State of Colorado

Data Strategy

2010

OIT

Office of Information Technology

Table of Contents

Executive Summary 3
Introduction 8
Purpose 8
Authority 8
Document Overview 9
Background and Overview 10
The Importance of Data Management 10
Historical Perspective 11
Key Business Drivers 13
Data Management Program 14
Data Management Principles 15
Organization 16
Strategic Vision 22
Program Goals and Strategic Objectives 23
Enterprise Data Management Framework 24
Data Governance 27
Data Governance Maturity Models 29
Enterprise Data Tools 31
Design and Architecture 37
Concept of Operations 38
Security 44
Privacy 45
Data Set Consolidation 45
Implementation Roadmap and Milestones 47
Colorado's Vertical Domain Approach to Integrated Data Sharing 48
Establish: Create Data Governance and Management Program 50
Capture and Design 52
Integrate: Cross-Agency Data Sharing and Interoperability 53
Provide and Perform: Data as Utility and Knowledge Management 55
Timeline 57
Performance Metrics 57
Appendices 58
Appendix A – Glossary of Terms and Acronyms 58
Appendix B – HB 09-1285 65

Executive Summary

Since 2007, the Governor's Office of Information Technology (OIT) has produced a ground-breaking and progressive agenda for information sharing and information technology management in the State of Colorado. It is recognized by both the Governor and the Colorado General Assembly that in order to more effectively serve citizens, improve the efficiency and effectiveness of state government, and to inform policy making, a strong program of information sharing is required across all lines of business the state serves, as illustrated below:

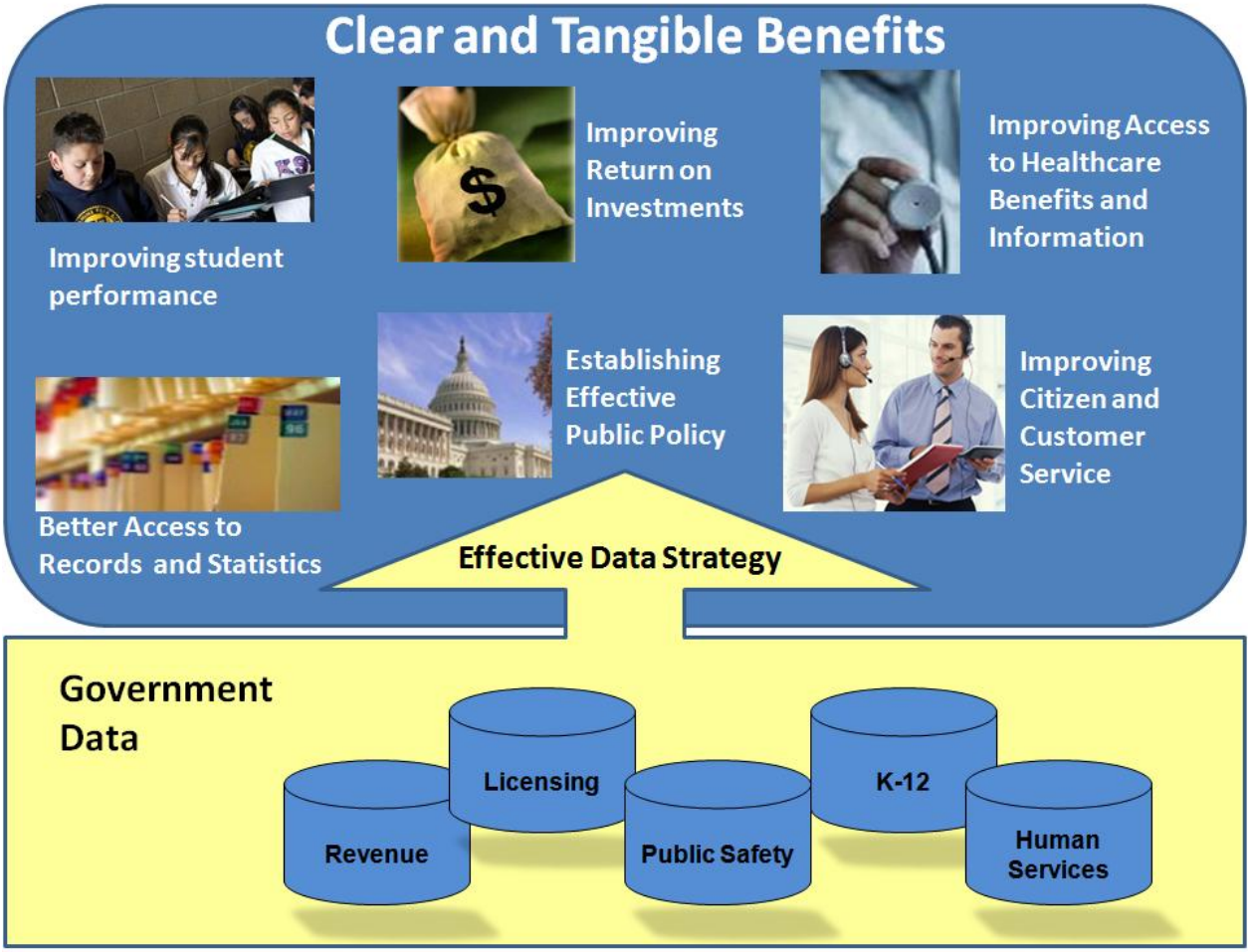


Figure 1.1: Realizing Benefits from an Effective Data Strategy

OIT's focus on agile application and service delivery means that strong concentration and emphasis on the underlying data required for employees, agencies, legislators, and others to do their work *must* be a top priority. The "businessization" of government requires a disciplined approach to dismantling data silos, implementing infrastructure to enable sharing across agencies, branches, and levels of government, and tools to build capacity for knowledge and performance management.

Specifically, a progressive data management strategy can provide the following benefits across the enterprise:

- *Education* - Ensures that a seamless education system from pre-school to graduate school is preparing our young people for the demands of the 21st Century by linking records over time (PreK-20), analyzing performance, and studying educational effectiveness.
- *Social Services* –Creates means to capture data once – regardless of point of entry into the state system - about a child, youth or family, and use that data across multiple state service programs to directly certify them for supplemental or additional services based on child or family eligibility. This program has a twofold benefit: ensuring that all children and families receive the benefits that they may not otherwise have applied for, and reducing fraudulent claims against the system by comparing the records.
- *eGovernment Services* – Provides single-sign on for citizens and businesses to have access to all of their current state account information (driver information, vehicle registration, tax information, benefits, etc.) through one portal. Digital signature services would be available to complete transactions electronically, end-to-end, with the state. The state would also be able to provide services such as address change that get completed once and shared with all agencies with which the individual does business.
- *Workforce and Economic Development* – Creates strategic, targeted and systemic responses to economic conditions and labor market changes. Information sharing can help support the development of timely, accurate information to identify key industries, examine the state of regional economies, explore the root causes of skills gaps, and promote strategic planning that addresses the needs of workers and employers alike.
- *Law Enforcement* – Improves state and community security and safety postures. All major reviews of the nation's response to the terrorist attacks of 9/11 maintain that integrated information technology and improved information sharing across agencies at all levels of government are vital to an effective homeland security strategy.
- *Policy Making* - Helps lawmakers and policy makers answer questions and predict program results to help ensure the best use of limited State resources and effectiveness of State programs.

OIT is the central authority for all information technology (IT) systems, resources, and budget in the state of Colorado, and is authorized to set standards, policies, and guidelines for how those IT assets operate, communicate and are managed. OIT has been examining the issue of data management for two years, bringing together cross-functional, multi-agency teams, interviewing stakeholders, benchmarking other states, and doing other critical research to determine the most effective way to establish a new data management program in an environment of historically siloed agencies and systems.

Two pieces of legislation - HB 08-1364 (Interdepartmental Data Protocol) and HB 09-1285 (Government Data Advisory Board – GDAB) - have been passed to address these issues. A formal Data Management program was officially established in July 2009 in OIT's Office of Enterprise Architecture, and is headed by the nation's first State Chief Data Officer. The goals of the Data Management program are to:

- Establish a sustainable data management and governance program.
- Identify and inventory all state data sets, and implement enterprise data standards.
- Create an integrated data sharing environment to provide a single, accurate, consistent source of data for all stakeholders.
- Optimize data and information assets to reduce costs.
- Ensure privacy, security and compliance requirements.

The Government Data Advisory Board (GDAB), seated in August 2009, is a multi-agency central governance authority, comprised of representatives of 12 state agencies, local governments, non-governmental organizations and research institutes, and a wide variety of education stakeholders. The GDAB's mission is to provide guidance and recommendations on how the state should govern and manage data and data management systems to improve the efficiency and effectiveness of state government, citizen service delivery and policy-making. The GDAB is one of the very few such Boards in any state in the country, established in legislation and appointed by the Governor, to provide the central governing structure for enterprise data sharing initiatives. Additionally, a Data Steward Action Council (DSAC) is being established to formalize and organize the stewardship activities and processes enterprise-wide based on information subject areas. The DSAC will also create a common baseline of information: a statewide foundation for data sharing, information discovery, and future architectures.

Moving the state to an enhanced data management future based on shared integrated data will improve the speed and quality of decision making and delivery of services to the state‘s constituents while reducing the duplicative cost associated with non-integrated systems. Figure 1.2 below illustrates the high-level business and technical overviews of the state’s data management system components across four stages of implementation (Establish, Capture and Design, Integrate, Provide and Perform).

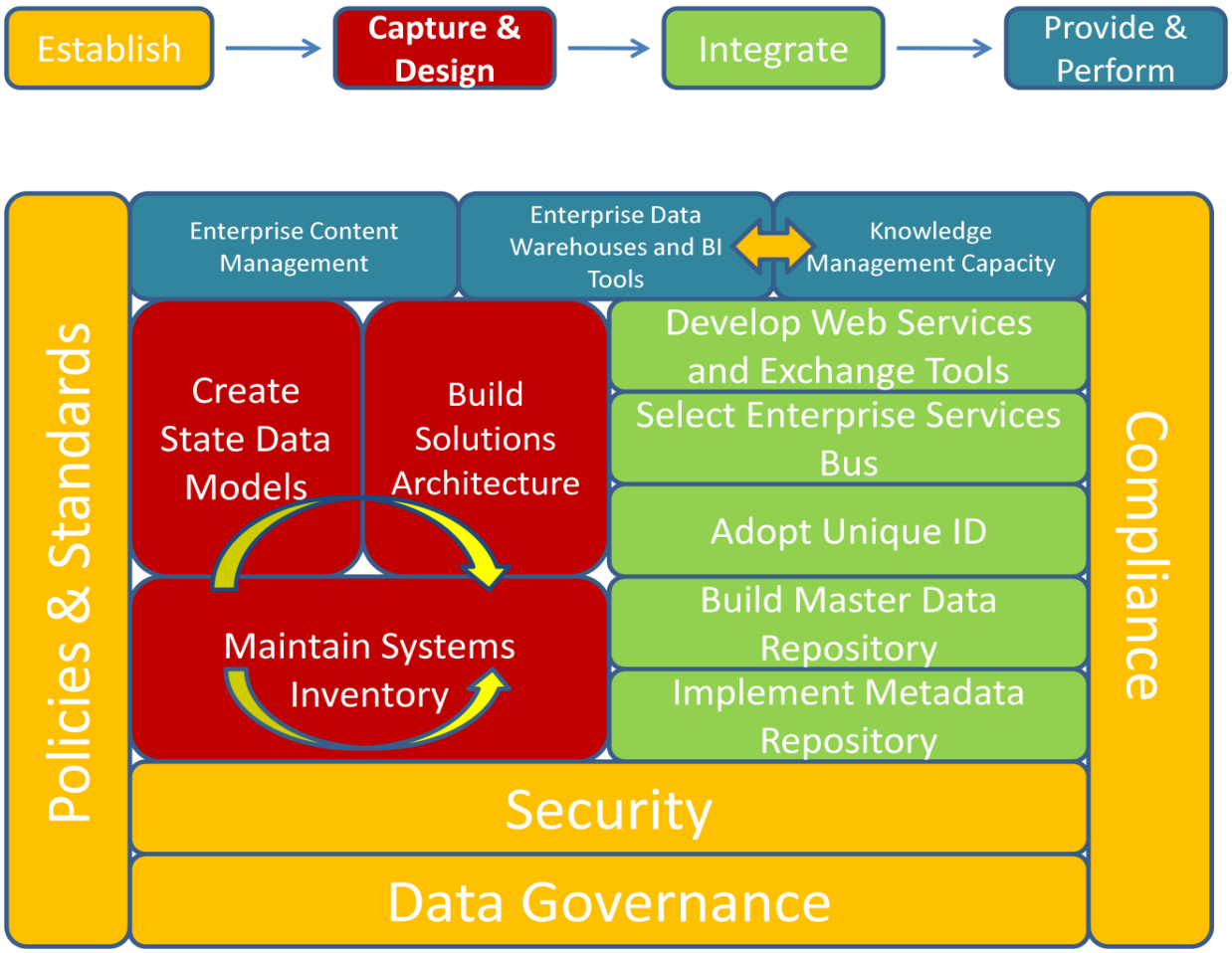


Figure 1.2: DM Program Strategic Articulation

This figure outlines how the state will implement and realize the longer term benefits from this strategy. The ambition of the state's vision for this program is commensurate with the dramatic scale of improvement in outcomes being pursued. This program includes initiatives related to understanding and modeling state data, integrating data across state systems, and providing the right data to the right audience at the right time for use in continuous improvement via knowledge management.

There are a number of data sharing efforts and data management projects underway in the state, and OIT is working with many initiatives and agencies already to begin establishing processes and standards. As an example, Figure 1.3 below is the state's conceptual model for the Statewide Data Longitudinal System (SLDS) proposal for the U.S. Department of Education, and showcases the architectural foundation for improving student outcomes and performance. OIT is involved in this multi-agency effort to help set architectural standards, data standards, and governance policies and procedures.

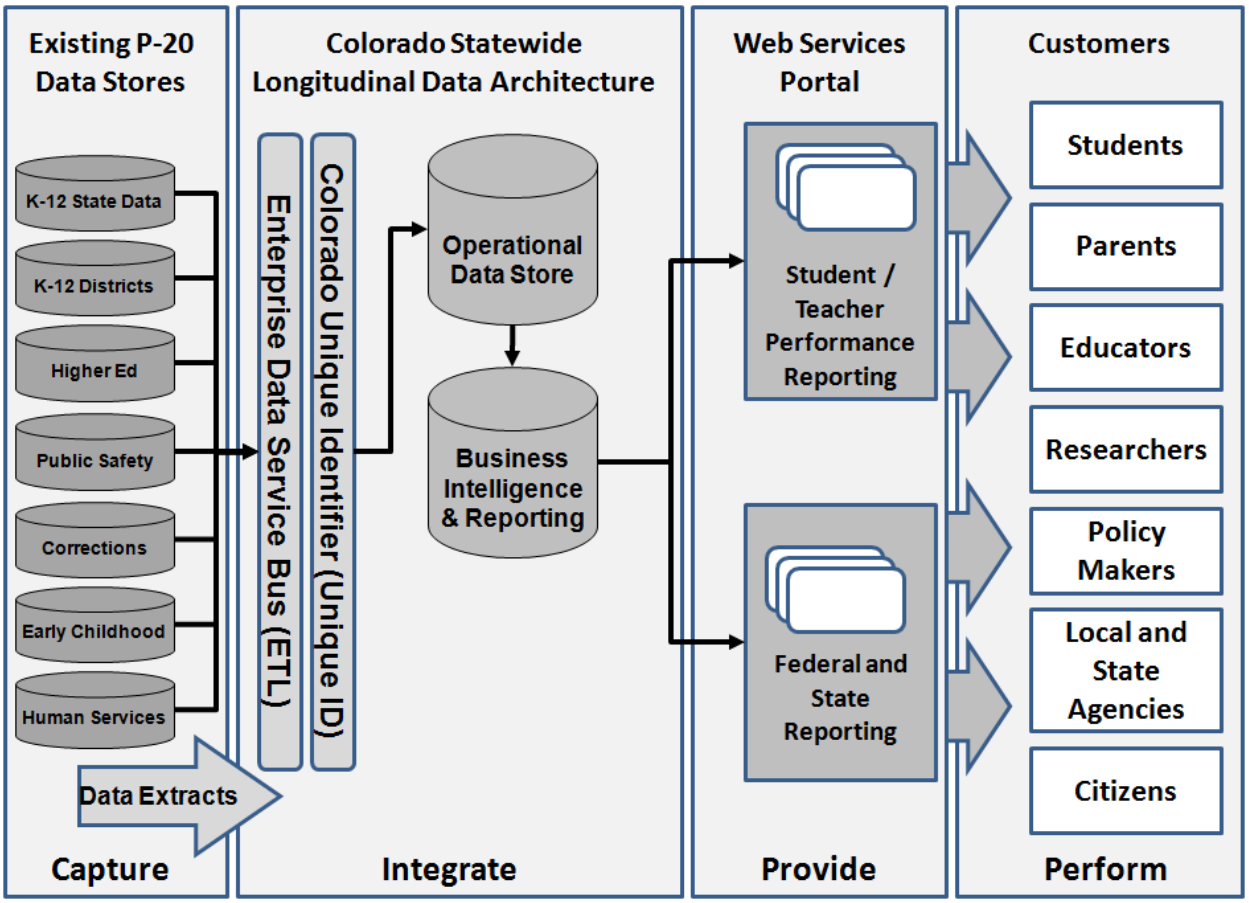


Figure 1.3: State Education Longitudinal Data System Architecture

OIT has successfully developed a new enterprise approach to delivering technology, leveraging private sector business models and public-private partnerships to modernize the infrastructure and thus enable the agile delivery of new applications and government services to state agencies and citizens as efficiently, cost-effectively, and sustainably as possible. OIT is pleased to present this data strategy as one in a number of steps towards breaking down government silos, improving citizen service delivery and strengthening the performance of state government.

Introduction

This is the state's first enterprise data strategy and it is an exciting time for data management in the State of Colorado. There is much work to be done, and it must be done strategically, surgically, with a focus on providing long term and scalable solutions for our citizens. There is a danger in trying to do too much with limited resources, and conversely in not getting enough accomplished to show value, commitment and progress. Necessarily, efforts will be undertaken in a modular approach, and prioritized appropriately. Certainly, the state must deal responsibly with its existing, sometimes aging, infrastructure, and agencies don't have the human capital to move as quickly as perhaps some would wish.

Purpose

This document outlines the state's data management and governance program. It describes the key business drivers behind the program, the status of the as-is environment of data today, the organizational alignment of the program, and the migration plan for moving forward with an enterprise approach. The target audiences for this document include Agency directors and program people, state information technology professionals, legislators, and others interested in how the state manages its data. Detailed data management processes, policies, and technical standards are *outside* the scope of this document, but instead will be addressed through the Office of the Enterprise Architect.

It is intended that this document be reviewed annually by the State Chief Information Officer (CIO), State Enterprise Architect, Chief Data Officer, and the Government Data Advisory Board (GDAB). It should be updated at least every other year to accommodate political, fiscal, regulatory, and technology changes that may impact the state's direction or implementation strategy.

Authority

The Governor, OIT, and the Colorado General Assembly determined that operational efficiencies and improved business performance could be achieved through the establishment of enterprise authority and governance over all information technology activities. Legislation was introduced in the 2008 legislative session to achieve these opportunities and benefits. Senate Bill 08-155 received unprecedented bipartisan support with a 92-2 vote and significantly transformed the management of technology in state government. This historic legislation created a single information technology entity effective July 1, 2008 by transferring all Executive Branch CIOs to OIT and shifting the reporting structure of agency IT employees from the Executive Directors of Executive Branch agencies to the State CIO.

Pursuant to its authority under § 24-37.5-105(9) C.R.S., OIT has responsibility to formalize strategic alignment of IT investments with state policy priorities, IT policy and standards, while reducing duplication and overlap. The Data Management program and strategy as described in this document shall govern every state

Agency ("Agency") as defined in C.R.S 24-37.5-102(4), and all of its respective officers, departments, divisions, commissions, boards, bureaus, and institutions.

Document Overview

The remainder of this document is organized into the following sections:

Section 3: Background and Overview

This section provides a brief history of data management within the state and the key business drivers.

Section 4: State of Colorado Data Management Program

This section provides the vision and strategy for the state's data management program. It also provides information regarding the governance and organization structure of the program.

Section 5: Design and Architecture

This section discusses the high-level design and architectural components necessary to achieving the goals and objectives of the data management program.

Section 6: Implementation Roadmap and Milestones

This section discusses the plan, activities and timeframes for each of the major objectives and initiatives. This section also highlights associated risks and issues for the state in rolling this program out.

Section 7: Appendices

This section provides a Glossary of Terms and a copy of HB 09-1285.

Background and Overview

Given the importance of data in state business, data and information must be treated as a highly valuable enterprise asset. Data is the representation of facts as text, numbers, graphics, images, sounds or video. Facts are captured, stored, and expressed as data. Data can be both structured - stored in structured formats such as databases, flat files, and tagged electronic documents – and unstructured - any document, file, graphic, image, text, report, form, video, or sound recording that has not been tagged or otherwise structured into rows and columns or records. Information is data in context, used to guide decisions.

Citizens should see and expect one state government, despite the fact that government decisions increasingly require more types of data from multiple and diverse government lines of business. Citizens should have one central entry point to the state to see all types of service and program information that they are participants in, as they would if they went to Amazon.com™ or Google™.

Management of the state's data as a unified program is essential for the state to evolve towards building or buying systems in the future that communicate seamlessly, that secure private and sensitive data, and that eliminate redundant data stores and functions. The value and associated risks of the enterprise's data information assets must be ascertained if they are to be properly managed, shared, and protected. Understanding the criticality, value or relative value of data will help to determine the level of investment in security, access, quality assurance, and recoverability. Managing data, information, and knowledge assets in this way is not strictly an IT initiative – this is an enterprise initiative demonstrating strong collaboration across business and technology.

