OIT
Office of Information Technology
Increasing the effectiveness of government through information technology

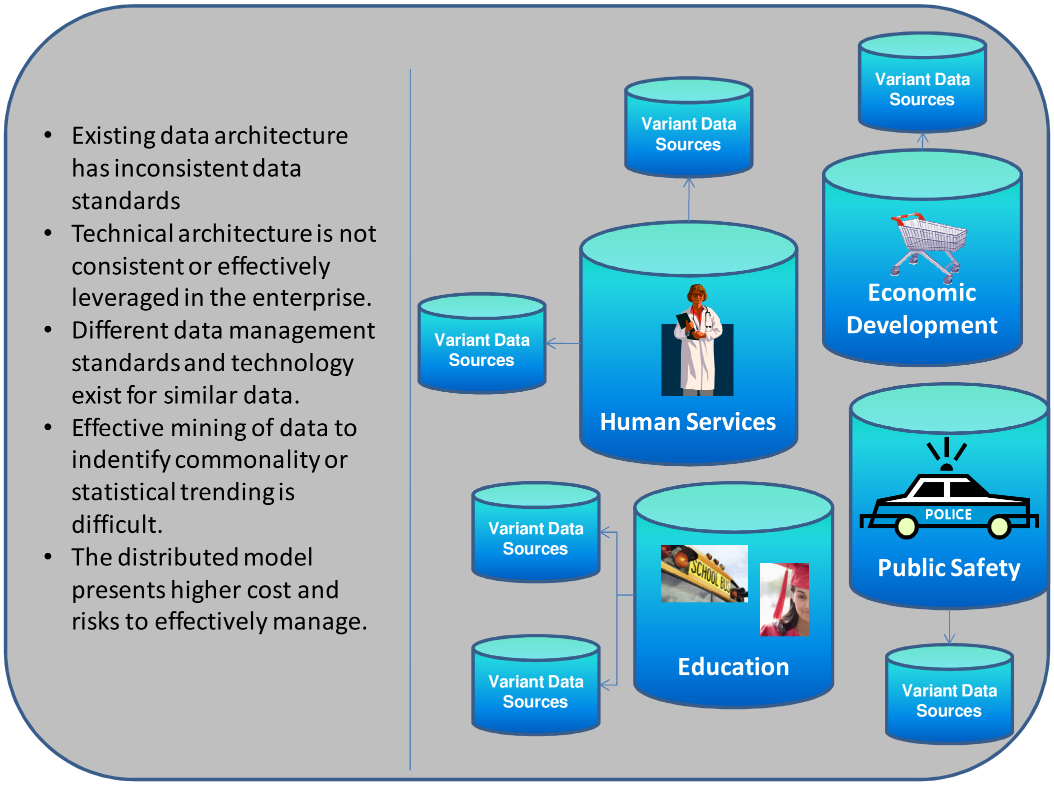
HB 08-1364 Interdepartmental Data Protocol Council Report

February 27, 2009

Table of Contents

Executive Summary

The demand for trusted information continues to spiral upward. The State of Colorado owns significant data resources, but turning those data resources into information assets that can be managed for effective decision making is not currently a mature process at the enterprise level as illustrated below:

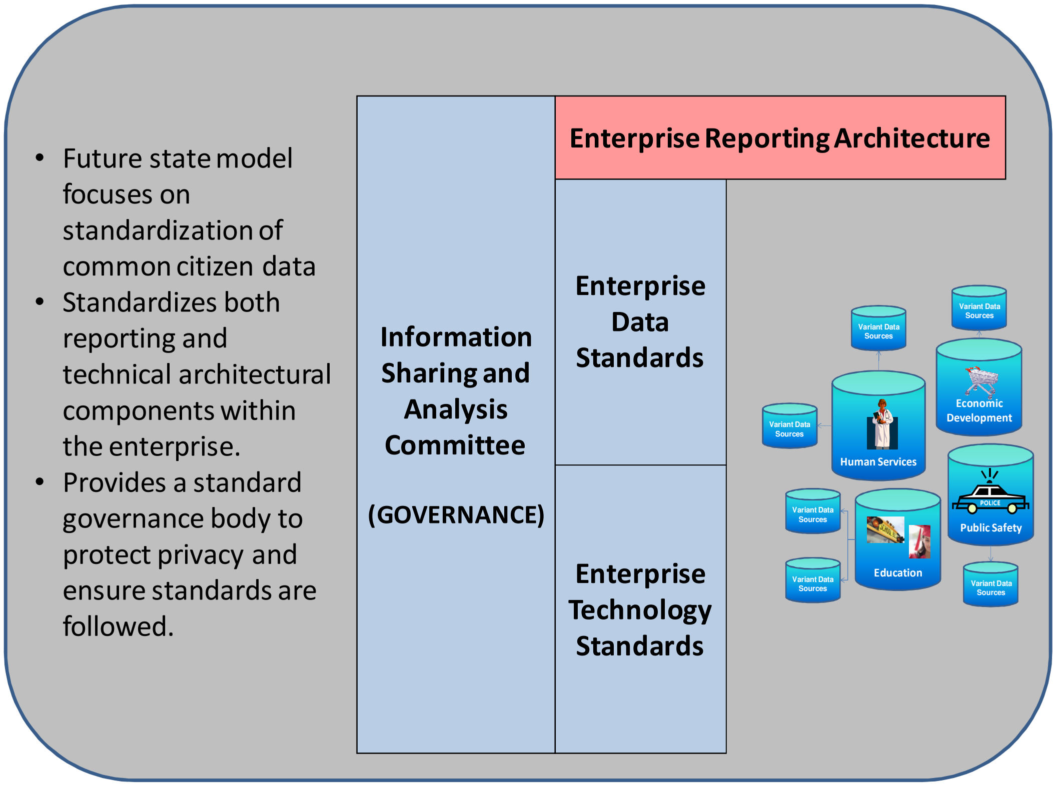


- Existing data architecture has inconsistent data standards
- Technical architecture is not consistent or effectively leveraged in the enterprise.
- Different data management standards and technology exist for similar data.
- Effective mining of data to indentify commonality or statistical trending is difficult.
- The distributed model presents higher cost and risks to effectively manage.

Figure ES-1

The quality of data and information impacts the quality of the decisions made by the State. Positive citizen outcomes are the ultimate objectives and this requires government to properly manage data, information and knowledge assets to make informed, accurate, and timely policy and resource decisions. The timely availability of accurate information will also enhance service deliveries across State agencies.

While enhanced data management and information sharing will improve decision making, accelerate service delivery and reduce the costs associated with non-integrated systems, the State is also faced with rapidly escalating expectations for data privacy from citizens. As citizens become increasingly aware of highly publicized data breaches in both the private and public sectors, expanded data sharing will have to address the protection of non-public private data that is included in many unit records associated with various state services.



- Future state model focuses on standardization of common citizen data
- Standardizes both reporting and technical architectural components within the enterprise.
- Provides a standard governance body to protect privacy and ensure standards are followed.

Information Sharing and Analysis Committee

(GOVERNANCE)

Enterprise Reporting Architecture

Enterprise Data Standards

Enterprise Technology Standards

Figure ES-2

Government performance depends on accurate and timely information. Real value creation comes from organization performance improvements based on cross-agency information sharing, as illustrated in the examples below from varying vertical sectors:

- *Policy Making* - Helps lawmakers and policy makers answer questions regarding the best use of limited State resources and effectiveness of State programs.
- *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.
- *Education* - Ensures that a seamless education system from pre-school to graduate school is preparing our young people for the demands of the 21st by linking records over time (PreK-20), analyzing performance, and studying educational effectiveness. Linking systems and including higher education data enables policymakers, teachers and school administrators to answer questions such as:
  - Which schools and programs add the most value to student learning?
  - Which schools are consistently high performing so their best practices can be studied?
  - By 10th grade, what are the best educational predictors of college and workforce readiness – to allow enough time to ensure all students are fully prepared by high school graduation?
  - What high school performance indicators are the best predictors of student's success in college or the workplace?
  - What percentage of high school graduates needs remedial courses in college?
  - Which teacher preparation programs produce the graduates whose students have the strongest academic growth?
- *Social Services* –Creates means to directly certify students for supplemental nutritional services based on family eligibility for food stamps, as done in Michigan and other states. In Michigan, the

State Department of Human Services provided the Department of Education a file of families that are eligible for food stamps. This file is then compared with student data from school districts to identify eligible students that can be directly certified without additional application processes. This program has a twofold benefit: ensuring that all eligible students receive the benefits that their parents may not otherwise have applied for, and reducing fraudulent claims against the system by comparing the files.

- *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.

Objective

The Legislature thoughtfully approached the issue of enterprise data sharing with the passage of House Bill 08-1364. HB 08-1364 directed the Governor's Office of Information Technology (OIT) to convene a Data Protocol Development Council ("Council") to assist in designing and implementing an interdepartmental data protocol. The goal of the cross-departmental data protocol is to facilitate information sharing across agencies, and to assist in formulating and determining the effectiveness of state policies.

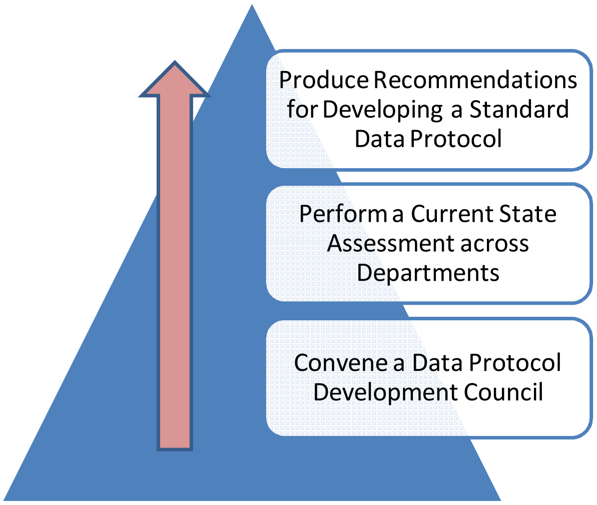

Produce Recommendations for Developing a Standard Data Protocol
Perform a Current State Assessment across Departments
Convene a Data Protocol Development Council

Figure ES-3

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 stated goals and objectives of HB-1364. The Council was comprised of representatives from Executive Branch agencies and interested parties as deemed necessary by OIT.

The focus of the Council was targeted to achieve the following objectives:

- To analyze and determine the effectiveness of State policies and resources by examining an issue across multiple State agencies;
- To formulate informed strategic plans for the application and use of State resources based on strong, accurate, reliable, multi-dimensional data;
- To enable more efficient collecting, storing, manipulating, sharing, retrieving, and releasing of data across State agencies.

Approach Summary

The Council was convened on August 21, 2008 and completed its work in February 2009. The Council restricted its study of the data sharing protocol to unit records. Unit records are defined as records pertaining to individuals. The Council reviewed and baselined the current data systems; data sharing practices and applications; governance policies and procedures; and, statutory and regulatory guidelines. Each Agency identified *one* representative data store that is of relatively high importance to that agency with regard to data sharing. Additional detail can be found in *Section 1, Background and Overview*.

Findings Summary

The current challenges to data sharing and data management can be traced to how applications and systems have evolved over the past decades: decentralized, stovepiped systems with little cross-agency communication. Agencies have not previously been incentivized to work together to share information or to develop common architecture and standards. Siloed applications are an increasing business problem for the State, as they carry a high cost of ownership; can't be changed easily to react quickly to ongoing business demands; often require a legacy skill that fewer and fewer people hold; do not meet compliance demands; and, finally, do not enable a 360° view of the State enterprise.

That said, the Council **recommends** that the State continue to move forward with an enterprise data sharing initiative based on the following:

1. Expansion of data sharing in service delivery channels

While neither Colorado nor any other state has demonstrated much experience with enterprise data sharing or validated the benefits, the desired goals established in HB 08-1364 remain compelling. The Council identified existing cross-agency data sharing programs that are already delivering value and confirmed that industry trends and research support accelerated investment in the processes and infrastructure to enable implementation of the protocols to expand shared data services.

2. Standardization can enable efficient and secure data sharing

The technology and security teams assigned to the council have validated that expanded data sharing goals are technologically achievable. In some cases, the implementation of expanded data sharing will require limited modification of existing architectures. Implementation of the current enterprise architecture, data governance organization, and business processes currently planned under the CIO's IT Consolidation strategy will strengthen the State's data sharing efforts.

3. Acceleration of IT consolidation will enable enterprise data sharing

Since the State of Colorado has just launched an IT Consolidation program that will ultimately include cross-agency operations and IT support, implementation of enterprise data sharing is a natural extension. However, the current state of IT operations is decentralized with siloed data. As with any significant change in organization and structure, the Council identified some risks and barriers to uniform implementation of a consistent protocol to share data across all agencies for all applications. The barriers vary in the form of diversity in application and data architecture, to inconsistent information security controls that may not be adequate to protect citizen privacy at either regulatory or State policy levels.

Justification

The Council identified the following justifications for expanding data sharing:

- Several State agencies are already sharing data to serve citizen needs in multiple business environments, including public safety, judicial, and citizen services.
- Departments identified more opportunities to share data in most business aspects of state government, including education, workforce development, and citizen services.

- Baseline analysis of the progress made by other states, as well as the research contributed by the National Association of State Chief Information Officers (NASCIO), provide early validation of anticipated benefits from enhanced data sharing has been identified.
- The State CIO is leading the State through a planned migration from decentralized IT operations to an enterprise approach. This migration is a pre-cursor to conducting cross-agency planning and oversight for integrated programs.

Framework for Implementing a Standard Framework for Data Sharing

According to the guidelines included in HB 08-1364, the Council identified the key elements for establishing the protocols and policies necessary to implement State-wide data sharing. Since an existing enterprise data architecture and associated data governance structure has not yet been deployed in the early stages of the new IT consolidation program, the following key elements must be addressed in future enterprise IT consolidation plans:

- Establishment of a data governance organization
- Adoption of a data sharing model
- Creation of a metadata registry and associated processes
- Improvement of enablers such as enterprise security and privacy policies

Risks and Barriers to Uniform Implementation

While the business case to expand data sharing has been established, the State is not uniformly prepared to implement cross agency data sharing due to the following risks and barriers.

- Lack of a State data model or central data identification source
- Lack of shared architectures
- Lack of enterprise processes
- Costs not yet well understood
- Ensuring privacy and security of data
- Establishing effective practices for implementing and managing data exchanges
- Establishing an enterprise culture of data sharing

Although these have been identified as risks and existing barriers, they are listed here not because they cannot be overcome, but to enable the State to better prepare for the challenges ahead.

Recommendations

To implement expanded data sharing, the Council recommends consideration of the following recommendations:

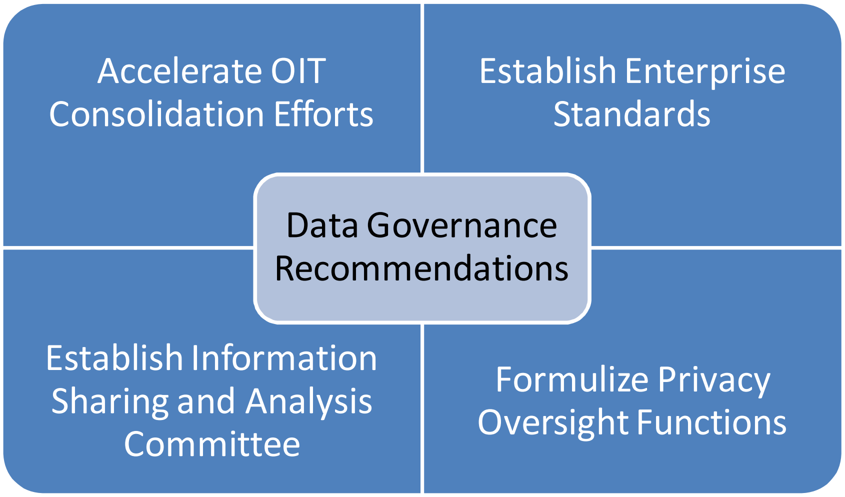

Accelerate OIT Consolidation Efforts
Establish Enterprise Standards
Data Governance Recommendations
Establish Information Sharing and Analysis Committee
Formulize Privacy Oversight Functions

Figure ES-4

1. ***Establish a Government Data Advisory Board (GDAB) to Provide Oversight and Leadership***
    a. Formalize a strategic advisory council with senior State leadership and integrate working groups of business users to provide oversight and leadership.
    b. Provide a forum to initiate requests, to establish guidelines, and to provide oversight to ensure programs are implemented efficiently.
    c. Establish a statewide MOU framework to enable more efficient planning across agencies to facilitate sharing data.
    d. Develop a list of priorities for rolling out expanded data sharing to those systems with the highest return and match the State's ability to both implement data sharing and protect the privacy of citizen data.
    e. Approve new policies to guide cross-agency data sharing since existing policies do not accommodate efficient data sharing.
    f. Develop metrics to track and measure the performance of this program.
2. ***Accelerate deployment of the CIO's IT Consolidation Program to institutionalize enterprise IT planning***
    a. Begin immediately with low cost objectives such as implementing a common data sharing protocol using the NIEM model and XML format.
    b. Formalize an enterprise architecture function that establishes central data identification and planning functions with skilled resources to select target programs and define protocols to share data.
    c. Leverage the current movement towards "banding" agency functions to align systems and data sharing needs into defined channels.
        i. Continue to inventory sources of unit records and data owners
        ii. Continue to inventory unit record locations and data architecture types.
    d. Establish a data governance function with a designated new organization to assess data privacy risks and design adequate new controls to protect citizen privacy.
    e. Formalize a new enterprise data architecture team within the central enterprise architecture group.

        i. Extend the data architecture framework to define specific data models to be implemented in new data sharing programs.
        ii. Develop a prototype design and testing platform to evaluate data sharing and consolidation strategies.
    f. Establish enterprise licenses for all tools needed to implement data sharing across the enterprise.

3. ***Formalize Privacy Oversight Functions***
    a. Establish a formal Privacy Office with a Chief Privacy Officer with authority and defined duties to guide citizen privacy protection.
    b. Develop a statewide privacy policy to replace internal agency policies.
    c. Formalize a review process with the Chief Information Security Officer to ensure that the statewide information security program will protect private citizen data maintained in shared unit records at levels required by interagency MOU.
    d. Review the Colorado Open Records Act and update the statute to allow for the protection of private citizen data.

The Council views this work as modular in nature. Identified new data sharing initiatives must fully comply with all policy and architectural recommendations. As systems are upgraded, they will have to come into compliance with the new standards. When new systems are developed, they too will have to fully comply with the program. This is a pragmatic, cost-effective approach to rolling out this program, while ensuring that over time all State IT systems and data sharing efforts come into compliance with the recommendations outlined by the Council.

The Council appreciates the opportunity to review this important issue and to provide its thoughts and recommendations to the State Chief Information Officer through this report.

Section 1 - Background and Overview

HB 08-1364

House Bill 08-1364 directed the Governor's Office of Information Technology (OIT) to convene a Data Protocol Development Council ("Council") to assist in designing and implementing an interdepartmental data protocol. HB-1364 was one of Governor Ritter's priority bills from the 2008 legislative session. The goal of the cross-departmental data protocol is to facilitate information sharing across agencies and assist in formulating and determining the effectiveness of State policies. This project examined what is currently in place today, and provides recommendations for moving forward to accomplish interagency data sharing in a uniform manner.

The mission of the Council was to provide guidance, policies and procedures for implementing a data sharing architecture across the State enterprise that will achieve the stated goal and objectives of HB-1364. The Council was comprised of representatives from Executive Branch Agency and interested parties as deemed necessary by OIT. Council meetings were open to the public.

HB 1364 was initially driven by distinct needs identified by the Governor's P-20 Education Coordinating Council to analyze longitudinal data regarding factors including improving teaching and learning, informing public policy, fostering a culture of evidence-based decision making, conducting research, evaluating system and program effectiveness, and providing reports to various stakeholder groups. The collective Colorado State government, as an entity that provides funding, resources and services to the citizenry of the State, has similar needs.

These include:

- The ability to analyze and determine the effectiveness of State policies and resources by examining an issue across multiple State agencies;
- To formulate informed strategic plans for the application and use of State resources based on strong, accurate, reliable, multi-dimensional data;
- To enable more efficient collecting, storing, manipulating, sharing, retrieving, and releasing of data across State agencies.

Why Data Sharing is Important

Given the importance of data in State business, data and information must be treated as a highly valuable enterprise asset. Data must be maintained to some level of quality if it is to be trusted or relied upon for decision making. Flawed data and information lead to flawed decision making.

Information must be integrated across the enterprise. Citizens should see one State government, despite the fact that government decisions increasingly require more types of data from multiple and diverse government lines of business. The State must properly manage its data resources within a data governance framework to enable effective decision making. Ultimately, this will allow the State to act as a single organization, with the ability to respond to environmental threats and opportunities faster and more effectively.

There are many benefits to the State for initiating an enterprise data sharing program. There are benefits for policy development and resource alignment, and also economic benefits, as it is cheaper to share and secure data than to recollect, store, maintain, secure in multiple, often redundant, data stores. There is greatly increased collaboration between the State and its Federal and Local partners, particularly in the areas of justice, public safety, and social programs. Some additional benefits include:

- Permitting cross-departmental analysis and forecasting by combining data from multiple sources;
- Allowing for validation of programs across agencies;

- Allowing for verification, refutation and refinement of findings;
- 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,
- Effective use of resources by avoiding unnecessary duplication of data collection, analysis, and reporting.

Incorporating data sharing in Colorado's long term planning is important today. It will assist agencies' program reporting and analysis. 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 eliminates 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, strategists and implementers, policy makers and citizens, career government employees and elected officials. All of these characteristics are founded on proper management of State government data, information and knowledge assets with the ultimate outcome – ***serving the citizenry***.

Council Approach and Scope

The Council was convened on August 21, 2008 and completed its work in February 2009. All meetings were open to the public. The Council restricted its study of the data sharing protocol to unit records. Unit records are defined as records pertaining to individuals. The Council reviewed and baselined the current data systems; data sharing practices and applications; governance policies and procedures; and, statutory or regulatory guidelines in place across Executive Branch agencies that maintain unit records. The Council was divided into three subcommittees to tackle these issues more efficiently. The subcommittees were Business, Legal, and Technology.

Due to the limited time and resources of this project, it was necessary to prioritize the review to a few selected unit record data stores. Each Agency identified *one* representative data store that is of relatively high importance to that agency with regard to data sharing. A preference was given to data stores that are already being shared. Examination of current solutions assisted in identifying key success factors. Selected data stores were benchmarked so that information such as (but not limited to) the following was collected:

- Data dictionary
- Data modeling tools
- Sample data inventory
- Application hardware and software
- Any existing data sharing done out of or in to that data store
- Existing agency policies/procedures/governance structures in place regarding sharing of that data
- Current statutory or regulatory (State or Federal) guidelines in place regarding the use of that data

The Council also benchmarked the work of other states in this area by examining governance structures and privacy policies for the cross-departmental data protocol. The Council identified an existing data sharing project in the State on which it can pilot, test, and adjust its recommendations so that the Council can get feedback in a real-world scenario.

Finally, the Council developed recommendations and an action plan for moving forward with developing and implementing the cross-departmental data protocol. *Procurement, development, and/or implementation of Council recommendations were outside the scope of work for this phase of the process. Additionally, the Council did not consider any business process re-engineering or change management work to be in-scope.*

Mission and Objectives

The goal of the cross-departmental data protocol is to facilitate information sharing across agencies and to assist in formulating and determining the effectiveness of State policies. The mission of the Council was to provide guidance, policies and procedures for implementing a data sharing architecture across the State enterprise that will achieve the stated goal and objectives of HB-1364.

In detail the goals and objectives of the Council (with subcommittee responsibilities in parentheses) are as follows:

Goal 1: Analyze the requirements of all State agencies that have a need to share unit record data with other agencies.

- Understand and document the unit record data captured, stored and maintained by State Executive Branch agencies (only one key data store needs to be documented by the end of February) (Technology);
- Understand and document the existing hardware, software, networking and communications systems that contain the unit record data (Technology);
- Understand and document the existing data sharing practices and applications employed by all State Executive Branch agencies (Business);
- Understand and document the existing governance policies and procedures employed by all State Executive Branch agencies with regards to collecting, storing, sharing, and destroying data (Business and Legal);
- Understand and document the existing statutory or regulatory guidelines in place at all State Executive Branch agencies with regards to collecting, storing, sharing, and destroying data (Legal).

Goal 2: Determine a data sharing protocol that meets the needs of State agencies.

- Assess existing national data sharing standards and benchmark the work of up to five other states in this area (Business);
- Develop an architecture for the development of the data protocol, including data normalization, identity resolution, and source data authority (Technology);
- Develop a governance structure, including processes and procedures, to be used by state agencies for sharing information with another State agency, with a political subdivision, or with a nongovernmental entity or an individual (Business);
- Establish the circumstances under which a State agency may release data to a political subdivision, a nongovernmental entity or an individual (Legal);
- Establish the format in which a State agency may release data to a political subdivision, a nongovernmental entity or an individual (Technology);

- Establish the retention and destruction policies of data that is shared by a State agency to a political subdivision, a nongovernmental entity or an individual (Legal);
- Ensure compliance with existing statutory and regulatory requirements (Legal);
- Create new or modify existing policies to ensure personal privacy and the protection of personal identifying information (PII) (Legal).

Goal 3: Develop recommendations and a strategy for moving forward.

