11-2001

Roadmap for Data System Integration

Dennis Tan  
*Edmund S. Muskie School of Public Service*

Eileen Griffin  
*Edmund S. Muskie School of Public Service*

Follow this and additional works at: [http://digitalmaine.com/dhhs_docs](http://digitalmaine.com/dhhs_docs)

**Recommended Citation**  
[http://digitalmaine.com/dhhs_docs/6](http://digitalmaine.com/dhhs_docs/6)

This Text is brought to you for free and open access by the Health & Human Services at Maine State Documents. It has been accepted for inclusion in Health & Human Services Documents by an authorized administrator of Maine State Documents. For more information, please contact statedocs@maine.gov.
EDMUND S. MUSKIE SCHOOL OF PUBLIC SERVICE educates leaders, informs public policy, and broadens civic participation. The School links scholarship with practice to improve the lives of people of all ages, in every county of Maine, and in every state in the nation.

Interdisciplinary, hands-on academic programs develop skills, expertise, and leadership in the public, private, and non-profit sectors. Collaborative partnerships with public agencies and private organizations generate innovative and effective approaches to complex challenges.

Research projects and public service activities inform policy makers and improve decision making through research, policy analysis, program evaluation, professional training, technical assistance, and demonstration projects. Professional development programs, policy conferences, public affairs broadcasts, and publications engage and inform the community, the state, and the nation.

### Academic Programs

**Master’s Degrees**
- Community Planning & Development
- Health Policy & Management
- Public Policy & Management

**Ph.D. Program**
- Public Policy & Management

**Certificate of Graduate Study**

**Joint Degrees with the University of Maine School of Law**

**Professional Training**

### Research and Public Service Programs

- Institute for Child & Family Policy
- Institute for Health Policy
- Institute for Public Sector Innovation

**National Centers**
- Community Partnerships for Older Adults
- New England Environmental Finance Center
- National Child Welfare Resource Center for Organizational Improvement

[www.muskie.usm.maine.edu](http://www.muskie.usm.maine.edu)
Roadmap for Data System Integration

Dennis Tan and Eileen Griffin

Institute for Health Policy
Edmund S. Muskie School of Public Service
96 Falmouth Street
P.O. Box 9300
Portland, Maine 04104-9300

November 2001

The Center for Health Care Strategies (CHCS), in Princeton, NJ, provided funding for the Olmstead Planning Grant. This grant was made possible through a separate grant to CHCS by the Robert Wood Johnson Foundation.
Acknowledgments

The authors would like to thank the many people who played a role in guiding and shaping the contents of this document. Colleagues at the Muskie School, including Maureen Booth, Paul Saucier, Elise Bolda, David Lambert, Larry Glantz, Gino Nalli, and Bob Keith at the very early stages of this project generously offered their time, expertise and historical memory to help frame the scope of this project and the ways that the State might use integrated data to better serve persons with disabilities. Members of the Steering Committee for Community-Based Living played a critical role in identifying the State’s objectives for inter-departmental coordination. Members of Maine’s Plan Development Work Group for Community-Based Living contributed their perspective as consumers, on the need for better integration of data and their concerns about the protection of that data. Staff from five departments within the State, including information systems staff, quality management staff, and program administrators generously provided information, as well as critiques and further guidance on the ways that integrated data may or may not serve their needs.

Funding for this effort was provided through a grant from the Center for Health Care Strategies (CHCS) to the Maine Department of Human Services. CHCS received its funding through the Robert Wood Johnson Foundation.

The views and opinions expressed in this report should not be attributed to the funders or to the University of Southern Maine.
# Table of Contents

Background ..................................................................................................................................... 1

Objectives for Coordination and Data Integration ........................................................................... 3

Data System Integration .................................................................................................................. 5
  - System Design
  - Status of Existing Integration Efforts
  - Factors Affecting System Integration

Core Data Set .................................................................................................................................. 17
  - Data Domains
  - Collecting Data
  - Using the Data
  - Data Elements
  - Applying the Measures

Next Steps for Data System Integration .......................................................................................... 38
Background

The 1999 U.S. Supreme Court decision, *L.C. v. Olmstead*, is the catalyst behind an effort to develop a strategy for integrating data across several of Maine’s departments. The *Olmstead* decision requires states to provide services to people with disabilities in the most integrated setting appropriate to their needs, in accordance with the Americans with Disabilities Act (ADA). The Supreme Court said that states can show they are complying with the ADA if they have a comprehensive, effective working plan for placing qualified persons with disabilities in less restrictive settings.

*Olmstead* applies to all people with a disability, as defined under the ADA, institutionalized or at risk of institutionalization (including in a nursing facility, treatment facility or other restricted settings). The ADA protects people with any type of impairment that “substantially limits one or more of the major life activities,” including a physical disability, a mental illness, a developmental disability, a chronic illness, an addiction disorder, blindness, and deafness. All state or other publicly funded institutional or home and community-based services offered to people with disabilities are subject to the *Olmstead* decision.

In January 2000, the U.S. Health Care Financing Administration (HCFA) wrote a letter to all state Medicaid directors encouraging them to take the lead in their states in developing a “comprehensive, effective working plan.” Beginning in February 2000, a Steering Committee for Community-Based Living was convened by the Maine Department of Human Services (DHS), comprising members from DHS, the Department of Behavioral and Developmental Services, the Department of Education (DOE), the Department of Labor (DOL), and the Department of Corrections (DOC). This Steering Committee was charged with developing Maine’s response to the *Olmstead* decision.