The Importance of Data Management

There are many benefits to the state for initiating an enterprise data management program. There are benefits to policy development, resource alignment, and collaboration, as it is cheaper to share and secure data than to recollect, store, maintain, secure in multiple, often redundant, data stores. Some additional benefits include, but are not limited to:

- Permitting cross-departmental enterprise data analysis and forecasting;
- Allowing for validation of programs across agencies;
- Allowing for verification, refutation and refinement of findings with metrics that provide validation;
- Promoting new research;
- Using evidence-based policy making;
- Encouraging diversity of analysis and testing of new or alternative hypotheses for policy making and results; and,
- Resourcing without duplication of data collection.

Historical Perspective

Unfortunately, many years of decentralized IT oversight, redundant IT software and hardware purchasing, a disjointed approach to infrastructure, and failed IT projects created an information technology patchwork that increased the cost of government and put mission-critical systems at risk. Figure 3.1 below is from an IT assessment commissioned early in Governor Bill Ritter Jr.'s administration, and depicts the large number of redundant infrastructures (software and hardware) across multiple departments. This includes the number of redundant systems collecting and storing data, in varying formats, and with varying privacy and security policies.

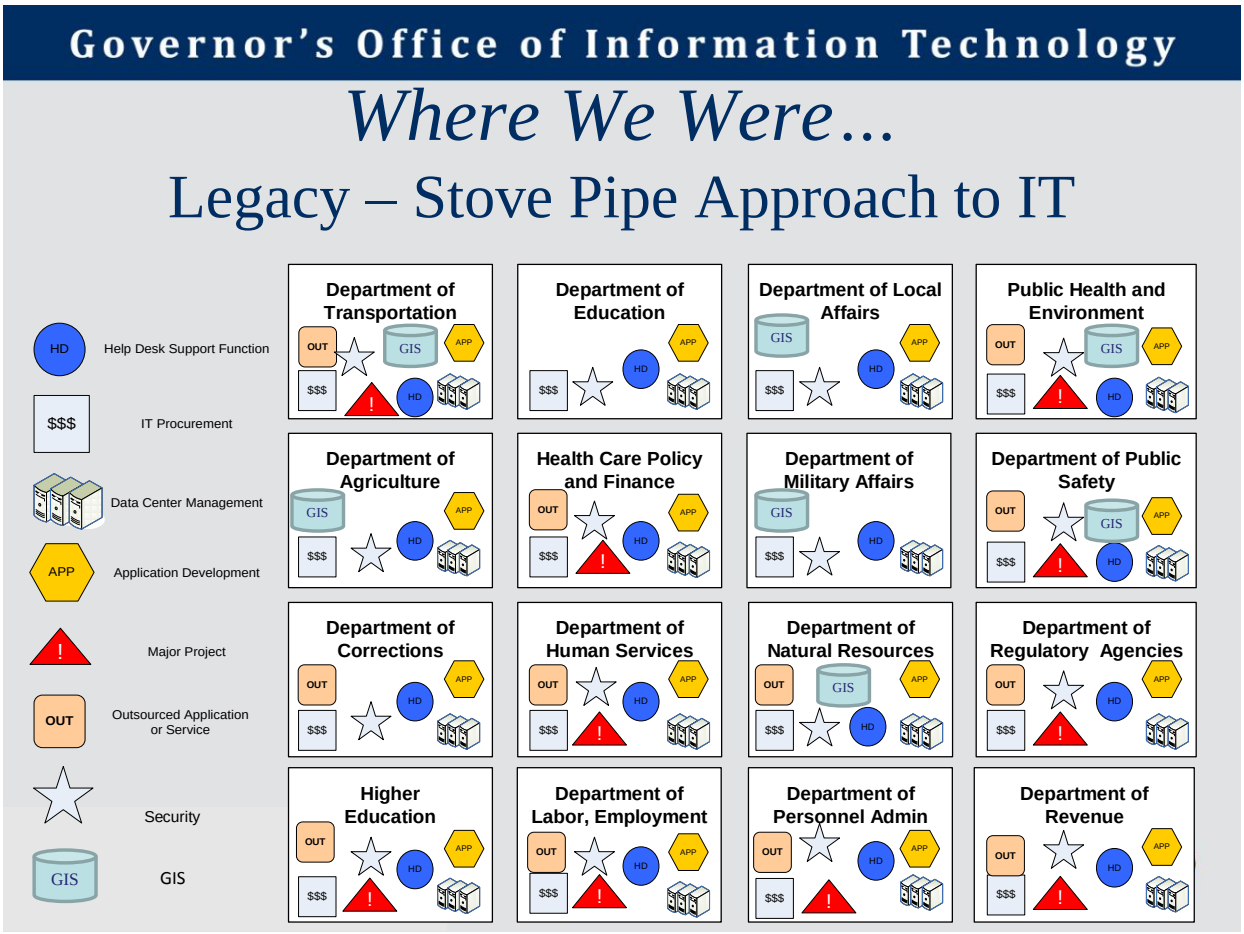


Figure 3.1: Colorado's Historical Siloed IT Environment,

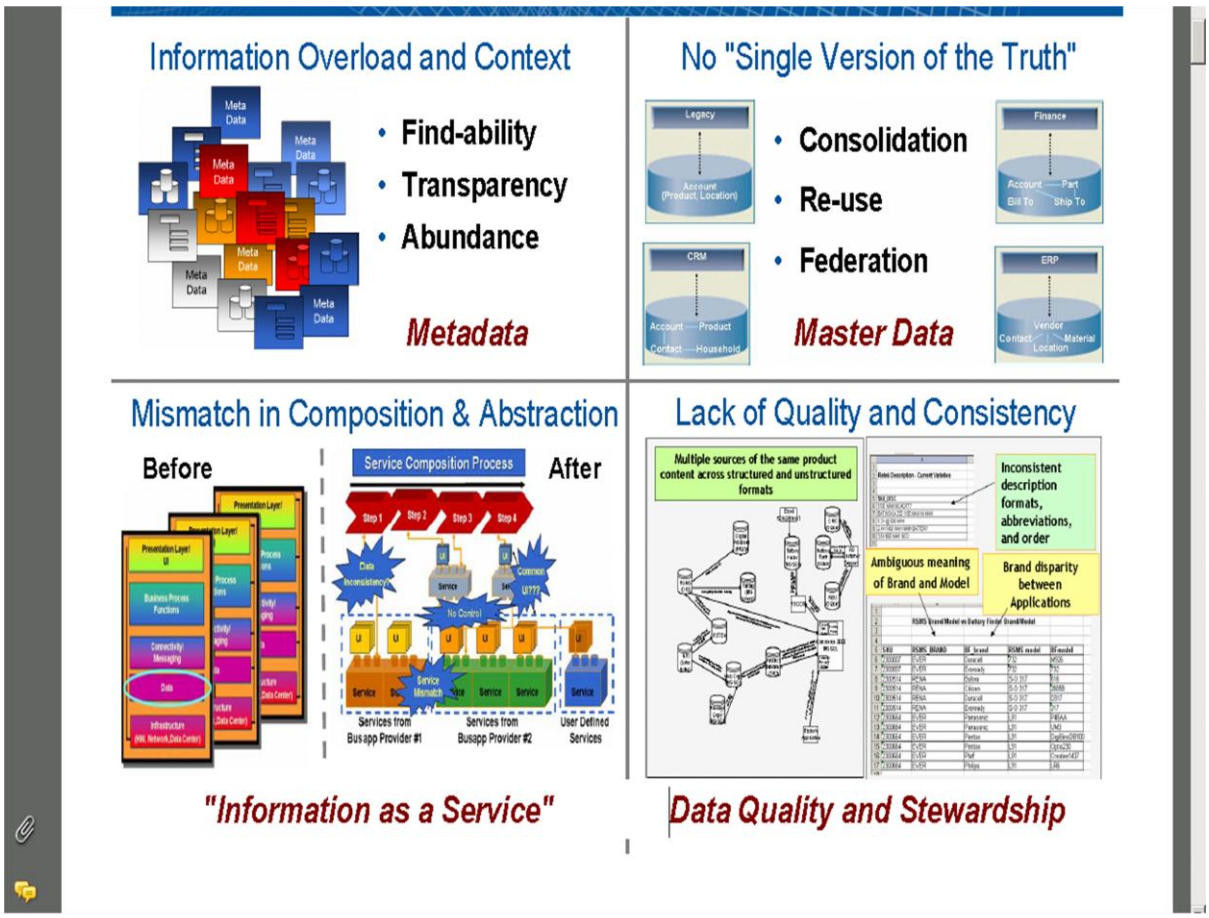
Source: C2P, The Colorado Consolidation Plan

There have never been enterprise standards for any technology infrastructure and project management processes have not been in place long enough for the business to define what they really need or to

coordinate like-programs across the enterprise. Enterprise data assets are stored and maintained to various levels of quality throughout state government. Application teams have worked in isolation and applications were built for immediate return. Project teams were incentivized to deliver results without considering the long term enterprise value and cost.

This decentralized approach has resulted in a myriad of data approaches, non-uniform naming conventions, formats, technologies, tools, staff expertise, and policies and governance approaches. There is very limited visibility into the actual types, format and governance of the data and data sharing activities that are occurring today in and across state Agencies. There is also no way currently to consistently and definitively apply the types of performance metrics that would be typical in a business environment.

The Gartner Group did a study to show the impact to the business when an organization has the types of inefficiencies described above. Figure 3.2 below is taken from that Gartner study:



Source: Gartner Group

Figure 3.2: Challenges with Siloed Systems Environment

It is clear that the state's current data environment isn't working and this problem will only get worse as digital record stores continue to grow. International Data Corporation (IDC) reports that the size of the digital record will grow by a compound annual growth rate of 60%, and by 2011 there will be more than 10 times the amount of electronic data that existed in 2006. We owe it to our citizens, and ourselves, to transform this situation. We're excited with the possibilities and efficiencies we see for the state by implementing this program.

Key Business Drivers

There are a number of influences that have converged in the past 12-36 months that are driving these changes in how the state is approaching data management. These include operational, legislative, financial and components. The Colorado General Assembly is requiring Agencies to be more efficient with their operations and reducing costs while delivering the same level of services. To that end, the Legislature has passed a series of bills over the past few years to address these issues, including HB 08-1364 (Interdepartmental Data Protocol) and HB 09-1285 (Government Data Advisory Board).

- The Data Protocol Development Council established under HB 08-1364 provided strong recommendation and guidance on establishing such a program. The mission of the Council was to provide guidance, policies and procedures for implementing a data sharing architecture across the state enterprise to achieve the data sharing goals and objectives of HB-1364. Two of the primary recommendations of the Council were to establish a formal enterprise architecture office and data management program with OIT and to establish a central governance authority to provide leadership and oversight for data sharing and data management activities. The central governance authority, the Government Data Advisory Board, was established in legislation in 2009 (HB 09-1285).
- The GDAB is a multi-agency central governance authority, comprised of representatives of 12 state agencies, local governments, non-governmental organizations and research institutes, and a wide variety of education stakeholders. The GDAB's mission is to provide guidance and recommendations on how the state should govern and manage data and data management systems to improve the efficiency and effectiveness of state government, citizen service delivery and policy-making. The GDAB is one of the very few such Boards in any state in the country, established in legislation and appointed by the Governor, to provide the central governing structure for enterprise data sharing initiatives.

There is also a renewed energy at the federal level to improve and expand on data sharing initiatives of the past, as well as provide improved access to government data sets for increased transparency. Several federal grant funding opportunities exist that have data sharing and integration as a key component and Data.gov and government Web 2.0 efforts continue to proliferate. Attention to data management as a formal program will position the state favorably as a recipient for these federal grant funding opportunities.

Data Management Program

Since 2007, the Colorado Governor's Office of Information Technology (OIT) has produced a ground-breaking and progressive agenda for information sharing and information technology management in the State of Colorado. It is recognized by both the Governor and the Colorado General Assembly that in order to more effectively serve citizens, improve the efficiency and effectiveness of state government, and to inform policy making, a strong program of information sharing is required, across all lines of business the state serves.

OIT's focus on agile application and service delivery means that strong concentration and emphasis on the underlying data required for employees, agencies, legislators, and others to do their work *must* be a top priority. The "businessization" of government requires a disciplined approach to dismantling data silos, implementing infrastructure to enable sharing across agencies, branches, and levels of government, and tools to build capacity for knowledge and performance management.

Data Management (DM) is an ongoing, centralized administrative function that consists of the planning and execution of policies, practices, and projects that acquire, control, protect, deliver, and enhance the value of data and information assets. The responsibilities typically assigned to this function include data and information strategy planning; data governance; data architecture management; data development; database operations management; data security management; reference and master data management; data warehousing and business intelligence management; document and content management; meta-data management; and, data quality management.

The purpose of the DM program within the Office of Enterprise Architecture (OEA) is to leverage data and information as enterprise assets and to establish standards and processes to enable more agile solutions and government services. The DM program has the following responsibilities:

- Ensure that data and information assets are known, usable, reusable, and can be accessed when and where needed.
- Develop processes and procedures to review Agency data documentation as deemed necessary.
- Review Agency data documentation and provide feedback to agencies in a timely manner.
- Establish common policies, procedures and standards to maximize the sharing and investment in data and information resources.
- Coordinate and leverage existing state investments in data and information resources.
- Provide metrics and dashboards on the state of the state's data management performance.
- Develop appropriate security and privacy policies to protect data assets, by working in conjunction with the Office of Cyber Security on data security management and the Attorney General's Office on privacy concerns.
- Establish collaborative and cooperative relationships with public and private sector organizations to invest strategically in data and information assets and promote reusability.

Our partners in this program are the state agencies, local government entities, non-profit and research organizations and the private sector. The Office of Cyber Security is especially critical for security and the Attorney General's Office for privacy protection and open records-related issues. Best practices, national standards, industry trends and subject matter experts will all be utilized as the program develops. Due to the state's strategic relationship and partnership with the federal government and fellow states, we will develop and adopt policies and standards that make it easier to facilitate communication flow across boundaries. This includes adoption of standards such as the National Information Exchange Model (NIEM).

Data Management Principles

The following are the seven key principles of the DM program, which support the enterprise vision, and to which all strategic and tactical data and information management efforts must be aligned:

1. **All data and information are valuable assets owned by the State of Colorado.**
    a. Information requires stewardship, and line-of-business data is managed within each Agency.
    b. The view of data is from a business process perspective.
    c. All data assets are inventoried, catalogued and known.
2. **Smart data collection, management and usage require enterprise architecture and data architecture standards to enable information sharing and reuse.**
    a. Data is defined once and used wherever applicable.
    b. Common policies, procedures and standards are established to maximize the sharing and investment in information resources.
    c. State investments in data and information resources will be coordinated and leveraged across the enterprise.
    d. National data standards shall be used when possible.
    e. National data exchange standards shall be used when possible.
    f. Each Agency implements the same standards for data quality, integrity and availability.
3. **Enterprise data and information emanates from institutionally identified authoritative data sources.**
    a. There will be a single authoritative source for each type of domain data.
    b. Data duplication that causes data reconciliation and differences in management reporting will be avoided.
    c. An enterprise metadata repository is developed and used by every Agency.
    d. Systems of record provide visibility of their data to other systems as needed and authorized.
4. **Data is of the highest quality, reliability and integrity.**
    a. Data quality controls are developed and implemented.
    b. Where possible, data is updated, modified and fixed at the source.
    c. Improvements are planned and resourced.

   d. Data governance performance is measured.
5. **Data and information must be secure and accessible.**
   a. Privacy policies must be strictly adhered to.
   b. Federal and state regulatory policies must be followed.
   c. Strong authentication and access control policies and systems will be implemented.
   d. Roles-based access methods should be implemented.
   e. Individually identifiable data should only be made available to authorized persons.
6. **The process of data governance in the state will be open, transparent, timely, and will incentivize cooperation and develop an environment of trust.**
   a. Agencies, citizens and other stakeholders will know and understand the processes by which data is governed and managed.
   b. Agencies, citizens and other stakeholders have a right to be heard in how data is governed and managed.
   c. An environment of collaboration and trust regarding data management and data sharing must be cultivated and maintained for success in this program.
7. **Throughout the state enterprise, people have timely access to data to improve their business.**
   a. The right data must be provided to the right people at the right time to serve the citizens of the state.
   b. An information exchange environment must be established to facilitate the efficient exchange of data and information.

Organization

In July 2009, the Office of the Enterprise Architect (OEA) was established to begin developing the state's enterprise architecture framework strategy in addition to setting policies, processes and standards in order to align the state's investment in IT assets with its business strategy, thus avoiding duplication or redundancy of infrastructure assets. Within the OEA, the state's Data Management (DM) program was established and the nation's first State Chief Data Officer (CDO) was named with the responsibility for developing and implementing the state's data strategy and chairing the Government Data Advisory Board (GDAB). OIT is formalizing the OEA and DM program through its rule-making authority, which gives its regulations the force of laaw.

The CDO and state CIO will be advised in this program by the GDAB. Additionally, a data stewards council will be established to bring together the people who know their data best to assist in developing the state's data processes, standards and policies. The scope of responsibility of each is described below.

Office of Enterprise Architecture

Enterprise Architecture (EA) is a comprehensive framework used to manage and align an organization's IT assets, people, operations, and projects with its operational characteristics. In other words, the

enterprise architecture defines how information and technology will support the business operations and provide benefit for the business. The purpose of the Office of Enterprise Architecture (OEA) within OIT is to deliver agile business solutions enabling connected government services.

The OEA has the following responsibilities:

- Develop a comprehensive EA framework and repository across the agencies to ensure the alignment between the business strategy and the supporting technical and data architectures.
- Develop processes and procedures to review Agency architectural documentation as deemed necessary.
- Review Agency architectural documentation and provide feedback to agencies in a timely manner.
- Set enterprise standards for all technical domains.
- Establish processes necessary for implementing new technologies and systems in response to the state's changing business needs.
- Coordinate and leverage existing state investments in information technology infrastructure.

The Colorado Enterprise Architecture Framework (CEAF) is a combination of various architectural approaches that have been tailored to meet the state's needs. As shown in Figure 4.1 below, the Enterprise (State of Colorado) has various Business domains (equating to Agency services) that are supported by Applications, Information, and Technology domains. Solution Sets and Security Services span these domains. Within each domain are various disciplines or subject areas.

The Business Domain is really the foundation for the other components of the EA. It is the high-level representation of the vision, mission, goals, objectives, and business strategies and functions that comprise the strategic business intent of government. That intent is then enabled through a variety of capabilities such as functions, processes, information, know-how, and technology critical to providing services to its citizens, agencies, bureaus, departments, business, vendors, branches and others with whom the government interacts.

The Information Domain and architecture aligns the business processes to the information systems that support these processes, promotes information sharing and facilitates cross-agency information exchanges. The objective is to manage the information of the enterprise. The DM program is centered in this domain.

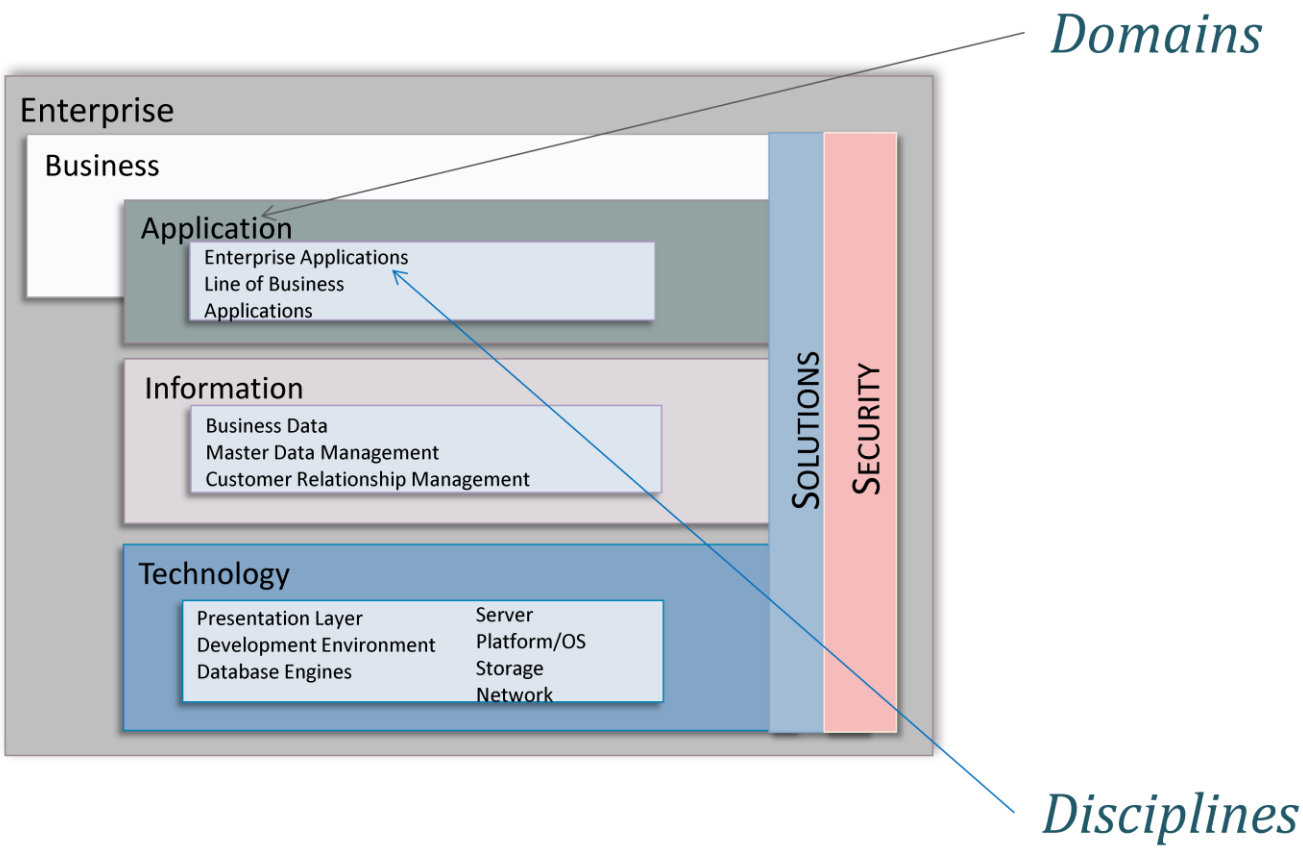


Figure 4.1: Colorado Enterprise Architecture Framework (CEAF)

Chief Data Officer

The Chief Data Officer (CDO) is responsible for developing the state's enterprise data strategy and data governance program. The position works within the OEA, and specifically furthers the enterprise's strategic data policies. Responsibilities for this position include:

- Acts as Director for driving and executing the enterprise data strategy.
- Develop the strategic plan for implementing a data governance and data management framework.
- Work with OIT management to implement the cross-agency data sharing protocol.
- Communicate and leverage the best practices for sharing and protecting citizen data.
- Develop enterprise internal and external data policies and procedures.
- Approves and communicates Board-level recommendations on education data sharing and management strategies.
- Develops and executes stakeholder communication plan.
- Coordinates legislative and policy actions as necessary for the data management strategy.