- Develop alternative and recommended solutions for implementing the data sharing protocol (All);
- Establish time lines for implementing the recommended solution across all State agencies (All);
- Identify high-level associated costs for the recommended solution (Technology and Business);
- Identify necessary statutory or regulatory changes (Legal);
- Identity critical gaps that must be addressed to ensure the success of this project (All);
- Identify next steps to ensure the project moves forward to the next phase (All).

Challenges Faced by Council

While the Council was able to accomplish much in its six months of work, there were a few challenges identified by the group that limited the Council's ability to develop a full-fledged solution:

- The broadness of the HB 08-1364 language meant that instead of focusing on one singular data sharing application with a known set of systems and constraints, the Council had to consider an unlimited possibility of agencies, systems, compliance environments, policies and procedures, etc. It was extremely important to scope the project down into something that would be manageable in the time frame in which the final report was due.
- The volunteer status of the Council participants necessarily meant that they were doing Council work in addition to their full-time jobs. While everyone was enthusiastic about the project, it was at times difficult for people to break away from their daily responsibilities.
- The limited time frame (six months) of the project made it impossible to complete a full assessment of agency systems, policies or regulatory environments. The end of project also coincided with the budget request cycle and 2009 legislative session.
- There was little readily-available documentation (systems, governance or legal/regulatory) that the Council could begin with, providing very limited visibility to the current state of the IT ecosystem.

Council Representation

HB 08-1364 mandated that OIT convene a Data Protocol Development Council (“Council”) to design and implement an interdepartmental data protocol. Executive branch agencies that collect unit record data were required to participate. The Chief Information Officer (CIO) could include additional persons on the Council if deemed necessary. Participating agencies included:

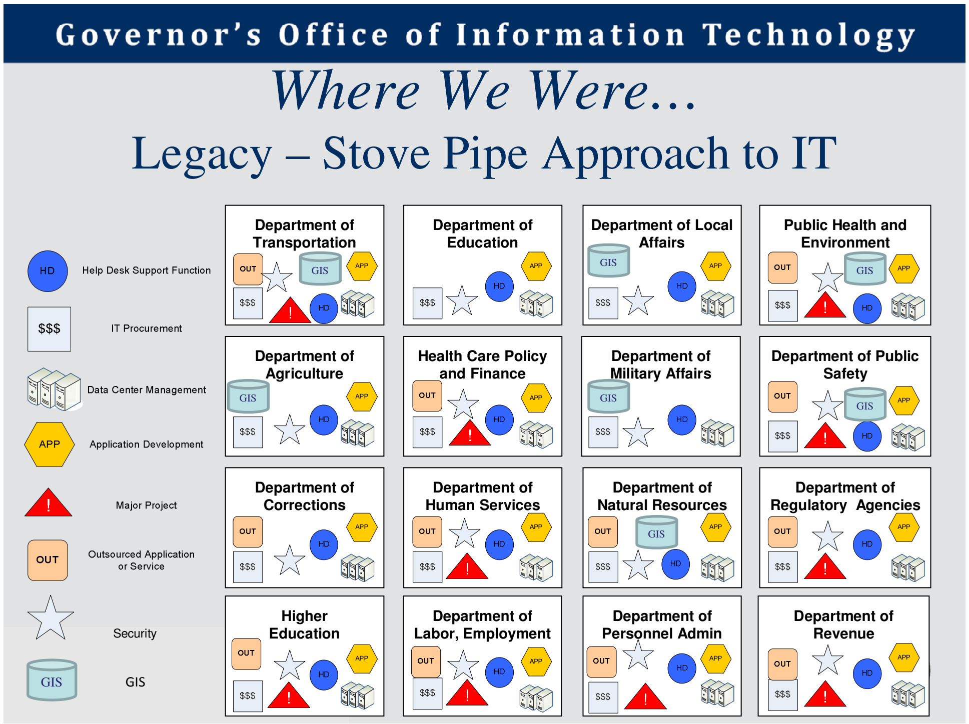
- Governor’s Office of Information Technology (OIT)
- Department of Agriculture (CDA)
- Department of Corrections (CDOC)
- Department of Education (CDE)
- Department of Health Care Policy and Finance (HCPF)
- Department of Higher Education (DHE)
- Department of Human Services (DHS)
- Department of Labor and Employment (CDLE)
- Department of Local Affairs (DOLA)
- Department of Natural Resources (DNR)
- Department of Personnel and 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)
- Office of Cyber Security (OCS)
- Secretary of State (SOS)
- Judicial
- Attorney General (DOL)

Additionally, members of four intra-agency groups also participated upon invitation from the CIO: the Data Governance Working Group (DGWG), the Statewide Traffic Records Advisory Council (STRAC), the Colorado Integrated Criminal Justice Information System (CICJIS), and the Colorado Children and Youth Information Sharing (CCYIS) initiative.

Section 2 - Current State of Data Sharing within State Executive Branch Agencies

The current challenges to data sharing and data management within the State can be traced to how applications and systems have evolved over the past decades. Enterprise information 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.

The following picture is from an IT assessment commissioned early in Governor Ritter's administration, and depicts the large number of redundant infrastructures (software and hardware) across multiple departments. This includes the number of systems capturing data and storing it redundantly, in varying formats, and with varying privacy and security policies.



Governor's Office of Information Technology

Where We Were…

Legacy – Stove Pipe Approach to IT

HD Help Desk Support Function

$$$ IT Procurement

Data Center Management

APP Application Development

Major Project

OUT Outsourced Application or Service

Security

GIS GIS

Department of Transportation

Department of Education

Department of Local Affairs

Public Health and Environment

Department of Agriculture

Health Care Policy and Finance

Department of Military Affairs

Department of Public Safety

Department of Corrections

Department of Human Services

Department of Natural Resources

Department of Regulatory Agencies

Higher Education

Department of Labor, Employment

Department of Personnel Admin

Department of Revenue

Figure 2-1

This decentralized approach has resulted in a myriad of data modeling approaches, naming standards, formats, technologies, tools, staff expertise, policies and governance approaches, and meta data standards. The result is an environment of stovepiped information characterized by unnecessary redundancy, inconsistency, contradictory data, and inconsistent methods for modeling data. 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.

There is very limited visibility into the actual level, format, and governance of the data sharing that is occurring today across State agencies. This keeps the State in a heightened state of vulnerability - not exactly what IT has or how it is being protected. This means the State is much less agile than it could be in reacting to challenges or proactively and predictively developing solutions. The predictive nature of data should be the intended long term capability – not only to manage risk, but to anticipate, to uncover and to prepare for opportunities and threats.

The Council conducted an analysis of the participating agencies, and determined the following:

1. Information Accessibility
    - Enterprise information assets are stored and maintained to various levels of quality throughout the State government.
    - The stovepiped approach has resulted in a diverse set of data modeling tools, approaches, naming standards, formats, technologies, polices and governance approaches, and meta data standards.
    - The environment is one of very siloed information characterized by unnecessary redundancy, inconsistency, contradictory data, and inconsistent methods for modeling data.

2. Productivity and Agility
    - The sample system inventory showed that although we are one State we have many different ways to store, transmit, and access data using a variety of vendors and technologies. The more varied the technical landscape the more challenging data sharing becomes.
    - There is limited visibility today into the actual level, format, and governance of the data sharing that is occurring today across State agencies. This means the State is much less agile than it could be in reacting to challenges or proactively and predictively developing solutions.
    - State system-of-records are often not well understood by their data owners, and often do not have the most accurate information.

3. Security
    - Though there are basic cyber security policies in place, there are varying degrees of implementation of these policies across the agencies.

4. Fiscal
    - The siloed approach to data storage and data analytics is expensive.
    - The individual acquisition, deployment and maintenance of decision support systems are repetitive and costly.
    - No enterprise licensing agreements are in place to support more cost effective implementations of data sharing or business intelligence initiatives.

5. Governance
    - There are varying data sharing policies and processes in place, with no coordinating body.
    - There is no statewide authority for how to best collect, store, maintain, share or dispose of data.
    - MOUs are used in limited situations and are not standardized.

6. Legal
    - State laws often contain language limiting the sharing of data.
    - Agency policies sometimes go beyond legal requirements to restrict data sharing.
    - Agencies are generally aware of their compliance environments.
    - There is not currently a statewide privacy office or statewide policies for protecting privacy.
    - There is no standard definition of PII (personally identifiable information).

7. Cultural
    - Siloed approaches to data warehousing have created a culture of vertical data ownership rather than enterprise data stewardship.

The remainder of this section describes in further detail the current environment in three primary areas: Governance, Technology and Systems, and Legal and Regulatory. Some existing data sharing initiatives at the Executive Branch level are then described.

Governance

Governance relates to decisions that define expectations, grant power, or verify performance. It consists either of a separate process or of a specific part of management or leadership processes. "Data Governance" in particular refers to the organizational bodies, rules, decision rights, and accountabilities of people and information systems as they perform information related processes. Data governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, and using what methods.

Colorado currently follows a decentralized data sharing process that puts each individual agency in charge of its related data. Certain data has restrictions on how it may be shared, if at all. The State also has basic cyber security policies on how sensitive data needs to be stored and transmitted. Many, but not all, data sharing arrangements are documented in a multitude of memoranda of understanding between agency executive directors.

There is no statewide authority on how best to store, maintain, share or dispose of that type of general data that is repeated in each respective data base. For example, an individual's name, address, date of birth, Colorado Driver's License number, Social Security Number, etc. are stored a multitude of times in disparate State applications.

Processes

The Council's research of current non-technical (i.e., business related, data sharing processes) within the State of Colorado consisted of participating agencies completing two process-related templates:

1. A Data Sharing Initiatives template where agencies recorded information on at least one application where data is shared,
2. A Compliance Mandates template where agencies recorded information on various types of mandates such as Federal and State statutory or regulatory requirements, rules, internal agency policies etc. on sharing data.

Although the information gathered on these templates does not represent a complete inventory of the State's applications, it does demonstrate a good sampling of processes Colorado agencies are using to complete their data sharing requirements.

Fourteen Executive Branch agencies plus offices of the Secretary of State, Law and Judicial participated. Several agencies participate with multiple other State agencies in large data sharing initiatives. Information on criminal justice data sharing, traffic data sharing and child and youth data sharing programs in Colorado can be found later in this report. Virtually all agencies reported that they had written data sharing agreements, agency policies and/or statutory instructions on some or most of their data sharing situations.

It is clear from reviewing the Data Sharing Initiative and Compliance Mandate templates that each agency that has primary responsibility for various types of data is aware of the legal and policy requirements that apply to sharing the related data. There are many mandates about what data can be shared, for what purpose, and with whom. Agencies reported only isolated cases where modifications or clarifications of the data sharing rules might be needed.

Technology and Systems

One of the tasks of the Council was to determine the current state of the unit data systems regarding the commonality and consistency of the systems containing unit data. To assist in this determination, a sample inventory of the systems was gathered across the departments, as well as a sample definition of the data within a representative system. The response to the sample inventory included metadata (data and information about data) collected from 12 agencies, including over 50 systems (one agency provided reports on 40 systems). While this is just a sample of the unit data systems and an even smaller percentage of the overall State systems, it gave the Council an insight into the current technological profile of Colorado.

Current Systems Profile

The sample system inventory showed that although there is one State government, it has many different ways to store, transmit, and access data using a variety of vendors and technologies. Since each agency has been independently funded and has also acquired its information systems independently over the years, this was not a surprise and is not unusual based on research from other states. For example, ten different database management systems were being used just in the sample data, a very wide variety of mainframes and servers, and over 50 vendors supplying information technology hardware and software products. While these may be the best choice for each system's use, high-level coordination and communication are required to share data between them. The more varied the technical landscape, the more challenging it becomes to share data.

Current Data Profile

The survey also requested metadata about the specific data elements from a representative system. This was an extremely important survey because in order to share, compare, and combine data, the State must know the definitions of each data element in each system, a means to identify it, what is in each field and how the data is related.

Since many, if not most, of the State systems are third-party vendor systems, the Council actually expected a wide variety of data definitions. This was indeed the case. Results showed that each system/vendor had its own data structures, data names, and element definitions. No systems shared any data names at all.

This sampling highlighted the complexity of finding data in the State's information systems. While it can and is accomplished, many hours of research are needed to ensure the correct and accurate data. There is no one place to go learn in which data system the data is stored, which copy of the data is the 'data of record', which one is correct, or even how many systems use and store the data. Typically, for each new data sharing initiative, hundreds of man-hours must be exerted at the beginning of the project just to answer these questions.

Data Modeling and Repository Tools

The Council also conducted an informal survey of data modeling, data dictionary, and repository registry tools used across the State. Only nine of the agencies responded as having models or data dictionaries. This was not unusual in the past when third-party vendors and custom-off-the-shelf (COTS) systems were purchased, as data dictionaries and logical/physical models usually needed to be specified as part of project deliverables, or they were not received by the agencies.

The results of the survey showed that, for the most part, robust modeling and data dictionaries have not yet been used much in the State. No tools focusing on data dictionaries or metadata repository were found. The three tools discovered that have actual data modeling abilities are ERStudio (Embarcadero Technologies) ER, ERwin Data Modeler (Computer Associates), and PowerDesigner (Sybase). These tools also have some metadata tracking and reporting abilities.

Other State Baselining Discoveries

- Consistent Data Values - In one very simple data sharing initiative, a Colorado agency automatically collected technical information from network devices around the State. The raw data had over 4000 ways of stating 32 state agency names. Although it is a relatively straightforward task to electronically scrub data to a standard format, part of any model needs to include the acceptable values for any data sharing initiative.

- Common Data Retention Periods - Typically, data associated with people in an individual program like children in a Head Start program or unemployment information is only retained for a limited time. This is due to both cost and privacy concerns. Yet, if the goal is to track the effectiveness of government programs across the life of an individual, longer retention times will be required. It may take decades to develop data sources that are long-lived enough to track people from the education system to the work force or criminal justice system. At this time, data in each agency is held for varying lengths of time, depending on need, statute, and policy. This needs to be consistently determined and applied across the departments to achieve meaningful data analysis results.

- Privacy and Security - For the State to fully protect the data of the citizens and to track the effectiveness of their tax dollars on social programs, a greater investment in security and privacy controls is required. Currently, most Colorado agencies do not meet minimum data protection standards and funding for security initiatives has fallen short of the amount required to comply with existing regulations.

- Federal Funds and Data Reporting – The National Information Exchange Model (NIEM) and its Extensible Markup Language (XML) component have become the defacto standard for communicating with Federal agencies, and has actually become the required method with many Federal agencies. Colorado needs to become familiar and fluent with the NIEM, as well as XML, to effectively communicate with the Federal government.

Legal and Regulatory

The Legal subcommittee reviewed the laws, regulations and policies (collectively, “laws”) set out in the baseline documents prepared by each participating agency, for each agency system selected for this program. Due to the time constraints of this project, it was not feasible to review the compliance and regulatory environment of every data system within every agency.

At this time, each agency collects data for its unique purposes. Each set of data is governed by a different set of compliance requirements. Health information, for example, is governed by the Health Insurance Portability and Accountability Act (HIPAA), but driver’s license information is governed by the Driver’s Privacy Protection Act (DPPA). Each agency that collects data not specifically protected from disclosure maintains the data in its own databases and for its own purposes.

Some agencies baselined a program that has no restriction regarding the sharing of information. While this may seem like a good thing, it also underscores the problem that these programs are not considering the nature of personal identifying information (PII), and are not currently requiring safeguards. Other agencies stated that no information may be shared under any circumstances regardless of the intended purpose. This is the opposite end of the spectrum with regard to protecting PII. It appears likely that the best practice will lie somewhere in the middle. PII must be protected, but it is also valuable to share information from a business perspective, to allow the greatest use and leverage of the information currently stored at the State level.

In order for extended data sharing to continue, all State and Federal laws and regulations and agency or agencies' rules and policies will need to be reviewed to determine what restrictions on sharing exist, and where those restrictions can be modified.

Existing Data Sharing Initiatives

There are many data sharing initiatives already in place across the State enterprise. Many are simple "point-to-point" data exchanges involving two agencies and a small data set. At the other end of the spectrum (though much fewer in number) are large, complex programs that have been operating for several years with a formal governing body, technology infrastructure, and defined policies and processes. Some examples of existing or new data sharing initiatives follow.

Colorado Integrated Criminal Justice Information System (CICJIS)

CICJIS is a collaborative program designed to facilitate information sharing at key decision points in the criminal justice process across the boundaries of organizations and jurisdictions among the State criminal justice agencies to enhance public safety, improve decision making, increase productivity and improve access to information.

CICJIS is an independent program that relies on the equal participation of the five CICJIS agencies – Colorado Bureau of Investigations (CBI), Colorado District Attorneys Council (CDAC), Judicial, CDOC and Division of Youth Corrections (DYC). CICJIS provides information sharing services that provide access to data across all the agencies and eliminate the need for redundant data entry using automated systems. The automated systems are designed to provide one-time entry of data and efficient access to justice information to all agency consumers and their customers using secure, role-based authenticated methods. These services were designed with a broad consensus that integrated justice information could potentially save lives, time and money.

The Colorado legislature mandated the development of CICJIS beginning in 1995. House Bill 95-1101 (amended by HB 05-1078) defined the composition of CICJIS to include the Departments of Public Safety, Corrections, and Human Services, and the Colorado Judicial Branch. It directed the executive directors of each agency to cooperate in the development of a strategic plan for the implementation and maintenance of an integrated criminal justice information system. The General Assembly adopted the strategy outlined in that plan, formally included the Colorado District Attorneys Council, and funded the effort through Senate Bill 96-221. System design was approved on September 9, 1996, and development began immediately thereafter. The initial phase of CICJIS, data transfers, was implemented on May 4, 1998. Query functions were implemented beginning in the summer of 1999.

The final phase of the CICJIS data-sharing implementation was felony disposition matching. CICJIS defines felony disposition matching as the percentage of court dispositions that match to an arrest posted on the CBI/CCIC Record of Arrest and Prosecution (RAP) sheet. The Colorado State Legislature set felony disposition matching as a critical measure of the program's success.

When CICJIS began, the felony disposition match rate was approximately 12%. Working closely with court clerks, local law enforcement and the district attorneys, the rate has improved to over 97% as of August 2008. Numerous counties have match rates approaching or exceeding 95%. This rate improves as CICJIS staff continues to address the local criminal justice business practices among the 64 counties in Colorado.

State Traffic Records Virtual Data Warehouse

In 2005, Congress enacted the Federal Highway Funding Reauthorization bill, the *Safe, Accountable, Flexible and Efficient Transportation Equity Act – a Legacy for Users* (SAFETEA-LU) which, under the highway safety provisions, provides significant additional funding to each state for the improvement of traffic records information systems. There are conditions. Each state must have an updated Strategic Plan for Traffic Records, a current Traffic Records Assessment and a statewide Traffic Records

Coordinating Committee (TRCC) with certain roles and responsibilities. In Colorado, the State Traffic Records Advisory Committee (STRAC) has served as the TRCC since the 1970's. Membership is comprised of six principal state agencies, along with local government representatives, universities, researchers and others. The six State agencies are CDOT, CDPS, DOR (Motor Vehicle Division), CDPHE, CDHS, and Judicial. Additionally, several local jurisdictions are involved in this cooperative data sharing effort.

Most databases still function as "islands of information" with limited data sharing and transfer. Data is inconsistent from one database to another. The quality of some data is questionable and accessibility, particularly for managers, is also limited. It is vital for managers to have reliable data upon which to make decisions concerning policy formulation and the allocation of resources for improvement in all areas of traffic records. These include accident records, driver and vehicle records, roadway information, emergency medical services pre-hospital records, citation and court records. The event-based virtual Virtual Data Warehouse will include, at least initially, statewide information on motor vehicle crashes and traffic citations. Crashes and citations will be tracked through the system from the time the event begins until its ultimate resolution.

The STRAC has recently completed the award of a Request for Proposal (RFP) issued in mid-2008 to start work on the Virtual Data Warehouse project.

Colorado Children and Youth Information Sharing (CCYIS)

The Colorado Prevention Leadership Council (PLC) and the Collaborative Management Program State Steering Committee (CMP-SSC) are interagency collaborative groups addressing coordination, collaboration and integration of children and youth prevention, intervention, and treatment services. One of the priority areas identified by both groups and their respective memberships is the need for improved data support and data sharing agreements across state departments to enhance long-range, integrated and comprehensive planning around common priorities at the State and local levels. Work in these areas is believed to support five important goals:

1. Improved monitoring and response to emerging social issues;
2. Data-driven resource allocation and utilization;
3. Local data utilization for needs assessment, strategic planning and monitoring;
4. The assessment of service impacts as reflected in changes in social and health indicators; and,
5. Improving the exchange of appropriate and necessary information across systems to improve coordinated services for children, youth, and families.

To this end, an expanded group called the Colorado Data Sharing and Utilization Group (CDSUG) has been formed to create sustainable data sharing and utilization infrastructure in service of these goals. Seven State departments are represented, including CDE, CDHS, CDPHE, CDPS, Judicial, HCPF and DYS. Local partners are also involved to discuss common issues related to coordination, collaboration and integration as related to services for children, youth and families.

As a part of CDSUG, the Colorado Children and Youth Information Sharing (CCYIS) initiative has just gotten formally underway, and is beginning work on defining governance, policies and procedures, and information technologies to achieve its goals of sharing data in the aggregate between agencies and systems and the appropriate sharing of client level information among agencies serving children and youth.

Comparison of Colorado to Other States

Council members studied best practices and lessons learned in the data sharing community (e.g., other states, higher education, research foundations) and other organizations that share data to perform various types of analysis, publication and transactions. The research studies revealed the following:

- The most successful data sharing initiatives had one central authority to process data sharing requests for everyone in the enterprise.
- These successful instances all have a standard written agreement template that documents how or if data will be shared between entities and what governance process will be followed all the way through the data sharing process.
- Data quality is a persistent attribute if studies and transactions are to be accurate. Responsibility for data quality needs to be assigned and audited.
- There are numerous technology approaches to sharing data; some are better than others depending upon the requirements of the sharing entities. Regardless of what technology is utilized, the processes must be well documented, practiced and standardized. The most advanced data sharing technologies store master unit data in one logical place.
- In any successful data sharing situation between disparate systems, there must be a robust method to ensure that unit records in one system match those of the other. The most advanced data sharing technologies match records on the fly every time the data is merged.
- The total amount of data sharing varies among public entities. Those who share more data between agencies and applications tend to have the most experience in successful data sharing scenarios.

The Business Subcommittee undertook a study of other states, and developed a benchmarking methodology that scored the six different process and initiative areas outlined above. One benchmark was assigned to each of the six major findings from meetings and research of written materials available on the Internet.

The project team researched various state websites. The results of our research indicate that the following five states have made progress in several or all benchmark areas: Arkansas, California, Kansas, Kentucky, and Virginia. The six benchmarks used were:

- Benchmark 1: Fully empowered, statewide central clearing authority and processes for data sharing requests.
- Benchmark 2: Comprehensive data governance and sharing agreement in place.
- Benchmark 3: Established audit system assessing data quality, validity, and reliability.
- Benchmark 4: Established data technology processes.
- Benchmark 5: Robust identifier (record matching) system.
- Benchmark 6: Extent of data sharing with other states, agencies within the State and sub-political entities.

Each benchmark area has a total of 60 points possible, for a total of 360 points per State. Five of the benchmarks have three questions and the other benchmark has two questions:

Benchmark 1:

1. Who has final authority for data sharing? (Centralized agency versus local agencies). Local 1-10 points; centralized 11-20 points. This question is looking at how organized and how centralized the data sharing effort is.

2. Who is responsible for initiating data sharing? (Centralized agency versus local agencies) Local 1-10 points; centralized 11-20 points. This question looks at enterprise efforts as opposed to individual agency efforts.

3. Level of authority for data sharing (i.e., legislative versus rules versus MOU). Handshake: 1-5; bi-lateral MOUs: 6-10; rule: 11-15; statute: 16-20. This question looks at how much authority is held.

Benchmark 2:

1. Level of the agreement(s) governing data sharing (legislation, MOU, verbal agreement). Local 1-10 points; centralized 11-20 points. This assesses the extent to which there is a contractual agreement.

2. How comprehensive is the agreement? (e.g., keyword definitions, purpose, parties and signatures, permitted/non-permitted uses, rules for access, limitations on disclosure, security requirements, retention and disposition of data). Four (4) points for each on list above; maximum of 40 points.

Benchmark 3:

1. Level of authority for identifying and correcting data integrity issues (centralized versus local). Local 1-10 points; centralized 11-20 points. This looks for the amount of authority for data integrity and data correction.

2. Validation processes (automated, human, etc.). Human 1-5 points; automated 6-10 points; combination of human and automated 11-20 points.

3. Comprehensiveness of processes (frequency, breadth, etc.). Not defined 0 point; annual review 1-5 points; quarterly 6-10 points; monthly 11-15 points; weekly or on going 16-20 points.

Benchmark 4:

1. Defined technology processes for sharing data (e.g. data transmission, data dictionary, change management, data retention and deletion technical solutions). 5 points for each on list above. This element measures the extent of progress in data sharing technology.

2. Data inventory (what data does the State have) process. Little to no inventory 1-5 points; early stages of inventory development: 6-10 points; basic inventory developed 11-15 points; comprehensive data inventory developed and in use 16-20 points. This element assesses the degree to which the State documents their data.