The members of the Steering Committee agreed that the *Olmstead* decision was an opportunity for their departments to develop a coherent vision, across departments, for providing services to people with disabilities. Seeing the *Olmstead* decision as a catalyst for a comprehensive, thorough evaluation of the State’s capacity to maximize community integration, Maine is creating a plan for improving home and community services.

In September 2000, the Maine Department of Human Services was awarded a grant from the Center for Health Care Strategies (CHCS) to fund the development of a strategy for integrating data across multiple departments. This document is the product of that grant. Through a series of meetings and discussions, it has been written with the guidance of members of the Plan Development Work Group for Community-Based Living (comprising Steering Committee members and consumer representatives), the Steering Committee, and state program management, information technology, and quality management staff.

---

1 Now called the Centers for Medicare and Medicaid Services (CMS).
2 Formerly the Department of Mental Health, Mental Retardation and Substance Abuse Services.
Developing an integrated data system is not an Olmstead "requirement." Rather it is one tool the State can use to determine how well it is complying with Olmstead and to identify ways for improving its performance. As Maine considers how to expand and build its system of home and community-based services, the State needs integrated data to guide its decisions on allocating scarce resources. To paraphrase one state bureau director, if Olmstead requires Maine to launch an expansive (and expensive) array of home and community services — the equivalent of an aircraft carrier — integrated data provides the navigation system. Without data linked across programs, the State does not have a comprehensive view of how many people it serves, what services are provided, to whom and by whom. Without this information, the State cannot comprehensively evaluate the needs and unmet needs of the people it serves, the effectiveness of its services, or whether its allocation of resources are maximizing sustainable community living.

This document is intended as a "roadmap" for future efforts at integrating data. The roadmap is written for program administrators who need to understand how an integrated system might work and what their added role and responsibilities would be. While written upon the foundation of technical expertise, the goal was to write a Roadmap accessible to those without technical expertise. This document describes:

- Maine’s objectives for inter-agency coordination and data system integration;
- the design and status of data system integration efforts in Maine;
- a roadmap for collecting and using data elements that will support the state’s objectives for data system integration; and
- a set of recommended “next steps” for data system integration.

Subsequent efforts funded under CMS’ Real Choices initiative will more fully define the operational features of integrated data. This grant will bring Maine one step closer to developing the technical architecture for managing and monitoring the full scope of services for persons with disabilities.

While this document is written within the context of the Olmstead decision, its application clearly reaches beyond Olmstead to many initiatives in which the value and need for integrated data has already been identified (e.g., the Children’s Cabinet Maine Marks program, the Children’s Mental Health Oversight Committee). The flexibility of the design for integrating data can meet many of these overlapping needs.
Objectives for Coordination and Data Integration

Steering Committee members agreed that they need to coordinate their efforts to provide services to people with disabilities. The five departments included in this initiative are:

- the Department of Human Services (which includes the state Medicaid agency, the agency providing services to the elderly and physically disabled, the public health agency, the child welfare agency, and the agency with responsibility for income support);
- the Department of Behavioral and Developmental Services;
- the Department of Education (special education services);
- the Department of Labor (rehabilitation services, assisted living services); and
- the Department of Corrections.

These departments have identified the following objectives for inter-departmental coordination:

- coordinated provider or resource development;
- coordinated service structure (e.g., eliminating gaps and overlaps);
- coordinated service delivery (e.g., single point of entry, integrated case management);
- coordinated response to judicial or legislative directives mandating coordination of services to people with disabilities (including the *Olmstead* decision, Chapter 790, the *French* settlement, consent decrees).
- coordinated monitoring of system performance (including quality of care, quality of life measures for the people served); and
- coordinated approach toward funding and allocating resources.

For the purposes of the first phase of data integration, we have not attempted to support all of these objectives. First, the core data set does not directly support the objective of coordinating care at the service delivery level. That is, this effort at system integration does not address the State’s interest in integrating or coordinating case management across programs; nor does it directly support a single point of entry. The focus of this core data set is on coordinating the management of the system at the aggregate, or system-wide, level. This level of coordination may have an indirect impact on the coordination of the delivery of services, but it is not the principal objective. In addition, this first phase of developing a core data set does not try to measure performance of the service system. That is, it does not try to measure the impact that the service system has on a desired outcome. Identifying a set of desired outcomes and deciding how to measure performance against those outcomes is outside the scope of this project, although an important next step for data system integration.

In an effort to support the other objectives identified by the Steering Committee, during this first phase, a series of preliminary measures were developed. These measures are descriptive rather than evaluative, providing the State with basic information to support compliance with *Olmstead* currently beyond its reach: an unduplicated count of the numbers served, the number in need but not being served, the services provided, etc. Across departments and programs, these measures are:
- who the State serves (including basic characteristics like age, gender, etc);
- the strengths, preferences and needs of the people the State serves;
- the services received by the people the State serves;
- which departments are serving whom;
- which providers are serving whom;
- which of the identified needs and preferences of the people the State serves are unmet;
- how many people are on waiting lists for which services;
- how long people have been waiting on waiting lists;
- the setting in which people receiving services reside;
- who or who is not being served in the most integrated setting appropriate to their needs;
- the resources the State has available for providing services;
- how much it costs to serve each person; and
- barriers to access.

These measures guide the scope of the core data set described in the second half of this document. The design for data system integration, described in the next section, is independent of these "Olmstead measures."

To develop these measures across departments, the data must be linked by personal identifiers. Once linked, identifiable information need not be accessible to end-users. For the purposes of these measures, the State may only need non-identifying information to analyze system-wide performance. The integrated system described in this document can prevent access to identifying information, or allow it for limited purposes.
Data System Integration

For many reasons, and with legislative and judicial actions prompting heightened interest, the State has already begun the process of building an integrated data system. In this section we:

- describe the overall design of the integrated system;
- identify the current status of system integration efforts; and
- review some of the factors that will affect data system integration.