The CDO has responsibility for the Information Domain within the CEAF.

Government Data Advisory Board

The GDAB is a multi-agency central governance authority, comprised of representatives of 12 state Agencies, local governments, non-governmental organizations and research institutes, and a wide variety of education stakeholders. The GDAB's mission is to provide guidance and recommendations on how the state should govern and manage data and data management systems to improve the efficiency and effectiveness of state government, citizen service delivery and policy-making.

The Board has the following responsibilities:

- Advise on the development, maintenance, and implementation of the data sharing protocol;
- Advise on the best practices for sharing and protecting citizen data;
- Review, advise, and provide input into the strategic plan for improving data governance;
- Advise on compliance, privacy and security data requirements;
- Advise on internal and external data policies and procedures;
- Advise on financial and budgetary components required for implementation; and,
- Specifically recommend education data sharing and management strategies.

Finally, the Board will develop recommendations with time frames and priorities for developing and implementing the cross-departmental data protocol. The Board and OIT and the CDO are aligned as shown in the Figure 4.2 below:

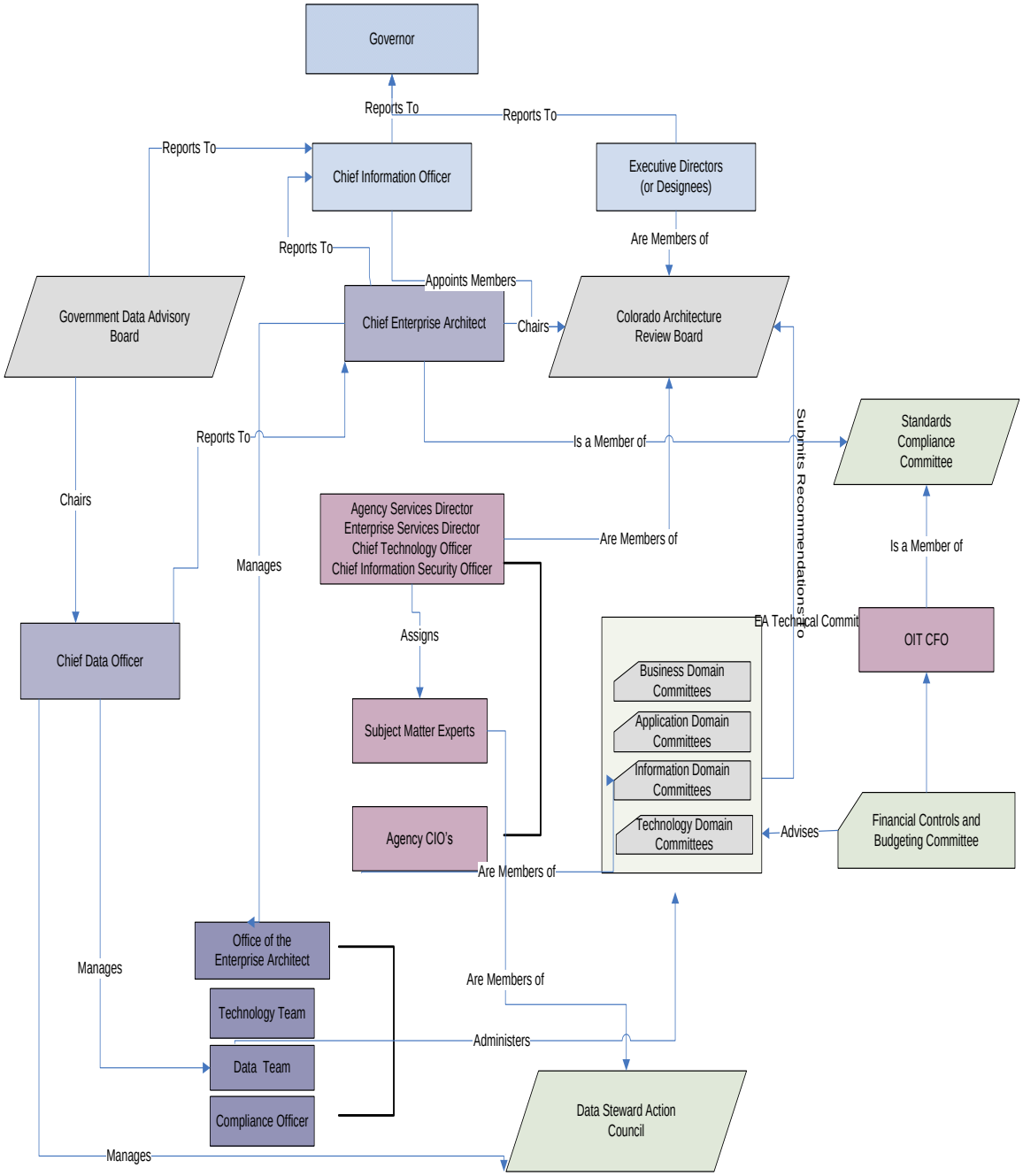


Figure 4.2: Enterprise Architecture and Data Management Reporting Structure

Data Steward Action Council

The Data Steward Action Council ("DSAC") is being established to formalize and organize the stewardship activities and processes statewide based on information subject areas. The DSAC will also

create a common baseline of information: a statewide foundation for data sharing, information discovery, and future architectures. The DSAC has the following responsibilities:

- Establish statewide business data definitions, data quality rules, domains, code descriptions and value sets, validation and resolution policies and methods.
- Define and verify technical data physical definitions.
- Define metadata repository input and retrieval methods.
- Provide input to and refine/define the Enterprise Data Model including the statewide Agency information subject data domains and sub-domains.
- Identify and define Stewardship standard processes, documents, and forms.
- Lead the inventory of the state's defined business functions.
- Lead the inventory of the state's data systems.
- Identify business rules and contextual security requirements regarding both the internal and external use and sharing of the data.
- Spearhead the effort identifying relationships between the Federal Enterprise Architecture (FEA), the Colorado Enterprise Architecture Framework, National Information Exchange Model (NIEM), recognized subject domain area models and internal business models.
- Establishing metrics for measuring the DSAC's success.

The DSAC will operate through the sponsorship of the CDO in the OEA and is led by an enterprise data architect.

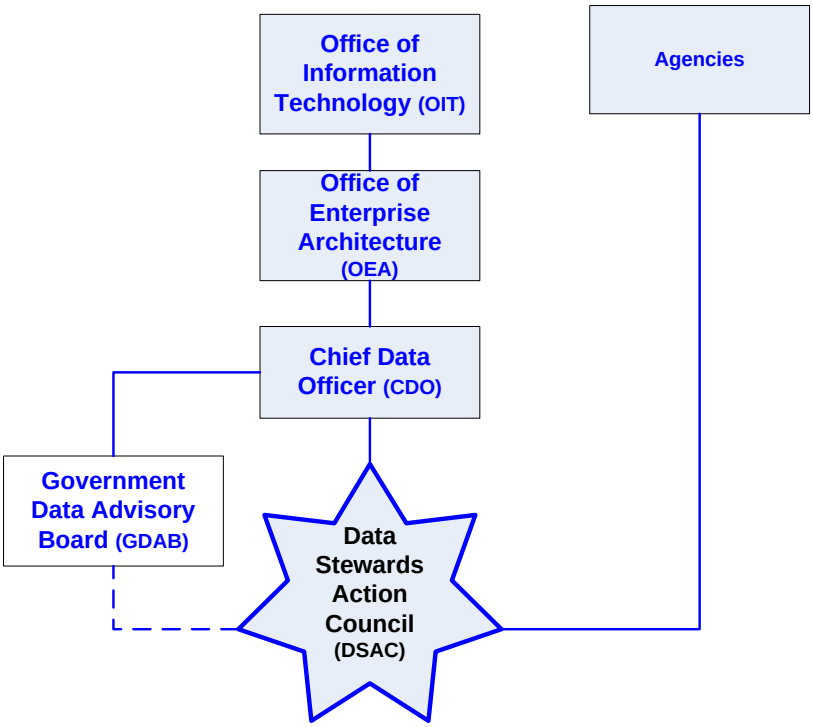


Figure 4.3: Data Stewards Action Council Organizational Alignment

Strategic Vision

The State of Colorado's vision is to be *one of the most innovative, admired organizations for our approach to data governance and enterprise data management.*

The mission for ***enterprise data management*** is to foster collaboration, innovation and agility in delivering government services to the citizens of Colorado through the seamless, efficient, strategic exchange of core data sets resulting in increased effectiveness of government operations. Figure 4.4 shows the alignment of the vision and mission with the program's strategic objectives and key initiatives.

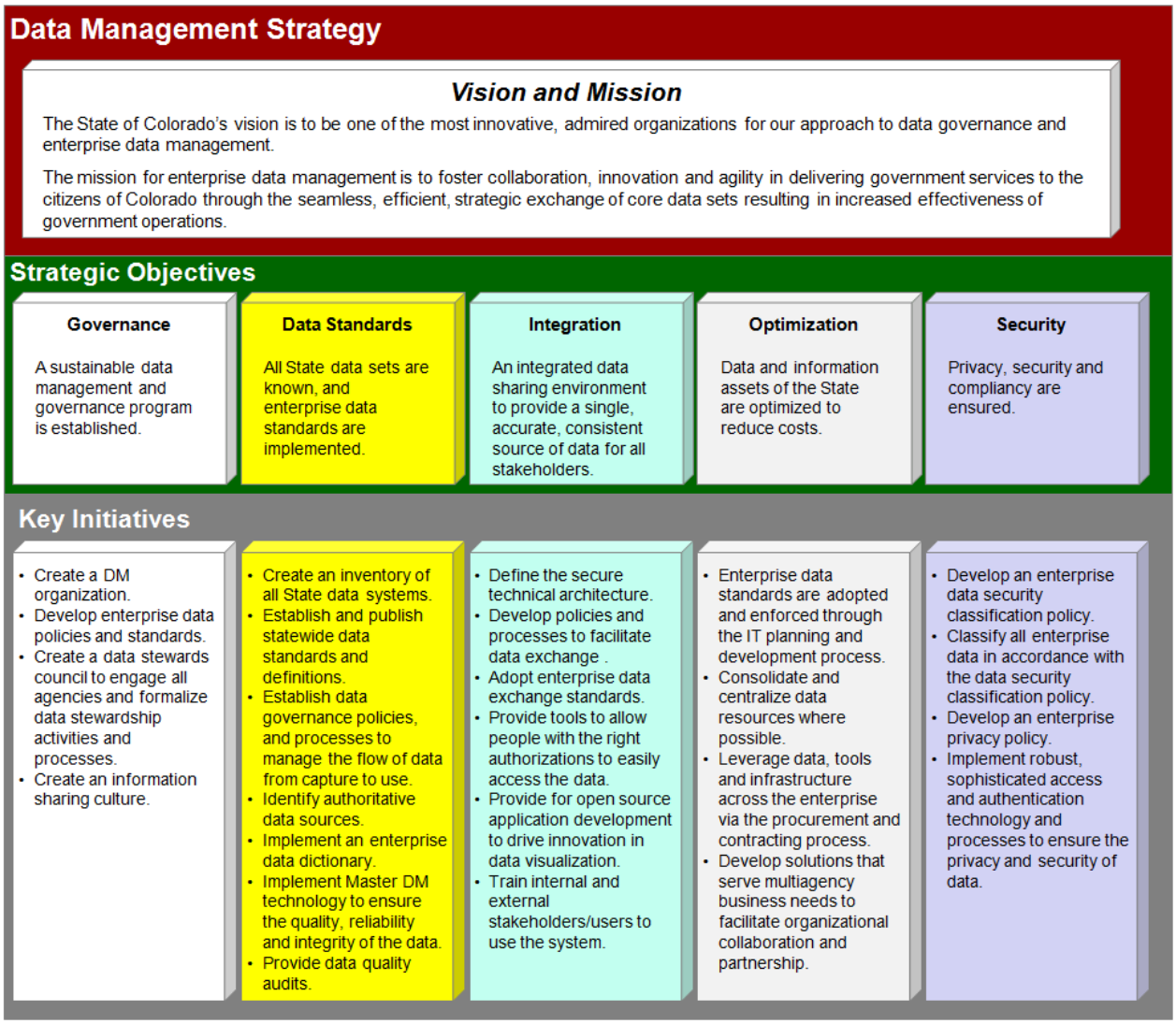


Figure 4.4: Strategic Vision and Program Goals

Program Goals and Strategic Objectives

The following are the five goals and associated objectives for the DM program. The initiatives and time frame for operationalizing these are outlined in **Section 6, Implementation Roadmap and Milestones**.

Goal 1: A sustainable data management and governance program is established.

Objective 1.1 – Create a data management organization.

Objective 1.2 – Develop enterprise data policies and standards.

Objective 1.3 – Create a data stewards council to engage all agencies and formalize data stewardship activities and processes.

Objective 1.4 - Create an information sharing culture.

Goal 2: All state data sets are known, and enterprise data standards are implemented.

Objective 2.1 – Create an inventory of all data and data systems in the state.

Objective 2.2 – Establish and publish statewide data standards and definitions for all data elements.

Objective 2.3 – Establish data governance policies, processes and standards to manage the flow of data from capture to use.

Objective 2.4 – Identify authoritative data sources for all data types.

Objective 2.5 - Create and implement an enterprise data dictionary and taxonomy.

Objective 2.6 – Implement Master Data Management technology to ensure the quality, reliability and integrity of the data.

Objective 2.7 - Provide data quality audits as part of the ongoing monitoring of data quality.

Goal 3: An integrated data sharing environment is created to provide a single, accurate, consistent source of data for all stakeholders.

Objective 3.1 – Define the secure technical data exchange architecture.

Objective 3.2 – Develop policies and processes to facilitate data exchange across state government agencies.

Objective 3.3 – Adopt enterprise data exchange standards.

Objective 3.4 – Provide tools to allow people with the right authorizations to easily access the data.

Objective 3.5 - Provide for open source application development to drive innovation in data visualization and encourage sharing of both information and technology among all interested education stakeholders.

Objective 3.6 - Train internal and external stakeholders/users to use the system.

Goal 4: Data and information assets of the state are optimized to reduce costs.

Objective 4.1 – Enterprise data standards are adopted and enforced through the IT planning and development process.

Objective 4.2 – Consolidate and centralize data resources where possible.

Objective 4.3– Leverage data, tools and infrastructure across the enterprise via the procurement and contracting process.

Objective 4.4 – Develop solutions that serve multi-agency business needs to facilitate organizational collaboration and partnership.

Goal 5: Privacy, security and compliancy are ensured.

Objective 5.1 – Develop an enterprise data security classification policy.

Objective 5.2 – Classify all enterprise data in accordance with the data security classification policy.

Objective 5.3 – Develop an enterprise privacy policy.

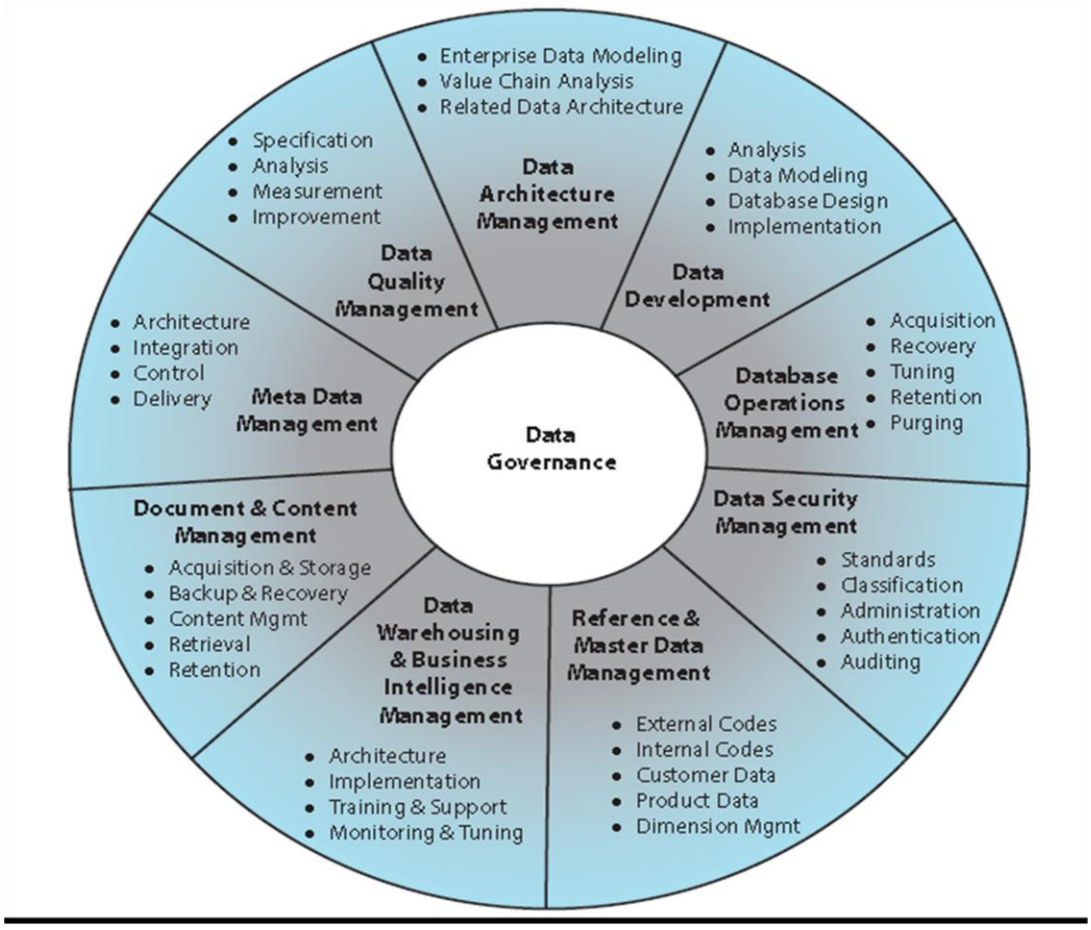
Objective 5.4 – Implement robust, sophisticated access and authentication technology and processes to ensure the privacy and security of data.

Enterprise Data Management Framework

Data Management (DM) is an ongoing, centralized administrative function that consists of the planning and execution of policies, practices, and projects that acquire, control, protect, deliver, and enhance the value of data and information assets. The responsibilities typically assigned to this function include data and information strategy planning; data governance; data architecture management; data development; database operations management; data security management; reference and master data management; data warehousing and business intelligence management; document and content management; meta-data management; and, data quality management.

Like any asset, data has a lifecycle. Data is created or acquired, stored and maintained, used, and eventually destroyed. Many things happen with data over the course of its lifecycle, and each of those activities must be managed and governed appropriately. In order to facilitate communication and dialogue about data management, a standard language must be used, and this comes in the form of the data management process model.

The Data Management Associations' (DAMA) Functional Framework is an organizing structure that promotes consistency to meet organizations' planning and process needs. The overall data management function encompasses 10 major component functions, as shown in Figure 4.5 below:



Source: Data Management Association

Figure 4.5: Data Management Framework

1. Data Governance: The exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets. Data Governance is high-level planning and control over data management.
2. Data Architecture Management: Defining the data needs of the enterprise, and designing the master blueprints to meet those needs. This function includes the development and maintenance of enterprise data architecture, within the context of all enterprise architecture, and its connection with the application system solutions and projects that implement enterprise architecture.
3. Data Development: Designing, implementing, and maintaining solutions to meet the data needs of the enterprise. The data-focused activities within the system development

lifecycle, including data modeling, data requirements analysis, and design, implementation, and maintenance of database' data-related solution components.

4. Data Operations Management: Planning, control, and support for structured data assets across the data lifecycle, from creation and acquisition through archival and purge.
5. Data Security Management: Planning, development, and execution of security policies and procedures to provide proper authentication, authorization, access, and auditing of data and information.
6. Reference and Master Data Management: Planning, implementation, and control activities to ensure consistency with a "golden version" of contextual data values.
7. Data Warehousing and Business Intelligence Management: Planning, implementation, and control processes to provide decision support data and support for knowledge workers engaged in reporting, query and analysis.
8. Document and Content Management: Planning, implementation, and control activities to store, protects, and access data found within electronic files and physical records (including text, graphics, images, audio, and video).
9. Metadata Management: Planning, implementation, and control activities to enable easy access to high quality, integrated meta-data.
10. Data Quality Management: Planning, implementation, and control activities that apply quality management technique to measure, assess, improve and ensure the fitness of data for use.

This is the framework by which OIT will develop its governing processes, procedures, standards and guidelines to manage the data of the State of Colorado. All policies, processes and standards will be available on the OIT website (www.colorado.gov/oit) as they are published. This is the first such formal, enterprise program to be developed within the Executive Branch, and there is much work to do. Early efforts will be focused on the following areas (see Figure 4.6 below), primarily to support enterprise data sharing efforts per statute. Eventually every piece of the circle will have a full set of guidances that OIT, agencies and data stewards will utilize to manage state data assets and maximize their use.

Figure 4.6: Initial DM Target Policy Areas

Data Governance

Data governance is the core function of the DM framework shown above. There are many definitions of data governance, but to use the DAMA definition: "…it is the exercise of authority and control (planning, monitoring and enforcement) over the management of data assets. The data governance function governs how all the other functions are performed." Data governance, once established, must be an on-going program, and one of continuous improvement.