3. Master data system (store common data attributes one place in the State for use by all, e.g. name, address, date of birth etc.). No plans for master data 0 points; plans under consideration 1-5 points; plans developed 6-10 points; master data system implemented 11-20 points. This element was looking for progress in storing certain personal information in one place for all agencies to use.

Benchmark 5:

1. Defined technology processes for ensuring data from disparate systems is accurately matched. No Process defined 0 points; plan developed 1-10 points; plan implemented 11-20 points. This question is looking for progress in identifying technology.

2. Automated robust dataset linkage tool in place to match data on the fly. No process 0 points; manual 1-10 points; automated 11-20. This was asked to discern progress in automating record matching.

3. Process in place to protect unique personal identification (e.g. masking, 3rd party holding data, encryption keys). 5 points for each of the above. 5 points total for other processes. This element looks for ways to protect personal information.

Benchmark 6:

1. Number of agencies in the State currently sharing data. None 0 points; 2 points for each agency; with a maximum of 20 points. This question looks for implementation progress.

2. Number of applications sharing data. None 0 points; 2 points for each application; with a maximum of 20 points. This element looks for implementation progress.

3. Number of other states, sub-political entities and Federal government sharing state data. None 0 points; 2 points for each application; with a maximum of 20 points. This question looks for implementation progress.

Several members of the Council participated in teleconferences with the four states identified above. There were three main teleconference goals:

1. To establish contact with selected other states for possible future conversations and collaboration.
2. To ensure that the Committee interpreted each state's web site presentations correctly and found most of the pertinent information.
3. To have a personal conversation with data sharing teams of selected other states and drill down on key specifics of those states' progress and plans.

The subcommittee's goal was not to rank the success of the states participating in the benchmarking but rather to compare Colorado's progress in establishing important data sharing processes. Ultimately, the Committee desired to collaborate with personnel in other states who are interested, and to share various processes and templates.

Benchmarking Results

Results of the benchmarking process revealed that although all five states and Colorado had various degrees of early progress in data sharing, most continue to operate with bi-lateral agreements between agencies who share their data to achieve some specific requirement. There has been little to no progress in data sharing initiatives that seek to improve efficiencies and citizen service delivery by establishing master data processes: i.e. one stop shopping. **Figure 2-2** below shows the results of the five benchmark scores.

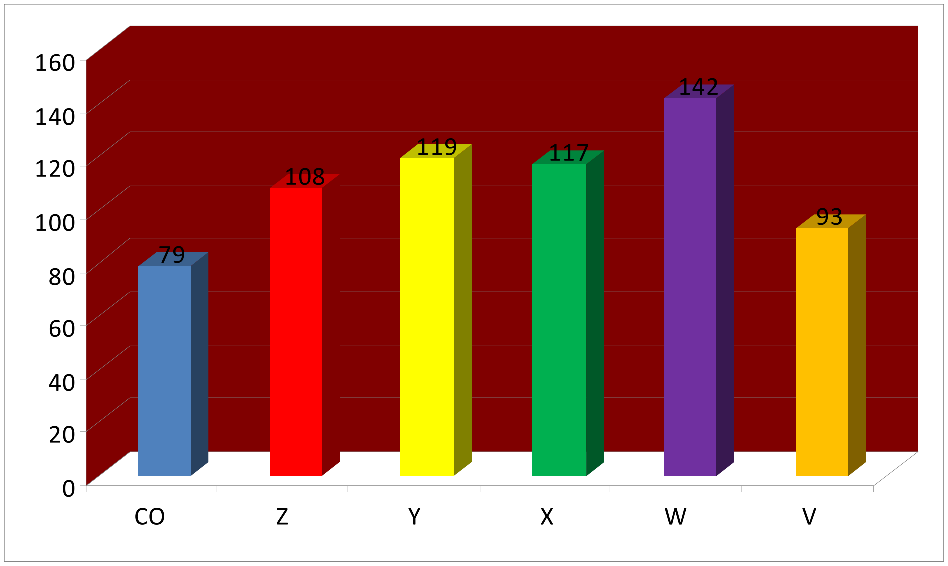


Figure 2-2

Although Colorado and the benchmarked states are among the leaders in data sharing, they all still scored between the 25th and 40th percentiles. Scoring the states in relation to the maximum possible score, 360 points, is illustrated in **Figure 2-3** below.

Most states are just in the early stages of experimenting with tools and technologies that allow unit records to be accurately merged from disparate data systems.

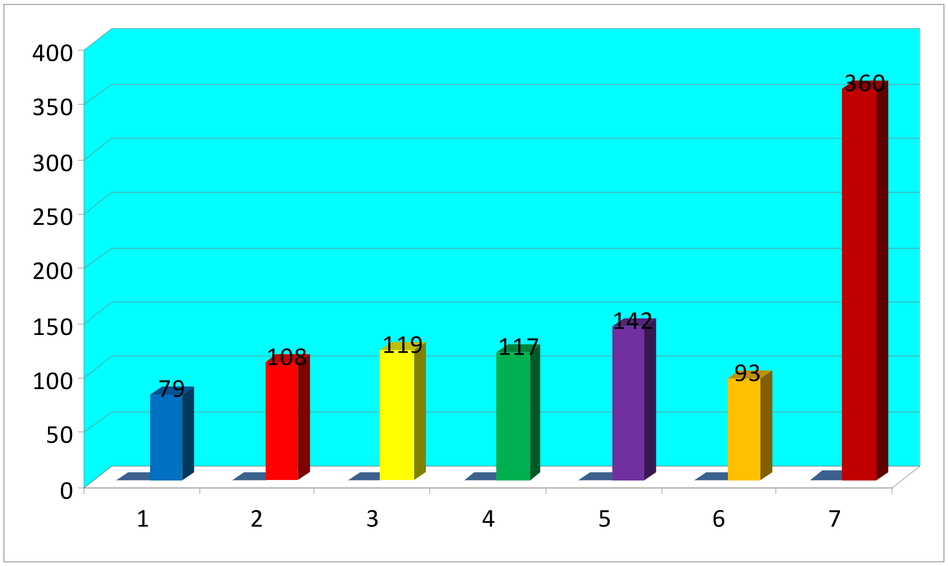


Figure 2-3

Specific findings for those states benchmarked can be found in the appendices.

Section 3 - Best Practices for Data Sharing and Integration

Data sharing often leads to increased efficiency and better coordination of services and management for all partners involved in the data exchange. Critical for success is a cultural change that views data as information and knowledge, and that these are the primary, and most critical, assets of state government. Managing information as an enterprise asset requires effective data governance. While the proper technology is important to actually execute information sharing across the enterprise, the key to creating and operationalizing a successful, sustainable information sharing program is a strong governance structure and program. This report will discuss governance in much detail throughout the rest of this report, as it is so intrinsically inherent to all aspects of a data sharing initiative, whether agency-to-agency, multi-agency, or across the enterprise.

In researching and speaking with thought leaders and other states about enterprise data sharing, several best practices quickly came to the surface and were echoed across the board. The following describes these best practices in some detail.

Managing Information as an Enterprise Asset

It is an evolutionary process to go beyond simply capturing data for line of business reporting to being able to integrate data across an enterprise for more accurate, efficient, and precise decision and policy making. It requires a sophistication of thought that looks beyond the walls of an agency or department to see the enterprise holistically and with a common purpose. It requires information sharing to be integrated into strategic planning and all agencies coming to the table to determine guiding principles and how data governance looks across the enterprise.

It also means identifying shared benefits. While the missions of state agencies are often materially and consequently different, they share a common purpose: serving the citizens of Colorado. Sharing data to guide effective policy decisions can improve the lives of the citizens of this State. State leaders must foster cross-sector trust, buy-in and cooperative work towards data sharing systems development by keeping all stakeholders and constituents focused on the benefits of sharing. Cross-cutting State priority issues, such as economic growth, workforce development and adult literacy, can serve as galvanizing forces. A shared data management strategy will improve the alignment of initiatives across departments.

These concepts will be worked through and refined as data sharing efforts continue across the State.

Enterprise IT Consolidation

An interagency, enterprise-wide approach that directs each agency to adopt common standards and architecture allows agencies and jurisdictions to more easily share information. Consolidating IT infrastructure investment decisions under the oversight of the State CIO allows this coordination and enforcement of standards. Additionally, centralized IT review boards should be created to act on requests for funds to support implementation of integration projects, particularly projects that fully support enterprise initiatives. Moving IT investment decisions up to a central authority helps free agency directors from having to choose between projects that benefit only their lines of business, and projects that will benefit the entire enterprise.

Fortunately, Colorado has already begun to implement strong measures along these lines. In 2007, Governor Bill Ritter, Jr. announced a multi-year information technology consolidation plan that folds State government's decentralized operations into the Governor's Office of Information

Technology (OIT). This was formalized through SB 08-155, the "*IT Consolidation Bill*". The plan calls for centralized information technology management, purchasing, spending, and planning. The plan will also create a statewide enterprise structure compared with today's department-by-department model. OIT is well down the path of centralizing information technology resources to create an IT organization that is streamlined, efficient, and optimized to deliver the critical tools necessary to fulfill the missions of each department and to ensure the services and information those departments provide are able reach to the citizens of Colorado.

Governance

As described briefly in Section 2, enterprise data sharing cannot be done without a strong governance model. Data governance refers to the operating discipline for managing data and information as a key enterprise asset. Data governance is primarily a business, not an IT, function. The operating discipline includes organization, processes, and tools for establishing and exercising decision rights regarding valuation and management of data. Key aspects of data governance include decision-making authority, compliance monitoring, policies and standards, data inventories, preservation, data quality, data classification, data security and access, data risk management, data valuation, full lifecycle management, content management, and records management. Additionally, ongoing partnership agreements, processes for handling data requests, and the ability to reconcile technical differences across agencies are important.

Data governance is essential to ensuring that data is accurate, appropriately shared, and protected. The quality of data and information will certainly impact the quality of the decisions that consume it. Positive citizen outcomes are the ultimate objective, which requires government to properly manage data, information and knowledge assets in order to make informed, accurate, and timely policy and resource decisions.

While there is not yet a formal enterprise data governance strategy in place in the State, OIT has been working towards this goal since 2008, and plans to finalize this in 2009. Additionally, the Office of Cyber Security is initiating work on the development of enterprise data security classification policies based on the FIPS 199 security policy.

Understanding the Geopolitical Complexities Across Agencies

The complexity government agencies face in creating an effective data sharing model and data interoperability appears to increase proportionally with the number of boundaries crossed, the number and type of information resources to be shared, and as the number of technical and organizational processes to be changed or integrated increases. These difficulties result from the reality that sharing information involves large parts, if not the whole, of an enterprise or policy domain.

Trust also needs to be established among agencies, and it is a fundamental, but often overlooked, piece of any data sharing initiative. If I give you my data are you going to use it appropriately? How can I trust that you will secure it properly? How can I trust that you will allow only authorized people to view the data? And, equally as important, how can I trust that you won't use my data against me?

The *Information Sharing Complexity Matrix* (see **Figure 3-1**) provides a mechanism for characterizing a cross-boundary interoperability initiative and identifying the level of complexity to be expected in creating the interoperability and information sharing capability necessary for transformation. The first dimension refers to the focus of the initiative, which can be meeting a specific need or problem, or building systemic capacity. The second dimension takes into consideration the associated level of organizational involvement with three categories of involvement: intra-organizational, inter-organizational, and inter-governmental.

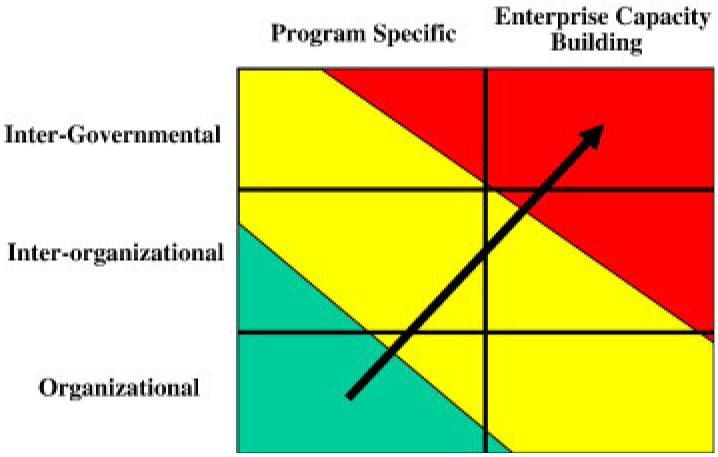


Figure 3-1 Information Sharing Complexity Matrix

With respect to improving interoperability and data sharing, the ability to understand the level and nature of the complexity early on and before investments are made is important. The *Information Sharing Complexity Matrix* provides a simple but clear conceptual model to help government managers identify the types of "boundaries" that will be crossed and some of the associated barriers and challenges they might face within a specific interoperability initiate. Of course, acknowledging the complexity of these "future challenges" is only a beginning. Government leaders need to move from understanding to action. It is critical to bring together all stakeholders at the beginning to receive input.

While the State is in its infancy with regards to enterprise data sharing, it has become increasingly apparent to the Council that in moving beyond point-to-point data sharing, the addition of each agency or data store that is added doesn't just simply increase the complexity of the system and processes to be developed, it ***exponentially*** increases the complexity. Many data sharing initiatives go beyond the State to include Federal and local agency partners. Their needs, processes, and compliance and regulatory parameters must also be considered, and adds to the complexity of the development of these programs.

Understanding the Data

The State needs to understand the characteristics of data and take steps to make its data ready for sharing. This includes the areas of accessibility, availability, quality, auditability, and security.

1. **Accessibility -** Data that is ready for data sharing is easily accessible to any community of interest (COI), both from a data discovery standpoint (common data definitions, descriptions, and locations are identified) as well as from an actual data content standpoint (reports or tools are available to use). The processes and security must be in place to support those activities. Accessibility also means decoupling data (all data types, structured and unstructured) from siloed agency applications and beginning to manage data independently, across the organization.
2. **Availability** - Available data can be defined as data that is available for consumption by the user in the timeframe needed. This may be historical data or current data, it may be master data (data that is basically static and changes infrequently) or transactional data (data that is volatile). The timeframe may vary from needing a report for tomorrow's budget meeting, to a historical analysis of program data over time, to monthly change reports. Coordinated governance processes will facilitate timely exchange of data to deliver it in the timeframe needed. Available data satisfies the data consumer's needs in both content and quality in the timeframe where it is needed and effective.

3. **Quality** - Data quality refers to the reliability and effectiveness of data measured by a defined state of completeness, validity, consistency, timeliness and accuracy that makes it appropriate for a specific use or able to satisfy a given need. Data quality assurance (DQA) is the process of verifying the reliability and effectiveness of the data. Maintaining data quality requires reviewing or auditing the data periodically and assessing it for quality and, if the quality has degraded, cleansing it. Cleansing data typically involves standardizing the definitions and values, updating it to achieve standardization and completeness, and removing duplications. This is done across systems regardless of physical implementation technology.
4. **Auditability** - The data, as well as reports from the data, need to be certifiable in many cases. This means that the data must be able to be tracked from source to destination and verified for security and accuracy, and that it follows State data administration guidelines, State disclosure policies, and controls. This also means the reporting tool must not allow alterations of the data once it has been reported.
5. **Security** - Securing access to data, whether static or in transit, is a critical piece to the integrity of both the data as well as the State. All data should have a designated security level (or classification) indicating who can view, use, and redistribute it. This classification will follow the data as it moves through the system, and ensure proper handling. The proper security needs to be enforced during data extraction and data movement and loads, as well as during the reporting process. This may be enforced by encryption methods, through robust access management policies, to the data stores, to the data itself, or based on the actual values of the data. Disclosure and compliance documents should verify adherence to the designated security level.

The State today has very limited visibility into the actual state of data across its agencies, which is not unusual for an organization that has traditionally been decentralized, as Colorado's IT infrastructure has been. It is one of the areas that requires the largest amount of work, and longest time period to implement, as the State moves forward with enterprise data sharing and data governance.

Managing Identities and Assuring Privacy

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. Is the John E. Smith with a date of birth of 10/22/64 the same as John Edward Smith with a date of birth 10/22/64? Or, is it John Eugene Smith, date of birth October 22, 1964? Is R. Matt Baker the same as Robert Matthew Baker or Roger Matt Baker?

Duplicative identity data stored in multiple locations within the same agency, as well as across agencies, causes a negative impact on accuracy. If there is a lack of commonly used standards, it makes appropriate cross-function collaboration difficult, thus impacting time-sensitive mission needs. It also complicates an individual's attempt at redress and can reduce personal privacy. Additionally, privacy protection efforts that vary in complexity across agencies*,* often slows or prevents data sharing initiatives from occurring.

The Council's vision is for a much more substantially organized identity management framework that will be worked through as the Council's work continues. The ideal situation would be creating a single "version of the truth" with regards to identity demographic information, and anonymizing personal information where appropriate.

Transparency

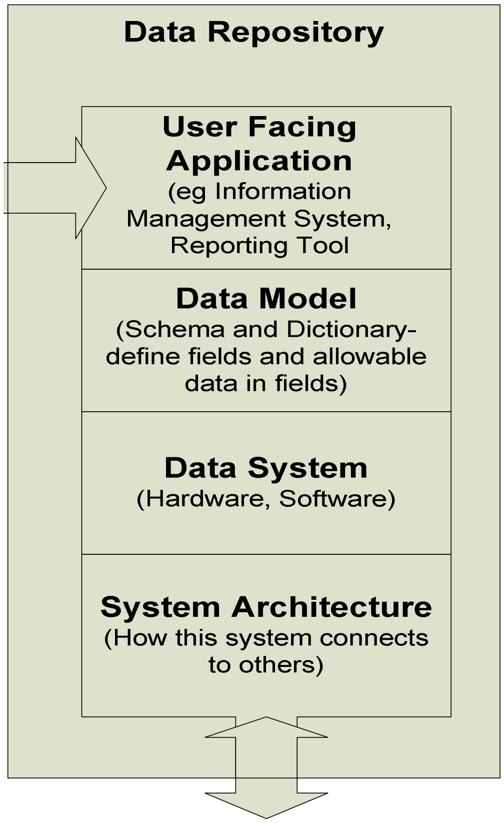
Finally, and universally, transparency in the program’s purpose, the reason why information is shared, how it will be used, who will have access to the information, how it will be secured, and whether individuals can access and correct their personal information is critical to establishing confidence in the program from all stakeholders.

OIT remains committed to transparency and openness, as long as system security is not compromised and vulnerabilities within the enterprise are not created.

Section 4 - Technical Approaches

The mechanics of populating data repositories to share data happens at four separate layers: the application layer, the data model layer, the data system (or platform) layer, and the systems architecture layer. Technical approaches to enable data sharing at every level can be defined.

Figure 4-1


Data Repository
User Facing Application
(eg Information Management System, Reporting Tool
Data Model
(Schema and Dictionary- define fields and allowable data in fields)
Data System
(Hardware, Software)
System Architecture
(How this system connects to others)

Components of a Data Repository

As **Figure 4-1** shows, people usually interact with a data system through an application, or computer program. Reporting or analytical applications might have names like the Human Resources Data Warehouse or an operational system might be the Colorado Benefits Management System.

Data models consist of logical and physical data models. The logical data model contains the overall data architecture of the enterprise data regardless of the system where it resides. The logical data model, housed in a metadata registry or metadata dictionary, contains at a minimum business sanctified data names, data definitions, an enterprise logical model and links to the physical data models where the actual data resides.

The physical data model defines how the data is organized, including the size and types of fields in a data base, how the fields inter-relate to each other, and the allowable data and form for the data in each field. The database management system (DBMS) definitions of field names and the common words used to define them are usually called a data dictionary or data catalog.

The data system implemented within a database management system is part of the application, which includes hardware (e.g., servers) and software (e.g., database applications). When a series of applications use the same software and compatible hardware, this is called a platform. This application system can stand alone or it may be part of a larger data architecture connecting it to other systems and other data repositories.

The systems architecture is the collection of all of the applications systems, databases, and data repositories of the enterprise.

Although the technology exists to connect any system to any other across the internet, and to translate from one program and database to another, developing this translation software is time-consuming, expensive, and unique to every two systems it connects. To technically enable data sharing between governments and government organizations, a common system architecture and logical data model are needed. New data sharing initiatives and new applications and database systems can be designed to this standard from the beginning to facilitate future data sharing. Legacy systems can be converted over time as systems are upgraded.

Common Data Sharing Architectures

The following are approaches being considered by the State in implementing data sharing.

- Physical Data Warehouse
- Federated, Virtual Data Warehouse
- Point-to-Point Data Exchange
- Hybrid Data Warehouse Collaborations

These are described in more detail below.

Physical Data Warehouse

The simplest way to share data is to have a single integrated system where data is entered and processed and a single warehouse environment where data is accessed for reporting, analysis, and data mining. In the case of unit data, the ideal situation would be a single input system for all the unit master data (name, address, phone, birth date, etc.) that was highly secure with each agency relating this to its own data.

The advantages of a physical data warehouse are the ease of data integration and sharing, and the centralized control, making security and privacy compliance, compliance with a standard data dictionary, and control of the data quality easier and more efficient. The primary disadvantage of this model is that it often takes a large and expensive up-front design and effort.

Federated or Virtual Data Warehouse

In this concept, data is identified and 'viewed' from a central point, but not actually moved until the data is needed. Different organizations maintain their own data in a common format and physical data base or data warehouse. They agree on technical protocols to physically share data. At the central virtual data warehouse location reports and views of the data are created as they are requested by pulling what is needed from the other database. The virtual database does not store anything permanently.

The federated approach is best when data is not to be kept for any length of time, when reporting on current data only and not doing historical analysis, and with small amounts of data. The cost is at the data sharing level but can potentially be leveraged for reuse if there is a single data sharing clearinghouse governing the processes and defining the implementation architecture.

Point-to-Point Data Exchanges

In this method, predominant in the State today, each agency, government, or citizen creates a data project for each data sharing effort. The parties involved negotiate agreements independently, create individual reporting architectures, and create individual data sharing and translation processes.

An electronic data exchange must use a common type of technology for file transfers and formats. Formats like XML are becoming widespread because they allow file transfer through web services and they contain meta-data about each field, meaning it carries the data fields and the field characteristics. This allows more flexibility in reformatting the data at each end into a format that suits the user.

The primary advantage of this method is that it works well for small, well-defined situations and requires minimal upfront design. The primary disadvantage is that when more agencies are involved, the complexity of single agreements and exchange rules rapidly becomes overwhelming.

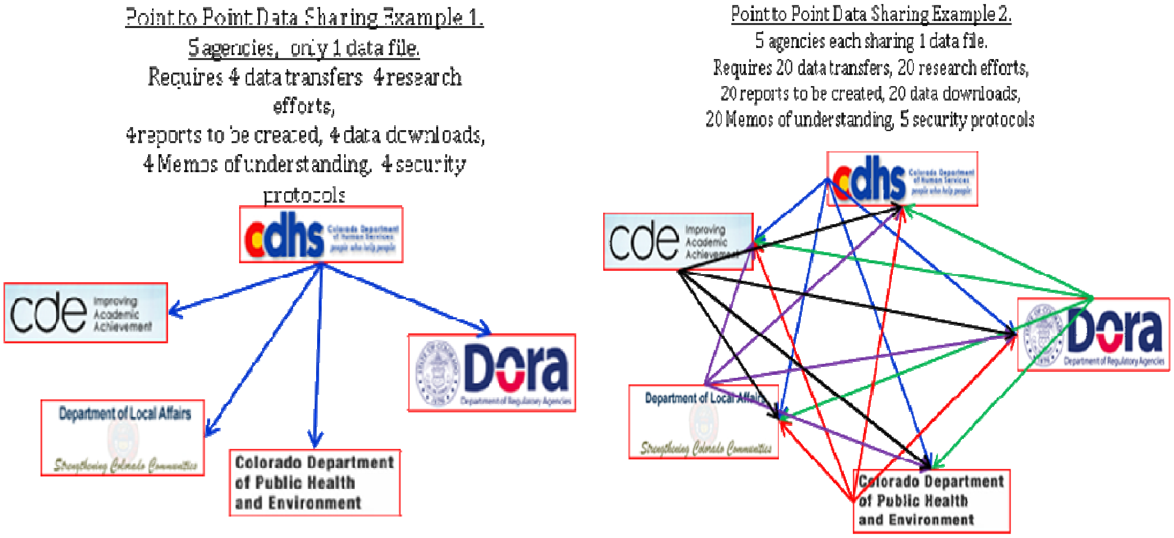

Point to Point Data Sharing Example 1.
5 agencies, only 1 data file.
Requires 4 data transfers 4 research efforts,
4 reports to be created, 4 data downloads,
4 Memos of understanding, 4 security protocols
cdhs Colorado Department of Human Services people who help people
cde Improving Academic Achievement
Dora Department of Regulatory Agencies
Department of Local Affairs Strengthening Colorado Communities
Colorado Department of Public Health and Environment
Point to Point Data Sharing Example 2.
5 agencies each sharing 1 data file.
Requires 20 data transfers, 20 research efforts,
20 reports to be created, 20 data downloads,
20 Memos of understanding, 5 security protocols
cdhs Colorado Department of Human Services people who help people
cde Improving Academic Achievement
Dora Department of Regulatory Agencies
Department of Local Affairs Strengthening Colorado Communities
Colorado Department of Public Health and Environment

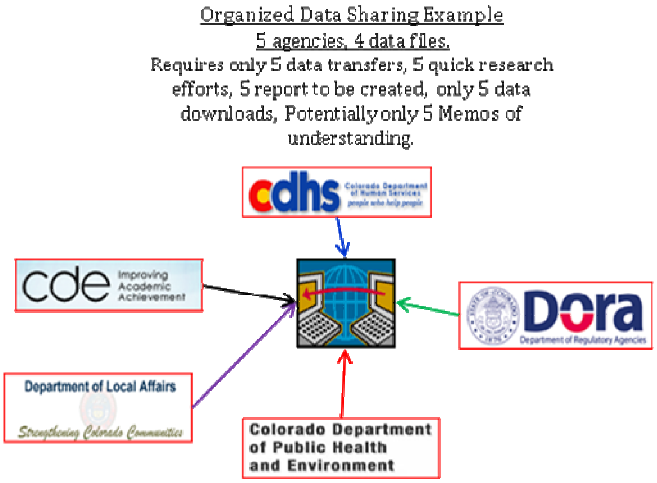
Figure 4-2

To address this problem, several data sharing initiatives have developed common data models so that each agency signs one agreement with a central management agency. Again, this works in well defined situations where there are either few data exchanges, or many that can all use the same format.