System Design

The SCHEMATIC on page 16 provides a visual framework for discussing a design that would support integration across the diverse sets of data resources needed here. The basic components of that design include:

- the guiding principles for system integration;
- the sources of data;
- integration functions; and
- report extraction function;

Guiding Principles for System Integration

Each of the departments and bureaus providing source data for an integrated data system has created that source data to meet its own needs. Each department is also at different stages in developing its information systems infrastructure. In some cases, the different needs and stages of development for data system infrastructure are driven by the services or programs administered through that department. In other cases, external factors, such as budget constraints, federal reporting requirements, confidentiality or political factors, may have determined which data is collected and how it is stored. Integration of these data systems must take into account not only these pre-existing constraints on what is available, but also the likelihood that any or all of these conditions may change over time. In addition, the integrated system must recognize that the department or bureau collecting the data must maintain the authority to manage and change that data as necessary to meet the needs specific to the programs the data supports and to respect confidentiality provisions inherent to their functions.

To that end, the design of the integrated data system is built upon the following principles:

---

3 The design described in this document is the product of many years of work, guided by the vision and direction of Dennis Tan, consultant, Fran Finnegan, former Director of the Bureau of Medical Services, Sawin Millett, former Associate Commissioner of the Department of Mental Health, Mental Retardation & Substance Abuse (DBDS), and Walter Lowell, Agency Technology Officer of DBDS. This system design is described here to provide a common understanding of the direction undertaken, the constraints that influence that direction, and to provide further opportunity for input.

4 This description is intended to provide a vocabulary and framework for discussion for a lay audience. A technical appendix for information system staff is planned.

Edmund S. Muskie School of Public Service 5
- the source location of the data (i.e., computer system location as well as database structure location);
- a hierarchy for determining data quality (i.e., guidelines and rules for resolving conflicting values for the same data element, such as gender); and
- transformation rules that convert the source data to the integrated data (e.g., if the departments all have different ways of defining race and ethnicity, the transformation rule will convert all data sources to one common structure and code set).

Thus, when an end-user wants to know what needs have been identified for people served across departments, tables stored in the meta-data database will identify where that information is located within the source data sets for each department. The first layer of software will follow the roadmap to that data, returning it to the integrated system. The second layer of software accesses the meta-data roadmap to find out how the source data is to be converted to integrated data.

Mapping the data through the meta-data functions minimizes the disruption resulting when the source data changes. As an example, if DBDS proceeds with its plans to implement a behavioral health managed care program, the business needs for a managed care program will require changes to DBDS’ existing information system, changing the conditions (i.e., parameters for collecting and editing data) and availability (i.e., electronic accessibility) of data to be contributed to the integrated database. Existing data formats might change, new data elements might be added, required fields might become optional. In addition, new specialized managed care needs will require new information systems for enrollment, changing a primary care physician, or calculating the capitation rate. The system design for the integrated database must be capable of accommodating these changes with a minimum of computer programming and down-time for the integrated system. To achieve that goal, DBDS will alert the integrated system administrator of the change to the data source. The system administrator will then revise the data maps stored in the meta-data database using the online meta-data administration screens to reflect the updated data sources.

**Report Extraction**

A meta-data function will also be used for supporting the end-user in managing the complexity of an integrated database derived from multiple sources. The meta-data function will build in rules to aid in the interpretation of data, and the underlying business rules driving the collection of data from multiple sources. In some cases, the meta-data function will enforce standards for interpreting the data, mitigating erroneous query results and misinterpretation. The meta-data layer will also be used to automatically disable certain linkages which might appear to be logical but are inconsistent with the business practices of the data source.

At this layer, the data will be grouped by domain, or category of information. For example, all data across departments related to the person (e.g., name, identifiers, date of birth) will be grouped. All assessment data across sources will be grouped. Grouping the data in these broad, logical categories provides a framework for managing the data. These domains will be supported by software and online screens. (For more discussion of the domains, see the CORE DATA SET discussion below.)
At least two audiences for extracted reports/information are anticipated:

- **Authorized state staff.** Reports for authorized state staff will likely be in two forms: aggregate information without identifiers and case specific information with identifiers disclosed. Each department will determine its own standards for who is authorized to access identifying information through the integrated data system. The systems controls will not only limit which end-user has access, it will limit what data elements each end-user is entitled to view. Access to identifying integrated data will be denied in the absence of a justified need.

- **The general public.** Making standardized reports, or possibly even data sets (with all identifying information eliminated) available to the general public would enhance the community’s understanding of the resources and needs of people with disabilities. To permit public access to a data set, special protections would be necessary. Even with identifying information eliminated, the system would need to be able to adjust the rules depending on the queries asked to make sure that identity could not be discovered. For example, a public user would not be able to ask for information that might identify people within a small geographic areas or served by a specific provider, since the identity of some people might be recognized. The system would allow the same queries over broad geographic regions or broad categories of providers, to obscure identity.

The report extraction component will support analysis of the integrated data in a number of ways. Examples of the way the reporting function can support analysis of the integrated database include:

- customized reports;
- automated grouping;
- enforced limitations; and
- security.

**Customized reports.** The end-user will be able to customize the reporting information available, responding to the specialized needs of the end-user or set of end-users. For example, the value for date of birth will be translated into a set of age groupings customized to the needs of the department analyzing the data. BMS might need only to know whether a person is a child or an adult, while DOE will need a much finer age breakdown for just children. When an end-user enters a query, only the pre-defined age groupings for that department will be produced. These customized reports will also be capable of accepting various flexible parameters entered by the users when they request a report. These parameters will be tailored automatically to the data sources selected for the report.