Figure 4.7 below from the Data Governance Institute shows the data governance activities, participants, and processes:

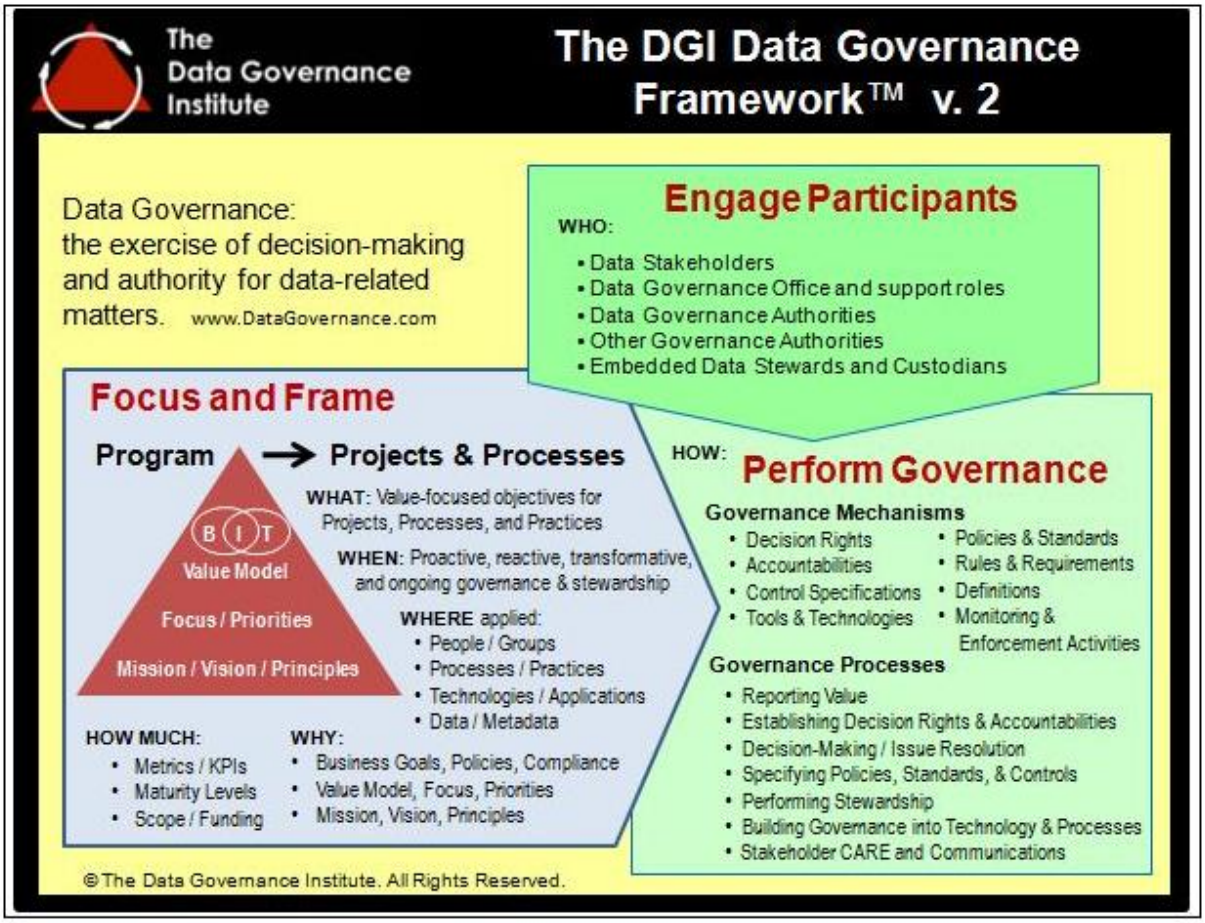


Figure 4.7: Data Governance Framework

Data governance is truly the heart of the intersection between business and technology. Information technology does not exist for its own sake. It exists to meet the needs of the business. The business side of the organization drives requirements that the IT organization implements on their behalf. These requirements include data element definitions, business process, priorities, compliance requirements, privacy issues, rules, and more. While the IT organization implements these requirements through technology, it is not the IT organization that defines the requirements (see Figure 4.8 below).

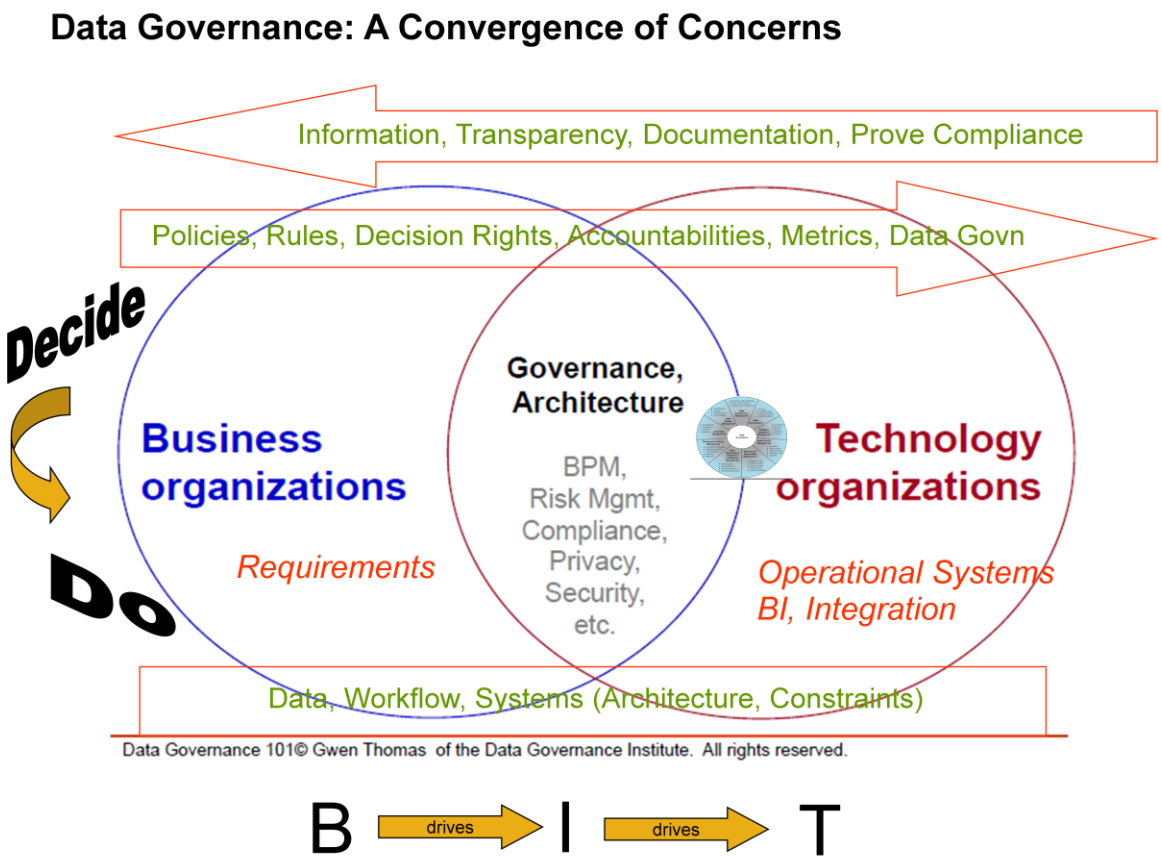


Figure 4.8: Governance Convergence

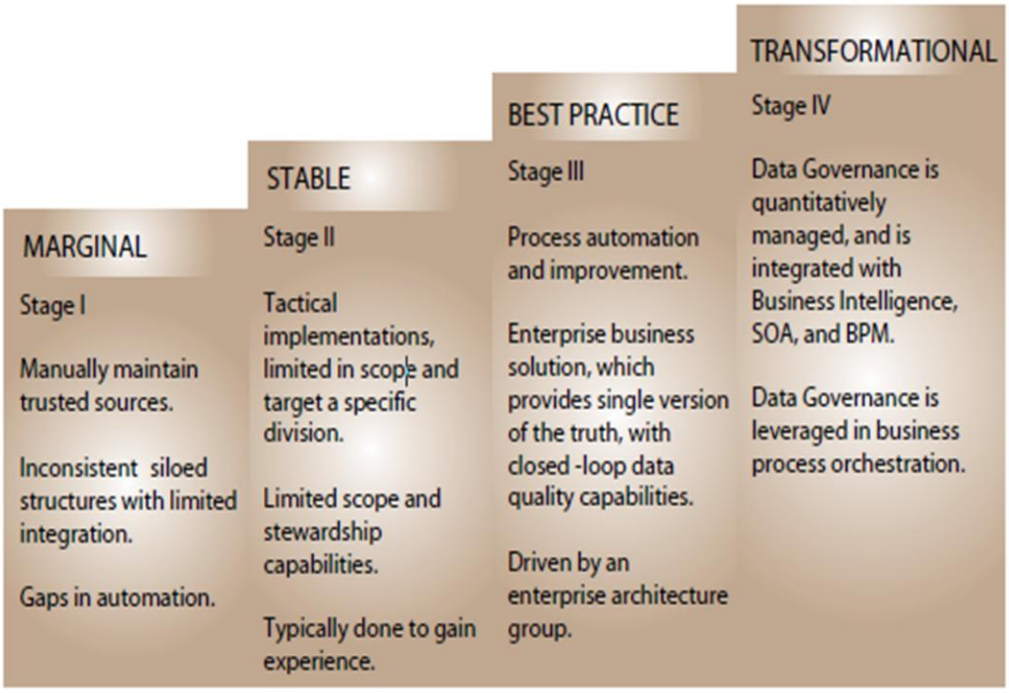
Data governance then sets the policies, standards, principles, rules, etc. that become uniform across all lines of business, then enforces those rules, manages conflicts, and deals with any new issues. Data governance in Colorado will be managed by the CDO, in conjunction with the GDAB and the DSAC.

Data Governance Maturity Models

The National Association of State Chief Information Officers (NASCIO) has published several documents on data governance. "Data Governance Part II: Maturity Models" provides a thorough discussion of how enterprises progress on their path of data governance once they've decided to move forward with a formal program. The maturity model describes the journey from the AS IS to the SHOULD BE regarding the management of data, information and knowledge assets. In parallel to this journey regarding data governance is the journey that describes a maturing enterprise architecture operating discipline. State government must understand where it is today and where it needs to go. This is an important step in *planning the journey* in managing information as an enterprise asset. Data governance maturity models provide the means for gauging progress. By presenting intermediate milestones as well as the desired

end state, maturity models assist in planning HOW state government will reach the next level of effectiveness, as well as WHEN and WHERE within state government.

Most maturity models are relatively similar in theme and progressive steps. The following example is from Oracle Corporation and shows the transformation an enterprise will go through as it moves from siloed and decentralized with no central governance, to a strong, robust data governance program with buy-in from all key stakeholders.



Source: Oracle Corp.

Figure 4.9: Oracle Corp. Data Maturity Model

The state is just beginning its DM program and so ranks “Marginal” according to the maturity model scale. Over the next few years, as the program takes hold and objectives are realized, we will move up the scale towards “Transformational”.

Enterprise Data Tools

The DM program will provide five primary tools to the enterprise to facilitate data discovery, integration, sharing, re-use, and systems development activities. These are:

- Enterprise Systems and Data Inventory
- Business Data Value Matrix
- Master Data Repository
- Metadata Registry
- Enterprise Data Models

Enterprise Systems and Data Inventory

The inventorying of state data systems began with the 1364 Council in 2008. When complete, the inventory will provide the baseline and visibility that the CDO and DM program need to successfully implement this program. Information is being gathered on:

- Systems – the hardware, software, applications, networking and communications systems containing and surrounding the data
- Data – the data elements, data dictionary, data models, etc. for all state systems
- Compliance Environment – existing statutory or regulatory laws with regard to collecting, storing, sharing and destroying data

Data.Gov Efforts

President Obama has a made a commitment to open government that is unparalleled. One demonstration of this commitment is the launching of the Open Government Initiative. This initiative is founded on three principles:

- Transparency – to enable greater accountability, efficiency, and economic opportunity by making government data and operations more open.
- Participation – to create early and effective opportunities to drive greater and more diverse expertise into government decision making.
- Collaboration – to generate new ideas for solving problems by fostering cooperation across government departments, across levels of government, and with the public.

The difference now is the availability and provision of raw, machine-readable data, structured to allow manipulation and analysis electronically. This allows the combining of multiple public data sets using mashups to create new information, data and services. Data within the context of these discussions is unabridged data; data that has not been aggregated, summarized or interpreted. (Source: NASCIO)

State government has always been a source or supplier of public data, however it has often been difficult to discover and not available in a raw format to be consumed and reused. “Democratization of

data" is often used to describe the process of making government data more available. This is now being accomplished through federal, state and local "data.gov" portals.

Colorado has begun initial efforts with its data.gov initiative, and will continue to prioritize this work in 2010 through the DM program, as the inventory of the state's data sets builds. A site within the OEA portion of the OIT website will be launched to support this effort.

Business Data Value Matrix

The *Business Data Value Matrix* will cross-reference each business Agency with the subject data area that is important to and supports that Agency. This matrix forms the high-level communication base that acts as a foundation for economical and quality data sharing and integration.

The state business data value matrix will include agencies across the top of the matrix, and the subject domain and sub-domains of the data along the left side of the matrix with an indicator in the intersecting squares indicating if data is created or used by that Agency. These matrices are used to quickly see where the impact of changes to data will be reflected across the organization and may be produced in report from the metadata repository.

The DSAC will maintain the state business data value matrix (see an example matrix in Figure 4.10 below) and a lower-level two-dimensional element usage matrix (EUM) that contains CRUD information (Create, Read, Update, and Delete) that cross references the data elements of the state to the Agencies and applications that define and use that data.

Subject Areas	Subject Area Major Domain sub-area	CDA	CDOC	HCPF	DHE	CDHS	CDLE	DOLA	DMVA	DNR	CDPS	CDPHE	DORA	CDOT	CDE	DPA	SOS	OIT	DOR
Case			X				X	X											
Filing		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
Financial		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
License													X	X					X
Location																			
	Geographical	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
	Address	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Organization (Internal, External)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Person (Employee, Client)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Name	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	Demographics	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Program/ Service		X					X			X									
	Education				X										X				
	Corrections			X				X	X										
	Health			X								X							
	Public Safety								X		X								
	Social					X							X						
	Transportation		X											X					
Project																			
Resources (depletable, non-depletable)		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Figure 4.10: Business Data Value Matrix

The subject data model and associated Business Data Value Matrix depicted above are *initial drafts*, meant to be refined and validated over time by the data stewards and their agencies. There will be a data steward needed in each area where the subject area intersects the state Agency. In most cases, these will be the same people within an Agency but in some cases the specialization will require separate data stewards. The coordinating data steward will be responsible for the subject area and be the responsible DSAC member.

The data stewards will work with data architects to refine the model to support the various state businesses and the Agencies' individual missions and obtain the information to add the create, read, update, delete information at the detailed application systems and data element level.

Master Data Management

MDM is a modern approach that provides data management and specialty services out of the box designed to deal with the data fragmentation and data quality issue. It also provides a different approach with regard to data integration. Traditionally data integration is done on a point to point basis

creating a tremendous complexity that increases exponentially with the number of systems that are participating in the integration. MDM leverages modern service oriented architecture (SOA) integrations based on a hub and spoke approach that reduces the integration efforts. MDM takes advantage of the canonical objects that are the common representation of the citizen attributes across all the parties involved in the integration. Due to the nature of the systems the concept of "citizen" will vary depending on the intent the system was created. The concept of "citizen" for a call center application is different from an HR system or from a financial system.

As an example, there is a need to define "citizen" so there is a common set of attributes across all the system within the state and that is the canonical object. The benefits of the MDM solution resides in the ability to consolidate, clean and share and standardize citizen information across these heterogeneous Agency systems throughout the state and it does have a huge impact in the operations , the services and downstream analytics and reporting. MDM can deliver the right data to an analytics framework. MDM can deliver cross reference ids across agencies for a particular citizen.

MDM has to have a robust and extensible repository at a heart of the solution. MDM collects information from all the systems so as to have a composite record that includes attributes from all of these systems for a complete picture. MDM also provides services for consistency across systems such as:

- Consolidate and maintain unique, complete and accurate master information across the state for data elements in one central location
- Cleanse data centrally
- Assign unique identifiers to persons in a centralized fashion
- Distribute this information as a single source of truth as a service to all operational and analytical applications just in time

These capabilities provide the state with the ability to get better information from existing IT assets and infrastructure; to deliver consistent definitions across the organization; to lower data management costs; and, to enhance reporting capabilities and to provide a foundation for agility and innovation.

MDM must seamlessly integrate with modern SOA and leverage standards based Identity Management Systems in order to manage the master data across the many systems across the state agencies that are responsible for data entry, and bring the clean citizen master data to the applications and processes that run the business. MDM becomes the central source for accurate fully cross-referenced real time master data. It must seamlessly integrate with data warehouses and the Business Intelligence (BI) systems, designed to bring the right information in the right form to the right person at the right time. In addition to supporting and augmenting SOA and BI systems, the MDM application must support data governance and enables orchestrated data stewardship across the enterprise.

MDM provides services with the ability to:

- *Consolidate Information* into one master repository from disparate systems and organizational units; this step involves the identification and cross-referencing of people by their external ids or their representation in other systems. The initial creation of a best version is when the unique id (Colorado Unique Personal ID, CUPID) is created. It also involves the storing of all of the source information, known as the Source Data History, so that a record of where and how the best version or golden record was created. Basically it is a historical view of all the merges that occurred in creating the best version so that a data steward can go back and recreate the steps that led to the current best version. Survivorship entails the weighting of certain fields by source in the creation of the golden record. For example, the call center would have a higher confidence level for the address fields than the web site because call center personnel validate the person's address by asking, "Do you still live at 123 Elm Street?"
- *Cleanse and Enrich* data centrally is the application of the various data quality components to your information.
- *Distribute Data* as a single point of truth for a consistent enterprise view, sharing the information either in batch or just-in-time publishing.
- *Leverage Master Data* to service consuming applications, enterprise business processes and decision support systems in support of key Agency and state level business processes.

Metadata Repository

The cornerstone of data sharing is metadata. Metadata is data about data. It can be employed in a variety of ways to enhance the value and usability of data assets. Traditionally it is used to define data structures and relationships to support development of applications. It will also enable discovery of data assets.

An enterprise metadata repository will be implemented as part of the DM program. The purpose of the enterprise metadata repository is to establish a central point for data identification, standardization, and sharing. This ensures consistent use of the data assets and data resources across the state, facilitates easy mapping of data between computer systems, and lowers the costs of migrating to new systems, including service-oriented architecture. Users and application developers will be able to quickly discover data assets by searching the registry, and data quality will improve because all departments will use the same definitions and standards.

Enterprise Data Models

An enterprise data model is an integrated, subject oriented model that identifies the subject areas and business entities (classes of real business things and concepts) for the state, providing the highest level of communication. The business rules governing the relationships between business entities and some of the essential business data attributes are also defined here to better understand the relationships and core data that may be shared. It consists of three levels: the Subject Area View, the Conceptual View, and the Logical View models (see Figure 4.11 below). Each level becomes more detailed and complex, with the Subject Area View providing data subject areas and the business areas or agencies

that use them, the Conceptual View identifying the most significant business entities and their relationships, and the Logical View containing business attributes added to the entities.

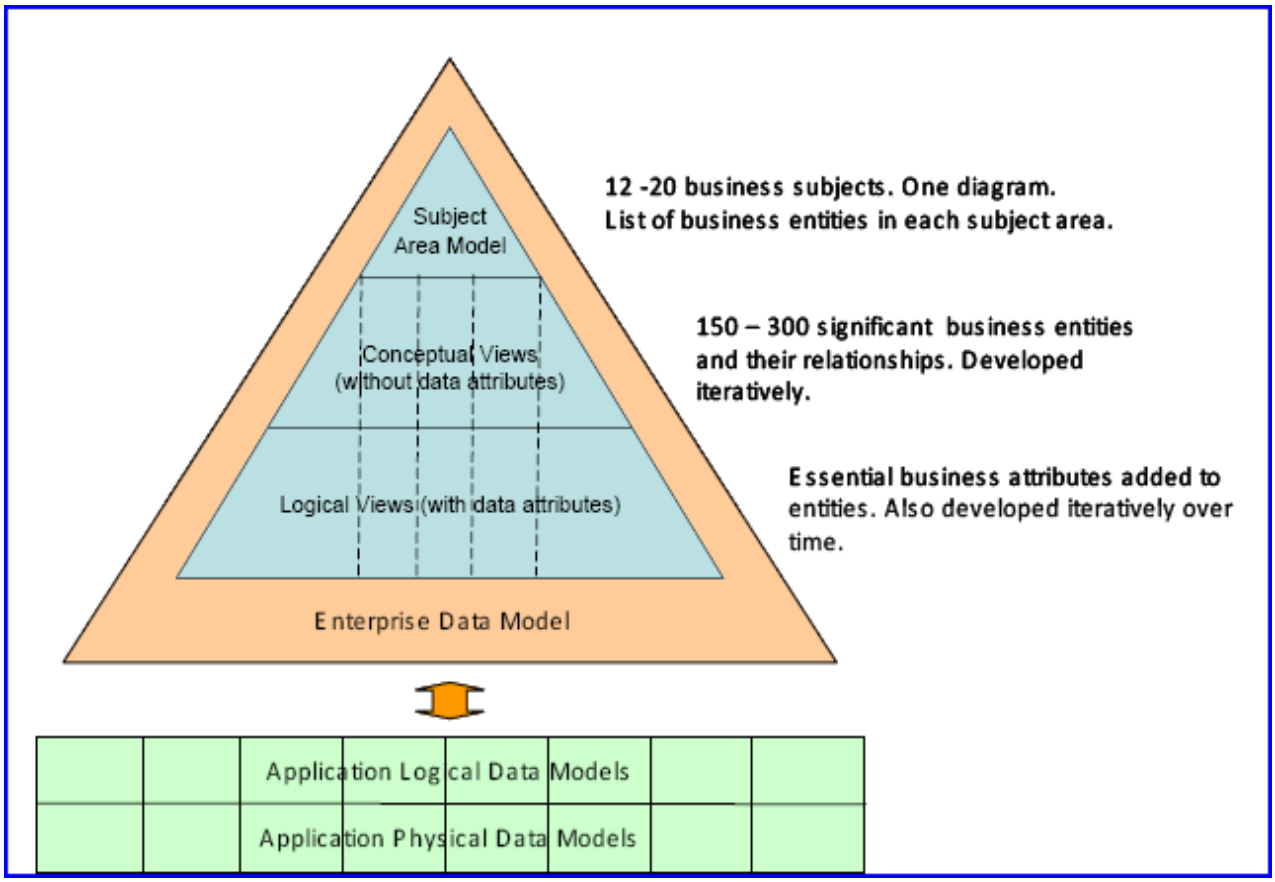


Figure 4.11: Enterprise Data Models

The Enterprise Data model is used to define and analyze data sharing, data integration and overall data requirements and to provide the foundation for designing logical and physical data structures that support these requirements. Model creation is iterative and is developed and refined over time.