Hybrid Data Warehouse Collaboration

This architecture uses a combination of the three approaches listed above. It starts with a data sharing foundation, which creates a metadata registry. A metadata registry is a listing of all critical data definitions, databases and repositories and their system characteristics, as well as a unifying data model.

As specific data sharing initiatives are identified, this data sharing foundation coordinates the development of reusable physical data models, and may actually identify the need to create a system of record for key pieces of data. For example, the Human Services Database may be the system of record for unit data on all people receiving benefits from the State. If a school wants information on a student whose family is also receiving food stamps, it would include the information on that family from the Human Services Database as listed in the metadata registry.



Organized Data Sharing Example

5 agencies, 4 data files.

Requires only 5 data transfers, 5 quick research efforts, 5 report to be created, only 5 data downloads, Potentially only 5 Memos of understanding.

cdhs Colorado Department of Human Services people who help people
cde Improving Academic Achievement
Dora Department of Regulatory Agencies
Department of Local Affairs
Strengthening Colorado Communities
Colorado Department of Public Health and Environment

Figure 4-3

The advantage of the hybrid approach is that it provides the foundation and flexibility to identify and reuse previous efforts, the opportunity to satisfy the user's needs in a more reasonable amount of time, and an ongoing forum for leveraging today's efforts for the future.

Common Data Models

In analyzing possible data models, two Federal data models stood out - the Federal Enterprise Architecture (FEA) model and the National Information Exchange Model (NIEM). These two models stood out due to the increased collaboration between the State and its Federal partners, which often requires compliance with Federal standards for projects that are grant funded.

The Federal Enterprise Architecture

The FEA is designed to ease sharing of information and resources across Federal agencies, reduce costs, and improve citizen services. The FEA project began in February of 2002. Spearheaded by the Federal Office of Management and Budget (OMB), the purpose of this effort was to identify opportunities to simplify processes and unify work across the agencies and within the lines of business of the Federal government. Its use has been widespread across the government sectors as it has evolved. It is a collection of reference models that create a common taxonomy and ontology for describing IT resources.

National Information Exchange Model

The National Information Exchange Model (NIEM) program was launched in February of 2005 as a partnership between the US Department of Homeland Security (DHS) and the US Department of Justice (DOJ). The NIEM framework provides standard vocabulary; guidance and processes that help promote effective and efficient information sharing capabilities across organizational boundaries. NIEM is widely adopted within State and local governments, with at least 39 of the 50 states reporting NIEM-conformant information exchange projects.

Of all the existing data sharing models, NIEM is the most widely used and provides the greatest open source library of schemas. The Global Justice XML Data Model and the Environmental Protection Agency Exchange network are both implementations of the NIEM model. The Colorado Integrated Criminal Justice Information System (CICJIS), pre-dates the NIEM model, but follows similar guidelines. This organization is moving towards the NIEM model in future implementations. By setting the NIEM open–source repository as the Colorado standard, the State takes the greatest advantage of development work done by other U.S. governments, including the Federal government, other states, and other Colorado implementations. No other central model provided these advantages.

Please refer to Appendix F for additional detail on the FEA and NIEM.

Common Steps to Data Sharing

The Technical Subcommittee has identified some common steps in developing any data sharing initiative. Regardless of the size of the initiative, most went through all of these key activities.

1. Defining the Business Requirements for the Data Sharing Effort

Data sharing takes effort, and upfront investment on the part of each agency that takes part in a data sharing project. Agency managers must be able to show how data sharing benefits their mission or service provided to citizens in order to devote resources to these projects.
Further, the best design or architectural approach will depend on why the data is being shared. Knowing these details upfront can significantly streamline and simplify the design process, reducing current and future costs and development time and increasing measurable benefits.

2. Locating and identifying the data to be shared

Data sharing will require specific, detailed information about each data source in order to design a data sharing agreement that allows each agency to meet their legal and business requirements as data custodians. Since most systems in the State were not developed with data sharing in mind, collecting a directory of data available for sharing is a significant task.

3. Designing Methodology to Physically Share the Data

Once the data to be shared has been identified and agreements as to the security and accessibility have been reached, architecture for sharing the data must be determined. For most data sharing, reporting, and analysis purposes the original data in the source or operational system where it is input and originally stored is not in the best format nor is it easily accessible.

Common Enablers and Barriers to Data Sharing

The following Data Decision Integrity Wheel (see **Figure 4-4** below) illustrates the many inputs into the decision to physically build an accessible, secure, cost-effective, potentially reusable data store. Industry best practice shows that there are common enablers for making these decisions across organizations. These include:

- Use of an enterprise metadata registry
- Common identity management and resolution processes
- Common cyber security policies and practices

- Common Access Management for Data

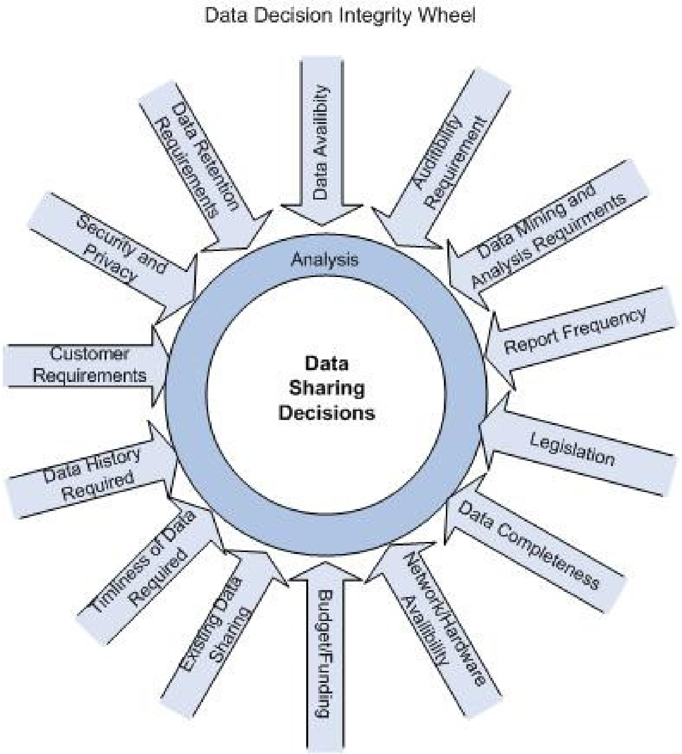


Figure 4-4

These are discussed in more detail below.

Enterprise Metadata Registry

It is recognized in the industry that a central point of data administration and data sharing can enable data sharing initiatives by reducing the cost of design and development and providing common, “open-source” models and systems.

The purpose of an enterprise metadata registry, sometimes known as a metadata dictionary, 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 data, facilitates easy mapping of data between computer systems, and lowers the costs of migrating to new systems including service-oriented architecture if desired. The entity with the responsibility for the registry will coordinate the agency data stewards and the input to the data dictionary to insure that at a minimum the following is accomplished:

- Each data element has a clear and unambiguous data element definition that is supported by the State agencies involved with the data.
- The origin of the data, i.e., the system of record, is identified for each element as well as the designated system for history, such as a data mart or data warehouse.
- The data stewards are identified for each data element.
- Each data element does not conflict with other data elements in the metadata registry.
- Each element defined as a code has clear enumerated value definitions.

- The data element defined is still being used.
- The data element is being used consistently in various computer systems.
- Version control of the data elements is in use.
- Adequate documentation on appropriate usage and notes for the data elements.
- The usage of the data element in each data sharing effort is documented.

Data quality will improve because all departments will use the same definitions and standards. Characteristics of data quality include security, timeliness, and accessibility.

Identity Management and Resolution

The Council's vision is for a much more substantially organized identity management framework. There must be a process in place for identity resolution, to ensure that the information being delivered as part of a data sharing exchange is data on the correct person. There are several potential strategies for doing this, including managing identities across disparate systems through a State-level unique identifier or creating a centralized repository of demographic data for all agencies to tap into, thereby creating a single "version of the truth" with regards to identity demographic information. There are also several off-the-shelf tools that can be utilized for identity resolution and scoring, identity validation and authentication, and identity fraud prevention. These strategies and tools will be investigated, and a final solution developed, as the Council's work continues.

Cyber Security Policies

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 2005, Colorado passed cyber security legislation to start to create a security baseline. This legislation, now contained in C.R.S. 24-37.5-401 through 406, establishes a Chief Information Security Officer (CISO) with authority to regulate cyber security for all State Departments excluding Higher Education. Rule 8 CCR 1501-5 ("Rules in Support of the Colorado Information Security Act") was passed in 2006. In addition, nineteen (19) Colorado Cyber Security Policies were established. These policies are listed in Appendix G, along with a summary of the impact of security policies on data sharing within Colorado.

While strong cyber security policies could enable data sharing, an immature program may create more barriers than it overcomes. Every program must have sufficient resources to meet the growing list of security requirements and stay current with standards, both in Colorado and at the Federal level if federally protected data is used.

Access Management

Improper access and authentication of users can result in direct and dire consequences to an application, system, and organization. OIT has developed the "Identity and Access Management: Assurance and Authentication Guidelines" (for additional information, please refer to www.colorado.gov/oit, "Policies/Standards" tab) to assist users in selecting an appropriate level of authentication to resist threats to their data, users, and organizations that could result from unauthorized use of system transactions. This approach emphasizes the development of authentication requirements based on risk. It is designed to approach the task from a business perspective, identify organization risk, and then match those risks to the appropriate technical solution. This is accomplished through a risk assessment for each transaction. The assessment identifies:

- risks, and
- their likelihood of occurrence

This section outlines the steps agencies should take to conduct a risk assessment of the e-government system.

1. Data Governance Analysis
2. Impact Assessment
3. Likelihood Assessment
4. Calculate Risk Rating
5. Determine Security Level

From the Risk Assessment, agencies can then determine the appropriate Assurance Level for the data or transaction in question, as well as appropriate levels of Identity Proofing and related Authentication Technologies.

Conclusion

A hybrid approach represents the greatest flexibility to the State for data sharing efforts because it allows for *planned* reusable data sharing (i.e., the combination of the physical data warehouse, the federated/virtual database and the point-to-point architectures). Implementation will be based on specific business requirements, and will follow guidance from the State Enterprise Architect.

Section 5 - Recommendations

The Council appreciates the opportunity to examine this important issue and to provide its recommendations to the State CIO. This report is just the beginning, and there is still much work to be done beyond the work accomplished by the Council. Work needs to be continued and finalized in two main areas: **governance** and **technology**. Overall, a strategy must be developed around data sharing and data management. Is policy effectiveness the primary goal or is operational efficiency? Is Colorado trying to improve access to data or protect privacy at all costs? Each is valid, and not necessarily mutually exclusive. However, there are tactical and resource decisions that must be made in order to accomplish any of the above, and accomplishing one goal does not necessarily accomplish all of them.

Two things absolutely must be done above all else if enterprise data sharing is going to be successful within the State: the State must understand all the data it has, and there must be a formal governing body in place to oversee the policies and procedures that govern the management and sharing of this data.

In documenting and knowing the data owned by the enterprise, more effective data sharing projects can be implemented, more effective policy making can occur, operational efficiencies can be gained, and resources can be utilized more effectively. This work can be done in phases, as new data sharing initiatives are begun, as new systems are implemented, or as system upgrades are done. Systems must be documented in the same way, and that information fed to the Enterprise Architecture group and data team in OIT.

Building a formal and binding data governance framework must include people, policies and processes, technologies, and, standards and definitions:

- People – data council, data owners, data stewards, analysts (business and data), developers, architects
- Policies and Processes – data quality management, data security, privacy, exception handling, stewardship guidelines
- Standards and Definitions – data definitions and context, technology standards, enterprise data models, master reference data
- Technologies – data integration, data profiling, data cleansing, metadata management, data modeling, feedback loops for quality control

The remainder of this section contains recommendations in the following areas: governance and process; technical; and legal.

Governance and Process

The scope of enterprise data sharing is large and there are many stakeholders, including the Executive Branch of State Government, State agencies headed by elected officials, the Judicial and Legislative branches of State Government, local governments, Federal government partner agencies, educational institutions, private businesses, non-governmental organizations, and the citizens of the State. With the breadth of responsibilities these entities have, governance over the exchange of data is critical in order to ensure that each entity and the subject of the data are protected when personal data is shared outside of the collecting agency.

A formal data governance process will describe the "rules of engagement" by which all will play regarding data management and data sharing. Setting expectations as to how data is handled once it is shared is essential. There are numerous mandates from Federal and State governing bodies that determine how data should be handled. The Health Insurance Portability and Accountability Act (HIPAA) is but one example.

The goals and objectives of Governance are as follows:

- To develop ‘Data Usage Policies’ for sensitive personal data received from other entities.
- To develop performance standards for how entities receiving data shall receive, process, store and dispose of the data.
- Initiate, modify and terminate data sharing agreements between entities based on either entities ability to meet the security, performance, and usage policies agreed to by the entities
- Provide a central point of adjudication and/or advice on data sharing proposals between entities.
- Involve and understanding the needs of the stakeholders
- Ensure consistent application of rules, guidance, requirements, safe guards, and resources to Data Sharing initiatives

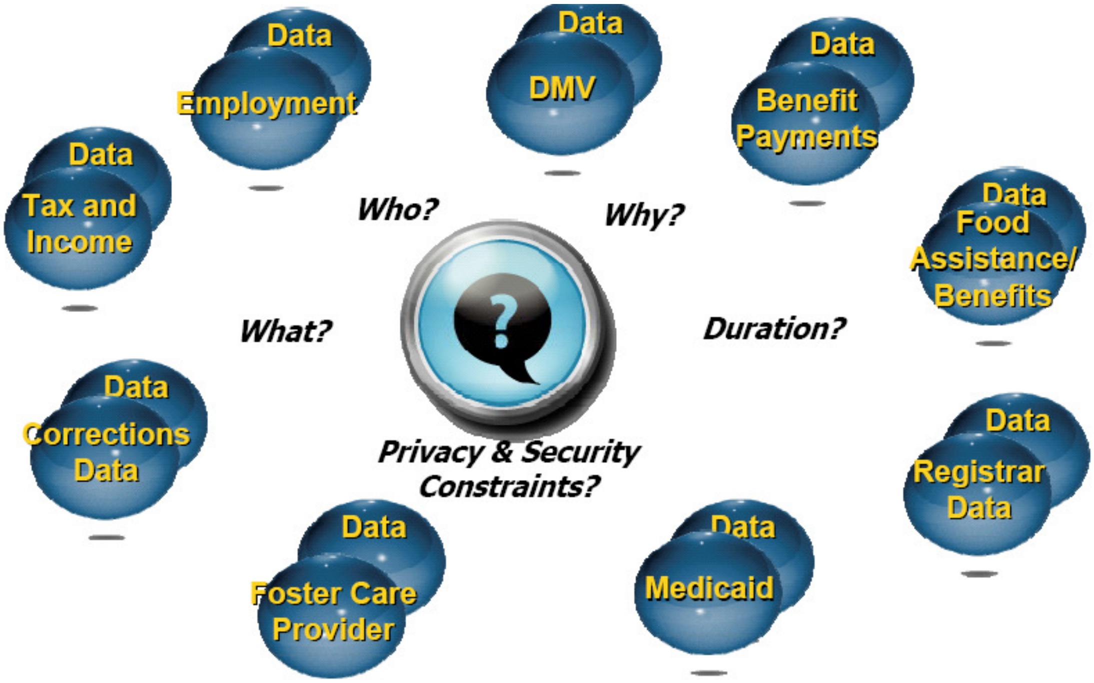


Figure 5-1: With no Centralized Approach to Data Governance

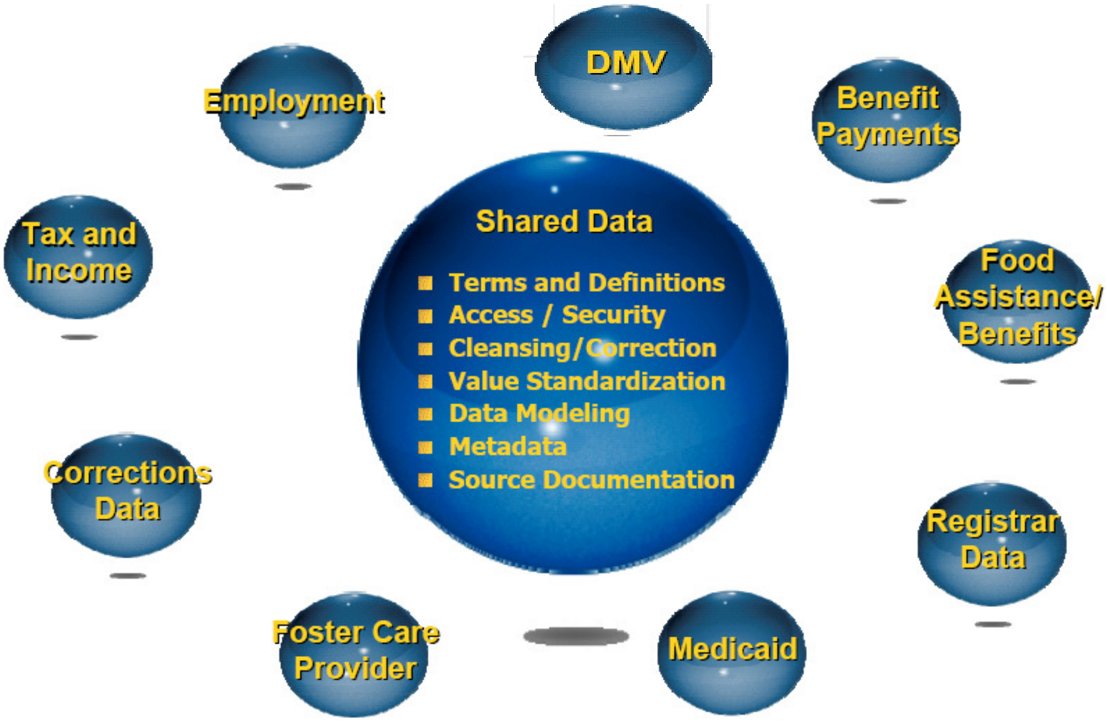


Figure 5-2: With a Centralized Approach to Data Governance

Establish An Advisory Board

The Council recommends that a formal advisory council be established to provide oversight, policies, and guidance to the enterprise. The advisory council would recommend to the State CIO and Enterprise Architect policies, direction, and priorities for the State's data sharing initiatives. Specific activities which the advisory council would engage in include (but are not limited to):

- Creation of formal business documents to set the direction for data sharing in the State.
  - Includes strategic vision document, communications plan, policies, goals and processes.
- Formalization of the request process for data sharing between agencies.
  - Ensure that all data sharing initiatives are inventoried and managed under the same set of guidelines throughout the Executive Branch
  - Includes escalation and adjudication processes
  - Includes retention and destruction policies
- Creation of an umbrella Memorandum of Understanding (MOU) for all Executive Branch agencies
  - Determination and resolution of ownership and authoritative source issues
  - What agencies are responsible for what systems and data of record?
  - Where are the official sources of certain unit data and what are the rules around sharing such data?
- Adoption of standards for all business processes pertaining to data sharing among agencies.
  - Definitions of all key terms

    - Data management and handling policies
    - Metrics should be developed to track the progress and success of the State's data sharing initiatives.
- Planning for organizational change and transformation/change management. This includes developing a plan to staff and operate the State's data sharing initiative.

Appropriate and efficient data sharing cannot occur without a strong data governance model in place. The most effective data governance correctly aligns people, processes, and technology to convert data into strategic information and knowledge assets for the State government enterprise. The creation of a permanent data governance advisory council would be a strong move forward in this direction for the State.

An example of a data governance structure is below.

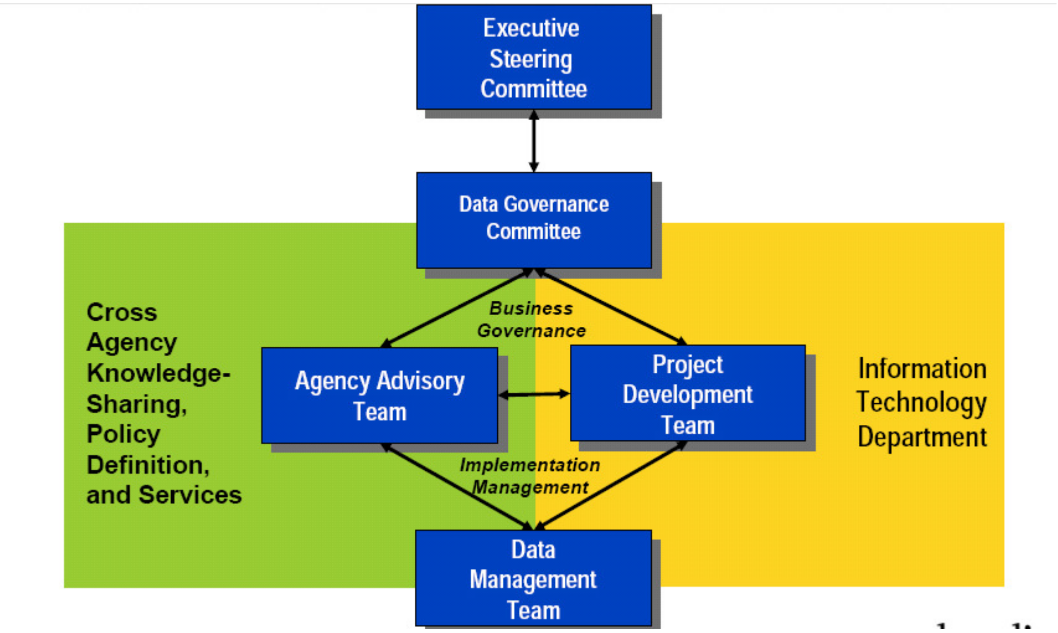

Executive Steering Committee
Data Governance Committee
Cross Agency Knowledge-Sharing, Policy Definition, and Services
Business Governance
Agency Advisory Team
Project Development Team
Information Technology Department
Implementation Management
Data Management Team

Figure 5-3

Within OIT, an advisory board should be established which will act as the overall steering committee for the program. The Council should be comprised of Executive Branch agency representatives, as well as representatives from other stakeholders, including local governments and citizens. Participating State agencies should have business, technical and policy/compliance/legal subject matter experts serving in this role. Executive leadership must be engaged as a critical component of success in order for collaboration and consensus to occur. The lines of business managers have the primary responsibility for their systems, not the technology teams. Those individuals who are closest to the data on a day-to-day basis maintain great responsibility for the information.

A Data Governance Committee should be established within the Information Sharing Advisory Council to establish the policies and procedures described above. This group will work with both the business side and technical side for this work.

Master Data Sharing Agreements

Partnership agreements officially recognize the exchanges that will be taking place. Well articulated agreements, which take into account legal issues and create a shared understanding of the data and its analyses, lead to productive and long-standing data partnerships. The agreements should be as comprehensive as possible, covering keyword definitions, the purpose/description of the project, parties and signatures, permitted/non-permitted uses, rules for access, limitations on disclosure, security requirements, the retention and disposition of data, and the penalties for misuse.

Enterprise Processes for Data Requests

The Council recommends that common policies be established across all Executive Branch agencies, where possible, for the collection, management, maintenance and exchange of information. Some of this work was done by the Council, but this should review and documentation should continue. As processes and data stores are documented, there are long-term opportunities to increase efficiencies and analytical capabilities, and to provide more efficient constituent services through business process re-engineering.

Well-established and documented procedures for accepting and handling data requests help to streamline request processes and remove ambiguity about the process. Processes must be documented and available to the public. It is advantageous to consistently assign data requests to specific offices and staff to ensure coordination and accuracy. Data responses will be timelier because data requests will be processed more efficiently under a single set of business processes, by staff who are well-versed in those processes. Request logs should be kept to track requests, requestors and fulfillments.

Data Ownership and Administration

In reality, the data in the State's systems are owned by the State government. Each department has responsibility for a subset of the State's data. A data administration program should be created to develop common data definitions and identify the stewardship of the State's unit data. Data stewards, or often called data custodians, in each organization should be identified from the business side of each agency, and should become part of a committee whose responsibility is to create common definitions and establish stewardship bounds. Data stewardship resides with the data content experts within each agency and a method for stewardship resolution must be created when data is duplicated and stewardship issues arise. In cases of duplication and overlap a method to establish the source of record or hierarchy of sources must be determined.