**Enforced standards for querying and using the data.** The meta-data function will allow the source department to define the business rules and policies associated with specific data or groupings of data. The integrated system will reflect these limits on the data. For example, Medicaid’s primary diagnosis has a very specific meaning and specific limitations. Similarly, the behavioral health procedure codes for physician services have very specific meaning and
specific limitations. Consistent with the business rules governing these elements, the integrated system will allow some queries, but will block others. Similarly, the system will block queries for certain combinations of elements, limit their use, or, if use is permitted, return the results of a query with a warning explaining the limitations of those results. For example, when a query targets Medicaid’s primary diagnosis as part of its criteria, an appropriate warning message will be displayed and printed, listing the Medicaid billing policies and rules governing primary diagnosis. In the case of changes to the source data, the end-user will be informed of the data changes and be advised of the impact these changes will have on previously created queries and the appropriateness of future queries.

Security. There is a close relationship between the meta-data and the security functions of the integrated database. The system will need to be able to identify and differentiate among users of the system. The security system must also map the organizational structure for each user, enabling restrictions to be applied at the organizational level, in addition to the individual level. This ability goes beyond the simple concept of user identification and access to screens and reports. The combination of the meta-data and security must work jointly to filter what fields are visible to which user, as well as what types of information are available for each permitted field. (For example, a user may have access to diagnosis information but only to those clients that are “registered” or “enrolled” with that user.)

Status of Existing Integration Efforts

Within the State there has been great interest in integrating data. Different departments and bureaus have made progress toward linking their information systems. For many integration efforts, progress has been slowed by existing infrastructure; if data integration is an afterthought, not built into the architectural design of the system, the ability to integrate data is usually very limited. As a result, rather than a “seamless,” or invisible (and realtime) interface between data systems, integrated data must be accessed through batch, manual data exchange.

BMS and DBDS have been working together to build integration into their information system architecture. BMS and DBDS jointly manage the behavioral health portion of Maine’s $1.2 billion Medicaid program. This shared responsibility has necessitated close integration and cooperation, including for information systems.

During the last two years, DBDS has embarked on the phased implementation of an Enterprise Information System (EIS) that will seamlessly integrate all business areas (mental health, mental retardation, children’s services and substance abuse). The core of this system is scheduled to be implemented by January 2002. The EIS system will also be integrated with BMS’ claims processing systems and the Maine Medicaid Decision Support System (MMDSS). Additionally, the EIS will enable integration with DBDS’ state facilities (state hospitals, etc.).

Over the last five years, BMS has moved forward on a strategy of modular/functional replacement of its Medicaid legacy mainframe systems. This strategy includes the development and successful implementation of the following large information systems:

- Maine Medicaid Decision Support System (MMDSS)
Maine Enrollment and Capitation System (MECAPS)
Maine Long-term care Assessment System (MECARE)
Maine/New Hampshire Joint Immunization Registry (IMMPACT)
Maine Behavioral Health System – jointly with DBDS (EIS)

The final component of the Medicaid legacy mainframe system consists of the Maine Claims Management System (CMS). This system will create infrastructure that will enable every Medicaid provider throughout the State of Maine to have online, real-time access to submit Medicaid claims and encounters, and to also receive and collect integrated data. A vendor for this project will be selected by late summer of 2001. DBDS is a joint partner on the CMS project.

Factors Affecting System Integration

HIPAA

Over the next two years, the State will need to take steps to ensure its compliance with the Health Insurance Portability and Accountability Act of 1996 (HIPAA). HIPAA has significant implications for data integration.

*HIPAA Overview.* HIPAA reaches this initiative via two avenues. First, HIPAA establishes standards for "administrative simplification." It establishes standard formats for certain electronic transactions (e.g., claims, eligibility, payment), and adopts standard code sets (procedures, diagnoses, etc). Under HIPAA, unique identifiers for health plans, providers, employers, and individuals are to be used for all transactions within the scope of HIPAA.

HIPAA also addresses security and privacy, requiring covered entities to identify who needs access to private health information (PHI) and what level of access. HIPAA governs routine disclosures and non-routine disclosures of PHI, requiring a covered entity to verify an individual's authority to access PHI before the PHI is disclosed. Every time a non-routine disclosure is made, it must be documented; HIPAA requires that covered entities be able to provide individuals with an accounting of disclosures going back six years.

HIPAA reaches health plans, health care providers, health care "clearinghouses," and "business associates" that perform functions requiring the disclosure of private health information. The U.S. Department of Health and Human Services (HHS) has promulgated rules under HIPAA and the Office of Civil Rights, within HHS, will be responsible for enforcing compliance. Those within the reach of HIPAA will need to come in compliance with the administrative rules and the privacy rules over the next two years.

*Applied to this initiative.* HIPAA applies to the Medicaid program and will reach to the operations of many other government programs, including some of those party to this initiative. Other bureaus within DHS, along with DBDS, will need to come into compliance with HIPAA. The Bureau of Rehabilitation Services, within the Department of Labor, is very likely to have
private health information within the reach of HIPAA. HIPAA also reaches health care provided to prisoners in Maine’s correctional facilities. Education records, including special education records, are specifically excluded from the reach of HIPAA, although health care providers in schools will still need to comply with HIPAA when engaging in a HIPAA transaction (which would include submitting a claim for Medicaid reimbursement).