It will be used as the cornerstone to support data discovery activities, identify data needed for state business processes, locate data for data sharing and integration efforts, and will be leveraged to enable quicker data identification for development or purchase of solutions to business needs.

Design and Architecture

Since the state's primary driver at this point in time is cross-agency data sharing, the initial efforts of the program are targeted to support these mandates and data sharing initiatives. This section describes the functional components necessary to provide the state's vision of a robust, fully-optimized data sharing environment. The core infrastructure solutions (see Figure 5.1 below) proposed in this ambitious program include:

- SOA Infrastructure and Enterprise Service Bus that are standards-based and that have pre-built capabilities to expose the Master Data Management structure as a canonical object model.
- A "Shared Network Cloud" for application development, data sharing, integration, master data management, searching and retrieving. The cloud will provide virtual or physical access to any number of applications or data assets (e.g., databases, document storage, and registries).
- An "Identity as Service" unique identifier system to be hosted and mapped to Agency-specific IDs through Master Data Management (MDM) with Data Quality and Data Matching capabilities.
- A metadata repository to establish a central point for data identification, standardization, and sharing. This ensures consistent use of the data assets and data resources across the state, facilitates easy mapping of data between computer systems, and lowers the costs of migrating to new systems.
- Robust access and authentication system to protect privacy and ensure that access to data on a roles-based, need-to-know system.
- An Extract Transform Load (ETL) product that has the capability to reach out to various data sources where applicable (various databases, SOA capability, mainframes [via SOA connectors], flat files, etc.).

There will be strong governance and change management in place to ensure the delivery of expected results based on business requirements and goals.

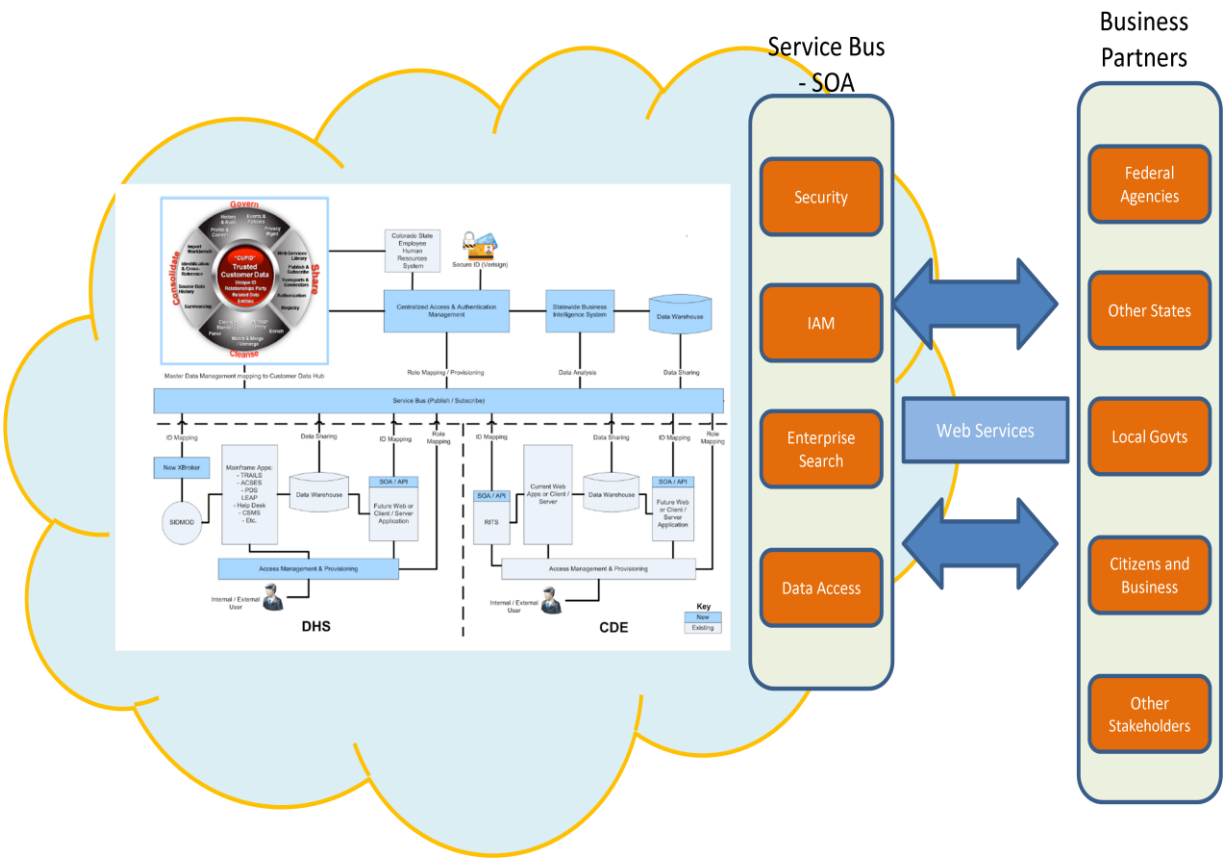


Figure 5.1: Data Services Architecture

Concept of Operations

This section explains each of the major architectural components in more detail. Responding rapidly to changes in competition, market dynamics, and regulatory mandates, with timely information, is critical for the effective functioning and overall success of the state. To meet rapidly changing market demands, many businesses have become increasingly service-driven, both in the ways they interact with customers and partners, and in how they design and build their IT infrastructure.

Service-Oriented Architecture (SOA) has emerged as the leading IT agenda for infrastructure reformation, to optimize service delivery and ensure efficient business process management. Part of the paradigm shift of SOA are fundamental changes in the way IT infrastructure is designed—moving away from an application infrastructure to a converged service infrastructure. Service-Oriented Architecture enables discrete functions contained in enterprise applications to be organized as layers of

interoperable, standards-based shared "services" that can be combined and reused in composite applications and processes.

In addition, this architectural approach also allows the incorporation of services offered by external service providers into the enterprise IT architecture. As a result, enterprises are able to unlock key business information in disparate silos, in a cost-effective manner. By organizing enterprise IT around services instead of around applications, SOA helps companies achieve faster time-to-service and respond more flexibly to fast-paced changes in business requirements.

The state is applying these architectural concepts and principals with its approach to cross-agency data sharing and integration.

Enterprise Service Bus

The core of SOA success depends on an Enterprise Service Bus (ESB) that supports dynamic synergy and alignment of business process interactions, continual evolution of existing services and rapid addition of new ones. To realize the benefits of SOA, it is imperative that IT organizations include a robust and intelligent service intermediary that provides a layer of abstraction to mask the complexities of service integration in heterogeneous IT environments, typical in today's enterprises. While an intermediary layer of abstraction previously implied a platform for customizing enterprise applications, today it implies toolkits for service customization and scalable infrastructures that support loosely coupled service interactions with a focus on service mediation.

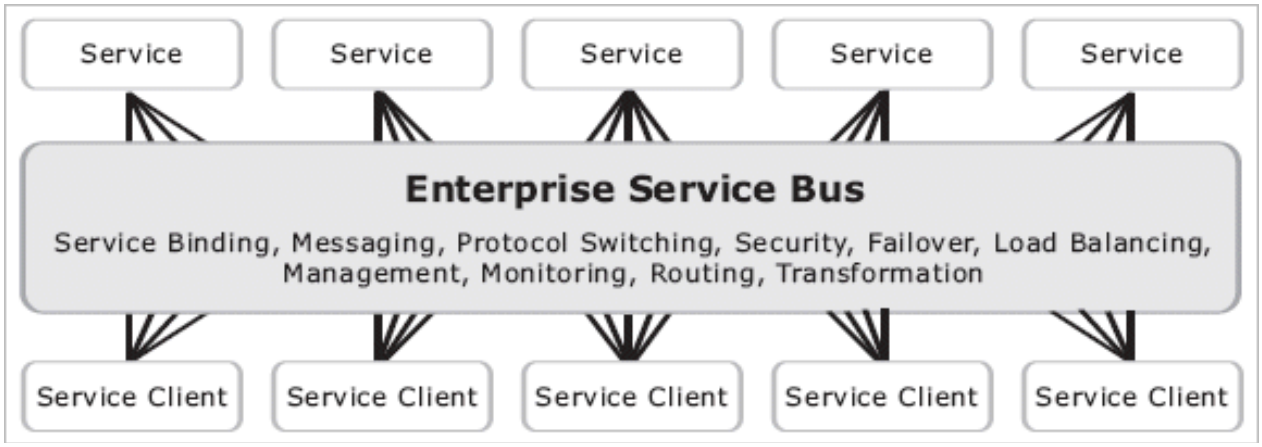


Figure 5.2: Enterprise Service Bus

ESBs have been instrumental in the evolution of integrated middleware infrastructure technology by combining features from previous technologies with new services, such as message validation, transformation, content-based routing, security and load balancing. ESBs use industry standards for

most of the services they provide, thus facilitating cross-platform interoperability and becoming the logical choice for companies looking to implement SOA.

An ESB provides an efficient way to build and deploy enterprise SOA. ESB is a concept that has gained the attention of architects and developers, as it provides an effective approach to solving common SOA hurdles associated with service orchestration, application data synchronization, and business activity monitoring. In its most basic form, an ESB offers the following key features:

- Web services: support for SOAP (simple object access protocol), WSDL (web services definition language) and UDDI (universal description, discovery and integration), as well as emerging standards such as WS-Reliable Messaging and WS-Security.
- Messaging: asynchronous store-and-forward delivery with multiple qualities of service.
- Data transformation: XML to XML.
- Content-based routing: publish and subscribe routing across multiple types of sources and destinations.
- Platform-neutral: connect to any technology in the enterprise, e.g. Java, .Net, mainframes, and databases.

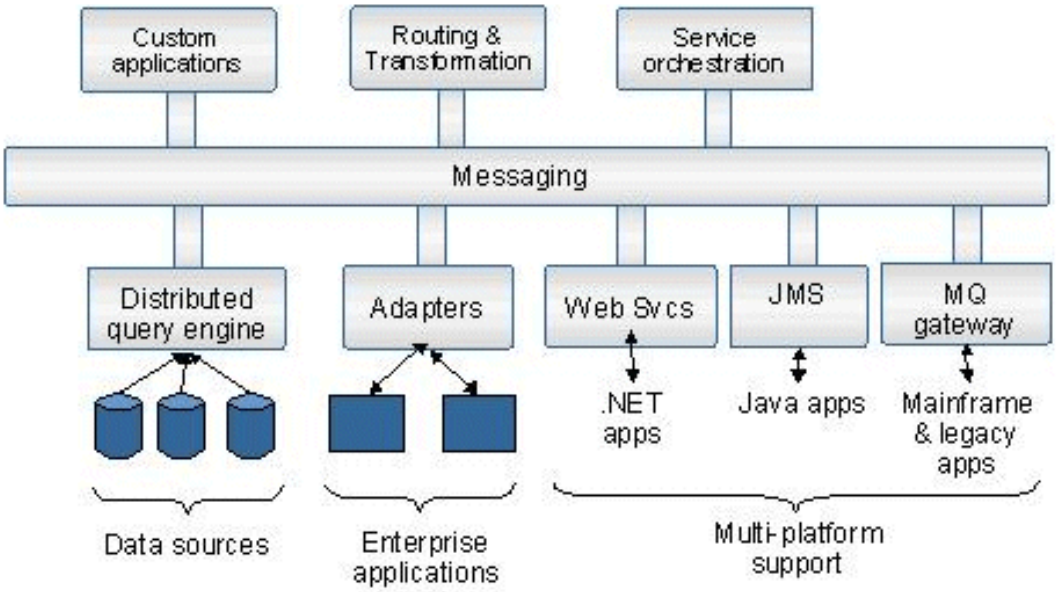


Figure 5.3: ESB Architecture

A robust SOA suite offers:

- Adapters, to enable connectivity into packaged and custom enterprise applications, as well as leading technologies.
- Distributed query engine, for easily enabling the creation of data services out of heterogeneous data sources.
- Service orchestration engine, for both long-running (stateful) and short-running (stateless) processes.
- Application development tools, to enable the rapid creation of user-facing applications.
- Presentation services, to enable the creation of personalized portals that aggregate services from multiple sources.

Using ESBs offers greater flexibility for enterprises to connect heterogeneous resources, by eliminating the need for brittle high-maintenance point-to-point connections. Adding an ESB intermediary between service consumers and service providers, shields them from the implementation details of underlying service end-point interfaces, reducing or eliminating the redevelopment and redeployment impacts at the service-consumer level.

Shared Network Cloud

According to IBM, cloud computing is a style of computing whose foundation is the delivery of services, software and processing capacity using private or public networks. The focus of cloud computing is the user experience, and the essence is to decouple the delivery of computing services from the underlying technology. Beyond the user interface, the technology behind the cloud remains invisible to the user, making cloud computing incredibly user-friendly. Cloud computing is an emerging approach to shared infrastructure in which large pools of systems are linked together in private or public networks to provide IT services. The need for such environments is fueled by dramatic growth in connected devices, real-time data streams and the adoption of service-oriented architectures and Web2.0 applications, such as mashups, open collaboration, social networking and mobile commerce.

Data and identity services will live in the cloud, which will be developed incrementally, again, vertical by vertical, or by priority cross-agency needs. Information regarding shared data will be published and available for users and applications when they need it.

Identity as Service

The state has just finished a study and has developed the architecture for a Colorado unique identifier. The concept for this "**Identity as Service**" strategy is to leverage a unique master ID for each citizen to

be able to integrate information from multiple state and local agencies. This will support data sharing and interoperability between agencies, build a governance framework ensuring quality and integrity of statewide identities, link identifiers between siloed or disparate systems, enable Colorado to track and examine performance of state services over time, promote Interstate data sharing efforts, protect privacy, and, improve access management and security.

The ability to have a master ID – Colorado Unique Personal Identifier (CUPID) - mapped to Agency specific IDs, serves as the basis for providing services across agencies, sharing data across agencies (state and local), and enables longitudinal analysis across all agencies and the entire lifecycle of a person (e.g., cradle to grave). Coupled with business intelligence reporting tools, the state can conduct analysis on the correlations between programs conducted by different agencies.

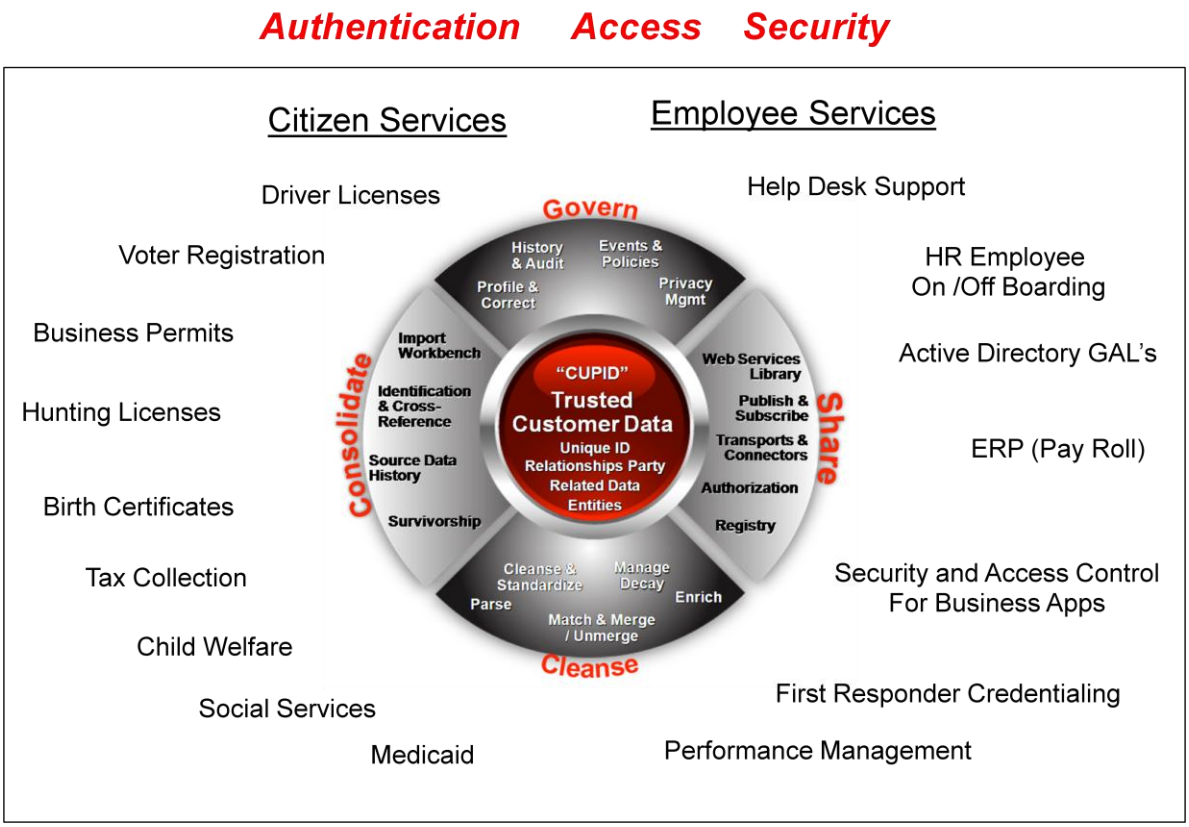


Figure 5.4: Colorado Identity As Service Model

With this concept, we recognize that a person only has one identity, but can have multiple attributes and privileges (e.g., driver, voter, receiver of benefits, employee, first responder, patient) assigned to her or him. The assigning of attributes assigned will remain with the agencies and programs that are serving individuals.

Finally, identity verification and fraud management tools will be provided through the CUPID services. These services can help agencies address such questions as:

- Is this identity a valid identity (e.g., is this Social Security Number valid? is this SSN/Date of Birth combination valid?)
- Does this identity belong to the person claiming the identity? (e.g., is someone else claiming that Social Security Number? Is this identity deceased?)
- Is the address a legitimate address? (e.g., is this a residential address or a business address? Does this address really exist?)

Again, attribute assignment and access control rights remain at the Agency level, but these centralized services will assist the agencies in provide the best, most trusted service delivery possible.

Metadata Repository

The purpose of the enterprise metadata repository is to establish a central point for data identification, standardization, and sharing. This ensures consistent use of the data assets and data resources across the state, facilitates easy mapping of data between computer systems, and lowers the costs of migrating to new systems, including service-oriented architecture. Data quality will improve because all agencies will use the same definitions and standards.

Users and applications tag data assets with metadata to enable discovery of data assets. Metadata information to be included with each data element will include at a minimum:

- Steward responsible for the data element
- Data domain and the data stewards responsible
- Data sub domain and the data stewards responsible
- Data semantic definition
- Data physical definition in existing system(s)
- Application systems, data warehouses, data marts where the data resides including completeness and timeliness
- Recognized authoritative reports that already contain the information and their location and ownership
- Data sharing initiatives that already exist for this data and their contacts
- Data exchanges that already exist for this data and their ownership
- Security information for that data, internal and external
- Business rules regarding data domain values and relationships and the integrity and validity in each system
- Security Classification

Access and Authentication Management

The second part of the "**Identity as Service**" solution recommendation provides a flexible statewide access and authentication management strategy. Leveraging the CUPID, the state can now offer automated role-based provisioning and de-provisioning and single-sign-on access for citizens and employees across the state. The system interacts with state human resources systems to ensure only the correct levels of access to Agency systems are available to employees, therefore enhancing security. Any future data sharing projects will take advantage of the unique identifier to tie Agency systems together and to provide strong authentication and access controls to protect privacy. The incorporation of the unique identifier will also ensure that the agencies are accurately accessing the correct information about its citizens, without compromising or sharing Personally Identifiable Information (PII), thus further protecting the citizens' privacy. The Federal Identity, Credential, and Access Management (FICAM) framework will be used as guidance for interoperability with our federal and state partners.

Additionally, this program will serve as the core used to create trusted digital identity representations of those individuals, bind those identities to credentials that may serve as a proxy for the individual in access transactions, and leverage the credentials to provide authorized access to an Agency's resources. Imagine as a citizen being able to log-in to a state portal to get the most current information regarding taxes paid, vehicles registered or recreational permits issued; reserve a camp site, request a copy of your birth certificate for a passport application or fill out a change of address that immediately updates all relevant state systems; or, electronically complete and sign all paperwork for a new business license. All of this will be possible with the new infrastructure.

ETL Services

Several commercially available extract, transform, and load (ETL) tools can be used to share data between disparate systems where source system data is first gathered and transferred to a processing system and later transformed and loaded into the destination systems. The tools can be used to program data transformations and facilitate loading the source data into a central data store. After data has been loaded into the central processing systems, it is available for the consuming systems. Administrative capabilities of the tools enable administrators to monitor system processes, identify problems and manually back out data if needed. ETL services are primarily needed with high volumes of data.

Security

Strong, consistent cyber security policies throughout all state Agencies are an important enabler for data sharing, particularly when sharing sensitive or confidential data such as unit data that identifies a particular individual. When all systems meet the same basic standards for security and privacy, agencies can trust each other to handle sensitive data appropriately. When agencies do not have a common approach to cyber security, security must become an important factor in every data sharing agreement.

In the state, the Office of Cyber Security (OCS) is responsible for security risk management across state departments. The DM program will continue to follow the guidance and technical standards issued by OCS for all initiatives.