Metrics Development

A system of metrics should be developed to help monitor performance of the information sharing program across the enterprise. Categories of metrics could include:

- Reduction in redundant data collection
- Reduced collection burden on agencies
- Improve "one-version" data quality by implementing common definitions and standards
- Improve access to and security of data because processes will be standardized and policies will be explicit and consistent
- Turning data resources into information assets that can be managed for effective decision and policy making
- Better coordination of services and management

Some examples of metrics that could be implemented and monitored are below:

- *Measure 1*. Number of agencies that complete the “as is” enterprise information architecture benchmarks.
    - Measure type, Measure frequency, data source & calculation, baseline, target
- *Measure 2*. Number of agencies adopting statewide data exchange standards
    - Measure type, Measure frequency, data source & calculation, baseline, target
- *Measure 3*. Number of collaborative IT solutions deployed
    - Measure type, Measure frequency, data source & calculation, baseline, target
- *Measure 4*. Number of formal partnerships (multi-agency, local government to State government, State agency to higher education, private)
    - Measure type, Measure frequency, data source & calculation, baseline, target

Technical Implementation Strategy

The long term goal of OIT is to standardize statewide enterprise architecture (EA) as a means of connecting individual agency goals to a shared information technology strategy so that the State can realize the return on its IT investment. EA is a key governance discipline providing oversight of information technology investments, standards, processes, alignment of business and IT objectives. It is also responsible for planning and implementing the various architectures required to support business objectives. Since the passage of SB 08-155, a statewide Enterprise Architect has been named, and work has begun to align vertically-oriented organizations into an enterprise-focused organization. The Council recommends that all technologies needed for enterprise data sharing be vetted by the Enterprise Architect and any technology review board established by the EA team.

The following recommendations recognize the current fiscal climate and budget situation across State government. Some recommendations are near-term, and others are long-term. The Council recognizes that it could be years before all agencies to come into compliance with these standards, but suggests the following strategy for implementation:

- Any new data sharing initiative must be approved by OIT and must implement immediately the recommendations below as part of their project management and systems development.
    - The STRAC initiative will follow these guidelines, per discussions with OIT and the STRAC Board of Directors.
    - CCYIS will follow these guidelines per discussions with OIT and the CCYIS leadership.
- As existing major data sharing initiatives begin to examine system or architectural upgrades, they must come into compliance with these standards.
- The State’s standard MOU must be completed and executed by the participating agencies prior to the data exchange.

Step 1: Immediate Data Exchange Standardization

The Council recommends that the following be done immediately by OIT and all State agencies:

- Adopt the FEA to document systems; adopt NIEM as the data exchange mode; and, utilize XML as the data exchange format.
- Begin building a metadata repository and registry based upon the information documented for the Council.
- Supply baseline document templates to all agencies beginning or considering data sharing initiatives. Require that documentation be provided to OIT for central storage and management.
- Continue the review of all tools currently in place for data management and exchange. Recommend standards to the Enterprise Architect.

- OIT will provide oversight and guidance to new data sharing initiatives regarding the new policies and procedures.

The State should immediately move to a common data exchange model based on the FEA architecture, and use the NIEM models and the NIEM data exchange as a central coordination point. NIEM requires the use of XML as the format for data exchanges. This approach allows for the greatest flexibility initially, while still providing a centralized exchange structure to support multiple development initiatives. Until Colorado architectural standards are approved, the existing NEIM models should be examined by new data sharing initiatives and reused if possible.

Information Exchange Package Documentation (IEPDs) are available for reuse immediately, at no cost to the State. An IEPD is a complete definition of an Information Exchange Package (IEP), usually composed of schemas for data exchange, and documentation for understating the business context and usage of the data being exchanged. Today there are NIEM IEPDs that are available for reuse on the NIEM website, at http://niem.gov.

We recommend XML as the data exchange format because it is the most commonly used and most universal of all the languages. Additionally, most Federal government grant programs require the use of XML. Many State and local government agencies use legacy applications; most of these will have XML capabilities built in, whereas some of the newer and more advanced mark up languages may require more work to implement in older databases and applications. XML files can be transmitted through web services or as flat files through a file transfer mechanism. This flexibility is important for lower cost, ease of implementation, and flexible security options.

In addition, to lay the foundation for a reusable and more cost effective state architecture, all users of software products should be required to supply a metadata dictionary to OIT.

Step 2: Establish an Enterprise Data Team within OIT

The Council recommends that the following be done in the near-to-short term:

- Begin building a dedicated OIT enterprise data architecture team
- Orchestrate the fulfillment of approved data sharing requests
- Negotiate enterprise licenses for recommended tool sets to realize economies of scale in pricing
- Issue Requests for Information (RFIs) for any tools needed to facilitate enterprise data sharing that are not currently in use by any State agency
- Develop data matching methodology
- Continue to work on data inventories
- Begin building the logical data models, starting with demographic data

Additionally, the Colorado Cyber Security Rules and Policies provide the first step in enabling secure data sharing across State agencies. To make these policies the critical data sharing enabler they need to be, Colorado must:

- Develop a Data Classification Policy to assign security levels to the State's data.
- Ensure that Colorado standards for confidential and PII data are consistent with the other applicable regulations.
- Fund security enhancements statewide to bring Colorado IT systems into compliance with its own standards and all applicable regulations.

Colorado has a good start on cyber security and needs more support and development to sustain future data sharing initiatives and growth, but cyber security remains a moving target, and the State must remain on constant vigilance to protect its systems.

Step 3: Implement a Reusable Data Sharing Architecture

The Council recommends the following for a longer-term strategy:

- Begin identifying possible data for consolidation (e.g., name, address)
- Assist in making final decisions regarding physical models
- Coordinate with EA team for architecture decisions
- Implement strong identity management policies and controls

Legal and Regulatory

The Council's recommendations on the legal and regulatory front fall into two primary areas: statutory, regulatory and policy revisions; and, the creation of a State Chief Privacy Officer. These are discussed below.

Statutory Revisions

In order for there to be more data sharing throughout State government, the barriers that currently exist, and that can be removed, must be. Some Federal laws, such as HIPAA and FERPA, cannot be changed. Interpretation of these laws may allow the sharing of certain data, but only within the guidelines of the interpretation provided by the Federal government. It should be possible to get an opinion on each of these Federal policies with regards to data sharing from the State Attorney General that would be applicable equally to all impacted State agencies. This would eliminate any conflicting agency interpretations. There are other Federal statutory conflicts that could be resolved by legislative intervention, but this is not likely.

Limitations on data sharing due to *State law restrictions* may more easily be addressed. The statutory limits include, but are not limited to, provisions in some laws that information collected may be used only for the specific purpose for which it is collected and may not be used for any other purpose without the express written permission of the individual who provides the information. These limitations can be expanded upon by revising statutes to allow collection of information for legitimate *State purposes*, rather than a specific individual purpose.

Finally, some rules and regulations can be revised by the responsible agency to allow data sharing while keeping in mind the need to protect PII from unauthorized disclosure. Internal agency policies, likewise, can be revised to allow sharing of data while still protecting the privacy of individuals and the integrity of the data.

Chief Privacy Officer

The position of a State Chief Privacy Officer (CPO) should be formalized. The CPO's responsibilities should include, but not be limited to, creating policies and procedures to assure PII is protected during all stages of data sharing and data use. The position of CPO should reside in the Office of the Attorney General, as this position is legal in nature. The CPO is responsible for working with electronic data, as well as working with agencies regarding data maintained on mediums other than electronic, including paper.

The CPO's mission should be to minimize the impact on the individual's privacy, particularly the individual's personal information. The CPO will not merely look to protect electronic PII but all PII held by all agencies, regardless of form. The CPO will be the State's authority regarding compliance with the letter and spirit of State and Federal laws promoting privacy and should have authority for the following:

- To require compliance with State and Federal privacy laws and to create policies and guidelines for agencies to follow regarding protection of PII;
- To require agencies to perform Privacy Impact Assessments;
- To hold agencies accountable for failure to comply with CPO policies, guidelines and directives; and,
- To provide education and outreach to build a culture of privacy and adherence to responsible protection of PII.

As an initial recommendation, all agencies should continue to assess their data, and identify laws, regulations and policies that currently regulate or hinder data sharing by their agency. This assessment should consider all data they maintain, paying particular attention to PII and electronically stored PII, but it should not be limited to these.

Section 6 - Next Steps

The Council has identified the following items that can begin immediately with little to no cost to the State, but continue the momentum begun by the HB 08-1364 initiative.

Legislation

As of this report, a follow-up piece of legislation, HB 09-1285, has recently been introduced in the 2009 legislative session. The purpose of this bill is to establish the Government Data Advisory Board (GDAB) that will advise the State CIO on:

- The ongoing development, maintenance, and implementation of the interdepartmental data protocol;
- Best practices in sharing and protecting data in State government;
- 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;
- Rules and procedures for responding to data requests submitted by an entity outside of State government;
- A schedule of fees that OIT 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.
- On other issues pertaining to data sharing

This legislation will formalize the data sharing effort within the State.

Accelerate Deployment of the CIO's IT Consolidation Program

The leadership and authority of the Governor's Office are particularly critical to State support for improving the sharing of information. A culture of collaboration around information sharing and improved performance is key to the State's direction to improve services, reduce redundancy and further knowledge sharing and sustainable quality.

- Adopt the FEA to document systems; adopt NIEM as the data exchange mode; and, utilize XML as the data exchange format.
- Begin building a metadata repository and registry based upon the information documented for the Council.
- Supply baseline document templates to all agencies beginning or considering data sharing initiatives. Require that documentation be provided to OIT for central storage and management.
- Establish enterprise licenses for all tools needed to implement data sharing across the enterprise.

Potential Issues and Risks

The Council identified several potential issues and risks but did not have time to investigate how to manage or resolve these. These are items that will be turned over to the GDAB to continue work on, and include the following:

- Culture and change management
    - Removing internal agency barriers for sharing data
    - Redeveloping cultural barriers existing at the agency level and the “it’s mine” mentality
    - Achieving buy-in from Executive Directors by explaining “what’s in it for you”
- Funding
    - Even when funds were/are available, without culture change, funding for many of the recommendations and needs around data sharing, data quality, data definition and registry may not be available or may not be a priority.
    - In a challenging economic climate, such as the one that exists at the time of this writing, funding may be out of the question.
    - Because it is difficult to justify or measure the projected impact of some of these recommendations, and/or difficult to explain to non-technical managers, analysts, officials, and decision-makers, the importance of these recommendations may be lost or not fully understood.
- Resource staffing
    - OIT is consolidating IT staff from all executive branch agencies. Not only is that a culture shift for those staff members, an environment that fosters data sharing and cooperation within OIT must also be created.
- Separation of powers
    - Unless legislation and other regulations/protocols provide the mechanisms to enable cooperation, and are binding, across all branches of State government, data sharing, quality, definition and registry may be difficult to achieve.
- Liability
    - What is the State’s liability if PII and other private information is lost, compromised, or otherwise mishandled?
    - Does the State currently meet the Federal and State privacy and security standards for data and criminal justice information? If not, would its liability increase if data sharing were pursued without first remedying that situation? Can it be remedied? What are the challenges in order to do so?
    - Is the State able and willing to take the necessary steps to include, and enforce, strong measures in memoranda of understanding with its partners, and within its agencies?
    - Is it willing and able to bear the costs of securing and/or enhancing its systems, policies, and/or protocols in order to meet those measures?

Section 7 - Findings of the Student Unique Identifier Working Group

COLORADO DEPARTMENT OF EDUCATION

201 East Colfax Avenue • Denver, Colorado 80203-1799
303.866.6600 • www.cde.state.co.us

Dwight D. Jones
Commissioner of Education

Robert K. Hammond
Deputy Commissioner

Kenneth R. Turner
Deputy Commissioner

February 12, 2009

Michael Locatis, CIO
Governor's Office of Information Technology
1580 Logan St., Suite 200
Denver, CO 80203

Dear Mr. Locatis:

With the passage of HB 08-1364, a cross-state agency workgroup was appointed to 1) provide recommendations to the legislature on the assignment of a uniquely identifying student number to children who receive state-subsidized or federally subsidized early childhood education services; 2) adopt protocols by which the Colorado Department of Education (CDE), the Colorado Department of Human Services (CDHS), school districts, charter schools and the early childhood councils shall cooperate in assigning the numbers; and 3) to consider methods by which to encourage and facilitate the assignment of uniquely identifying student numbers to students who are receiving early childhood education services that are not subsidized by state or federal funding.

Pursuant to C.R.S. 22-2-134, I submit to you the attached report, which addresses each of the three tasks outlined in the legislation. The report considers the advantages and challenges of implementing an identifier and reviews existing state data systems that currently assign identifiers to children from birth to five years old. The workgroup has presented two diverse possible protocols to provide the means for the assignment of numbers. The report highlights the need for clarification of purpose and cost analysis of implementation before further action – including the promulgation of rules for the assignment of numbers – can be determined.

Sincerely,

Dwight D. Jones
Commissioner

Introduction

With the passage of HB 08-1364, a cross-State-agency workgroup was appointed to 1) provide recommendations to the legislature on the assignment of a uniquely identifying student number to children who receive State-subsidized or Federally subsidized early childhood education services; 2) adopt protocols by which the Colorado Department of Education (CDE), the Colorado Department of Human Services (CDHS), school districts, charter schools, and the early childhood councils shall cooperate in assigning the numbers; and 3) to consider methods by which to encourage and facilitate the assignment of uniquely identifying student numbers to students who are receiving early childhood education services that are not subsidized by State or Federal funding.

The workgroup included members from CDE, CDHS, the Colorado Department of Public Health and Environment (CDPHE), the Colorado Department of Heath Care Policy and Financing (HCPF), the Office of Information Technology (OIT), and Qualistar. The group met weekly for discussion and knowledge sharing, and to draft a report to present to stakeholders who included representatives from school districts, community health services, and community child care services. This is the final report, which addresses each of the three tasks outlined in the legislation.

Premise

A uniquely identifying student number, by itself, won't provide insight about a student's progress from preschool through higher education. The links connecting the data sets are what makes the number meaningful. The success of this legislation lies in the determination of what the data will be used for, and then building the systems and data-sharing protocols to support that mission. If the goal is to assign a number for the purpose of longitudinal analysis of a student's educational progress, then the number currently assigned for educational analysis, the State Assigned Student Identifier (SASID), should be utilized. If the goal is to assign a number for the purpose of tracking an individual's activities beyond the boundaries of public education, to include all State and Federal services that an individual might access, then a different number should be issued.

To ensure the program's success, the resulting identifier should elude single agency identity and have State ownership. Agencies investing resources in a data-sharing program should feel that the cost associated with implementation will pay off in the long run.

It is essential that the data-sharing program for which the number is to be assigned will be a program of which parents will want their children to be a part. The benefit of the research in which the number will be utilized should outweigh any privacy concerns parents might have, even though the results of outcome studies for their children will be applied to future generations.

Considerations of Implementation

Advantages of assigning a uniquely identifying student number

- A uniquely identifying student number would support the efforts of Governor Ritter's P-20 council to enable the longitudinal tracking of student progress from early childhood to postsecondary education and into the workforce.

- Longitudinal analysis of an individual's progress could show the effectiveness of taxpayer investments in the programs and services accessed by the individual.

- Longitudinal tracking of services accessed by individuals could produce quality comparisons of service providers and programs.

- Quality comparisons of service providers and programs could empower parents by providing them with the information needed to make the best choices for their children.

- Inter-agency data sharing could lead to better service and support for Colorado families due to increased efficiencies in transactions, a reduction in duplicate data collection efforts, and lower costs.

Challenges of assigning a uniquely identifying student number

- The cost associated with implementation of a new system that would assign numbers, or updating an existing system to accommodate assigning numbers, and integration of the datasets, would be material.

- If all information pertaining to an individual is housed under one number, the risk of disclosure of sensitive information is higher. In the event of a security breach, the number could reveal information about all services and education the individual has received.

- Some parents would not want their children to be tracked with a number, and would not want others to have the ability to use the number to access sensitive information about their child.

- The difficulty of matching the correct individual to records contained in multiple stand-alone systems could be significant.

- Currently there is no one entry point at which all children who receive early childhood education services, funded and non-funded, are assigned a uniquely identifying number. The development of a number would have to cross multiple agencies' systems.

Data Systems Currently in Place

A data integration project is in the early stages of implementation at CDHS, which will integrate data sets from the department's data systems, for the purpose of facilitating visibility and accountability across services. The aim is to achieve better social outcomes, ensure financial integrity, mitigate fraud, error and abuse, reduce delivery costs, make programs more accessible, and move from program-centered to client-centered service. [1] Using business intelligence and performance management software, the department will have access to real-time information that will allow them to make better business decisions, the ability to act with speed and efficiency, and the ease of compliance reporting using one source for information. Comparable advantages could be realized from a system designed to associate early childhood services with educational outcomes, especially if all services in which an individual and their family were enrolled were included, such as Medicaid, food benefits, financial assistance, child welfare services, correctional services, etc.

The workgroup identified several existing State data systems, presented in the accompanying chart on page 4, that currently assign uniquely identifying numbers to subsets of Colorado's population. Not all of Colorado's 430,838 children aged birth to five years[2] is included in any one of the individual systems, and

[1] 'Colorado's Approach to Using and Implementing BI Tools', Ron Ozga, OIT/CDHS CIO, PowerPoint presentation, Governor's Office of Information Technology

[2] According to the Colorado Division of Local Government, the population of children age 0 to 5 in 2008 was **430,838** http://dola.colorado.gov/demog_webapps/population_age_gender

none of the systems is focused specifically on all children who receive early childhood education services. Note that an individual child may have records in more than one of the systems represented. Children enrolled in Colorado’s Head Start program do not receive a unique number for that program, but could have numbers assigned through other State data systems.

The workgroup determined that although none of the systems reviewed currently includes all students who receive early childhood education services, it would be important to utilize an existing system for the assignment of a uniquely identifying student number, considering the significant cost of implementing a new system and the current economic environment. The workgroup looked at the data elements collected by each system to find common attributes and to determine places where the systems might overlap. The purpose of this was to assess which of the systems, if any, might be modified to assign uniquely identifying student numbers.

TABLE 1 HB 08-1364 Systems Index

State Agency	System	Total Representation of Age 0-5* as % of Population Age 0-5	ID#	SS#	Mother's SS#	Alien Reg. #	Local ID (LASID)	Sex	Last Name	First Name	Middle Name	Date of Birth	Title	Date of Death	System Interacts with the Following System(s)
CDPHE	Birth Certificate Database	70,700 births per year[3]x 5yrs 82%	X	O	X			X	X	X	X	X	X	X	Immunization registry, newborn hearing and metabolic screening
CDE	RITS	36,500 8.5%	X				X	X	X	X	X	X			Automated Data Exchange at CDE and school district student information systems
N/A	Head Start	8,800 2%													Head Start does not assign ID numbers to track children
CDHS	SIDMOD**	205,448 47.7%	X	O	O	O		X	X	X	X	X	X	X	Interwoven with every benefits system in CDHS and used by HCPF for Medicaid*
CDHS	CHATS*** (CCDBG)	16,124 3.7%	X	O	O	O		X	X	X	X	X	O		SIDMOD
CDHS	CBMS	140,741 33%	X	O	O	O		X	X	X	X	X	O	O	SIDMOD
CDHS	CO Trails	14,404 3.3%	X	O	O	O		X	X	X	X	X	O		SIDMOD
CDHS	ACSES	26,058 6.0%	X	O	O	O		X	X	X	X	X	O		SIDMOD
HCPF	MMIS	141,045 children in CHP+ or Medicaid 38%	X	For some clients, not all				X	X	X	X	X			CBMS

Key: X=Mandatory, 0=Optional

*The number represented in any system includes individuals who might no longer be active in the program, and who might no longer reside in Colorado
**SIDMOD generates ID numbers for CHATS, CBMS, CO Trails, and ACSES
***All of the Child Care and Development Block Grant participants (CCDBG) are included in CHATS

[3] Colorado Health Information Data Set: Birth Statistics http://www.cdphe.state.co.us/scripts/htmsql.exe/cohid/NatalityPub.hsql

Systems Description

- Birth Certificate Database, housed at CDPHE, issues unique birth certificate numbers for children who are born in Colorado, and also stores birth certificate information for children born out of state to Colorado residents.

- Record Integration Tracking System (RITS), housed at CDE, assigns SASIDs to Colorado's Pk-12 public school children; including students enrolled in the Colorado Preschool Program (CPP), special education early childhood services, and preschools located in public school buildings which include Title 1, Head Start, parent funded and school funded programs.

- State Identification Module (SIDMOD), housed at CDHS, assigns unique identifiers to children who enter any benefits system in CDHS and these identifiers are used also by HCPF for Medicaid

- Child Care Automated Tracking System (CHATS), housed at CDHS, is an eligibility and payment system for all childcare programs, including low income childcare, Colorado works childcare, employment first childcare, and child welfare childcare. CHATS uses the SIDMOD ID#.

- Colorado Benefits Management System (CBMS), housed at CDHS, is an eligibility and payment system for self-sufficiency programs including food benefits and financial assistance. CBMS uses the SIDMOD ID#.

- Colorado Trails, housed at CDHS, is the statewide case management system for Child Welfare Services, and is the Division of Youth Corrections case management system. CO Trails uses the SIDMOD ID#.

- Automated Child Support Enforcement System (ACSES), housed at CDHS, is utilized by the Colorado Division of Child Support Enforcement to conduct the business of providing child support services. This system is fully integrated into the business of child support enforcement, and is essential for the management of all the main processes required by the Federal Office of Child Support Enforcement (OCSE). ACSES uses the SIDMOD ID#.

- Medicaid Management Information System (MMIS), housed at HCPF, processes and reimburses Medicaid claims. MMIS uses the SIDMOD ID#.

Recommendations

The workgroup presents two possible scenarios for how CDE, CDHS, school districts and charter schools shall cooperate in assigning the uniquely identifying student numbers.

Scenario #1

The first scenario involves upgrading or replacing an existing system in order to accommodate the assignment of a uniquely identifying student number to all children aged birth to five who receive State subsidized or Federally subsidized services. The idea is to include as many children as possible, as early as possible, in order to capture the most history for the purpose of longitudinal analysis. SIDMOD is the most far-reaching of the systems, as it is interwoven with every benefits system in CDHS, and is used by HCPF for Medicaid. The group discussed the viability of updating SIDMOD for the purpose of assigning IDs to all children who receive early childhood education services. It was determined that SIDMOD would need to be upgraded or replaced. SIDMOD is a 22-year-old system that is no longer funded or maintained. An upgrade to SIDMOD would involve updates to all of the systems it touches. A feasibility study and a needs assessment should be conducted to determine the fiscal impact of an upgrade or replacement of SIDMOD. Even if it should be determined that SIDMOD will not be the system used to assign the numbers, it is clear that in order to support the business continuity of the systems for which it assigns numbers currently, SIDMOD must be maintained. If SIDMOD should fail, these systems would be crippled, impeding the ability to conduct comprehensive P-20 and beyond longitudinal data analysis.

Protocol

All systems would act as hubs, continuing to collect data specific to the needs of the programs they support, but would interface with SIDMOD to request the identifier number. All systems would use the SIDMOD ID as the key to their data sets, replacing any unique identifiers used previously by each system that were not issued by SIDMOD. Using one key identifier for all datasets would facilitate data sharing.

Pros	**Cons**
SIDMOD already includes 47.7% of Colorado children aged 0-5.	SIDMOD is an old system in need of upgrade or replacement, the cost of which would be significant.
An agreement between SIDMOD and vital statistics would allow SIDMOD to assign an ID to all new Colorado births, extending its reach further.	The stand alone systems don't include all of the 0-5 population, including Head Start, those born out of State to non-residents, and those who never received services from CDHS or HCPF, and who do not enroll in public education.
Access points for requesting SIDMOD IDs are already established at county departments of social services and would be considered a manageable number of hubs	County departments of social services might not have the capacity to accommodate increased business volume if asked to request SIDMOD IDs for programs in addition to the ones they currently handle the intake for.
	There would be costs associated with linking existing ID numbers to the SIDMOD ID.
	Existing systems would have to undergo data conversion without losing continuity.
	Identifiers used in historical data would have to be converted to the SIDMOD ID for the purpose of

	longitudinal analysis.

Scenario #2

The second scenario proposes that the agencies continue to assign their own program-generated identifiers to their data sets, and the P-20 data system would integrate the data received from each system.

Protocol

Each system would continue to issue separate, program-based ID numbers, but when the child enters public school, the data sets would be integrated by the P-20 data system. Agencies would submit data to the P-20 system. The SASID generated by the RITS would be the key attribute utilized by the P-20 data system to link the data sets.

Pros	Cons
One agency doesn't carry the burden of assigning a number to every Colorado child aged birth to 5.	Matching multiple IDs from the separate data sets to the correct student could be challenging, and would require 1-2 FTE to manage instances requiring manual review.
Wouldn't have to train a new group of end users, as they would already be familiar with their existing applications.	The stand alone systems don't include all of the 0-5 population, including Head Start, those born out of State to non-residents, and those who never received services from CDHS or HCPF, and who do not enroll in public education.
Existing ID numbers would remain in place.	It could be difficult to enforce data integrity standards across agency lines.

Summary

Scenario #1 involves upgrading or replacing the SIDMOD system for the purpose of assigning identifiers in exchange for the program-specific identifiers currently in use across State agencies. Scenario #2 suggests that the P-20 data system should utilize SASIDs for the integration of data sets from stand-alone systems, which would continue to assign their own program-based identifiers. With either scenario, SIDMOD will need to be supported, and it will be necessary to identify a funding source for ongoing maintenance and upgrades. For the assignment of unique student identifiers, the intended usage and estimated costs need to be determined before a scenario can be recommended.