HIPAA will have a positive impact on data integration. Establishing standard formats and code sets will assist in standardizing data across departments and programs. Data collected with a unique identifier consistent across all records will make linking data a much simpler task. HIPAA will also be a very powerful tool for addressing privacy concerns. By coming into compliance with HIPAA, the State will be doing much of the work that would need to be done for this initiative -- determining who has access, the level of access, etc.

Confidentiality and Data Protection

The ability to identify persons across programs is critical to integrating data across programs. However, for many uses, the value of integrated data does not depend on the disclosure of identifiable data. The State will need to determine when (and if) the disclosure of identifiable integrated information is justified. By responding to HIPAA, the State will already be doing much, but not all, of this work. While HIPAA sets a “minimum” standard, other state and federal statutes are stricter and there will likely be variation across departments in what standards apply. Each department (not the integrated system) will still need to determine the level of protection afforded their identifying data when accessed through the integrated system. The integrated system will need to enforce these department level protections regarding who has access to what information.

For the integrated system, developing a strategy for protecting identifying data will be a two prong process. First, the individual departments will have to develop a protocol for sharing identifying data and determining who will have access to the data, taking into account all of the constraints limiting access within their own departments, including HIPAA. This strategy should be recorded in a memorandum of understanding. In addition, all users of the integrated system will need to agree in writing to comply with confidentiality standards set for identifying data acquired through the integrated system. (These standards will need to be at least as rigorous as the standards of the departments.) In addition, the integrated system will need to develop the data protections to enforce the access limitations electronically.

For the purposes of this document, we assume that confidentiality and data protection are not a barrier to data system integration. We assume the departments can develop, or have developed, a strategy for sharing data across departments. Information systems technologies have matured to the point where the technical ability to limit access to person-level data to only those given permission to access that data by the client, can and have been accomplished. We also assume that as a general rule, the only data disclosed, or published, as a result of data system integration

---

9 The privacy protections and rights associated with education records are governed under the federal Family Educational and Privacy Act (FERPA).
10 For example, Massachusetts’ implementation of its CommBridge system and DBDS’ enterprise information system.
will be aggregate information with no associated identifying information. Access to identifying information will be limited, available only to authorized users with a justified need.

**Inter-Departmental Agreement, Commitment and Investment**

Ultimately, the progress of an integrated data system will depend on the interest and commitment across the departments participating in this initiative. At a minimum, an integrated data system will require:

- some level of effort on the part of departmental and program information systems staff in declaring data to the integrated system, and otherwise interfacing with the integrated system on an ongoing basis;
- inter-departmental agreement on the administrative infrastructure responsible for administering the meta-data function;
- inter-departmental commitment to financing the administration of an integrated data system;
- inter-departmental agreement on how to address confidentiality concerns;
- departmental investment in program managers and quality assurance staff in learning how to use and analyze inter-departmental data.

In addition, the State’s success at integrating data will depend on leadership committing resources, educating and communicating to program and information system staff the value of integrated data to the State, and cultivating the “cultural shift” required to increase the understood value of inter-departmental coordination and communication at all levels.

Because the architecture of an integrated system can be built in the absence of data to integrate, an incremental approach to department participation is possible. The State’s recognized need for integrated data -- reflected in both administrative and legislative initiatives, and furthered by judicial pressures -- might provide enough incentive to build the infrastructure around a few key programs, allowing integration to proceed at an incremental pace for others. A pilot would minimize risk and demonstrate “do-ability” for the participants and others. As the level of participation grows, the value of the integrated system will also grow, likely bringing with it greater participation.
Inter-Departmental Core Data Integration Project
Proposed Cooperative Processing Schematic.

META-DATA TABLES
META DATA ADMINISTRATION & MAINTENANCE FOR (A), (B), & (C)

LOAD HOLDING TABLES

DATA SOURCES

DEPARTMENTAL DATA SOURCES

STATE OF MAINE COMPUTER NETWORK (WAN)

LOAD

WWW/Internet

NON-IDENTIFYING INFORMATION

WEB-HOSTED REPORTS
INDICATOR REPORTS
ADVANCE QUERY TOOLS (e.g., Trend & Data Analysis, etc.)

DATA INTEGRATION PROCESSORS

SECURITY AND ACCESS MANAGEMENT APPLICATION LAYER

BACKUP, RECOVERY AND SYSTEMS OPERATIONS

SOFTWARE FOR COMMUNICATING BUSINESS, RULES/POLICIES & LIMITS OF DATA

CONSOLIDATED INTEGRATED INTER-DEPT DATABASE/WAREHOUSE

INTEGRATION OF DATA

General Public

'Authorized' Online Query 'Authorized' Online Query

NON-IDENTIFYING INFORMATION

Response to Queries

Drawing2 4/24/2001

Institute for Health Policy
Core Data Set

The integrated system will produce an integrated database. Given the various stages of development in each department’s information infrastructure, it is unlikely that the integrated data set will have access to all of the elements needed to support the measures identified at the beginning of this document. In addition, because of the diverse functions of each department, bureau, and program, inevitably (and, in many cases, appropriately) there will be inconsistencies among the departments in what is collected and how it should be interpreted. This section attempts to create a roadmap for the type of information that is needed to support the objectives for coordination, indicating where consistency across departments should be sought. This section also outlines many of the issues that must be confronted in collecting and using data across a diverse set of data sources. This section:

- describes an organizational framework for the data, or the data domains;
- identifies issues related to collecting the data;
- identifies issues related to using the data;
- identifies a set of core data elements that would support the objectives for inter-departmental coordination; and
- applies the inter-departmental measures to the data elements.

