stabilized (and eventually will erode), with contracted/negotiated purchases now made available through standard acquisition processes. Trained and available resources will grow over time. Kansas architects will strive to minimize the varieties and options so as to provide commonality and consistency in technology deployment.

Phase 3 Twilight. This phase is also referred to as technology retirement. In the natural course of product lifecycles, all elements will eventually be in this category. Candidates for this category come from current standards. Architecture teams in Kansas will be sensitive to the early retirement or pre-mature retirement of technology components.

Elements listed in this category throughout the KITA are discouraged from future planning, procurement or upgrades. This is not to imply that retired products need to be replaced. New IT systems being planned need to review technology elements that are classified as twilight and avoid any new purchases or inclusion of these options.

Sometimes, business situations may require the extension of services listed as twilight. These situations can be approved on an exception basis. Support costs and maintenance for these elements will increase over time. Vendors may offer fire-sales discounts for technology in the twilight category. Acquisitions from this category need to be avoided when possible.