Recommendations for Non-Subsidized Preschool Services

In either scenario, there will be children age 0-5 whose data won't be captured. The workgroup determined that there is no entry point at which students who are receiving early childhood education services, who are not born to Colorado citizens, and who aren't subsidized by State or Federal funding would be assigned a uniquely identifying student number. Early childhood education service providers would have no incentive to collect and submit student data, in order to assign a number, if not receiving State or Federal funding. If the number would be used to track the performance of early childhood education service providers, an incentive could be that the service providers would have access to the results of outcome studies, information which they could use to enhance their services and which could also be used for marketing their services to clients. However, these same results could be a disincentive for lower performing service providers, if they felt they would be disadvantaged by them. Further, families might opt out of having a number assigned to their children who are not receiving State or Federal funding, before being shown evidence of how the assignment of the number will be used in outcome studies to improve services. If the results of these studies would benefit future students and not their own, parents would have even less motivation to agree to the assignment of an identifier for their child.

The recommendation is not to attempt to assign identifiers to unfunded children, but instead to collect, through self-declaration, information about a child's previous enrollment in non-funded early childhood education services when the child does enter State or Federally subsidized education services.

Promulgation of Rules

Sections 2 and 3 of HB 08-1364 state that rules shall be promulgated as necessary with the State Board of Human Services and the State Board of Education for the assignment of uniquely identifying student numbers to students receiving early childhood education services. The workgroup is unable to make recommendations for the promulgation of rules at this time, for the following reasons. Further instruction is needed regarding the intended purpose of assigning uniquely identifying student numbers to children receiving early childhood education services. The purpose of assigning a number needs to be defined before a clear path forward can be determined. Secondly, two options have been presented in this report, and the promulgations of rules would depend on the decision by the legislative body regarding which scenario would best meet the needs of the State.

Conclusion

Two very different scenarios have been presented for the assignment of uniquely identifying student numbers to children who receive State or Federally subsidized early childhood education services. Pros and cons for each scenario have been presented, and protocols for sharing data have been suggested. The discussion would not be complete without addressing the following items, which are applicable to both scenarios.

The protocol for both scenarios should require all agencies to agree to the same level of confidentiality when sharing, accessing and storing data, and to abide by State and Federal guidelines, including but not limited to the Health Insurance Portability and Accountability Act (HIPAA), the Family Educational Rights and Privacy Act (FERPA), and Colorado's vital statistics regulation CRS 25-2-117. The agencies should agree to share data only when the request for data serves a legitimate State purpose. An information advisory council should be formed to serve as the authoritative body to determine the legitimacy of data requests and the appropriate use of shared data.

Data standardization and best practices should be adopted by all participating agencies, in order to maximize the efficiency of any system that would link data sets together, and to mitigate the assignment of multiple identifiers to the same individual. It is recommended that the use of the birth certificate for the validation of names and birthdates should be enforced.

The applicability of the shared data should have an all-inclusive focus. Whatever system is used to link the datasets together, and at whatever agency the system is housed, all agencies should feel that they are equally served by the system's capabilities to share and analyze data and produce information. Every agency that agrees to share data with other agencies should have an equal opportunity to benefit from their efforts, and the system shouldn't be built solely for the analysis of educational outcomes. Furthermore, parents should feel confident that the data-sharing program would promote better outcomes for all of Colorado's children, without discrimination.

There is much positive benefit to come from a unique identifier for preschool children, however, the purposes of assigning such a number must be defined and detailed cost analysis of various options must be conducted before Colorado can proceed with implementation.

Section 8 - Appendices

Appendix A - HB 08-1364 Act

CONCERNING INTERDEPARTMENTAL DATA PROTOCOLS.

Be it enacted by the General Assembly of the State of Colorado:
SECTION 1. Article 37.5 of title 24, Colorado Revised Statutes, is amended BY THE ADDITION OF A NEW PART to read: PART 7 INTERDEPARTMENTAL DATA PROTOCOL

24-37.5-701. Legislative declaration. (1) THE GENERAL ASSEMBLY HEREBY FINDS THAT:

(a) EACH AGENCY OF THE STATE, THROUGH THE PROCESS OF PROVIDING GOVERNMENTAL SERVICES, COLLECTS A SIGNIFICANT AMOUNT OF DATA WITH REGARD TO PERSONS WHO HAVE INTERACTIONS WITH GOVERNMENTAL AGENCIES;

(b) CREATING CROSS-DEPARTMENTAL DATA INTEROPERABILITY AND PROTOCOLS USED BY ALL STATE EXECUTIVE BRANCH AGENCIES WILL SIGNIFICANTLY INCREASE THE EFFICIENCY OF STATE GOVERNMENT AND ENHANCE THE ABILITY OF MULTIPLE STATE AGENCIES TO EFFECTIVELY AND EFFICIENTLY PROVIDE SERVICES TO INDIVIDUALS WITHIN THE STATE;

(c) THE DATA COLLECTED THROUGH THE PROVISION OF GOVERNMENTAL SERVICES, IF APPROPRIATELY COLLECTED AND SYNTHESIZED, WILL PROVIDE VALUABLE INFORMATION TO GUIDE MEMBERS OF THE GENERAL ASSEMBLY AND PERSONS WITHIN THE STATE EXECUTIVE BRANCH AGENCIES IN FORMULATING STATE POLICY AND IN DETERMINING THE EFFECTIVENESS OF STATE POLICIES;

(d) IT IS IMPERATIVE IN ESTABLISHING PROCEDURES AND PROTOCOLS FOR CROSS-DEPARTMENTAL DATA PROCESSING THAT THE STATE TAKE ALL POSSIBLE MEASURES TO ENSURE PERSONAL PRIVACY AND PROTECT PERSONAL INFORMATION FROM INTENTIONAL OR ACCIDENTAL RELEASE TO UNAUTHORIZED PERSONS AND FROM INTENTIONAL OR ACCIDENTAL USE FOR UNAUTHORIZED PURPOSES.

(2) THE GENERAL ASSEMBLY THEREFORE CONCLUDES THAT IT IS IN THE BEST INTERESTS OF THE STATE TO CREATE AN INTERDEPARTMENTAL DATA PROTOCOL TO ASSIST IN FORMULATING AND DETERMINING THE EFFECTIVENESS OF STATE POLICIES.

24-37.5-702. Definitions. AS USED IN THIS PART 7, UNLESS THE CONTEXT OTHERWISE REQUIRES:

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

(2) "COUNCIL" MEANS THE DATA PROTOCOL DEVELOPMENT COUNCIL CONVENED PURSUANT TO SECTION 24-37.5-703.

(3) "DATA" MEANS UNIT RECORDS.

(4) "INTERDEPARTMENTAL DATA PROTOCOL" MEANS AN INTEROPERABLE, CROSS-DEPARTMENTAL DATA MANAGEMENT SYSTEM AND FILE SHARING PROCEDURE THAT PERMITS THE MERGING OF UNIT RECORDS FOR THE PURPOSES OF POLICY ANALYSIS AND DETERMINATION OF PROGRAM EFFECTIVENESS.

(5) "PERSONAL IDENTIFYING INFORMATION" MEANS A PERSON'S FIRST NAME OR FIRST INITIAL AND LAST NAME IN COMBINATION WITH HIS OR HER SOCIAL SECURITY NUMBER OR DRIVER'S LICENSE NUMBER OR IDENTIFICATION CARD NUMBER.

(6) "POLITICAL SUBDIVISION" MEANS A MUNICIPALITY, COUNTY, CITY AND COUNTY, TOWN, OR SCHOOL DISTRICT IN THIS STATE.

(7) "STATE AGENCY" MEANS EACH PRINCIPAL DEPARTMENT WITHIN THE EXECUTIVE BRANCH, INCLUDING EACH BOARD, DIVISION, UNIT, OFFICE, OR OTHER SUBDIVISION WITHIN EACH DEPARTMENT, EACH OFFICE OR AGENCY WITHIN THE GOVERNOR'S OFFICE, EACH STATE-SUPPORTED INSTITUTION OF HIGHER EDUCATION, AND EACH LOCAL DISTRICT JUNIOR COLLEGE; EXCEPT THAT "STATE AGENCY" SHALL NOT INCLUDE ANY DEPARTMENT, AGENCY, BOARD, DIVISION, UNIT, OFFICE, OR OTHER SUBDIVISION OF A DEPARTMENT THAT DOES NOT COLLECT UNIT RECORDS.

24-37.5-703. Data protocol development council - convening. ON OR BEFORE SEPTEMBER 1, 2008, THE CHIEF INFORMATION OFFICER SHALL CONVENE THE DATA PROTOCOL DEVELOPMENT COUNCIL TO ASSIST IN DESIGNING AND IMPLEMENTING THE INTERDEPARTMENTAL DATA PROTOCOL. THE COUNCIL SHALL CONSIST OF NO MORE THAN TWO REPRESENTATIVES FROM EACH STATE AGENCY WHO HAVE RESPONSIBILITY FOR THE COMMUNICATION AND INFORMATION RESOURCES AND COMMUNICATION AND INFORMATION RESOURCES TECHNOLOGIES WITHIN THE STATE AGENCY. THE CHIEF INFORMATION OFFICER MAY INCLUDE ADDITIONAL PERSONS ON THE COUNCIL IF HE OR SHE DETERMINES ADDITIONAL PERSONS ARE NECESSARY TO FULLY REPRESENT ALL OF THE STATE AGENCIES. THE COUNCIL SHALL MEET AS OFTEN AS NECESSARY AT THE CALL OF THE CHIEF INFORMATION
OFFICER TO COMPLETE THE DUTIES SPECIFIED IN THIS PART 7.

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

(1) THE CHIEF INFORMATION OFFICER, WORKING WITH THE COUNCIL, SHALL CREATE 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 THE INTERDEPARTMENTAL DATA PROTOCOL, THE CHIEF INFORMATION OFFICER AND THE COUNCIL SHALL ESTABLISH TIME LINES BY WHICH THE STATE AGENCIES SHALL IMPLEMENT THE INTERDEPARTMENTAL DATA PROTOCOL.

(2) THE INTERDEPARTMENTAL DATA PROTOCOL SHALL BE DESIGNED TO ENABLE EACH STATE AGENCY TO ACCURATELY AND EFFICIENTLY COLLECT AND SHARE DATA WITH THE OTHER STATE AGENCIES. AT A MINIMUM, THE INTERDEPARTMENTAL DATA PROTOCOL SHALL BE DESIGNED TO ENSURE THAT DATA COLLECTED BY DIFFERENT STATE AGENCIES CAN BE MATCHED AND DISCREPANCIES IN THE DATA PROCESSING RECONCILED TO ACCURATELY IDENTIFY DATA PERTAINING TO THE SAME RECORD WITHOUT ALLOWING ANY PERMANENT SHARING OF PERSONAL IDENTIFYING INFORMATION AMONG STATE AGENCIES WITHOUT EXPRESS AUTHORIZATION FROM THE EXECUTIVE DIRECTORS OF THE ORIGINATING AND RECEIVING STATE AGENCIES.

(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:

(a) ESTABLISH THE CIRCUMSTANCES UNDER WHICH AND THE REASONS FOR WHICH A STATE AGENCY MAY SHARE INFORMATION WITH ANOTHER STATE AGENCY, WITH A POLITICAL SUBDIVISION, OR WITH A NONGOVERNMENTAL ENTITY OR AN INDIVIDUAL;

(b) ESTABLISH THE FORMAT IN WHICH A STATE AGENCY MAY RELEASE DATA TO A POLITICAL SUBDIVISION, A NONGOVERNMENTAL ENTITY, OR AN INDIVIDUAL;

(c) ENSURE 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, AND THE FEDERAL "HEALTH INSURANCE PORTABILITY AND ACCOUNTABILITY ACT OF 1996", 42 U.S.C. SEC. 1320d TO 1320d-8; AND

(d) ENSURE THAT A STATE AGENCY DOES NOT PERMANENTLY SHARE PERSONAL IDENTIFYING INFORMATION WITH ANOTHER STATE AGENCY WITHOUT EXPRESS AUTHORIZATION FROM THE EXECUTIVE DIRECTORS OF THE ORIGINATING AND RECEIVING STATE AGENCIES OR WITH A POLITICAL SUBDIVISION, A NONGOVERNMENTAL ENTITY, OR AN INDIVIDUAL, OTHER THAN THE INDIVIDUAL WHO IS THE SUBJECT OF THE INFORMATION.

(4) NOTWITHSTANDING ANY PROVISION OF THIS SECTION TO THE CONTRARY, THE INTERDEPARTMENTAL DATA PROTOCOL SHALL NOT NULLIFY ANY MEMORANDA OF UNDERSTANDING EXISTING AS OF JANUARY 1, 2008, NOR PROHIBIT THE CREATION OF MEMORANDA OF UNDERSTANDING AFTER SAID DATE, BETWEEN OR AMONG STATE AGENCIES CONCERNING DATA SHARING OR ANY OTHER DATA SHARING PRACTICES.

(5) NOTWITHSTANDING ANY PROVISION OF THIS SECTION TO THE CONTRARY, THE INTERDEPARTMENTAL DATA PROTOCOL SHALL NOT PROHIBIT THE RELEASE TO OR SHARING OF DATA WITH NONGOVERNMENTAL ENTITIES OR INDIVIDUALS IF THE RELEASE OR SHARING IS OTHERWISE REQUIRED, PERMITTED, OR ALLOWED BY THE PROVISIONS OF PART 2 OF ARTICLE 72 OF THIS TITLE OR OTHER STATE OR FEDERAL LAW, OR IF THE RELEASE OR SHARING OCCURS PURSUANT TO CONTRACT OR OTHER AGREEMENT WITH A STATE AGENCY.

24-37.5-705. Data sharing - authorization. (1) WITH THE IMPLEMENTATION OF THE INTERDEPARTMENTAL DATA PROTOCOL, EXCEPT AS SPECIFICALLY PROHIBITED BY STATUTE, EACH STATE AGENCY IS AUTHORIZED, IN ACCORDANCE WITH THE PROVISIONS OF THE INTERDEPARTMENTAL DATA PROTOCOL, TO SHARE WITH THE FOLLOWING ENTITIES DATA COLLECTED IN THE COURSE OF PERFORMING ITS POWERS AND DUTIES:

(a) OTHER STATE AGENCIES;

(b) AGENCIES WITHIN THE LEGISLATIVE AND JUDICIAL DEPARTMENTS;

(c) POLITICAL SUBDIVISIONS; AND

(d) NONGOVERNMENTAL ENTITIES AND INDIVIDUALS.

24-37.5-706. Interdepartmental data protocol cash fund - created. (1) THE CHIEF INFORMATION OFFICER IS AUTHORIZED TO SEEK AND ACCEPT GIFTS, GRANTS, OR DONATIONS FROM PRIVATE OR PUBLIC SOURCES FOR THE PURPOSES OF THIS PART 7. ALL PRIVATE AND PUBLIC FUNDS RECEIVED THROUGH GIFTS, GRANTS, OR DONATIONS SHALL BE TRANSMITTED TO THE STATE TREASURER, WHO SHALL CREDIT THE SAME TO THE INTERDEPARTMENTAL DATA PROTOCOL CASH FUND, WHICH FUND IS HEREBY CREATED AND REFERRED TO IN THIS SECTION AS THE "FUND". THE MONEYS IN THE FUND ARE

CONTINUOUSLY APPROPRIATED TO THE OFFICE OF INFORMATION TECHNOLOGY FOR THE DIRECT AND INDIRECT COSTS

ASSOCIATED WITH THE IMPLEMENTATION OF THIS PART 7. THE CHIEF INFORMATION OFFICER AND THE OFFICE OF INFORMATION TECHNOLOGY SHALL NOT BE REQUIRED TO IMPLEMENT THE PROVISIONS OF THIS PART 7 UNTIL SUCH TIME AS AT LEAST ONE HUNDRED THIRTEEN THOUSAND FIVE HUNDRED DOLLARS ARE CREDITED TO THE FUND. IT IS THE INTENT OF THE GENERAL ASSEMBLY THAT THE PROVISIONS OF THIS PART 7 BE IMPLEMENTED WITHOUT THE USE OF STATE MONEYS.

(2) ANY MONEYS IN THE FUND NOT EXPENDED FOR THE PURPOSE OF THIS PART 7 MAY BE INVESTED BY THE STATE TREASURER AS PROVIDED BY LAW. ALL INTEREST AND INCOME DERIVED FROM THE INVESTMENT AND DEPOSIT OF MONEYS IN THE FUND SHALL BE CREDITED TO THE FUND. ANY UNEXPENDED AND UNENCUMBERED MONEYS REMAINING IN THE FUND AT THE END OF A FISCAL YEAR SHALL REMAIN IN THE FUND AND SHALL NOT BE CREDITED OR TRANSFERRED TO THE GENERAL FUND OR ANOTHER FUND.

24-37.5-707. Interdepartmental data protocol - report. THE CHIEF INFORMATION OFFICER SHALL SUBMIT TO THE GOVERNOR AND THE STATE, VETERANS, AND MILITARY AFFAIRS COMMITTEES OF THE HOUSE OF REPRESENTATIVES AND THE SENATE, OR ANY SUCCESSOR COMMITTEES, A REPORT ON OR BEFORE MARCH 1, 2009, CONCERNING DEVELOPMENT AND IMPLEMENTATION OF THE INTERDEPARTMENTAL DATA PROTOCOL.

SECTION 2. Part 1 of article 2 of title 22, Colorado Revised Statutes, is amended BY THE ADDITION OF A NEW SECTION to read:

22-2-134. Unique student identifier - early childhood education - rules. (1) ON OR BEFORE SEPTEMBER 1, 2008, THE COMMISSIONER, IN COOPERATION WITH THE EXECUTIVE DIRECTOR OF THE DEPARTMENT OF HUMAN SERVICES, SHALL CONVENE A WORKING GROUP TO REVIEW THE ISSUES PERTAINING TO THE ASSIGNMENT OF A UNIQUELY IDENTIFYING STUDENT NUMBER TO CHILDREN WHO RECEIVE STATE-SUBSIDIZED OR FEDERALLY SUBSIDIZED EARLY CHILDHOOD EDUCATION SERVICES, INCLUDING BUT NOT LIMITED TO SERVICES PROVIDED THROUGH THE CHILD CARE DEVELOPMENT BLOCK GRANT AND HEAD START. IN CONVENING THE WORKING GROUP, THE COMMISSIONER AND THE EXECUTIVE DIRECTOR OF THE DEPARTMENT OF HUMAN SERVICES SHALL INCLUDE REPRESENTATIVES FROM THE DEPARTMENT OF EDUCATION AND THE DEPARTMENT OF HUMAN SERVICES AND EPRESENTATIVES OF SCHOOL DISTRICTS AND OTHER INTERESTED STAKEHOLDERS.

(2) THE WORKING GROUP SHALL ADOPT PROTOCOLS BY WHICH THE DEPARTMENT OF EDUCATION, THE DEPARTMENT OF HUMAN SERVICES, SCHOOL DISTRICTS, CHARTER SCHOOLS, THE EARLY CHILDHOOD COUNCILS, AS DESCRIBED IN SECTION 26-6.5-103.3, C.R.S., AND THE EARLY CHILDHOOD CARE AND EDUCATION COUNCILS, AS DEFINED IN SECTION 26-6.5-101.5 (6), C.R.S., SHALL COOPERATE IN ASSIGNING THE UNIQUELY IDENTIFYING STUDENT NUMBERS. THE WORKING GROUP SHALL ALSO CONSIDER METHODS BY WHICH TO ENCOURAGE AND FACILITATE THE ASSIGNMENT OF UNIQUELY IDENTIFYING STUDENT NUMBERS TO STUDENTS WHO ARE RECEIVING EARLY CHILDHOOD EDUCATION SERVICES THAT ARE NOT SUBSIDIZED BY STATE OR FEDERAL FUNDING.

(3) ON OR BEFORE FEBRUARY 1, 2009, THE COMMISSIONER SHALL REPORT TO THE HEAD OF THE OFFICE OF INFORMATION TECHNOLOGY THE FINDINGS AND PROTOCOLS ADOPTED

BY THE WORKING GROUP. THE HEAD OF THE OFFICE OF INFORMATION TECHNOLOGY SHALL INCORPORATE THE FINDINGS AND PROTOCOLS OF THE WORKING GROUP INTO THE REPORT MADE TO THE GOVERNOR AND THE STATE, VETERANS, AND MILITARY AFFAIRS COMMITTEES OF THE HOUSE OF REPRESENTATIVES AND THE SENATE, OR ANY SUCCESSOR COMMITTEES, PURSUANT TO SECTION 24-37.5-707, C.R.S. (4) FOLLOWING ADOPTION OF THE PROTOCOLS, THE STATE BOARD OF EDUCATION SHALL PROMULGATE RULES PURSUANT TO THE "STATE ADMINISTRATIVE PROCEDURE ACT", ARTICLE 4 OF TITLE 24, C.R.S., AS

NECESSARY FOR THE ASSIGNMENT OF UNIQUELY IDENTIFYING STUDENT NUMBERS TO STUDENTS RECEIVING EARLY CHILDHOOD EDUCATION SERVICES. THE STATE BOARD SHALL COLLABORATE WITH THE STATE BOARD OF HUMAN SERVICES IN PROMULGATING RULES AS PROVIDED IN THIS SUBSECTION (4) TO ENSURE THAT THEY DO NOT CONFLICT WITH ANY RULES PROMULGATED BY THE STATE BOARD OF HUMAN SERVICES PURSUANT TO SECTION 26-6-121, C.R.S.

SECTION 3. Part 1 of article 6 of title 26, Colorado Revised Statutes, is amended BY THE ADDITION OF A NEW SECTION to read:

26-6-121. Preschools - unique student identifying numbers - rules.

(1) ON OR BEFORE SEPTEMBER 1, 2008, THE EXECUTIVE DIRECTOR, IN COOPERATION WITH THE COMMISSIONER OF EDUCATION, SHALL CONVENE A WORKING GROUP, AS DESCRIBED IN SECTION 22-2-134, C.R.S., TO REVIEW THE ISSUES PERTAINING TO THE ASSIGNMENT OF A UNIQUELY IDENTIFYING STUDENT NUMBER TO CHILDREN WHO RECEIVE STATE-SUBSIDIZED OR FEDERALLY SUBSIDIZED EARLY CHILDHOOD EDUCATION SERVICES, INCLUDING BUT NOT LIMITED TO SERVICES PROVIDED THROUGH THE CHILD CARE DEVELOPMENT BLOCK GRANT AND HEAD START.

(2) THE WORKING GROUP SHALL ADOPT PROTOCOLS BY WHICH THE DEPARTMENT OF EDUCATION, THE DEPARTMENT OF HUMAN SERVICES, SCHOOL DISTRICTS, CHARTER SCHOOLS, THE EARLY CHILDHOOD COUNCILS, AS DESCRIBED IN SECTION 26-6.5-103.3, AND THE EARLY CHILDHOOD CARE AND EDUCATION COUNCILS, AS DEFINED IN SECTION 26-6.5-101.5 (6), SHALL COOPERATE IN ASSIGNING THE UNIQUELY IDENTIFYING STUDENT NUMBERS. THE WORKING GROUP SHALL ALSO CONSIDER METHODS BY WHICH TO ENCOURAGE AND FACILITATE THE ASSIGNMENT OF UNIQUELY IDENTIFYING STUDENT NUMBERS TO STUDENTS WHO ARE RECEIVING EARLY CHILDHOOD EDUCATION SERVICES THAT ARE NOT SUBSIDIZED BY STATE OR FEDERAL FUNDING.

(3) FOLLOWING ADOPTION OF THE PROTOCOLS, THE STATE BOARD SHALL PROMULGATE RULES PURSUANT TO THE "STATE ADMINISTRATIVE PROCEDURE ACT", ARTICLE 4 OF TITLE 24, C.R.S., AS NECESSARY FOR THE ASSIGNMENT OF UNIQUELY IDENTIFYING STUDENT NUMBERS TO STUDENTS RECEIVING EARLY CHILDHOOD EDUCATION SERVICES. THE STATE BOARD SHALL COLLABORATE WITH THE STATE BOARD OF EDUCATION IN PROMULGATING ANY NECESSARY RULES TO ENSURE THAT THEY DO NOT CONFLICT WITH ANY RULES PROMULGATED BY THE STATE BOARD OF EDUCATION PURSUANT TO SECTION 22-2-134, C.R.S.

SECTION 4. No appropriation. The general assembly has determined that this act can be implemented within existing appropriations, and therefore no separate appropriation of state moneys is necessary to carry out the purposes of this act.
SECTION 5. 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 6, 2008, if adjournment sine die is on May 7, 2008); 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, if approved by the people, shall take effect on the date of the official
declaration of the vote thereon by proclamation of the governor.

Appendix B - Project Charter

Project Description

House Bill 08-1364 directs the Governor's Office of Information Technology (OIT) to convene a Data Protocol Development Council ("Council") to design and implement an interdepartmental data protocol. HB-1364 is one of Governor Ritter's priority bills from the 2008 legislative session. The goal of the cross-departmental data protocol is to facilitate information sharing across agencies and to assist in formulating and determining the effectiveness of state policies. This project will examine what is currently in place today, and provide recommendations for moving forward with an architecture and processes to accomplish interagency data sharing in a uniform manner.