Data Domains

The data is organized into eight domains. These domains are generalized categories of information organized around the core business areas of the departments (recognizing that not all departments will have all core business areas.) These domains provide a logical framework for managing the integrated database and core data set. The eight domains are organized as follows:

- **People (or Client) management information.** This domain will group all identifying information and characteristics related to a person for whom information is collected (not always the same as a client). This information will include identifying information (e.g., name, personal identifiers, social security number) and demographics (date of birth, place of birth, gender, race and ethnicity).

- **Provider management information.** This domain will group all information about direct service providers and their capacity. Included might be locations, number of licensed beds, staff and their credentials, etc. This information would be collected by the departments through licensing, provider enrollment in the Medicaid program, or as part of the contracting process.

- **Client capacity and needs management information.** This domain will group all information related to a person’s capacity and needs. Needs may be identified and captured through a variety of tools or methods. Each program or department that determines needs will have its own tool or process for doing so, with its own internal logic, determined by the specific needs of the program administered.
Plan management information. This domain will capture all information related to a person’s service plan for addressing identified needs. The plans may be of varied form, depending on the program for which it was developed and may span varying durations.

Services. This domain will group all information related to who received a service, the provider that provided or billed for the service, the type of service received, the reason for the service, and when and where a service was received. Services will certainly include claim information but will also include non-claims sources (e.g., service encounters reporting, performance based contracting or budget reporting, financial or grant reporting, etc.). Referrals (to whom, by whom, and date) will also be included.

Financial management information. This domain will group all information related to the cost of services provided, who paid for it, etc. In addition, this section will group all information related to departmental pricing rates, and programmatic costs associated with the service. It is important to remember that the integrated database is not designed or intended to be a financial or accounting system. This domain will not accommodate complicated cost allocations and distributions. The financial information available will simply provide a ‘costing’ of services provided without allocation to funding sources.

Human Resources. This domain will group all information related to human resources but not direct service providers. For example, some departments have staff providing direct service and staff for administering and managing the program. Information related to administrative staff will be grouped here. Human resources can be related to any organization or entity (e.g., state department, provider, schools, even family resources). Often this information is relevant when analyzing support availability and adequacy, as well as other indicators.

Outcomes. This domain will group all data measuring system performance. Included will be both measurements captured by the individual departments and measurements developed and stored by the end-user. Keeping in mind the measures identified for inter-departmental coordination, the outcomes focused on here will be waiting lists and measures of barriers to access. It is anticipated that many departments currently capture non-client identifying outcome measures and performance indicators. The system design proposes to collect this information and make it available within the integrated database.
Collecting the Data

Collecting data across diverse sources raises multiple issues. This section:

- explains the relationship between departmental ownership of the data and the role for inter-departmental coordination for collecting the data;
- describes some of the data collection issues that cross all domains or data elements;
- describes some of the data collection issues specific to certain domains or elements.

Departmental Ownership and Inter-departmental Coordination

In general, data is collected as part of a “transaction” performed in administering a program or delivering a service. A transaction might be providing a referral, determining eligibility, performing an assessment, developing a plan of care, providing the service or submitting a claim.

The nature of each transaction must be determined by the department, bureau or provider administering or providing a program or service. Each program will have “business rules” for how a transaction is performed. For example, for each program, there will be standards for what type of information is required for determining eligibility or whether a claim should be paid. These business rules might be the product of a statutory or regulatory requirement, a matter of judgment or some other factor within or outside the control of the department creating the data source.

The integrated data set must be predicated on the transactions and business rules dictating the data collected, and the autonomy of each department in determining the transactions, business rules and resulting data. However, while respecting departmental ownership, there is a role for inter-departmental coordination.

Three tiers of data are identified, each tier requiring different levels of standardization across departments:

- **Required.** Elements required for linking and de-duplicating records should be standard and should be collected across all departments -- without this data there is no way to integrate the data. Which elements are required for linking and de-duplication will depend on the algorithm used for linking and de-duplication. Likely candidates include identifiers (including, but not limited to social security number, Medicaid number, provider’s unique number, etc.), name, date of birth, and gender.

- **Standard collection recommended across all departments.** Ideally, demographic information would be collected across all departments using standard measurements. Collecting consistent information about age, gender, address, etc., is not likely to interfere with program function and would greatly enhance the ability to describe the population served across departments.

- **Collected if relevant, standardization recommended where possible.** All other information should be collected as relevant to the program, where possible using
standard values. For example, claims and service data should or should not be collected depending on program needs. When service data is collected, ideally it would be collected using standardized forms (e.g., HIPAA form 837) to maximize the ability to integrate the data across departments.

For elements with a specific set of possible values, ideally, the departments will agree on a set of values so that the information will be recorded consistently across programs. For example, for elements such as race and ethnicity, the value recorded must be one of several options (e.g., Asian/Pacific Islander, Black/African American). For those elements, which have changed over time, the possible values recorded by one department might be different from the possible values allowed by another. These departments will need to resolve these inconsistencies. In general, when possible values are not dictated by program needs, the departments should look to national standards, such as those of the United States Census Bureau, to maximize comparability with other data sources.

In the absence of, standardized data elements, we recommend that each department:

- Avoid collecting data elements in groupings or “rollups” when a “continuous” value can be collected. For example, one program might need to know only that a person’s income falls within certain ranges. Rather than collecting actual income, the only data collected might be whether income falls within one of five ranges. Another department might collect a different set of ranges or the person’s actual income. Comparisons across these data sources would produce limited or unclear results. Without interfering with the operation of any program, the data could be collected in its least aggregated form (e.g., the actual amount), allowing the most flexibility for persons wishing to use the information across programs. At the very least, more discrete categories should be established allowing more options in aggregation.