Privacy

Privacy continues to be a top priority for the state in managing its systems and data. Privacy laws continue to be dynamic, and OIT works closely with the state's Chief Privacy Officer in the Office of the Attorney General on all privacy-related issues. When implemented correctly, technology can actually be an enabler of strong privacy practices.

A fundamental tenet of both data sharing and privacy protection is ensuring that the data reported about an individual as part of a data or information sharing request is actually data about the individual being requested. The unique ID is designed to map Agency identities to a master identifier that incorporates data cleansing and data quality functionality to guarantee an individual's identity against all other potential matches in the state. The unique ID is strictly architected to use the minimal set of attributes necessary to make de-duplication and uniqueness determinations on identities. Sensitive Agency data (everything from HIPAA-sensitive to educational and individual benefits data) is retained at the Agency level in Agency repositories. Additionally, the state will be able to provide de-identified record-level data to research organizations for evaluations of policies and programs within the state.

All state systems are protected behind strong firewalls, and the state will implement a strong access and authentication and authorization model to provide security and protect privacy by ensuring that access to data is on a need-to-know system. The provisioning system will leverage the unique ID directly, and use a roles-based method to control access to state systems. The provisioning systems will offer multiple authentication schemes depending on the level of protection required by the system and data.

The model is flexible and can accommodate low (username/password) to high (certificate-based) assurance levels. A provisioning tool will assign the corresponding authentication and authorization capabilities to users according to their role and the value of the information that is being protected. This process will address compliance with all federal and state regulatory requirements.

Data Set Consolidation

One of the goals of the state is to improve service delivery and efficiency while reducing operational costs. As OIT moves forward with DM efforts, opportunities will arise where it makes sense to centralize or consolidate data sets. Since most agencies utilize geospatial, identity and location/address data, these sets are likely candidates for centralized repositories that all agencies would then utilize for lines-of-business service delivery, as shown in Figure 5.5 below.

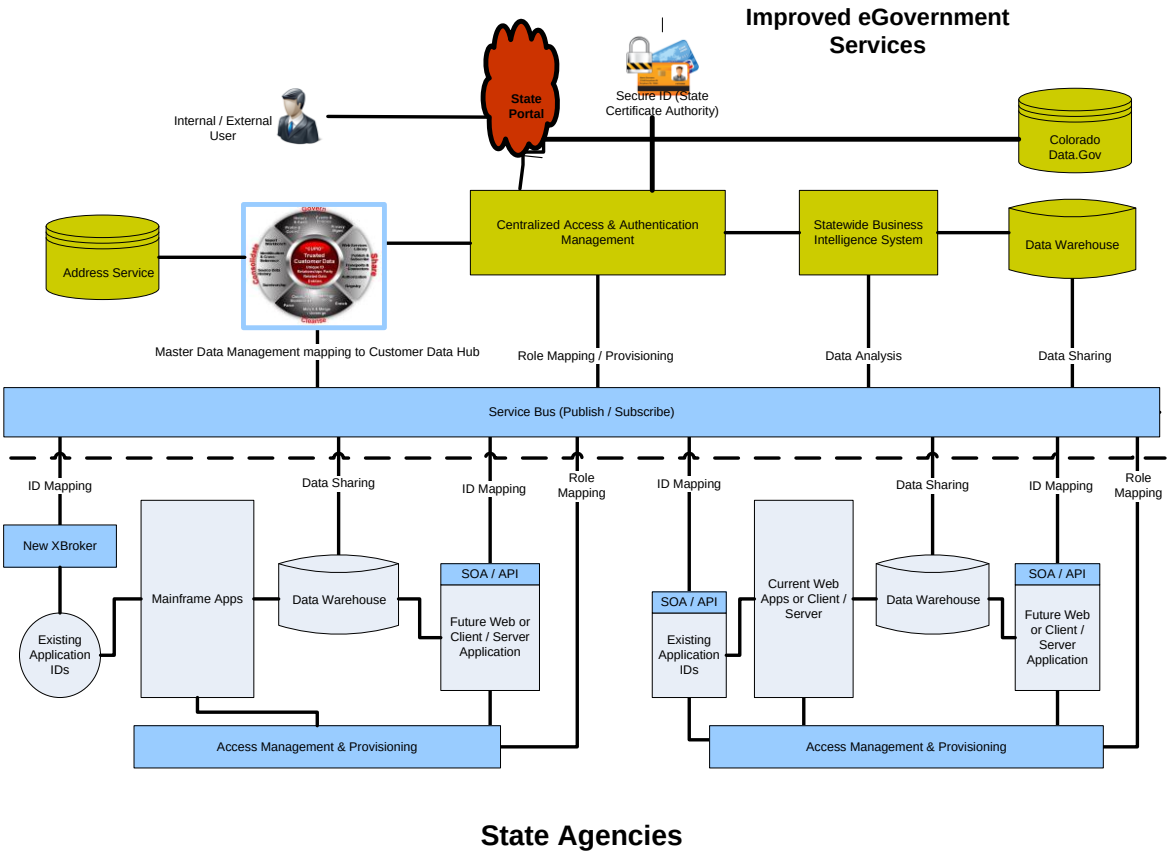


Figure 5.5: Improved eGovernment Service Delivery

Consolidating these data sets would also enable the state to expand its portfolio of eGovernment services offerings, both citizen-to-government (C2G) and business-to-government (B2G). Strong confidence in identity coupled with strong multi-factor authentication offerings through centralized services will enable the state to facilitate transactions online for citizens and employees to provide improved ease-of-use and improved customer experience. This includes basic services such as a name/address change service that populates all relevant systems, instead of separate changes for each Agency with whom a person is affiliated and a single view of all state services that a person is receiving.

As the Business Value Data Matrix continues to develop and be refined over the years, more data sets will be identified as candidate for centralization or consolidation. The multiple data warehouses across the state could also be consolidated. These are longer term, future efforts for the state.

Both of these efforts – centralization and consolidation – would improve the quality of the data and enable better analysis and knowledge management, since there would be more “single versions of the truth” and a reduction of inaccurate data scattered across multiple repositories.

The state will continue to investigate these options as we continue to build and develop the program, and as funding and other opportunities present themselves to take advantage of these options. It is important to note that there are regulatory and liability issues that must be considered prior to any of these efforts commencing.

Implementation Roadmap and Milestones

Moving the state to an enhanced data management future based on shared integrated data will improve the speed and quality of decision making and delivery of services to the state's constituents while reducing the cost associated with non-integrated systems. Figure 6.1 below illustrates the high-level business and technical overviews of the architectural components of the state's intended system.

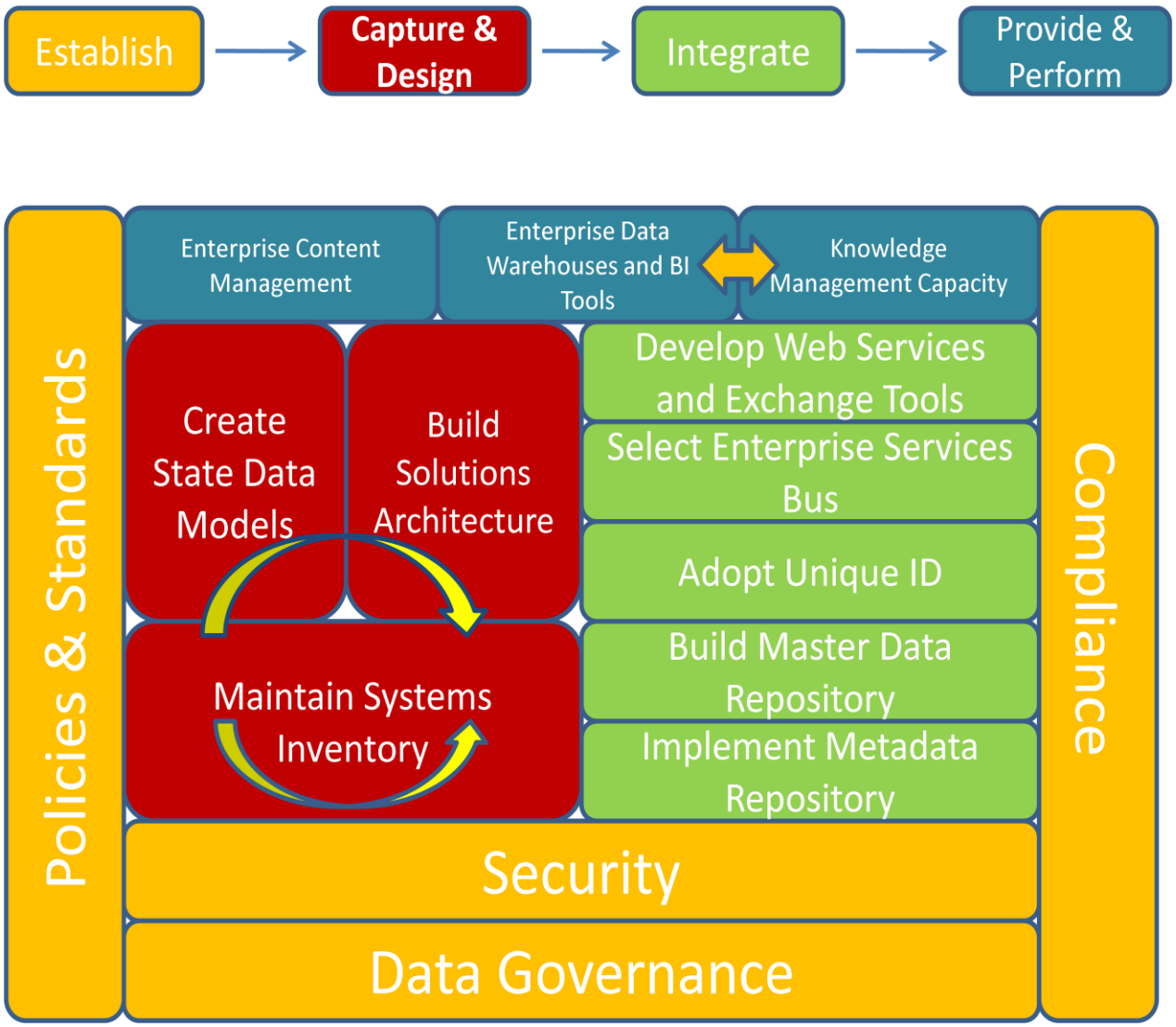


Figure 6.1: DM Program Strategic Articulation

There are four major categories of strategic objectives: *Establish*, *Capture and Design*, *Integrate*, *Provide and Perform*. They are color-coded and map to the initiatives within each. The *Establish* piece consists of all the governance activities necessary to creating the data management program, including policies, standards, and compliance checks. The *Capture and Design* activities will create a baseline of all state systems and data, as well as provide the enterprise models to be used in data sharing and integration efforts. Next, *Integrate* is building the architectural pieces to achieve cross-agency data sharing. Finally, *Provide and Perform* is delivering access to the data and people using information to use to make more informed decisions. Within each one of these major objectives will be several activities to support the successful outcome of that objective.

Colorado's Vertical Domain Approach to Integrated Data Sharing

The state will perform the work incrementally in stages, with a vertical business domain strategy. This work has begun across a number of business domains, including public safety, children's services, and education. Some examples are below.

All initiatives will follow OIT processes for capturing system and enterprise information, identifying data needs for business process purposes, and developing the technical architecture models. Legal, privacy and compliance issues will be worked through as part of the governance models. Data management policies and standards will be created through each phase to support the initiatives and the broader enterprise data management and sharing needs. Additional business segments (transportation, public safety, health, etc.) will be added based on priority and resource availability.

Education State Longitudinal Data System

The Colorado Department of Education (CDE), working in collaboration with OIT, the Governor's Office of Policy and Initiatives, and the Colorado Departments of Higher Education, Human Services and Labor and Employment, proposes to build a state longitudinal data system, Project SchoolView™, that meets the demands of the state's ground-breaking P-20 education reform agenda.

This agenda, which has been underway for several years, implements true P-20 education alignment across the state's education systems and is anchored by a common definition of postsecondary and workforce readiness to ensure students exit prepared for postsecondary education and workforce success. Project SchoolView™ envisions a flexible enterprise P-20 information and knowledge management system that will equip users to manage and use information for informed decision making ensuring postsecondary or workforce success.

Colorado Children Youth Information Sharing Initiative

This effort seeks to support the current Office of Juvenile Justice and Delinquency Prevention (OJJDP) effort in Colorado on juvenile justice data exchange. This effort is a multi-agency initiative that is

addressing the coordination, collaboration and integration of children and youth prevention, intervention, and treatment services. Eight state departments are represented, including Public Health and Environment, Education, Human Services, Division of Youth Corrections, Public Safety, Judicial, OIT, and Health Care Policy and Finance. Local partners are also involved to discuss common issues related to coordination, collaboration and integration as related to services for children, youth and families.

A major priority for this initiative is to share data between agencies and systems and to enhance long-range, integrated and comprehensive planning for children and youth around common priorities at the state and local levels. This effort has been underway for just under a year, and they are assisted in their efforts by the Center for Network Development (CND). The CND has recently completed the first version of the juvenile justice XML data model that is conformant with the National Information Exchange Model (NIEM). CND is in the process of testing and implementing key data exchanges with local and state agencies.

Early Childhood Service Delivery

Through both legislation and an Executive Order, the state is beginning an initiative that will ensure and advance a comprehensive service delivery system for children birth to eight using data to improve decision making, alignment, and coordination among federally funded and state funded supports and services targeted at young children and their families. A unified, interagency data system to promote sharing and use of common data for planning and accountability will be developed across the four service sector domains of health, mental health, early learning, and family support and parent education. Members of this multi-agency effort includes the Departments of Human Services, Public Health and Environment, Health Care Policy and Financing, Higher Education, OIT, local governmental groups, local school districts, providers of early childhood support services, non-profits, and the business community.

Risk Management

The following issues will be considered by OIT as this work progresses:

- Address privacy and security concerns of citizens
- Cultural and change management issues within state agencies
- Determine funding sources to implement the proposed enterprise program and system infrastructure
- Meet compliance standards set by federal and state statute and regulation
- Ensure that recommended statutory or regulatory changes can be met in a timely manner

Establish: Create Data Governance and Management Program

Objective: Establish sustainable data management and governance program.

Deliverable 1: Data Governance Program

Task 1.1 – Create a data management organization.

Action item 1.1.1 – Build DM team, to include architects, business analysts, and a project manager.

Action item 1.1.2 – Implement governance process with GDAB.

Task 1.2 – Create a Data Stewards Action Council to engage all agencies and formalize data stewardship activities and processes.

Action item 1.2.1 - Create a common baseline of information: a statewide foundation for data sharing, information discovery, future architectures.

Action item 1.2.2 – In FY 2010, the state will identify data stewards for statewide data elements to promote the accurate and reliable collection of data by a single source and the exchange of such data with approved business partners.

Action item 1.2.3 – The state shall encourage and support the formalization of interagency communities of interest (COI) that have as their goal coordinating, exchanging, and improving the quality of information and services available to citizens and constituents. COIs shall include representation from all impacted state agencies, and, where possible, local and federal government agencies.

Task 1.3 - Create an information sharing culture.

Action item 1.3.1 – Deliver ongoing communication to all stakeholders.

Action item 1.3.2 - Establish venues in which the experience and expertise of individuals is shared and leveraged to facilitate interagency collaboration and partnership.

Action item 1.3.3 – Implement a shared information repository to facilitate information sharing.

Action item 1.3.4 – Leverage the Colorado Information Managers Association to encourage the sharing of knowledge, experience and solutions.

Action item 1.3.3 – Develop processes to engage state Agency business leadership in IT strategic planning activities to ensure that OIT is meeting business requirements.

Deliverable 2: Policies and Standards

Task 2.1 – Develop enterprise data policies and standards.

Action item 2.1.1 – Identify universe of policies and standards to be developed, and assign to appropriate team member(s).

Action item 2.1.2 – Develop document library for internal and external publication of policies and standards.

Action item 2.1.3 – Continually identifies, review and refine policies.

Deliverable 3: Data Security Classification Policy

Task 3.1 – Develop an enterprise data security classification policy, in conjunction with the Chief Information Security Officer.

Action item 3.1.1 – Publish policy.

Task 3.2 – Classify all enterprise data in accordance with the data security classification policy.

Action item 3.2.1 – Purchase tool to assist with classification.

Task 3.3 – Develop an enterprise privacy policy, in conjunction with the Chief Privacy Officer.

Action item 3.3.1 – Publish policy.

Task 3.4 – Implement access and authentication technology and processes ensure the privacy and security of data.

Action item 3.4.1 – Set up provisioning/deprovisioning for various systems within an Agency.

Action item 3.4.2 – Create enterprise and agency specific roles for provisioning/deprovisioning of resources.

Action item 3.4.3 – Maintain associations between users and their physical assets.

Deliverable 4: Compliance Program

Task 4.1 - Develop enterprise compliance policies and standards.

Task 4.2 – Work with the Colorado Architectural Review Board to implement policies and standards.

Capture and Design

Objective: All state data sets are known and enterprise data standards are implemented.

Deliverable 5: Existing Systems Inventory

Task 5.1 – Create an inventory of all information systems in the state to develop a baseline of all state data systems identified with the business function they support.

Action item 5.1.1 – Complete 'as-is' enterprise information architecture at the system level.

Task 5.2 – Create an inventory of all significant data in the state to have a baseline of all state data assets

Action item 5.2.1 – Complete 'as-is' enterprise information architecture at the data element level.

Deliverable 6: Create Enterprise Data Models

Task 6.1 – Establish and publish statewide data standards and definitions for all data elements

Task 6.2 – Create enterprise conceptual and logical models.

Action item 6.2.1 – Purchase tool to contain information architecture.

Action item 6.2.2 -- Validate Subject Area Model with agencies via data stewards (DSAC).

Action Item 6.2.3 – Create Conceptual View of Enterprise Data Model with major data entities.

Action item 6.2.4 -- Create Logical View of Enterprise Data Model with major data entity keys and attributes.

Task 6.3 - Data governance policies, processes and standards are established to manage the flow of data from capture to use.

Task 6.4 – Identify authoritative data sources for all data types.

Task 6.5 - Create and implement an enterprise data dictionary and taxonomy.

Action Item 6.5.1 – Cross Reference data in systems with NIEM naming and model.

Deliverable 7: Create Solution Architectures

Task 7.1 – Develop solutions that serve multiagency business needs to facilitate organizational collaboration and partnership.

Action Item 7.1.1 – Identify data modelers for the state.

Action item 7.1.2 – Identify projects with data sharing opportunities.

Integrate: Cross-Agency Data Sharing and Interoperability

Objective: Stakeholders are provided with data of the highest quality, reliability and integrity in a timely manner to promote trust in the system and use of the system.

Deliverable 8: Implement Metadata Repository

Task 8.1 – Develop a standard set of Metadata components required for state data elements.

Action Item 8.1.1 Create metadata component standard.

Action Item 8.1.2 Collect and document the metadata components for data elements.

Task 8.2 – Utilize metadata repository for new development efforts.

Action Item 8.2.1 Create metadata access methods and procedures.

Deliverable 9: Build Master Data Repositories

Task 9.1 - Master Data Management technology is implemented to ensure the quality, reliability and integrity of the data.

Action Item 9.1.1 Unique ID is implemented with a Master Data strategy and technology.

Action Item 9.1.2 Identify state master data target areas.

Task 9.2 - Data stewards provide data quality audits as part of the ongoing monitoring of data quality on a daily basis via a user interface in the Master Data Management application.

Action Item 9.2 Develop procedures and methods for data quality monitoring and evaluation.

Deliverable 10: Implement Unique ID

Task 10.1 – Finalize and publish CUPID architecture and cross-agency requirements.

Task 10.2 – Write and publish solicitation.

Task 10.3 – Implement centralized CUPID architecture and application.

Task 10.4 – Begin implementation of CUPID interfaces at initial state Agencies.

Deliverable 11: Select Enterprise Services Bus

Task 11.1 – Define the secure technical data exchange architectural requirements.

Task 11.2 – Write and publish solicitation.

Task 11.3 – Begin implementation.

Deliverable 12: Develop Web Services Exchange Tools

Task 12.1 – Enterprise data exchange standards are adopted.

Action item 12.1.1 – NIEM is adopted as the data exchange standard.

Action item 12.1 2 – NIEM training is provide to all agencies.

Task 12.2 – Tools are provided to allow people with the right authorizations to easily access the data.

Task 12.3 – Leverage data, tools and infrastructure across the enterprise.

Task 12.4 - Train internal and external stakeholders/users to use the system.

Deliverable 13: Data Assets Are Consolidated and Centralized Where Possible

Task 13.1 – The state will identify opportunities and priorities to consolidate and exchange core data set with state, local, federal and business partners to further increase efficiencies, analytical and predictive capabilities, improve policy-making and citizen services.

Action Item 13.2 Potential data for sharing is identified across the agencies in the metadata repository.

Task 13.2 - Provide for open source application development to drive innovation in data visualization and encourage sharing of both information and technology among all interested education stakeholders.

Provide and Perform: Data as Utility and Knowledge Management

Objective: Stakeholders effectively use information to guide development, policy, programs and practice.

Deliverable 14: Enterprise Content Management

Task 14.1 – Determine requirements for an enterprise content management system to manage all unstructured data assets.

Task 14.2 – Implement enterprise content management system.

Deliverable 15: Enterprise Data Warehouses and Business Intelligence

Task 15.2 -- Develop statewide standards for data warehouse platforms, including open source software, security, and data storage/data base management standards.

Action Item 15.2.1 Develop and publish standards.

Task 15.3 – Develop statewide standards for business intelligence tools, including reporting, data mining, and web-based platforms.

Action Item 15.3.1 Develop and publish standards.

Task 15.4 -- Develop statewide open and flexible logical and physical architectural standards for data warehouse and data mart creation.

Action Item 15.4.1 – Identify open and flexible alternatives and target platforms for future use.

Task 15.5 – Determine opportunities for consolidating data for analysis and reporting efforts or replacing closed or high-cost systems with open or lower-maintenance systems.

Action Item 15.5.1 – Inventory current data warehouse and data mart environments.

Action Item 15.5.2 – Identify systems to be replaced during normal development or expansion needs.

Deliverable 16: Knowledge Management Capacity

Task 16.1 - Provide ongoing and cost-effective technical assistance (user support) to internal and external stakeholders in modifying their systems to meet new reporting and interoperability requirements.