The mission of the Council is to provide guidance, policies and procedures for implementing a data sharing architecture across the State enterprise that will achieve the stated goal and objectives of HB-1364. These guidances in the form of a final report and recommendations will be delivered to the State Chief Information Officer for presentation to the Governor and Legislature. The Council is comprised of representatives from Executive Branch Agency and interested parties as necessary.

Project Drivers

HB 08-1364 was initially driven by distinct needs identified by the Governor's P-20 Education Coordinating Council to analyze longitudinal data regarding factors such as improving teaching and learning; informing public policy; fostering a culture of evidence-based decision making; conducting research; evaluating system and program effectiveness; and, providing reports to various stakeholder groups. The collective Colorado State government, as an entity that provides funding, resources and services to the citizenry of the State, has similar needs.
These include:

- The ability to analyze and determine the effectiveness of State policies and resources by examining an issue across multiple State agencies;
- Formulate informed strategic plans for the application and use of State resources based on strong, accurate, reliable, multi-dimensional data;
- Enable more efficient collecting, storing, manipulating, sharing, retrieving, and releasing of data across State agencies.

Project Scope

The Council will convene on August 21, 2008 and complete its work by February 26, 2009. The Council will restrict its study of the data sharing protocol to unit records. Unit records are defined as records pertaining to individuals. The Council will review and baseline the current data systems; data sharing practices and applications; governance policies and procedures; and, statutory or regulatory guidelines in place across Executive Branch agencies that maintain unit records. Due to the limited time and resources of this project, it is necessary to prioritize the review to a few selected unit record data stores.

Each Agency will identify one representative data store that is of relatively high importance to that agency with regard to data sharing. A preference should be given to data stores that are already being shared. Examination of current solutions will assist identification of key success factors. Selected data stores will be benchmarked so that information such as (but not limited to) the following is collected:

- data dictionary
- application hardware and software

- any existing data sharing done out of or in to that data store
- existing agency policies/procedures/governance structures in place regarding sharing of that data
- current statutory or regulatory (state or federal) guidelines in place regarding the use of that data

The Council will develop guidelines for agencies to continue the benchmarking work on its own and report the results back to OIT beyond the time frame given in HB-1364.

The Council will benchmark the work of up to five other states in this area. The Council will develop a governance structure and privacy policy for the cross-departmental data protocol. The Council will also identify an existing data sharing project in the State on which it can pilot, test, and adjust its recommendations so that the Council can get feedback in a real-world scenario.

Finally, the Council will develop recommendations, time frames, and an action plan for moving forward with developing and implementing the cross-departmental data protocol. *Procurement, development, and/or implementation of Council recommendations are outside the scope of work for this phase of the process.*

Goals and Objectives

The goal of the cross-departmental data protocol is to facilitate information sharing across agencies and assist in formulating and determining the effectiveness of state policies.

In detail the goals and objectives are as follows:

Goal 1: Analyze the requirements of all State Agencies that have a need to share unit record data with other agencies.

- Understand and document the unit record data captured, stored and maintained by State Executive Branch Agencies (only one key data store needs to be documented by the end of February);
- Understand and document the existing hardware, software, networking and communications systems that contain the unit record data;
- Understand and document the existing data sharing practices and applications employed by all State Executive Branch Agencies;
- Understand and document the existing governance policies and procedures employed by all State Executive Branch Agencies with regards to collecting, storing, sharing, and destroying data;
- Understand and document the existing statutory or regulatory guidelines in place at all State Executive Branch Agencies with regards to collecting, storing, sharing, and destroying data.

Goal 2: Determine a data sharing protocol that meets the needs of State agencies.

- Assess existing national data sharing standards and benchmark the work of up to five other states in this area;
- Develop an architecture for the development of the data protocol, including data normalization, identity resolution, and source data authority;
- Develop a governance structure, including processes and procedures, to be used by state agencies for sharing information with another state agency, with a political subdivision, or with a nongovernmental entity or an individual;
- Establish the circumstances under which a state agency may release data to a political subdivision, a nongovernmental entity or an individual;

- Establish the format in which a state agency may release data to a political subdivision, a nongovernmental entity or an individual;
- Establish the retention and destruction policies of data that is shared by a state agency to a political subdivision, a nongovernmental entity or an individual;
- Ensure compliance with existing statutory and regulatory requirements;
- Create new or modify existing policies to ensure personal privacy and the protection of personal identifying information (PII).

Goal 3: Develop recommendations and a strategy for moving forward.

- Develop alternative and recommended solutions for implementing the data sharing protocol;
- Establish time lines for implementing the recommended solution across all State agencies;
- Identify high-level associated costs for the recommended solution;
- Identify necessary statutory or regulatory changes;
- Identity critical gaps that must be addressed to ensure the success of this project;
- Identify next steps to ensure the project moves forward to the next phase.

Project Deliverables

The deliverables of the cross-departmental data protocol project include:

- Templates and procedures to capture all agency baseline data;
- A comprehensive reporting structure to store and maintain the reported agency baseline data;
- A comprehensive report with the recommendations and strategy to be delivered to the State Chief Information Officer.
- Identified statutory, regulatory, and organizational changes necessary to the success of the data sharing protocol;

Risks

There are a number of risks associated with this project. Below is a summary of those known risks at the time of writing the Project Charter.

- The short time frame in which to complete this project;
- Obtaining agreement on the data sharing standards, governance, policies, and procedures from the diverse set of State Agency users;
- Developing a protocol that meets the diverse requirements of the State Agencies;
- Ability to secure the necessary funding to implement the cross-departmental data protocol.

Issues

The following issues need to be resolved before the project can move forward to be fully executed:

- How identity resolution will be done;
- Determine a funding source to implement a enterprise system vs. a agency specific system;
- 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.
- Satisfy privacy and security concerns of citizens.

Draft Timeline

Task	Due Date
Finalize Project Charter	9/1/08
1364 Council Kick-Off Meeting	8/21/08
Bi-Monthly (twice-a-month) Meetings	September 2008 – February 2009
Develop templates for data collection (**Goal 1**)	8/21/08
Baseline data due – system, data, compliance, data sharing	10/30/08
Determine requirements (**Goal 1**)	12/31/08
Identify necessary statutory changes to implement the protocol	12/31/08
Identify possible solution and recommendations (**Goal 2**)	1/31/09
Prepare final report (**Goal 3**)	2/26/09
State CIO Report Due to Governor & Legislature	2/27/09

Project Communications

The following table summarizes the communications for the project

Role	Type	Frequency	Author
Executive Sponsors	Overall Progress Report	Monthly	Pgm Mgr
Executive Sponsors	Risk/Issue Updates	Weekly	Pgm Mgr
Executive Stakeholders	Progress Reports	Monthly	Pgm Mgr
Project Staff	Overall Progress Report	Monthly	Pgm Mgr
Project Staff	Risk/Issue Updates	Weekly	Pgm Mgr
Project Staff	Agency Progress Reports	Monthly	Pgm Mgr

Sponsors and Stakeholders

The following stakeholders have been identified at this point in the project initiation.

Executive Sponsorship

- Mike Locatis, State Chief Information Officer, Governors Office of Information Technology
- Matt Gianneschi, Senior Policy Analyst for Education, Office of Governor Bill Ritter, Jr.

Stakeholder Agencies

- Governors Office of Information Technology (OIT)
- 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 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)
- Office of Cyber Security (OCS)
- Secretary of State (SOS)
- Judicial
- Attorney General (DOL)

Executive Stakeholders

- Department of Agriculture
  - John Stulp, Executive Director
  - Tony Jones, Chief Information Officer
- Department of Corrections
  - Ari Zavaras, Executive Director
  - Paul Lewin, Chief Information Officer
- Department of Education
  - Dwight D. Jones, Commissioner
- Department of Health Care Policy and Finance
  - Joan Henneberry, Executive Director
  - Andy Graziano, Chief Information Officer
- Department of Higher Education
  - David Skaggs, Executive Director
  - Dr. Julie Carnahan, Chief Information Officer
- Department of Human Services

  - Karen Beye, Executive Director
  - Ron Ozga, Chief Information Officer
- Department of Labor and Employment
  - Don Mares, Executive Director
  - Joe Lambert, Chief Information Officer
- Department of Local Affairs
  - Susan Kirkpatrick, Executive Director
  - Brian Morrow, Chief Information Officer
- Department of Natural Resources
  - Harris Sherman, Executive Director
  - Leah Lewis, Chief Information Officer
- Office of Information Technology
  - Mike Locatis, State Chief Information Officer
  - John Conley, Deputy Chief Information Officer
- Department of Personnel & Administration
  - Rich Gonzales, Executive Director
  - David Kaye, Director, Division of Human Resources
- Department of Public Health and Environment
  - James Martin, Executive Director
  - Bill Ferguson, Chief Information Officer
- Department of Public Safety
  - Peter Weir, Executive Director
  - Jim Lynn, Chief Information Officer
- Department of Regulatory Agencies
  - Rico Munn, Executive Director
  - Mike Whatley, Chief Information Officer
- Department of Revenue
  - Roxanne Huber, Executive Director
  - David Loewi, Chief Information Officer
- Department of Transportation
  - Russell George, Executive Director
  - Kim Heldman, Chief Information Officer
- Secretary of State
  - Mike Coffman, Secretary of State
  - Trevor Timmons, Chief Information Officer

- Department of Law
  - John Suthers, Attorney General
  - Susan Lin, Assistant Attorney General, Chief Privacy Officer
- Judicial Branch
  - Justice Mary Mullarkey, Chief Justice
  - Bob Roper, Chief Information Officer

The following people have expressed an interest in this project and are willing to assist in the concept development, analysis, design and implementation.

- Debi Erpenbeck, Department of Agriculture
- Marty Fry, Department of Agriculture
- Chuck Noll, Department of Corrections
- Dan Domagala, Department of Education
- Andy Graziano, Department of Health Care Policy and Finance
- Beth Martin, Department of Health Care Policy and Finance
- Ryan Allred, Department of Higher Education
- Jim Broyles, Department of Higher Education
- Marc Makert, Department of Human Services
- Prasanna Bennabhaktula, Department of Human Services
- Jim Yuhas, Department of Labor and Employment
- David Gestner, Department of Labor and Employment
- Mark Krudwig, Department of Local Affairs
- Leah Lewis, Department of Natural Resources
- Marc Fine, Department of Natural Resources
- Rob Lloyd, Department of Natural Resources
- Mike Amelon, Office of Information Technology
- Susan McMillan, Office of Information Technology
- Micheline Casey, Office of Information Technology, 1364 Council Program Manager
- Andrew Putnam, Department of Public Health and Environment
- Bob O'Doherty, Department of Public Health and Environment
- Jane Crisman, Department of Public Safety
- Rose Ramirez, Department of Public Safety
- Lisa Bradley, Department of Regulatory Agencies

- Brian Van Sickle, Department of Regulatory Agencies
- David Loewi, Department of Revenue
- Joan Vecchi, Department of Revenue
- Neil Tillquist, Department of Revenue
- Steve Hooper, Department of Revenue
- Mike Armbruster, Department of Transportation
- Guy Mellor, Department of Transportation
- Bob Roper, Judicial
- Chad Cornelius, Judicial
- Stacey Kirk, Judicial
- Samir Nanavati, Judicial
- Trevor Timmons, Secretary of State
- Mike Shea, Secretary of State

Roles and Responsibilities

Executive Sponsors

- Policy development
- Exploration and development of funding sources
- High level project objective development
- Championing the project amongst business staff and other Cabinet members

Executive Stakeholders

- Providing senior level approval and direction
- Championing project among agency staff
- Ensuring funding availability

Agency Executive Directors

- Staffing of the Agency Subject Matter Expert role
- Supporting Subject Matter Expert with necessary resources (time, authority, access, information, etc.)
- Providing staff to participate in requirements, use case development, testing and training.

Agency Level Project Managers

- Individual Agency Implementation Projects
- Agency Stakeholder Coordination
- Scope Management within Agency
- Issue Management within Agency

Program Manager

- Overall Program Management
- Coordination with Executive Sponsors
- Coordination with Executive Stakeholders
- Program Communications and status reporting
- Scope Management
- Budget Management
- Risk Management and mitigation
- Communication plan
- Issue Management plan

Appendix C - List of Council Members

The following people actively participated on the HB 08-1364 Council and assisted with analysis, concept development, and recommendations.

- Debi Erpenbeck, Department of Agriculture
- Marty Fry, Department of Agriculture
- Chuck Noll, Department of Corrections
- Victoria Etterer, Department of Corrections
- Dan Domagala, Department of Education
- Stacie Demchak, Department of Education
- Andy Graziano, Department of Health Care Policy and Finance
- Beth Martin, Department of Health Care Policy and Finance
- Diane Dunn, Department of Health Care Policy and Finance
- Ryan Allred, Department of Higher Education
- Jim Broyles, Department of Higher Education
- Marc Makert, Department of Human Services
- Prasanna Bennabhaktula, Department of Human Services
- Jim Yuhas, Department of Labor and Employment
- David Gestner, Department of Labor and Employment
- Alexandra Hall, Department of Labor and Employment
- Mark Krudwig, Department of Local Affairs
- Leah Lewis, Department of Natural Resources
- Marc Fine, Department of Natural Resources
- Rob Lloyd, Department of Natural Resources
- Mike Amelon, Office of Information Technology
- Susan McMillan, Office of Information Technology
- Micheline Casey, Office of Information Technology, 1364 Council Program Manager
- Sue Huang, Department of Personnel and Administration
- Mark Rothman, Department of Personnel and Administration
- Andrew Putnam, Department of Public Health and Environment
- Bob O’Doherty, Department of Public Health and Environment
- Jane Crisman, Department of Public Safety
- Rose Ramirez, Department of Public Safety

- Lisa Bradley, Department of Regulatory Agencies
- Brian Van Sickle, Department of Regulatory Agencies
- David Loewi, Department of Revenue
- Joan Vecchi, Department of Revenue
- Neil Tillquist, Department of Revenue
- Afshin Ghazvini, Department of Revenue
- Mike Armbruster, Department of Transportation, 1364 Council Data Architect and Technology Subcommittee Chair
- Guy Mellor, Department of Transportation, Business Subcommittee Chair
- Susan Lin, Department of Law, Legal Subcommittee Chair
- Stacey Kirk, Judicial
- Samir Nanavati, Judicial
- Trevor Timmons, Secretary of State
- Mike Shea, Secretary of State
- Chris Wallner, CDPS, representing CICJIS
- Steve Hooper, DOR, representing STRAC
- Rick Dakin, CoalFire Systems, representing DGWG
- Meg Williams, CDPS, representing CCYIS

HB 08-1364 Section 2 and 3 Workgroup Representation

Department of Education:

- Jan Rose Petro – Director of Data Services
- Sharon Triolo-Moloney – Assistant Director Early Childhood Initiatives
- Lori Goodwin Bowers – Supervisor of Colorado Preschool Program
- Anne Bygrave – Consultant Student Identifier Management Unit
- Alex Waltrip –IT Professional
- Nick Ortiz – Data Consultant Early Childhood Education

Office of Information Technology/Department of Human Services:

- Galina Krivoruk – Applications Director
- Richard (Skip) Flewelling – Applications Director
- Prasanna Bennabhaktula – Technical Manager

Department of Human Services:

- Patricia Logan – MSW, LSW - Division of Childcare

Department of Health Care Policy and Financing:

- Diane Dunn – Section Manager of Claims
- Beth Martin – Senior Data Analyst

Department of Public Health and Environment

- Alyson Shupe – Section Chief Health Statistics

Qualistar:

- Paula Neth – Chief Operating Officer

Local Stakeholder Representation

Aspen Family Services

- Marsa Williams - President

Colorado Children’s Campaign

- Jon-Paul Bianchi – Early Childhood Initiatives Director
- Kenny Smith – Early Childhood Education Policy Analyst

Denver Public Schools

- John Crawford – CPP Coordinator

Appendix D - List of Agency Baseline Templates Returned

Agency	CDA	DOC	CDE	HCPF	DHE	DHS	CDLE	DOLA	DNR	OIT	CDPHE	CDPS	DORA	DOR	CDOT	SOS	Judicial
Template																	
System	X	X	X	X	X	X	X	X	X	X	X	X	X		X		
Data	X	X	X	X	X	X	X	X	X	X	X	X	X		X		
Compliance	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	
Data Sharing	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X
	4	4	4	4	4	4	2	4	4	4	4	4	4	2	4	2	1

Executive Branch Agencies

TOTAL DUE	**60**
TOTAL IN	**56**
PERCENT IN	**93.3%**

With Non-Executive Branch Agencies

TOTAL DUE	**68**
TOTAL IN	**59**
PERCENT IN	**86.8%**

Note: A few agencies completed the System and Data baselines for two systems, with CDLE completing its System and Data baselines for all internal agency systems.

Appendix E - Benchmarking Results from Other States

Arkansas

The Arkansas State Education Agency (SEA) has a longstanding partnership with the National Offices of Research, Measurement, and Evaluation Systems (NORMES) at the University of Arkansas to expand its capacity to report student and school achievement results. State student assessment data reside at NORMES which undertakes Adequate Yearly Progress (AYP) calculations, developing school and district report cards, and reporting National Assessment of Educational Progress (NAEP) results—all available on their Web site. NORMES also conducts research and evaluation studies for the SEA and the Legislature. When grade inflation became a concern in Arkansas, NORMES took the ACT student results and compared them with student's composite student grade point averages. Grade inflation indices were calculated to predict the percent of students who might need a remediation course in college.

In addition to its agreement with the university, the SEA has a long-standing partnership with the Arkansas Department of Health and Human Services to simplify the free and reduced lunch eligibility and the Medicaid application process. Within the past two years, Arkansas has made contract agreements with higher education institutions to track K-12 students through college and with the State's Workforce Services Department to track students from K-12 into employment.

California

Priorities:

- Further development of Internet- and technology-based channels for the delivery of State information and services for the convenience of the public.
- A need for consistent and accurate data that will interface with other systems as necessary.
- The assurance that confidential information and valued assets are secure.
- The ability to easily access information and services while ensuring that such access is allowed only to those intended.
- Availability of appropriate tools for executive oversight, management decisions, and program implementation.
- Efficient and cost saving means to deliver services.
- Need to respond and transact quickly.
- Need to maintain systems and services in adequate working order throughout their life cycles and to replace or retire them when support is no longer possible.

Goals:

- Make Government services more accessible.
- Implement common business applications and systems to improve efficiency and cost-effectiveness.
- Ensure State information assets are secured and privacy protected.
- Lower costs and improve the security, reliability and performance of the State's IT infrastructure.
- Strengthen our technology workforce.
- Establish a technology governance structure.

"Objective 2 - Leverage Services between State Agencies, Federal and Local Government and Promote Interagency and Intergovernmental Data Sharing

- The State will pursue opportunities to collaborate with Federal and local agencies and within State government to leverage e-Government services. The State will coordinate interagency and intergovernmental data collection and management, to improve data sharing capabilities and reduce costs of acquiring and managing data.

- Many Federal, State and local government programs are interrelated or interdependent. Working together, governmental agencies can deliver better services to citizens and reduce the overall cost of implementing and maintaining service delivery systems. System and database designs often prescribe unique definitions and program-focused restrictions, inhibiting the use of data for other purposes, and resulting in duplication and incompatibility of data. The State can do a much better job of sharing data through collaborative planning efforts."

Actions

- By January 2007, the State CIO will establish a Federated Identity Management Steering Committee responsible for establishing a State vision, policies and standards regarding identification and authentication of State system users and data.
- By April 2007, the Federated Identity Management Steering Committee will identify selected State agencies as the owners of identity data for persons, businesses and other entities consistent with the intent to maximize secure and reliable collaborative data collection, management, and data sharing.
- By July 2007, the Federated Identity Management Steering Committee will develop policy, privacy, and data sharing rules recommendations to support statewide business needs for identity management.
- By July 2007, one or more State agencies will work with the Social Security Administration to develop data sharing capabilities using scalable enterprise technologies. The State will use the architecture of these initial systems as a foundation and model for enterprise data sharing.
- The State CIO will work with Agency Information Officers and Chief Information Officers to identify other opportunities for expanding interagency data sharing consistent with privacy interests and fair information practices.

Forthcoming Policy Releases

- Safeguarding Against and Responding to a Breach of Personal Information
- Personal Information Breach Notification: Requirements and Decision Making Criteria for State Agencies (SIMM 65D)
- Requests for and Approval to Release Personal Information for Research

Data Exchange Agreement Workgroup

- Charter – develop general approach, recommendations, guidance and tools for the development of agreements between government entities on the use of data
- 21 representatives from various government entities participating
- Timeline for completion – October 2008

http://www.oispp.ca.gov/government/documents/ppt/whats_new_july08.ppt

Recommended Practices on California Information-Sharing Disclosures and Privacy Policy Statements: http://www.oispp.ca.gov/consumer_privacy/pdf/infosharingdisclos.pdf

- Introduction
- Privacy Notice Laws
- Privacy and Customer Trust
- Benchmark Study
- Recommended Practices
- Information-Sharing Disclosures
- Disclosure Document
- Customer Choice Notice
- Notice of Information-Sharing Disclosure
- Privacy Policy Statements
- Notes

- Appendices
- Appendix 1: Advisory Group Members
- Appendix 2: “Shine the Light” Law
- Appendix 3: Online Privacy Protection Act

Engineering for Data Protection and Accountability: www.oispp.ca.gov/government/events/documents/PracticalPrivacy-Security-Panel.ppt - 2007-10-15

- Basic Considerations- Where should technology fit?
- The Dynamic Data Environment
- Business/ Organization Goals for Data Protection
- Engineering Goals for Data Protection
- Tools and Techniques for Data Governance and Accountability
- Secure Storage and Encryption

Protecting the Realm: Confronting the Realities of State Data at Risk: http://www.nascio.org/publications/documents/NASCIO-ProtectingRealm.pdf

The Government Online for Responsible Information Management (Go RIM) Web page provides a central location for information security standards, authority, guidance, forms, tools, and definitions related to California information security policy. These components augment the State Administrative Manual (SAM) security policies identified in SAM Sections 5300-5390 by providing State agencies with access to:

- Baseline security standards that support these policies, as well as other standards when applicable to a specific policy area;
- Laws, regulations, and other related Federal and State policies that provide the authority for the State's policy requirements;
- Guidance documents that provide directions, instructions, and best practices to aid in policy compliance
- Standardized and required forms associated with meeting policy requirements;
- Tools that include samples, templates, and other important resources to help a State agency implement a particular policy or standard
- Definitions for clarification in the meaning of terms, words or phrases referred to in the policy or standards.
- http://www.oispp.ca.gov/government/go_rim/default.asp

Kansas

- Has had a State chief data architect in place for over five years.
- Data sharing is still accomplished on an agency by agency basis.
- GIS data has a centralized clearinghouse.
- Is establishing a field by field data standard.
- Is conducting a proof of concept to validate a GIS data validation service.
- Has an initiative for Social Security one-stop shopping for citizens.
- Is conducting a GIS address matching initiative.
- Is researching some automated dataset linkage tools.

Kentucky

- Had established an Enterprise Architecture Standards Committee and initiated the centralization of IT some years ago.
- The current administration has slowed the pace of data sharing initiatives to reduce cost.
- State employee staff pursuing data sharing has been reduced to one Full-Time Equivalent.
- Is researching an IBM tool for data matching.

Virginia

- The Virginia Statewide Education Authority has partnerships with Virginia Commonwealth University, Southern VA Higher Ed Center, Joint Legislative Audit and Review Commission, W.E. Upjohn Institute for Employment Research, National Student Clearinghouse, CNA Corp. and the University of Virginia
- Most partnerships focus on program evaluation such as studying the impact of math specialists on student achievement and assessing the impact of pre-school programs on at-risk children.
- Other studies analyzed data on the post-high school activities of students served by limited English proficiency programs, and another merged literary screening data with State assessment dates to assess how literacy skills are associated with third grade performance.
- The Virginia Department of Education has a comprehensive Restricted Data Use Agreement which has many contract provisions that could possibly be used in a proposed Colorado Data Sharing Agreement (see Appendix B of "The Third Wave of Longitudinal Data Systems: Data Partnerships").
- Virginia (and Minnesota) has options for charging to fulfill data requests whereas most states do not.
- Virginia has outsourced some data sharing initiatives to Northup Grumman.
- Master data is part of Virginia's future vision.
- By and large, agencies who own the data, manage any data sharing arrangements.
- Is making significant progress with election data matching.
- Has established a portal for GIS and thirteen agencies share substance abuse data.

Appendix F – Technology Background

Approaches to Data Sharing

There are various ways to approach data sharing initiatives. This appendix will provide a brief overview of those approaches.

The simplest way to share data is to have a single integrated system where data is entered and processed and a single warehouse environment where data is accessed for reporting, analysis, and data mining. In the case of unit data, the ideal situation would be a single input system for all the unit master data (name, address, phone, birth date, etc.) that was highly secure with each agency relating this to its own data. However, in most companies and governments, this environment does not exist due to history of data processing development in individual departments and agencies. This means that data sharing requires its own efforts and programs.

Locating and identifying the data to be shared
A coordinated and managed data sharing effort needs a firm foundation based on defined methods, processes, and tools to identify, locate, define, classify, secure, and organize its information about the data to be shared.