In addition to standardized data elements, we recommend that the departments:

- Establish consistent data collection rules. The departments need to establish rules for collecting data consistently (not necessarily perfectly) within their own programs. For example, a program should consistently apply a standard for determining when a person becomes a person for the purposes of collecting data. The department might decide that it needs to collect data on all persons applying for services, or only when a person is determined to be eligible for services, or only when a person starts to receive services. If a department or program adheres to consistent data collection practices, the integrated system can reflect that department’s data collection rules and the data can be interpreted in the context of those rules. The absence of consistent data collection practices will increase the risks of misuse and misinterpretation of the integrated data.

Data Collection Issues Across All Domains

Some data collection issues cannot be remedied by inter-departmental coordination. Because of the diverse nature of the programs administered or the services provided, or by the very fact of
integration, different problems will emerge that must be resolved for integration to occur. This section outlines rules for managing some of the problems or matters of judgement anticipated.

**Linking and “de-duplicating.”** The fundamental first step in data system integration is linking person-level records across programs and eliminating duplications. The ability to accurately link and de-duplicate person-level records across programs will depend on the consistency across data sources of the rules for ensuring that information is collected accurately; and the availability of unique identifiers to make it possible to match people across data sources.

- **Consistent data collection.** In some cases, people will receive services but no identifying information or other data will be collected. For example, services are provided to a person in crisis, whether or not identifying information is collected. A record is created but incomplete or inaccurate data is collected.

In other cases, people will not be receiving services but identifying information and other data will be collected. For example, a family member might not be receiving services, but if the family member is closely connected to the services provided he or she might become part of the case records kept for the primary client. Or information is recorded for a person requesting a service, even if the need for the service is never determined or the service is never provided.

For the purposes of linking and de-duplicating records in the integrated data set, all persons for whom a record is opened will be counted as a person. It will be up to the end-user to determine the status of each individual. For example, an end-user will have to decide how to treat the unmatched records of persons for whom insufficient identifying information was collected.

- **Unique identifiers.** Linking and de-duplicating records is complicated by the lack of a unique identifier that can match people across programs. Identifying information (e.g., name, social security number) is subject to duplication and some information is subject to change. In addition, some departments do not require the collection of unique identifiers (e.g., substance abuse services).

Some alternative identifiers include:

**HIPAA universal client identifier.** Under HIPAA, universal identifiers will be created that are to be carried across programs. When these identifiers are available it will be much easier to link people across programs since most if not all programs will use the same identifiers. However, even with the HIPAA ID, the need for alternative identifiers is not eliminated. For people whose entry into the system pre-dates HIPAA and for whom historical data is needed, there will still be a need to link back to historical identifiers.

**Other unique client identifiers.** In the absence of a universal identifier, various state departments and programs, as well as providers, assign their own ID numbers. These identifiers are limited in their usefulness, since people who do not participate in the program or do not go to that provider will not have been given an ID. Also, if the integrated system were to rely on the identifiers
issued by one department or bureau, it would need to rely on that department or bureau to enforce the uniqueness of that identifier.

Name. In addition to the fact that names can be misspelled or more than one person can share the same name, other factors make names an imperfect identifier. For example, name changes, cultural differences in the order that names are recorded, false names, spelling errors, nicknames, all create opportunities for error. A child might be receiving special education services through a school system and behavioral health services through DBDS. If that child’s name is recorded incorrectly in either system, or her name has changed, linking based only on a child’s name will mean the records will not be matched; it will appear two different people are being served when only one is receiving services. Or if records are linked based only on names, two different children sharing the same name might be treated as the same person.

Social Security Number (SSN). The Social Security Number is not an ideal identifier since not all people served will have an SSN (e.g., immigrants, newborns). In addition, in some cases more than one person share the same SSN or the SSN has been re-cycled to a new person. Also, people have the right to refuse to share their SSN.

In the absence of a unique identifier, the State will have to develop rules for deciding when information is related to the same person. DBDS currently uses a combination of the client’s first, middle and last name, date of birth, and gender as one method to link and de-duplicate. Other options include using a combination of the Medicaid ID number, SSN, or any other unique identifier (including provider specific identifiers) when available, name, gender, race and ethnicity, date of birth and zipcode.

Data Collection Issues By Domain

This section outlines some of the rules for managing data collection problems specific to certain elements or domains. When confronting these data collection issues, the challenge will be to design a solution on the integration side, rather than setting standards for data collection. By designing solutions rather than setting standards, the integrated system will be flexible and adaptive, making it better able to withstand the test of time.

People. Linking people across data sources produces a number of data collection issues:

- **Race/Ethnicity.** The method for capturing race and ethnicity is still evolving. As definitions change over time, ideally departments will be conscious of the need to coordinate changes with other departments to assure that, if not identical, the definitions are at least consistent (e.g., the sub-groups used in one department can be rolled up into the major groupings used by the other).

- **Address.** Multiple data sources will be capturing address information over time. In addition, some people will have more than one address. Ideally, all the departments will develop consistent approaches to collecting address information, capturing and identifying a person’s mailing, legal and seasonal address information as applicable. For persons who are homeless, the departments could agree to identify the town as
that person’s address. The integrated system will have to track start date and end date for each address to allow the end-user to look back in time.

- **Which person enrolled in which program.** There will be wide variation across departments in determining when a person is “enrolled” in a program, for the purposes of that program. Some programs might have a formal enrollment process; others might consider people “enrolled” based on the fact of treatment. Or a person might be receiving services through a program but is not “enrolled” (i.e., not entitled to the full range of services available for enrollees). Understanding who is connected to a program is necessary to understanding program costs and many other performance indicators. For purposes of this document, the terms “enrolled” and “program” are intended to capture the relationship of each person to the department collecting the data. Ideally, each department will develop a consistent method for declaring “enrollment” status within its program.