Action Item 16.1.1 – Partner with the agencies to assist with information and systems development efforts, including new and updated system and application design projects, data

sharing programs and mandates, data identification efforts, data reporting efforts, metadata definition projects, and federal and local data reporting efforts.

Additional tasks and activities will be identified and recorded as each phase's requirements are fully developed.

Timeline

Figure 6.2 below depicts the high-level project plan for completing the previously described deliverables.

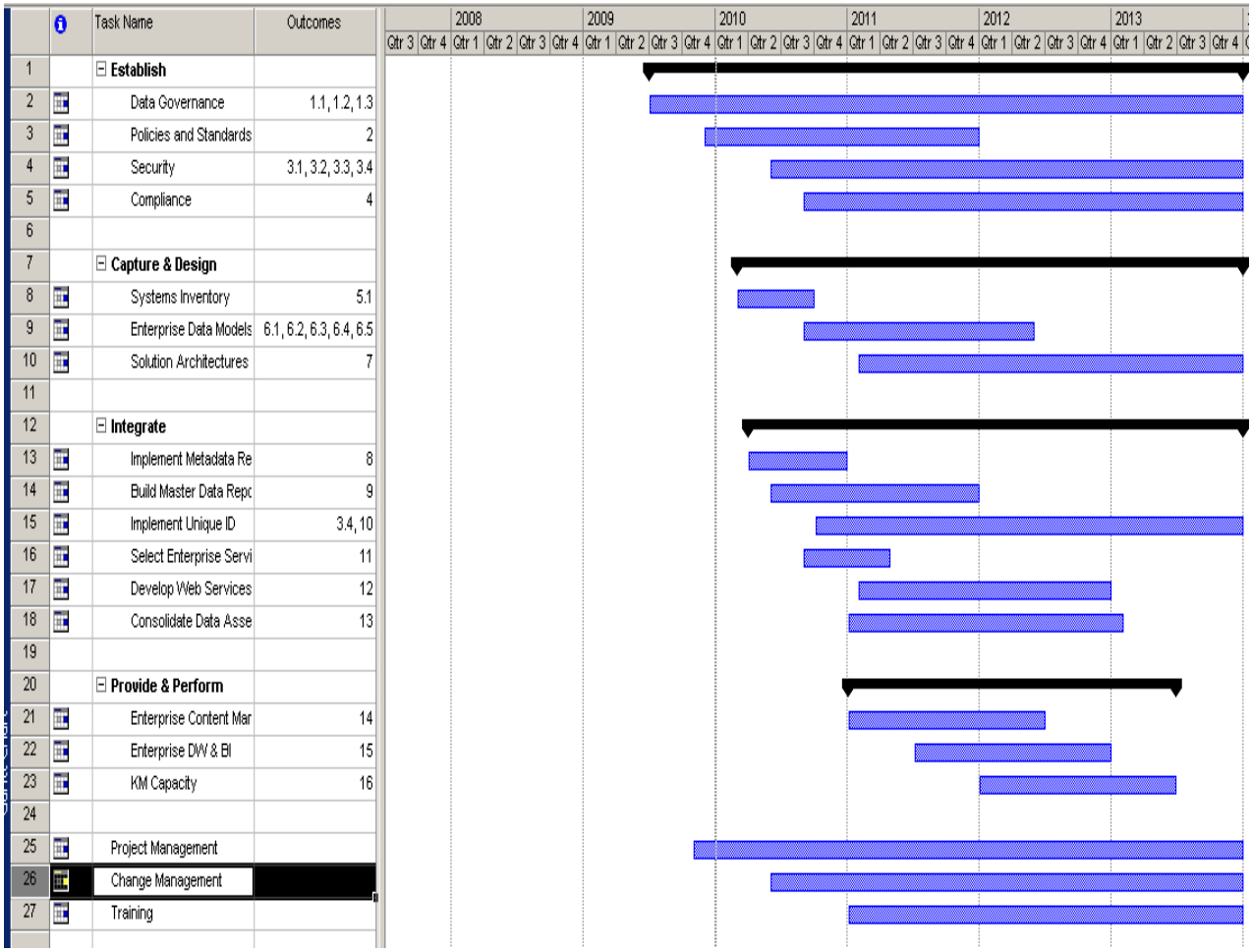


Figure 6.2: Timeline

Performance Metrics

The CDO, in conjunction with the GDAB, DSAC and OEA, will develop performance metrics to measure and monitor state and Agency performance and compliance in each of the goal and objective areas. These metrics will be published via a dashboard format and made publicly available through OEA's website.

Appendices

Appendix A – Glossary of Terms and Acronyms

This glossary of terms and acronyms is intended to serve as a communication vehicle for reading and understanding publications produced from the Data Management Program.

ASCII – Acronym for the American Standard Code for Information Interchange, which is a code for information exchange between computers.

AUP - Acronym for Acceptable Use Policy, which is a set of regulations that govern how a service may be used.

Authentication - A process for verifying that a person or computer is who they say they are.

Business Data Steward - A recognized subject matter experts working with data management professionals on an ongoing basis to define and control data. They will be more simply referred to as the data stewards.

Business Domains - Business domains are the natural divisions of the business architecture and are based on either functional or topical scope. Business domains represent the highest level of the state's business architecture blueprint.

Business Reference Model – The Business Reference Model (BRM) provides a framework facilitating a functional (rather than organizational) view of the federal government's lines of business (LoBs), including its internal operations and its services for citizens, independent of the agencies, bureaus and offices performing them. The BRM describes the federal government around common business areas instead of through a stovepiped, agency-by-agency view. It thus promotes agency collaboration and serves as the underlying foundation for the FEA and E-Gov strategies.

CIO – Acronym for Chief Information Officer.

CISO – Acronym for Chief Information Security Officer.

CMP-SSC - Acronym for the Collaborative Management Program State Steering Committee.

Conceptual Model - A layer of modeling that defines business entities and the relationships between these business entities. Business entities are the concepts and classes of things, people, and places that are familiar and of interest to the State.

Consolidated Reference Model - The FEA Consolidated Reference Model Document Version 2. Published in October 0f 2007, contains four of the five models (Performance Reference Model (PRM), Business Reference Model (BRM), Service Component Reference Model (SRM), Technical Reference Model (TRM), that make up the Federal Enterprise Architecture. The Data Reference Model, DRM, is referenced but not repeated in this document due to its complexity and volume. Abbreviated as CRM.

Coordinating Data Steward - The data steward responsible for coordination of data stewardship activities across an information subject area. This person is responsible for insuring the integrity,

quality, security, and coordination of associated metadata across the subject area and will lead a data stewardship team.

COPPA - Acronym for the Children's Online Privacy Protection Act.

COTS - Acronym for Commercial Off-The-Shelf software.

CPO - Acronym for Chief Privacy Officer.

CRM – See "Consolidated Reference Model".

Cyber Security – A branch of security dealing with digital or information technology.

Data – Data is the representation of facts as text, numbers, graphics, images, sounds or video. Facts are captured, stored, and expressed as data.

Data Context – Data context refers to any information that provides additional meaning to data. Data context typically specifies a designation or description of the application environment or discipline in which data is applied or from which it originates. It provides perspective, significance, and connotation to data, and is vital to the discovery, use and comprehension of data.

Data Dictionary - As defined in the *IBM Dictionary of Computing*, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format."[

Data Element - A precise and concise phrase or sentence associated with a data element within a data dictionary (or metadata registry) that describes the meaning or semantics of a data element.

Data Governance - Data governance refers to the operating discipline for managing data and information as a key enterprise asset.

Data Management - Data management is the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets.

Data Mining - The process of extracting hidden patterns from data. Data mining identifies trends within data that go beyond simple data analysis. Through the use of sophisticated algorithms, non-statistician users have the opportunity to identify key attributes of processes and target opportunities.

Data Modeling – A structured method for representing and describing the data used in an automated system. Data modeling is often used in combination with two other structured methods, data flow analysis and functional decomposition, to define the high-level structure of business and information systems.

Data Reference Model - The Data Reference Model (DRM) is a flexible and standards-based framework to enable information sharing and reuse across the federal government via the standard description and discovery of common data and the promotion of uniform data management practices. The DRM provides a standard means by which data may be described, categorized, and shared. These are

reflected within each of the DRM's three standardization areas of data description, data context, and data sharing.

Data stewardship - The formal accountability for state business responsibilities through ensuring effective definition, coordination, control and use of data assets.

Data Stewardship Teams - One or more temporary or permanent focused groups of business data stewards collaborating on data modeling, data definitions, data quality requirement specification, and data quality improvement, reference and master data management, and meta-data management, typically within an assigned subject area, lead by a coordinating data steward in partnership with a data architect.

Data Warehouse – A central repository for significant parts of the data that an enterprise's various business systems collect specifically designed for reporting. It is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management's decision making process, specifically providing data for Online Analytical Processing (OLAP) efforts.

DBA - Acronym for database administrator.

DQA - Acronym for Data Quality Assurance, which is a process of examining the data to discover inconsistencies and other anomalies. Data cleansing activities may be performed to improve the data quality.

EDE - Acronym for Electronic Data Exchange.

Enterprise - The State of Colorado Executive Branch Agencies.

ESID - Acronym for the encrypted state ID at the Colorado Dept. of Education.

ETL – Extract, Transform, and Load, which is a process to extract data from one source, transform (or cleanse) it, and load the result into another source. This is frequently part of populating a Data Warehouse.

Extensible Markup Language - Extensible Markup Language (XML) describes a class of data objects called XML documents and partially describes the behavior of computer programs which process them. XML is a subset of SGML, the Standard Generalized Markup Language. Among its uses XML is intended to meet the requirements of vendor-neutral data exchange, the processing of Web documents by intelligent clients, and certain metadata applications. XML is fully internationalized and is designed for the quickest possible client-side processing consistent with its primary purpose as an electronic publishing and data interchange format.

Federal Enterprise Architecture - The Federal Enterprise Architecture (FEA) consists 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 FEA in a common and consistent way. Through the use of this common framework and vocabulary, IT portfolios can be better managed and leveraged across the federal government.

FERPA – Acronym for the Family Educational Rights and Privacy Act.

Government Data Advisory Board (GDAB) – Advisory Board created by HB 09-1285 for the purpose of advising the State CIO on matters relating to data sharing.

HIPAA - Acronym for the Health Insurance Portability and Accountability Act.

Identity Management - Identity Management (IdM) means the combination of technical systems, rules, and procedures that define the owner-ship, utilization, and safeguarding of personal identity information. The primary goal of the IdM process is to assign attributes to a digital identity and to connect that identity to an individual.

Information – Information is data in context, used to guide decisions.

Information Architecture - The compilation of the business requirements of the enterprise, the information, process entities and integration that drive the business, and rules for selecting, building and maintaining that information.

Information Exchange Package Documentation - An Information Exchange Package Documentation (IEPD), is a specification for a data exchange and defines a particular data exchange. It is a set of artifacts consisting of normative exchange specifications, examples, metadata, and documentation encapsulated by a catalog that describes each artifact. The entire package is archived as a single compressed file.

Information Subject Area - Topical or functional categories of the business processes that are integral to the operations of the State and that span agencies statewide, such as Financial, Person, Geography, Organization, and Service.

Information Subject Sub-Area - A logical subset of an information subject area containing enough unique information to be addressed separately, such as within the subject area of person could be Customer (client/citizen) or Employee.

Knowledge – Knowledge is understanding, awareness, cognizance, and the recognition of a situation and familiarity with its complexity. Knowledge is information in perspective, integrated into a viewpoint based on the recognition and interpretation of patterns, such as trends, formed with other information and experience.

Knowledge Management – Knowledge management is the discipline that fosters organizational learning and the management of intellectual capital as an enterprise resource.

Logical Model - the logical data model diagrams add a level of detail for each subject area below the conceptual data model by depicting the essential data attributes for each entity. The enterprise logical data model identifies the data needed about each instance of a business entity. The essential data attributes included represent common data requirement and standardized definitions for shared data attributes.

Master Data – Data that is, for the most part, static, and changes infrequently.

Metadata – Metadata is data about data. An example is a library catalog because it describes publications. In this document, it is usually applied to databases.

Metadata registry – A metadata registry/repository is a central location in an organization where metadata definitions are stored and maintained in a controlled method. Included in the registry are approved enterprise data definitions, representations (models, XML structures), and links to physical constructs, values, exceptions, and data steward information.

Metadata – Metadata is "data about data." Metadata includes data associated with either an information system or an information object, for purposes of description, administration, legal and confidentiality requirements, technical functionality and security, use and usage, and preservation. Metadata gives us detail about both what the data means and how it's stated. Metadata is one of the greatest critical success factors to sharing information because it provides business users, developers and data administrators with consistent descriptions of the enterprise's information assets.

National Information Exchange Model - The National Information Exchange Model (NIEM) is a Federal, State, Local and Tribal interagency initiative providing a foundation for seamless information exchange. NIEM is a framework to bring stakeholders and Communities of Interest together to identify information sharing requirements, develop standards, a common lexicon and an on-line repository of information exchange package documents to support information sharing, provide technical tools to support development, discovery, dissemination and re-use of exchange documents; and provide training, technical assistance and implementation support services for enterprise-wide information exchange.

OMB – Acronym for the Federal Office of Management and Budget.

Online Analytical Processing - Online Analytical Processing (OLAP) is a reporting and data design approach intended to quickly answer analytical queries. Data to satisfy OLAP reporting and analysis needs are designed differently than data used for traditional operational use. Although OLAP can be achieved with standard relational databases, multidimensional data models are often used, allowing for complex analytical and ad-hoc queries with a rapid execution time.

Online Transaction Processing - Online Transaction Processing (OLTP) is a class of systems that facilitate and manage transaction-oriented applications, typically for data entry and retrieval.

P-20 - Education from pre-kindergarten through post-graduate college.

Performance Reference Model – Acronym PRM, is part of the FEA.

Personally Identifiable Information (PII) – PII refers to all information associated with an individual and includes both identifying and non-identifying information. Examples of identifying information which can be used to locate or identify an individual include an individual's name, aliases, Social Security Number, email address, driver's license number, and agency-assigned unique identifier. Non-identifying personal information includes an individual's age, education, finances, criminal history, physical attributes, and gender.

PLC – Acronym for the Prevention Leadership Council.

Repository - An information system used to store and access architectural information, relationships among the information elements, and work products.

SASID - Acronym for the State Assigned Student ID at the Colorado Dept. of Education.

SCRM – Acronym for the Service Component Reference Model; part of the FEA.

SIDMOD – Acronym for the State Identification Module at the Colorado Dept. of Human Services.

SIMU – Acronym for the Student Identifier Management Unit at the Colorado Dept. of Education.

State Enterprise Data Model - An integrated, subject oriented data model defining the essential data produced and consumed across the state. The purpose of a data model is to 1) facilitate communications as a bridge to understand data between people with different levels and type of experience and help us understand the business area 2) to formally document a single and precise definition of data and data related rules, and 3) to help explain the data context and scope of third-party software. The data model is composed of three layers for communication and best utilization: The subject area model, the conceptual model, and the logical model.

Structured Data – Structured data is data stored in structured formats such as databases, flat files, and tagged electronic documents.

Technical Data Steward - The information systems professional responsible for assuring integrity of the information captured, for proper handling of the information (not the content), and for assuring the information is available when needed. They are the custodians of the data assets and perform technical functions to safeguard and enable effective use of State data assets.

Transaction Data - Transaction data is data describing an event (the change as a result of a transaction) and is usually described with verbs. Transaction data always has a time dimension, a numerical value and refers to one or more objects (i.e. the reference data). Typical transactions are: financial: orders, invoices, payments; work: plans, activity records; logistics: deliveries, storage records, travel records, etc.

Unit Records - Records containing data that pertain directly to an individual.

Unstructured Data – Unstructured data is any document, file, graphic, image, text, report, form, video, or sound recording that has not been tagged or otherwise structured into rows and columns or records. Unstructured data is found in different kinds of electronic formats, including word processing documents, electronic mail, flat files, spreadsheets, XML files, transactional messages, report, business graphics, digital images, microfiche, video records, and audio recordings. Unstructured data also includes paper files. Typically, 80% of an organization's data assets reside in relatively unstructured formats.

XML – See Extensible Markup Language.

State Agency Acronyms

Colorado Bureau of Investigations (CBI)

Colorado Children and Youth Information Sharing (CCYIS)

Colorado Data Sharing and Utilization Group (CDSUG)

Colorado District Attorneys Council (CDAC)

Colorado Integrated Criminal Justice Information System (CICJIS)

Data Governance Working Group (DGWG)

Department of Agriculture (CDA)

Department of Corrections (DOC)

Department of Education (CDE)

Department of Health Care Policy and Finance (HCPF)

Department of Higher Education (DHE)

Department of Human Services (DHS)

Department of Labor and Employment (CDLE)

Department of Law (DOL)

Department of Local Affairs (DOLA)

Department of Natural Resources (DNR)

Department of Personnel & Administration (DPA)

Department of Public Health and Environment (CDPHE)

Department of Public Safety (CDPS)

Department of Regulatory Agencies (DORA)

Department of Revenue (DOR)

Department of Transportation (CDOT)

Division of Youth Services (DYS)

Governor's Office of Information Technology (OIT)

Office of Cyber Security (OCS)

Secretary of State (SOS)

Statewide Traffic Records Advisory Council (STRAC)

Appendix B – HB 09-1285

CONCERNING THE GOVERNMENT DATA ADVISORY BOARD, AND, IN CONNECTION THEREWITH, CREATING THE EDUCATION DATA SUBCOMMITTEE TO MAKE RECOMMENDATIONS FOR CREATION OF A STATEWIDE COMPREHENSIVE P-20 EDUCATION DATA SYSTEM.

Be it enacted by the General Assembly of the State of Colorado:

SECTION 1. 24-37.5-702 (1) and (2), Colorado Revised Statutes, are amended, and the said 24-37.5-702 is further amended BY THE ADDITION OF A NEW SUBSECTION, to read:

24-37.5-702. Definitions. As used in this part 7, unless the context otherwise requires:

(1) "ADVISORY BOARD" MEANS THE GOVERNMENT DATA ADVISORY BOARD CREATED IN SECTION 24-37.5-703.

(2) "CHIEF INFORMATION OFFICER" MEANS THE HEAD OF THE OFFICE OF INFORMATION TECHNOLOGY APPOINTED PURSUANT TO SECTION 24-37.5-103.

(3.5) "EDUCATION DATA SUBCOMMITTEE" MEANS THE SUBCOMMITTEE OF THE ADVISORY BOARD CREATED IN SECTION 24-37.5-703.5 TO PROVIDE POLICIES AND PROTOCOLS REGARDING SHARING EDUCATION DATA AMONG LOCAL AND STATE EDUCATION PROVIDERS.

SECTION 2. 24-37.5-703, Colorado Revised Statutes, is REPEALED AND REENACTED, WITH AMENDMENTS, to read:

24-37.5-703. Government data advisory board - created – duties - repeal. (1) (a) THERE IS HEREBY CREATED IN THE OFFICE OF INFORMATION TECHNOLOGY THE GOVERNMENT DATA ADVISORY BOARD, WHICH SHALL CONSIST OF THE MEMBERS SPECIFIED IN THIS SUBSECTION (1).

(b) ON OR BEFORE OCTOBER 1,2009, THE GOVERNOR SHALL APPOINT FOUR MEMBERS OF THE ADVISORY BOARD AS FOLLOWS:

(I) AN EMPLOYEE OF A CITY, COUNTY, OR CITY AND COUNTY THAT COLLECTS AND MAINTAINS UNIT-LEVEL RECORDS, WHICH EMPLOYEE HAS EXPERTISE IN DATA SHARING AND INFORMATION TECHNOLOGY;

(II) A PERSON WHO IS SERVING ON A SCHOOL DISTRICT BOARD OF EDUCATION IN THIS STATE;

(III) AN EMPLOYEE OF A SCHOOL DISTRICT IN THIS STATE WHO HAS EXPERTISE IN DATA SHARING AND INFORMATION TECHNOLOGY; AND

(IV) A PERSON FROM AN INSTITUTION OF HIGHER EDUCATION OR A NONGOVERNMENTAL ORGANIZATION THAT, IN THE COURSE OF CONDUCTING RESEARCH, ROUTINELY REQUESTS DATA FROM

GOVERNMENT AGENCIES, WHICH PERSON HAS EXPERTISE IN DATA SHARING AND INFORMATION TECHNOLOGY.

(c) THE CHIEF INFORMATION OFFICER, OR HIS OR HER DESIGNEE, SHALL SERVE AS AN EX OFFICIO MEMBER AND CHAIR OF THE ADVISORY BOARD.

(d) (I) THE REMAINING MEMBERSHIP OF THE ADVISORY BOARD SHALL CONSIST OF A PERSON FROM EACH OF THE FOLLOWING DEPARTMENTS WHO IS EITHER AN EXPERT IN INFORMATION TECHNOLOGY OR RESPONSIBLE FOR DATA ADMINISTRATION WITHIN THE MEMBER'S RESPECTIVE DEPARTMENT AND WHO IS SELECTED BY THE HEAD OF THE MEMBER'S RESPECTIVE DEPARTMENT TO PARTICIPATE ON THE ADVISORY BOARD AT THE INVITATION OF THE CHIEF INFORMATION OFFICER:

(A) THE DEPARTMENT OF EDUCATION;

(B) THE DEPARTMENT OF HIGHER EDUCATION;

(C) THE DEPARTMENT OF HUMAN SERVICES;

(D) THE DEPARTMENT OF HEALTH CARE POLICY AND FINANCING;

(E) THE DEPARTMENT OF PUBLIC HEALTH AND ENVIRONMENT;

(F) THE DEPARTMENT OF LABOR AND EMPLOYMENT;

(G) THE DEPARTMENT OF PUBLIC SAFETY;

(H) THE DEPARTMENT OF CORRECTIONS; AND

(I) THE DEPARTMENT OF REVENUE.