Each need to consolidate and analyze data across agencies, every internal agency or citizen search for data for reporting or analysis, and each requirement by the Federal government to locate and transmit data requires time and resources to find and identify the correct data to be used. In many cases the data has been entered into more than one system across one or more agencies. Finding this data requires time to be spent determining which data is the ‘source data of record’ and identifying the exact definition and content of the data fields.

The options in this area are clear. Either each data sharing effort starts from the beginning each time with the entire process of data identification, definition conflict resolution, discovery, cleanup, etc. or a foundation is created using a metadata registry containing the information needed. Creating a metadata registry is not a one-time effort; rather it is a program for ongoing data sharing and possibly identifying system consolidation efforts. The payback in time and resource savings for data discovery, system consolidation, and reuse of designs, data, and reports grows as each system is documented and each data sharing effort completed.

Physically sharing data
Once the data to be shared has been identified and agreements as to the security and accessibility have been reached, architecture for sharing the data must be determined. For most data sharing, reporting, and analysis purposes the original data in the source or operational system where it is input and originally stored is not in the best format nor is easily accessible.

Normally the operational input system is referred to as an OLTP (online transaction processing) system where the data structures were designed for quick single transaction input and retrieval. In contrast, OLAP (online analytical processing) data structures for reporting and analysis are organized differently, often in relational structures, summary tables, and multidimensional cubes.

To be effective, a reporting or analytic architecture needs to be determined. The total system architecture depicting the relationship between the logical model, the physical model of the operational system, and the physical model of the analytic architecture (data warehouse) is illustrated in figure F-1.

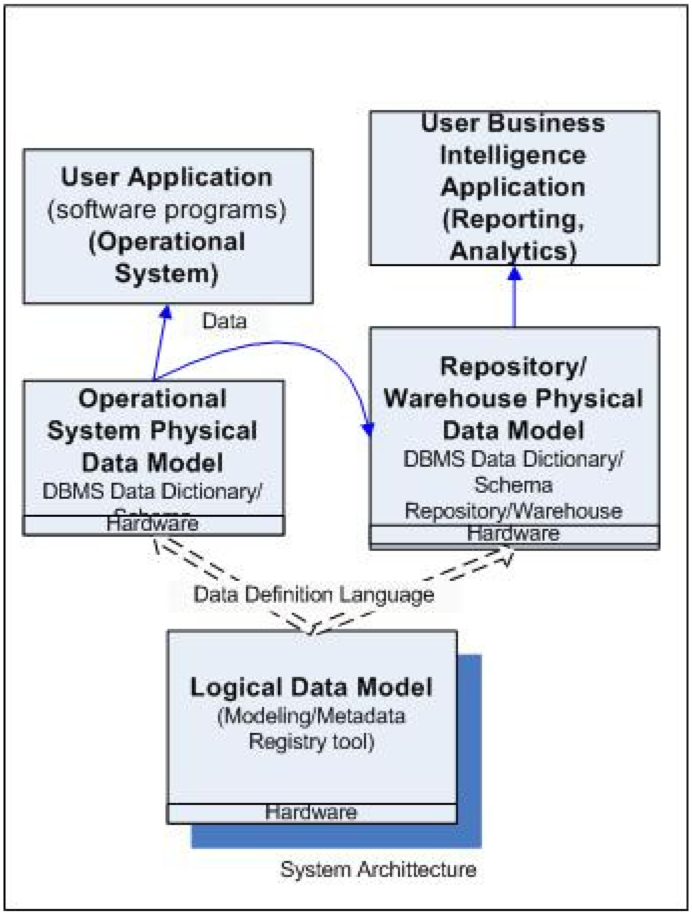


Figure F-1

Data Models

The Federal Enterprise Architecture (FEA)

Executive Order 13011, *Federal Information Technology*, established the Chief Information Officers (CIO) Council as the principal Interagency forum for improving practices in the design, modernization, use, sharing, and performance of Federal information resources. The Clinger-Cohen Act of 1996 assigned the CIOs with the responsibility to develop information technology architectures (ITAs). The Office of Management and Budget (OMB) M-97-02, *Funding Information Systems Investments*, October 1996, requires that Agency investments in major information systems be consistent with Federal, Agency, and Bureau ITAs. The CIO Council began developing the Federal Enterprise Architecture Framework in April 1998 to promote shared development for common Federal processes, interoperability, and sharing of information among the Agencies of the Federal Government and other Governmental entities.

In serving the strategic needs and direction of the Federal Government, the CIO Council seeks to develop, maintain, and facilitate the implementation of the top-level enterprise architecture for the Federal Enterprise. The Framework consists of various approaches, models, and definitions for communicating the overall organization and relationships of architecture components required for developing and maintaining a Federal Enterprise Architecture. The Framework allows the Federal Government to accomplish the following:

- Organize Federal information on a Federal-wide scale

- Promote information sharing among Federal organizations
- Help Federal organizations develop their architectures
- Help Federal organizations quickly develop their IT investment processes
- Serve customer needs better, faster, and cost effectively

The CIO Council chose a segment architecture approach that allows critical parts of the overall Federal Enterprise, called architectural segments, to be developed individually, while integrating these segments into the larger Enterprise Architecture. Federal Agencies can use the same or a modified approach to develop their ITAs in response to the Clinger-Cohen Act. In either case, the Framework can help with architecture development efforts at Federal organizations.

The architecture will serve as a reference point to facilitate the efficient and effective coordination of common business processes, information flows, systems, and investments among Federal Agencies and other Governmental entities. The Federal Enterprise Architecture is a strategic information asset base that defines the business, information necessary to operate the business, technologies necessary to support the business operations, and transitional processes for implementing new technologies in response to the changing needs of the business.

The Federal Enterprise Architecture Framework is a conceptual model that begins to define a documented and coordinated structure for cross-cutting businesses and design developments in the Government. Collaboration among the Agencies with a vested interest in a Federal segment will result in increased efficiency and economies of scale. In time, Government business processes and systems will operate seamlessly in an enterprise architecture that provides models and standards that identify and define the information services used throughout the Government.

http://www.whitehouse.gov/omb/egov/a-1-fea.html
http://www.cio.gov/Documents/fedarch1.pdf
http://xml.coverpages.org/ni2005-12-28-a.html
http://www.cio.gov/
http://www.whitehouse.gov/omb/egov/documents/DRM_2_0_Final.pdf

National Information Exchange Model (NIEM)

NIEM is the National Information Exchange Model program, launched in February of 2005 as a partnership between the US Department of Homeland Security (DHS) and the US Department of Justice (DOJ). The NIEM framework provides standard vocabulary; guidance and processes that help promote effective and efficient information sharing capabilities across organizational boundaries. NIEM is widely adopted within State and local governments, with at least 39 of the 50 states reporting NIEM-conformant information exchange projects. An important aspect of an electronic data exchange is the data model or schema used to transfer the data itself. The schema, or data-sharing model, describes the characteristics of the data fields themselves, including field name, type, length, and acceptable values.

The NIEM model is actually a library of open source models that government agencies can reuse or extend on a case by case basis. By setting this open–source repository as the Colorado standard, takes the greatest advantage of development work done by other US governments, including the Federal government, other states, and other Colorado implementations. No other central model provides these advantages; therefore, the Council recommends adoption of the NIEM model within the State.

Identity Resolution

Identity resolution is an operational intelligence process, typically powered by an identity resolution engine or middleware stack, whereby organizations can connect disparate data sources with a view to understanding possible identity matches and non-obvious relationships across multiple data silos.

It analyzes all of the information relating to individuals and/or entities from multiple sources of data, and then applies likelihood and probability scoring to determine which identities are a match and what, if any, non-obvious relationships exist between those identities.

Identity resolution engines then apply rules, based on common sense logic, to identify hidden relationships across the data. For example, perhaps Muammar al-Quaddafi and Mo'ammar Gadhafi are not the same individual, but rather two distinct people who share common attributes such as address or phone number.

More powerful identity resolution engines, such as IBM Corporation's Entity Analytics Solutions (EAS) and Infoglide Software Corporation's Identity Resolution Engine (IRE), also include a rules engine and workflow process, which apply business intelligence to the resolved identities and their relationships. These advanced technologies make automated decisions and impact business processes in real time, limiting the need for human intervention.

IBM’s Entity Analytic Solutions

IBM’s Entity Analytic Solutions (EAS) is unique identity disambiguation software that provides public sector organizations or commercial enterprises with the ability to recognize and mitigate the incidence of fraud, threat and risk. This IBM EAS offering provides insight on demand, and in context, as to "who is who," "who knows who," and "anonymously."

For most businesses and government agencies, it is important to figure out when a person is using more than one Identity Package (that is, name, address, phone number, social insurance number and other such personal attributes) intentionally or unintentionally. Identity resolution software can help determine when two or more different-looking identity packages are describing the same person, even if the data is inconsistent. For example, by comparing names, addresses, phone numbers, Social Security Numbers and other personal information across different records, this software might reveal that three customers calling themselves Tom R., Thomas Rogers, and T. Rogers are really just the same person.

Sometimes organizations want to share information across nontraditional boundaries - such as between a business and a government agency - and that poses serious privacy challenges. IBM’s Anonymous Resolution allows you to "disguise" sensitive data before you share it with others for purposes of identity resolution and relationship detection.

Identity Resolution - How does it work?

Identity Resolution software performs the entire identity resolution process in less than one second. IBM's ability to combine numerous value attributes beyond simply name and address data produces a previously unattainable level of precision and accuracy in identity recognition initiatives.

Identity Resolution acquires data from source systems in real time, and then performs a five-step entity resolution process:

- Name standardization: determines and applies root names (for example, Rob, Bob and Bobby become Robert).
- Address verification and correction: compares, verifies and corrects addresses with U.S. and international address databases.

- Data quality: applies data-driven quality rules to addresses, phone numbers and date of birth, social insurance numbers and other significant numbers.
- Data enhancement: identities are enhanced by adding outside data through IBM's extensive partner network.
- Entity resolution: compares new information with existing information to determine whether the new information is about a new or existing entity.

Infoglide Identity Resolution Engine (from the Infoglide web site)

Infoglide provides identity resolution solutions to retail, financial services, government, and information providers. Using the same powerful technology chosen by the U.S. Department of Homeland Security to protect air travel and fight terrorism, our Identity Resolution Engine™ (IRE) enables better decision making about the people with whom organizations interact and the schemes in which they operate. The IRE architecture is specifically designed to meet the unique requirements posed by multiple and hidden identities encountered by businesses and governments.

IRE glides across multiple, disparate data sources, applies sophisticated similarity search techniques to resolve multiple identities, and presents a single view of an individual. It then applies sophisticated algorithms and rules to uncover non-obvious relationships between that individual and other individuals or entities that indicate potential fraud or risk. The results of this operational business intelligence are then fed back into business systems and acted on.

Other Identity Resolution Products
Informatica Identity Resolution

Current State Implementations

CDPS

Automated Fingerprint Identification System (AFIS)

CDHS

SIDMOD - State Identification Module (Colorado) for client identity resolution

CDE

- Student Identifier Management Unit (SIMU) formed in May of 2002 to develop a system that uniquely identifies students for the purpose of longitudinal analysis and to produce achievement reports as required by the Federal government.

- Colorado legislation requires longitudinal analysis of the Colorado Student Assessment Program (CSAP) test. Tracking students over time and across district lines requires a State Assigned Student ID (SASID) to be effective.

  Data Elements
  The Record Integration Tracking System (RITS) uses five data elements to match students to SASIDs.
  - Last Name
  - First Name
  - Middle Name
  - Birth Date
  - Gender

  Matching Engine
  The matching engine is designed to attempt to match submitted student records to existing student records, resulting in one of the following outcomes:
  - The student record does not already exist and a SASID is automatically issued

- The student record does exist and the existing record is automatically updated
- The system is unable to determine whether the student record exists and the record is submitted for manual review.

Data Security

- SASIDs are converted to an Encrypted State ID (ESID) before entering CDE's Data Warehouse
- Names and sensitive non-directory information is stripped from data collections
- All public reporting and internal CDE use corresponds to the ESID to protect against accidental disclosure

Judicial

- SID from fingerprinting through CBI
- CBI also uses a name tool called Soundex that puts a percentage value on the likelihood of a match on name and DOB is used to search records when doing recommendations to the court based on criminal history.

Appendix G – Relevant State Cyber Security Policies

The policy documents highlighted below are part of the State of Colorado Cyber Security Policies mentioned in the Technology of this document, the *System Access and Acceptable Use Policy* and the *Online Privacy Policy*. All of the State's Cyber Security Policies were created to support the State of Colorado Chief Information Security Officer (CISO) in achieving the goals of the Colorado Information Security Act (C.R.S. 24-37.5, Part 4). Rules were passed in 2006 resulting in the 19 Colorado Cyber Security Policies listed below and found at:

http://www.colorado.gov/cs/Satellite/Cyber/CISO/1207820732279?rendermode=preview

All Agencies within the scope of these Policies must support and comply with the Requirements section of this document. Additional best-practice guidance is outlined in the Guidelines Section which has been designed to help an Agency achieve the objective of these Policies. For the purposes of this document, an "Agency" includes organizations as defined in C.R.S .24-37.5-102(5).

Rules in support of the Colorado Information Security Act

P-CCSP-001	Cyber Security Program R01
P-CCSP-002	Incident Response R01
P-CCSP-003	IT Risk Management R01
P-CCSP-004	Disaster Recovery R04
P-CCSP-005	Vendor Management R02
P-CCSP-006	Network Operations R01
P-CCSP-007	Systems and Applications Security Operations R01
P-CCSP-008	Access Control R01
P-CCSP-009	Change Control R01
P-CCSP-010	Physical Security R01
P-CCSP-011	Data Handling and Disposal R02
P-CCSP-012	Personnel Security R01
P-CCSP-013	**System Access and Acceptable Use** R01
P-CCSP-014	**Online Privacy** R01
P-CCSP-015	Security Training and Awareness R01
P-CCSP-016	Self-Assessment R01
P-CCSP-017	Security Metrics and Measurement R01
P-CCSP-018	Mobile Computing R01
P-CCSP-019	Wireless Security R01

These 19 policies establish technical standards and a program of compliance, evaluation and security performance monitoring required for effective data sharing. Although all of the policies work together to form a secure basis for data sharing, two of the polices listed in bold may impose some additional limitations. Both the Online Privacy and System Access and Acceptable Use policies (**see below**) require public statements about how collected or reported data may be used. These statements are generated uniquely for each system and may impose additional limits on data sharing beyond those in other regulations. While these statements can be changed and standardized, it is important to note this additional step in any data sharing agreements.

Security controls can be expensive both in terms of the hardware and software required and the additional staff time required for administrative procedures to implement the controls. Not all data, or data sharing efforts need the same level of security. Classifying data, as in the Federal Enterprise Architecture Model and the Federal Information Processing Standards, into categories based on the level of protection required, could reduce the cost of appropriate security controls and simplify complex management and implementation issues. If Colorado is to move forward with large scale data sharing efforts, a strong data classification policy that is consistent with Federal and financial guidelines will be required.

Authority

C.R.S. 24-37.5-401(1), C.R.S. 24-37.5-403(2)(b)-(c), C.R.S. 24-37.5-404(2)(b).

Scope

This policy document applies to every State agency ("Agency") as defined in C.R.S. 24-37.5-102(5). "State agency" means every State office, whether legislative, executive, or judicial, and all of its respective officers, departments, divisions, commissions, boards, bureaus, and institutions. "State agency" does not include State-supported institutions of higher education, the department of higher education, the Colorado commission on higher education, or other instrumentality thereof.

System Access and Acceptable Use Policy

Users of State systems can introduce errors or compromise critical data through intentional or unintentional acts. This policy requires all Agencies to develop an Acceptable Use Policy (AUP) that governs user behavior when accessing State systems. An effective Acceptable Use Policy mitigates risks to State data and systems introduced by system users.

Policy

All Agencies shall ensure users of private State systems abide by a common set of minimum criteria and acknowledge that they understand these criteria and agree to comply with them prior to obtaining access to such systems.

Definitions

For the purposes of this document, please refer to C.R.S. 24-37.5-102, C.R.S 24-37.5-402 and the Colorado Cyber Security Program Policy Glossary for any terms not specifically defined herein.

Roles and Responsibilities

Executive Director – is responsible for:

- Designating the responsibility of collecting and storing acknowledgement of Acceptable Use Policies.
- Delegating the responsibility of enforcing sanctions for Acceptable Use Policy violations.
- Ensuring an Acceptable Use Policy is developed and disseminated in accordance with this Policy.

Agency Chief Information Officer (CIO) – is responsible for ensuring an End User System Access and Acceptable Use policy statement is provided to staff, contractors and visitors that use private State systems prior to granting them access.

Agency Information Security Officer (ISO) – is responsible for:

- Conducting periodic audits of Acceptable Use Policy (AUP) acknowledgements submitted by IT system users in accordance with the Access Control Policy.
- Performing periodic audits for rogue or unapproved software.

Agency Staff – is responsible for reading, understanding and adhering to the policy and cooperating with the Agency ISO or CISO in investigations.

Agency Staff Supervisor – is responsible for ensuring that his/her subordinates have read, understood, and have agreed to the AUP and all other security policies as a condition of employment or a condition for granting access.

Requirements
Each Agency shall develop an End User Acceptable Use Policy (AUP) specific to its organization's needs. The following requirements apply:

- AUPs must identify roles and responsibilities for managers, employees, and system administrators.
- AUPs must contain sanctions for non-compliance.
- AUPs must not supersede State or Federal regulations
- AUPs must require compliance with the Colorado Cyber Security Program and Agency Policies.
- Agency employees must provide acknowledgement of the terms and conditions outlined in the AUP prior to using private Agency systems.
- Vendors, contractors, and visitors must provide acknowledgement of the terms and conditions of an AUP prior to using Agency systems.
- Agencies must record the user's acknowledgement of the AUP and maintain such acknowledgement as long as the agency systems are in use by that user.
- AUPs must define a security incident and instruct the user how to report suspected and actual incidents.
- AUPs must address usage of e-mail, Internet, telephone, remote access, and State applications.
- AUPs must state that the user has no right to privacy when using agency or State systems and that all electronic communications on State systems are monitored.
- AUPs must require the end user to use agency and State systems in a responsible, lawful, and ethical manner.
- AUPs must address responsibilities for managers with regard to hiring, transfer, and termination procedures for employees or contractors for whom they are responsible.
- AUPs must state that the use of State e-mail addresses in non-business related forums such as newsgroup postings, discussion boards, or instant messaging is expressly prohibited.
- AUPs must state that only approved software may be deployed on Agency IT systems, including P2P software, Internet Browser plug-in software, screen savers, PDA synchronization software, and encryption software, and must address procedures to request such software.
- AUPs must address the proper handling of State Data based on sensitivity (see CCSP Data Handling, and Disposal Policy P-CCSP-011).
- AUPs must specify the appropriate use of agency-owned and personally-owned removable media or external devices.
- AUPs must state that the disabling of security controls is a violation of policy.
- AUPs must specifically restrict intentional attempts to compromise State systems or data, to include network scanning, vulnerability scanning, security testing, or password cracking unless specifically authorized.

Each Agency shall identify in their Cyber Security Program Plan the methods of monitoring for AUP violations, for enforcing sanctions, and for updating the Cyber Security Program Plan to include enhancements to user training and awareness.

Guidelines
See Sample Agency Acceptable Use Policy, G-CCSP-012-1.

References
Sample Agency Acceptable Use Policy, G-CCSP-012-1.

Online Privacy Policy

Policy

Agencies shall provide a specific statement regarding information privacy to each end-user of public systems administered or operated by the Agency.

Definitions

For the purposes of this document, please refer to C.R.S. 24-37.5-102, C.R.S 24-37.5-402 and the Colorado Cyber Security Program Policy Glossary for any terms not specifically defined herein.

Requirements

At a minimum, a link to a privacy statement shall be provided to the user on each “page” that requests private data from a citizen. All notices must address:

- Rationale for the collection of any personal information and a description of how it will be used and who it will be shared with.
- Assertion of security and integrity protection provided to the information collected to prevent unauthorized disclosure or loss.
- End-user options regarding restriction on the collection and use of their personal information.
- End-user access to collected information and options for identifying and correcting errors.
- Options for reviewing records of the State of Colorado’s compliance with its privacy policy and information security practices.
- Options for recourse if the user data is misused. In the event a site or application interface does not collect private data from users, no privacy notice is required.

Responsibilities

Executive Director – responsible for:

- Enforcing the policy within his/her department and mitigating operational risks associated with State of Colorado’s commitment to online citizen information privacy principles.
- Designating a Point of Contact for reviewing complaints and authority to resolve issues.

Agency Chief Information Officer (CIO) – is responsible for ensuring all “Internet-facing” systems intended for use by the public are deployed with this notice.

Agency Information Security Officer (ISO) – is responsible for ensuring the implementation of online privacy safeguards as required by this policy.

Guidelines

This section describes best practices for meeting the objective of this policy. Online Privacy Statements shall implement and be in compliance with the following: COPPA; World Wide Web Platform for Privacy Preferences Project (W3P3P) 1.0.

Statement of Citizen Privacy and Online Access

To assist each State of Colorado citizen in understanding the rules governing electronic account access, a Privacy Policy Statement is provided on the State of Colorado’s Web site. This statement addresses and documents the applicable laws and regulations, State of Colorado’s online privacy principles, specific citizen rights and options, and citizen alternatives for accessing information. A link to this statement is to be provided from each public-access system.

Online Processing Privacy Notice

A link to view the State of Colorado's Privacy Policy Statement is to be provided when a citizen accesses any Internet site doing business on behalf of or in representation of the State Government.

References

- COPPA – Children's Online Privacy Protection Act
- World Wide Web Platform for Privacy Preferences Project (W3P3P) 1.0
- Better Business Bureau (BBB) Online

Appendix H - Acknowledgements

The Council wishes to acknowledge the following organizations and individuals who provided it assistance and time with this project:

- National Association of State Chief Information Officers – Doug Robinson, Executive Director; Eric Sweden, Senior Enterprise Architect
- National Center for Higher Education Management Systems – Peter Ewell, Vice President; Patrick Kelly, Senior Associate
- State of Arkansas – Claire Bailey, Chief Technology Officer
- State of California – Christy Quinlin, Chief Deputy, Office of the Chief Information Officer; Dale Alvarez, Data Processing Manager, Enterprise Architecture
- State of Kansas – Bill Roth, Chief Information Technology Architect; Bryan Dreiling, Information Technology Architect
- Commonwealth of Kentucky – Glenn Thomas, Director of Enterprise Governance, Commonwealth Office of Technology
- Commonwealth of Virginia – Virginia Enterprise Application Program: Peggy Feldman, Director, Virginia Enterprise Applications Program (VEAP) / Chief Applications Officer; Nadine Hoffman, VEAP Data Management Lead; and, Will Goldschmidt, VEAP Policy and Strategy Lead

Appendix I - References

1. “Data Governance – Managing Information As An Enterprise Asset, Part I, An Introduction”. National Association of State Chief Information Security Officers, April 2008.
2. “Data Governance Part II: Maturity Models – Setting the Vision”. National Association of State Chief Information Security Officers, January 2009.
3. The Federal Data Reference Model – see Glossary of Terms.
4. Mearian, L. “Study: Digital universe and its impact bigger than we thought.”, March 11, 2009, Computerworld.
5. “Data Governance: Changing Culture, Breaking Down Silos and Deciding Who Is in Control”. Data Quality Campaign, August 2008.
6. “Think Before You Dig: Privacy Implications of Data Mining & Aggregation”. National Association of State Chief Information Officers, September 2004.
7. “Data Governance: Is your Data Integration Technology Ready?”, Informatica, retrieved from Internet (www.informatica.com) on November 26, 2008.
8. “Optimizing State Investments for Justice Information Sharing”, National Governor’s Association Center for Best Practices, Economic and Technology Policy Studies Division, November 2002.
9. “Government Worth Having: A briefing on interoperability for government leaders”, Center for Technology in Government, October 2008.
10. “State Sector Strategies: Regional Solutions to Worker and Employer Needs”, National Governor’s Association Center for Best Practices, Social, Economic and Workforce Programs Division, November 2006.
11. “Improving Government Interoperability: A capability framework for government managers”, Center for Technology in Government, October 2008.
12. “En route to seamless statewide education data systems: addressing five cross-cutting concerns”, State Higher Education Executive Officers, August 2008.
13. “Using Data to Improve Student Achievement”, Data Quality Campaign, September 2008.

Appendix J - Glossary of Terms and Acronyms

Automated Fingerprint Identification System (AFIS) - Fingerprint identification system at the Colorado Department of Public Safety.

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 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 Federal Enterprise Architecture and E-Government 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

Consolidated Reference Model - The Federal Enterprise Architecture (FEA) consolidated reference model (CRM) document was published in October, 2007 and 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 FEA. The Data Reference Model, DRM, is referenced but not repeated in this document due to its complexity and volume. Abbreviated as CRM.

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 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 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.

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 – The Family Educational Rights and Privacy Act (FERPA) is a Federal law that protects the privacy of student education records. The law applies to all schools that receive funds under an applicable program of the U.S. Department of Education.

FIPS - Federal Information Processing Standard (FIPS), one of many standards set by the Federal government for exchanging or processing data.

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 - 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 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.

K-20 – Education from kindergarten through post-graduate college.

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), links to physical constructs, values, exceptions, and data steward information.

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.

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.

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.

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.

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

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

PLC – Acronym for the Prevention Leadership Council.

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

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

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

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

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

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.

XML – See Extensible Markup Language.

State Agency Acronyms

Attorney General (DOL)

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 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)