(II) NOTWITHSTANDING THE PROVISIONS OF SUBPARAGRAPH (I) OF THIS PARAGRAPH (d), THE GOVERNOR, AS HE OR SHE DEEMS APPROPRIATE, MAY DIRECT THE EXECUTIVE DIRECTOR OF ONE OR MORE OF THE DEPARTMENTS THAT ARE NOT SPECIFIED IN SAID SUBPARAGRAPH (I) TO SELECT A MEMBER FROM HIS OR HER DEPARTMENT WHO MEETS THE QUALIFICATIONS SPECIFIED IN SAID SUBPARAGRAPH (I) TO PARTICIPATE ON THE ADVISORY BOARD AT THE INVITATION OF THE CHIEF INFORMATION OFFICER.

(2) THE MEMBERS OF THE ADVISORY BOARD APPOINTED PURSUANT TO PARAGRAPH (b) OF SUBSECTION (1) OF THIS SECTION SHALL:

(a) SERVE FOUR-YEAR TERMS; EXCEPT THAT TWO OF THE MEMBERS APPOINTED PURSUANT TO PARAGRAPH (b) OF SUBSECTION (1) OF THIS SECTION SHALL INITIALLY SERVE TWO-YEAR TERMS;

(b) SERVE AT THE WILL OF THE GOVERNOR. IF A VACANCY ARISES DURING A MEMBER'S TERM, THE GOVERNOR SHALL APPOINT A PERSON MEETING THE SAME QUALIFICATIONS TO SERVE THE REMAINDER OF THE TERM.

(c) SERVE WITHOUT COMPENSATION AND WITHOUT REIMBURSEMENT FOR EXPENSES.

(3) (a) THE CHIEF INFORMATION OFFICER, OR HIS OR HER DESIGNEE, SHALL SCHEDULE THE FIRST MEETING OF THE ADVISORY BOARD AND SCHEDULE SUCCEEDING MEETINGS OF THE ADVISORY BOARD AS NECESSARY TO COMPLETE THE ADVISORY BOARD'S DUTIES SPECIFIED IN THIS SECTION.

(b) THE OFFICE SHALL PROVIDE TECHNICAL ASSISTANCE AND SUPPORT, TO THE EXTENT PRACTICABLE WITHIN EXISTING RESOURCES, TO ASSIST THE ADVISORY BOARD IN COMPLETING THE DUTIES SPECIFIED IN SUBSECTION (4) OF THIS SECTION.

(4) THE ADVISORY BOARD SHALL HAVE THE FOLLOWING DUTIES:

(a) TO ADVISE THE CHIEF INFORMATION OFFICER REGARDING THE ONGOING DEVELOPMENT, MAINTENANCE, AND IMPLEMENTATION OF THE INTERDEPARTMENTAL DATA PROTOCOL;

(b) TO ADVISE THE CHIEF INFORMATION OFFICER CONCERNING BEST PRACTICES IN SHARING AND PROTECTING DATA IN STATE GOVERNMENT;

(c) TO RECOMMEND TO THE CHIEF INFORMATION OFFICER RULES AND PROCEDURES THAT A STATE AGENCY SHALL FOLLOW IN REQUESTING, OR RESPONDING TO A REQUEST FOR, DATA FROM ANOTHER STATE AGENCY, INCLUDING BUT NOT LIMITED TO STRATEGIES FOR ENFORCING SAID RULES;

(d) TO ADVISE THE CHIEF INFORMATION OFFICER CONCERNING RULES AND PROCEDURES FOR RESPONDING TO DATA REQUESTS SUBMITTED BY AN ENTITY OUTSIDE OF STATE GOVERNMENT;

(e) TO RECOMMEND TO THE CHIEF INFORMATION OFFICER A SCHEDULE OF FEES THAT THE OFFICE MAY CHARGE TO STATE AGENCIES TO SUPERVISE AND ADMINISTER INTERDEPARTMENTAL AND EXTERNAL DATA REQUESTS, THAT A STATE AGENCY MAY CHARGE ANOTHER STATE AGENCY IN RESPONDING TO AN INTERDEPARTMENTAL DATA REQUEST, AND THAT A STATE AGENCY MAY CHARGE TO RESPOND TO A DATA REQUEST SUBMITTED BY AN ENTITY OUTSIDE OF STATE GOVERNMENT. IN RECOMMENDING THE FEE SCHEDULE, THE ADVISORY BOARD SHALL ENSURE THAT THE FEE AMOUNTS DO NOT EXCEED THE DIRECT AND INDIRECT COSTS INCURRED BY THE OFFICE OR BY THE STATE AGENCY THAT IS RESPONDING TO A DATA REQUEST.

(f) UPON REQUEST BY THE CHIEF INFORMATION OFFICER, TO ADVISE THE CHIEF INFORMATION OFFICER ON OTHER ISSUES PERTAINING TO DATA SHARING.

(5) THE ADVISORY BOARD SHALL ENSURE THAT THE RECOMMENDATIONS MADE PURSUANT TO SUBSECTION (4) OF THIS SECTION COMPLY WITH THE INTERDEPARTMENTAL DATA PROTOCOL.

(6) ON OR BEFORE JANUARY 15, 2010, AND ON OR BEFORE JANUARY 15 EACH YEAR THEREAFTER, THE ADVISORY BOARD SHALL SUBMIT TO THE CHIEF INFORMATION OFFICER ITS RECOMMENDATIONS FOR DEVELOPING AND IMPLEMENTING PROTOCOLS FOR SHARING DATA AMONG STATE AGENCIES AND ENTITIES AND WITH LOCAL GOVERNMENTS AND NONGOVERNMENTAL ENTITIES. THE CHIEF INFORMATION OFFICER SHALL REVIEW THE RECOMMENDATIONS AND TAKE THEM INTO ACCOUNT IN PREPARING A REPORT CONCERNING IMPLEMENTING PROTOCOLS FOR SHARING DATA AMONG STATE AGENCIES AND ENTITIES AND WITH LOCAL GOVERNMENTS AND NONGOVERNMENTAL ENTITIES. THE CHIEF INFORMATION OFFICER SHALL SUBMIT THE REPORT TO THE GENERAL ASSEMBLY ON OR BEFORE MARCH 1, 2010, AND ON OR BEFORE MARCH 1 EACH YEAR THEREAFTER.

(7) THIS SECTION IS REPEALED, EFFECTIVE JULY 1, 2019. PRIOR TO SUCH REPEAL, THE ADVISORY BOARD SHALL BE REVIEWED AS PROVIDED IN SECTION 2-3-1203, C.R.S.

SECTION 3. Part 7 of article 37.5 of title 24, Colorado Revised Statutes, is amended BY THE ADDITION OF A NEW SECTION to read:

24-37.5-703.5. Education data subcommittee - created – duties - repeal.

(1) THE EDUCATION DATA SUBCOMMITTEE IS HEREBY CREATED AS A SUBCOMMITTEE OF THE ADVISORY BOARD. THE EDUCATION DATA SUBCOMMITTEE SHALL CONSIST OF THE FOLLOWING MEMBERS:

(a) THE ADVISORY BOARD MEMBERS APPOINTED PURSUANT TO SECTION 24-37.5-703 (1) (b) (II) AND (1) (b) (III) TO REPRESENT SCHOOL DISTRICTS;

(b) A PERSON SERVING ON THE EDUCATION DATA ADVISORY COMMITTEE CREATED PURSUANT TO SECTION 22-2-304, C.R.S., WHICH PERSON IS APPOINTED BY THE GOVERNOR;

(c) THE ADVISORY BOARD MEMBER SELECTED FROM THE DEPARTMENT OF EDUCATION;

(d) THE ADVISORY BOARD MEMBER SELECTED FROM THE DEPARTMENT OF HIGHER EDUCATION;

(e) THE ADVISORY BOARD MEMBER SELECTED FROM THE DEPARTMENT OF HUMAN SERVICES; AND

(f) AT LEAST TEN MEMBERS APPOINTED BY THE GOVERNOR WITH EXPERTISE IN DATA SHARING BY EDUCATION AGENCIES, INCLUDING AT LEAST ONE REPRESENTATIVE FROM EACH OF THE FOLLOWING GROUPS:

(I) INFORMATION OFFICERS EMPLOYED BY THE SCHOOL DISTRICTS IN THE STATE;

(II) CHARTER SCHOOLS AUTHORIZED BY SCHOOL DISTRICTS PURSUANT TO PART 1 OF ARTICLE 30.5 OF TITLE 22, C.R.S.;

(III) THE STATE CHARTER SCHOOL INSTITUTE CREATED IN PART 5 OF ARTICLE 30.5 OF TITLE 22, C.R.S.;

(IV) THE BOARDS OF COOPERATIVE SERVICES CREATED PURSUANT TO ARTICLE 5 OF TITLE 22, C.R.S.;

(V) INFORMATION OFFICERS EMPLOYED WITHIN THE STATE SYSTEM OF COMMUNITY AND TECHNICAL COLLEGES ESTABLISHED PURSUANT TO SECTION 23-60-201, C.R.S.;

(VI) THE GOVERNING BOARDS OF THE STATE INSTITUTIONS OF HIGHER EDUCATION;

(VII) EARLY CHILDHOOD COUNCILS ESTABLISHED PURSUANT TO SECTION 26-6.5-103, C.R.S., AND EARLY CHILDHOOD CARE AND EDUCATION COUNCILS ESTABLISHED PURSUANT TO SECTION 26-6.5-106, C.R.S.;

(VIII) INSTITUTIONS OF HIGHER EDUCATION OR NONGOVERNMENTAL ORGANIZATIONS THAT, IN THE COURSE OF CONDUCTING RESEARCH, ROUTINELY REQUEST DATA FROM STATE AGENCIES;

(IX) NONPROFIT ADVOCACY GROUPS THAT WORK IN CHILDREN'S ISSUES AND ROUTINELY REQUEST DATA FROM STATE AGENCIES; AND

(X) STATEWIDE MEMBERSHIP ORGANIZATIONS OF EDUCATION PROFESSIONALS AND LOCAL BOARDS OF EDUCATION.

(2) THE GOVERNOR SHALL MAKE THE INITIAL APPOINTMENTS TO THE EDUCATION DATA SUBCOMMITTEE NO LATER THAN OCTOBER 1, 2009.

(3) THE MEMBERS OF THE EDUCATION DATA SUBCOMMITTEE APPOINTED BY THE GOVERNOR SHALL:

(a) SERVE TERMS OF FOUR YEARS; EXCEPT THAT, OF THE MEMBERS INITIALLY APPOINTED, THE GOVERNOR SHALL SELECT APPROXIMATELY ONE-THIRD OF THE MEMBERS TO SERVE TWO-YEAR TERMS AND APPROXIMATELY ONE-THIRD OF THE MEMBERS TO SERVE THREE-YEAR TERMS;

(b) SERVE AT THE WILL OF THE GOVERNOR. IF A VACANCY ARISES DURING A MEMBER'S TERM, THE GOVERNOR SHALL APPOINT A PERSON MEETING THE SAME QUALIFICATIONS TO SERVE THE REMAINDER OF THE TERM.

(c) SERVE WITHOUT COMPENSATION AND WITHOUT REIMBURSEMENT FOR EXPENSES.

(4) (a) THE ADVISORY BOARD MEMBER SELECTED FROM THE DEPARTMENT OF EDUCATION SHALL SCHEDULE THE FIRST MEETING OF THE EDUCATION DATA SUBCOMMITTEE. AT THE FIRST MEETING, THE EDUCATION DATA SUBCOMMITTEE SHALL ELECT A CHAIR FROM AMONG ITS MEMBERS TO SERVE FOR A TERM NOT EXCEEDING TWO YEARS, AS DETERMINED BY THE SUBCOMMITTEE. A MEMBER SHALL NOT BE ELIGIBLE TO SERVE AS CHAIR FOR MORE THAN TWO SUCCESSIVE TERMS.

(b) THE EDUCATION DATA SUBCOMMITTEE SHALL MEET AS OFTEN AS NECESSARY, AT THE CALL OF THE CHAIR, TO COMPLETE ITS DUTIES.

(c) THE OFFICE, TO THE EXTENT PRACTICABLE WITHIN EXISTING RESOURCES, SHALL PROVIDE TECHNICAL ASSISTANCE AND SUPPORT TO THE EDUCATION DATA SUBCOMMITTEE TO ASSIST THE SUBCOMMITTEE IN COMPLETING ITS DUTIES PURSUANT TO THIS SECTION.

(5) THE EDUCATION DATA SUBCOMMITTEE SHALL HAVE THE FOLLOWING DUTIES:

(a) TO RECOMMEND TO THE CHIEF INFORMATION OFFICER AND THE ADVISORY BOARD PROTOCOLS AND PROCEDURES FOR SHARING EDUCATION DATA AMONG CHARTER SCHOOLS, SCHOOL DISTRICTS, BOARDS OF COOPERATIVE SERVICES, THE DEPARTMENT OF EDUCATION, THE DEPARTMENT OF HIGHER EDUCATION, AND STATE INSTITUTIONS OF HIGHER EDUCATION;

(b) TO RECOMMEND TO THE CHIEF INFORMATION OFFICER AND THE ADVISORY BOARD APPROPRIATE INFORMATION TECHNOLOGY AND REPORTING FORMATS FOR EDUCATION DATA;

(c) TO RECOMMEND DATA ELEMENT STANDARDS FOR INDIVIDUAL STUDENT RECORDS FOR USE BY CHARTER SCHOOLS, SCHOOL DISTRICTS, BOARDS OF COOPERATIVE SERVICES, THE DEPARTMENT OF EDUCATION, THE DEPARTMENT OF HIGHER EDUCATION, AND STATE INSTITUTIONS OF HIGHER EDUCATION;

(d) TO RECOMMEND ELECTRONIC STANDARDS BY WHICH CHARTER SCHOOLS, SCHOOL DISTRICTS, BOARDS OF COOPERATIVE SERVICES, THE DEPARTMENT OF EDUCATION, THE DEPARTMENT OF HIGHER EDUCATION, AND STATE INSTITUTIONS OF HIGHER EDUCATION MAY SHARE DATA CURRENTLY BEING SHARED THROUGH OTHER MEANS, INCLUDING BUT NOT LIMITED TO INTEROPERABILITY STANDARDS, STANDARDS AND PROTOCOLS FOR TRANSFER OF RECORDS INCLUDING STUDENT TRANSCRIPTS, AND THE USE OF DATA-EXCHANGE TRANSCRIPTS;

(e) TO RECOMMEND THE DESIGN AND CONTINUING DEVELOPMENT OF A STATEWIDE COMPREHENSIVE P-20 EDUCATION DATA SYSTEM THAT MAY INCLUDE, BUT NEED NOT BE LIMITED TO, IMPLEMENTATION OF AN INTEROPERABILITY DATA FRAMEWORK AND PROTOCOLS AND STANDARDS FOR DATA INPUT AND FOR MAKING AND RESPONDING TO DATA REQUESTS TO ENSURE THAT PRESCHOOL THROUGH POSTSECONDARY EDUCATION ENTITIES THROUGHOUT THE STATE CAN SHARE EDUCATION DATA; AND

(f) UPON REQUEST BY THE CHIEF INFORMATION OFFICER, TO ADVISE THE CHIEF INFORMATION OFFICER ON OTHER ISSUES PERTAINING TO EDUCATION DATA SHARING.

(6) (a) THE EDUCATION DATA SUBCOMMITTEE SHALL ENSURE THAT ITS RECOMMENDATIONS CONFORM WITH THE INTERDEPARTMENTAL DATA PROTOCOL AND ARE IN COMPLIANCE WITH ALL STATE AND FEDERAL LAWS AND REGULATIONS CONCERNING THE PRIVACY OF INFORMATION, INCLUDING BUT NOT LIMITED TO THE FEDERAL "FAMILY EDUCATIONAL RIGHTS AND PRIVACY ACT OF 1974", 20 U.S.C. SEC. 1232g.

(b) THE EDUCATION DATA SUBCOMMITTEE SHALL ENSURE THAT ITS RECOMMENDATIONS FOR THE STATEWIDE COMPREHENSIVE P-20 EDUCATION DATA SYSTEM INCLUDE THE ELEMENTS REQUIRED IN THE FEDERAL "AMERICA COMPETESACT", 20 U.S.C. SEC. 9801 ET SEQ., IN ORDER TO QUALIFY FOR THE MAXIMUM AMOUNT OF FEDERAL FUNDING AVAILABLE THROUGH THE "AMERICAN RECOVERY AND REINVESTMENT ACT OF 2009", PUB.L. 111-5.

(7) ON OR BEFORE DECEMBER 1, 2009, AND AT LEAST EVERY SIX MONTHS THEREAFTER, THE EDUCATION DATA SUBCOMMITTEE SHALL SUBMIT TO THE CHIEF INFORMATION OFFICER AND THE ADVISORY BOARD ITS RECOMMENDATIONS PREPARED PURSUANT TO SUBSECTION (5) OF THIS SECTION.

THE CHIEF INFORMATION OFFICER SHALL REVIEW THE RECOMMENDATIONS AND TAKE THEM INTO ACCOUNT IN PREPARING A REPORT CONCERNING PROTOCOLS AND PROCEDURES FOR SHARING STUDENT DATA AMONG PRESCHOOL THROUGH POSTSECONDARY EDUCATION ENTITIES, INCLUDING BUT NOT LIMITED TO THE CREATION OF A STATEWIDE COMPREHENSIVE P-20 EDUCATION DATA SYSTEM. THE CHIEF INFORMATION OFFICER SHALL COMBINE THE REPORT WITH THE REPORT PREPARED PURSUANT TO SECTION 24-37.5-703 (6) AND SUBMIT THE COMBINED REPORT TO THE GENERAL ASSEMBLY ON OR BEFORE MARCH 1, 2010, AND ON OR BEFORE MARCH 1 EACH YEAR THEREAFTER.

(8) THIS SECTION IS REPEALED, EFFECTIVE JULY 1, 2019. PRIOR TO SUCH REPEAL, THE EDUCATION DATA SUBCOMMITTEE SHALL BE REVIEWED AS PROVIDED IN SECTION 2-3-1203, C.R.S.

SECTION 4. 24-37.5-704 (1) and the introductory portion to 24-37.5-704 (3), Colorado Revised Statutes, are amended to read:

24-37.5-704. Interdepartmental data protocol - contents.

(1) The chief information officer, working with the council, shall create ADVISORY BOARD, SHALL OVERSEE THE IMPLEMENTATION OF the interdepartmental data protocol, which at a minimum shall include protocols and procedures to be used by state agencies in data processing, including but not limited to collecting, storing, manipulating, sharing, retrieving, and releasing data. In designing IMPLEMENTING the interdepartmental data protocol, the chief information officer and the council shall establish ADVISORY BOARD SHALL MONITOR COMPLIANCE WITH THE timelines by which the state agencies shall implement the interdepartmental data protocol.

(3) In creating The protocols and procedures included in the interdepartmental data protocol by which state agencies may share data and by which a state agency may release data to a political subdivision or to a nongovernmental entity or an individual the council shall, at a minimum:

SECTION 5. 24-37.5-106 (1) (o), Colorado Revised Statutes, is amended, and the said 24-37.5-106 (1) is further amended BY THE ADDITION OF THE FOLLOWING NEW PARAGRAPHS, to read:

24-37.5-106. Chief information officer - duties and responsibilities - broadband inventory fund created - repeal. (1) The chief information officer shall:

(o) Supervise the chief information security officer appointed pursuant to section 24-37.5-403 (1); and

(q) IN CONSULTATION WITH THE GOVERNMENT DATA ADVISORY BOARD CREATED IN SECTION 24-37.5-703, ADOPT RULES AND PROCEDURES THAT A STATE AGENCY, AS DEFINED IN SECTION 24-37.5-702 (7), SHALL FOLLOW IN REQUESTING, OR RESPONDING TO A REQUEST FOR, DATA FROM ANOTHER STATE AGENCY;

(r) IN CONSULTATION WITH THE GOVERNMENT DATA ADVISORY BOARD CREATED IN SECTION 24-37.5-703, ADOPT RULES AND PROCEDURES FOR RESPONDING TO DATA REQUESTS SUBMITTED BY AN ENTITY OUTSIDE OF STATE GOVERNMENT; AND

(s) IN CONSULTATION WITH THE GOVERNMENT DATA ADVISORY BOARD CREATED IN SECTION 24-37.5-703, ADOPT A SCHEDULE OF FEES THAT THE OFFICE MAY CHARGE TO STATE AGENCIES TO SUPERVISE AND ADMINISTER INTERDEPARTMENTAL AND EXTERNAL DATA REQUESTS, THAT A STATE AGENCY MAY CHARGE ANOTHER STATE AGENCY IN RESPONDING TO AN INTERDEPARTMENTAL DATA REQUEST, AND THAT A STATE AGENCY MAY CHARGE TO RESPOND TO A DATA REQUEST SUBMITTED BY AN ENTITY OUTSIDE OF STATE GOVERNMENT. THE CHIEF INFORMATION OFFICER SHALL ENSURE THAT THE AMOUNT OF THE FEES DOES NOT EXCEED THE DIRECT AND INDIRECT COSTS INCURRED BY THE OFFICE OR BY THE STATE AGENCY THAT IS RESPONDING TO A DATA REQUEST.

SECTION 6. 2-3-1203 (3), Colorado Revised Statutes, is amended BY THE ADDITION OF A NEW PARAGRAPH to read:

2-3-1203. Sunset review of advisory committees. (3) The following dates are the dates for which the statutory authorization for the designated advisory committees is scheduled for repeal:

(ff) JULY 1, 2019:

(I) THE GOVERNMENT DATA ADVISORY BOARD CREATED IN SECTION 24-37.5-703, C.R.S.;

(II) THE EDUCATION DATA SUBCOMMITTEE CREATED IN SECTION 24-37.5-703.5, C.R.S.

SECTION 7. Act subject to petition - effective date. This act shall take effect at 12:01 a.m. on the day following the expiration of the ninety-day period after final adjournment of the general assembly that is allowed for submitting a referendum petition pursuant to article V, section 1 (3) of the state constitution, (August 5, 2009, if adjournment sine die is on May 6, 2009); except that, if a referendum petition is filed against this act or an item, section, or part of this act within such period, then the act, item, section, or part, if approved by the people, shall take effect on the date of the official declaration of the vote thereon by proclamation of the governor.