- **Unique identifiers.** Ideally, each department will maintain all of the unique identifiers attached to individuals by their program, department, or by providers. Keeping this information will enhance the system’s ability to match records across sources. The ability to unduplicate or link individuals is at the heart of data integration. This challenge is not unique to inter-departmental data integration but also exists within each department; in the course of developing an integrated system, each department should concurrently address its own challenges of linking data.

- **Pseudo- or unknown clients.** In some cases, incomplete identifying information will be collected. Ideally, each department will have consistent rules for handling unknown clients, allowing a record for a “pending” client to be created, with a “pseudo-client” identifier, and with an identified party responsible for completing the record with accurate identifying information. The “pending” client is usually a client entering the system in crisis, without opportunity for registering identifying information. Once identity is determined, a responsible party should complete the record. An “unknown” client is a client entering the system where the department is simply unable to collect identifying data; examples include some homeless individuals or ‘pending’ clients that have died, etc.

- **Family, household and professional relationships:** It is not clear whether an integrated system can develop rules for reconciling differences in family or household composition. For example, if one department says a child has two parents and the other says one, how can the system reconcile that conflict? If one says the mother’s name is Mary, the other Alice, who is right? Professional relationships (e.g., guardians, primary care physician, care coordinator) tend to be more easily resolved.

- **Eligibility & Legal Status.** Program eligibility and legal status (e.g., whether under court commitment, whether court-ordered treatment, etc.) will change over time. The system will need to keep track of the changes to allow the end-user to look back in time.
Provider.

- **Group versus individual providers.** Data should be collected on both the organization and institutional providers and the individual providers employed by the organization or practicing individually. The ability to track billing provider, fiscal agents, and ownership relationships is important for eliminating multiple counts of providers, a valuable capacity for addressing network adequacy, etc.

- **Consistent measures of provider capacity.** The method for capturing provider capacity data is evolving and will continue to have departmental variations. Provider staffing information, including specialties and sub-specialties are essential elements. Each provider organization will report to each department the number of their staff providing services for that department's program. For example, a provider might have 10 staff members altogether, seven of whom provide services under DBDS programs and eight of whom provide services through the school system. Ideally, the departments will develop a consistent approach to collecting provider capacity data.

- **Provider location:** It will not always be possible to match provider name and location, given the variation in the way that information is collected. The principal challenge is matching location addresses due to the free-form nature of addresses (e.g. "Main Street" versus "Main St.").

**Assessment.** Not all departments or program have formal assessment tools. Schools, for example, have a formal assessment process governed by rule. However, there is no standardized assessment tool (and hence no electronic assessment data) that crosses all schools. For those departments that do have standard assessment tools, there is wide variation in what those tools measure. All of these tools were developed for their specific purposes and do not lend themselves to standardization. The process of determining capacity, preferences and needs will vary with program. For the purposes of an integrated database, each department is responsible for the integrity of its assessment data, if collected. It is not the object of this project to make recommendations on what should be collected for whom. The end-user, with the aid of the meta-data rules, will have to analyze the data that is there.

Although the integrated system will not have standardized data to compare across programs, it can standardize along certain domains or "buckets" for general categories of information. These buckets provide a mechanism for grouping like, but different, data across programs, to provide a more complete picture of the circumstances governing each person's care.

Below is a set of recommended domains, built around the efforts already made by DBDS. These domains may evolve as further work with individual departments is done.

- **Family Information** Family composition, history, interactions, living arrangements, etc.
- **Basic Need.** Heat, food, electricity, plumbing, water, clothing, telephone, housing transportation, etc.
Medical. Diagnosis, physical functioning, cognition, behavior, medications, oral, dental, vision, nutritional patterns etc.

Mental Health.

Social and Recreational. Formal and informal social, family, community supports, etc.

Legal. Guardianship status, custody, involvement in criminal justice system, police involvement, attorney, etc.

Safety. Risk of harm to self or others, from self or others.

Educational/Vocational. Special education services, job supports, etc.

Spiritual and Cultural. Factors that may impact service plan delivery, acceptance, treatment, etc., such as religion, nationality, cultural practices, etc.

Crisis. Setting for crisis placements, frequency of use, etc.

Trauma. Sexual, physical abuse or neglect, death or loss, etc.

Drug and Alcohol History

Plan. Like assessments, the elements of a service or treatment plan will vary with the program. However, unlike assessments, plans do share a common structure. The components of a plan consist of:

- Needs. The needs identified in the assessment process to be addressed by the plan;
- Goals. The goals identified in the plan.
- Action to be taken. For each goal, the strategies for attaining the goals to be taken.
- Activity. Progress notes for steps taken toward goals.

Information within a plan can be correlated and standardized along domains or “buckets” for general categories of information, identical to those used for the assessment domains. Using the same domains for plans and assessments will permit the correlation of assessed needs with the plan of activities or services. Again, these domains may evolve as more is learned about the plans used in other departments.

Services. Collecting service data across a diverse set of systems presents a diverse set of issues:

- Common code sets and claims forms. Through HIPAA, a revised claims form (HIPAA 837) will be replacing the HCFA 1500 and UB 92 claims forms for professionals and facilities. The new HIPAA form, and underlying code sets will be more comprehensive than its predecessors, going beyond clinical services to cover a wider spectrum of the types of services likely to be covered. To the degree that there is discretion, departments should use and require this common claims form and the common code sets when appropriate and relevant.

- Group versus individual providers. Information on the provider staff member providing the service, in addition to the organizational provider with whom the staff member is associated, should also be recorded. Differentiating between the provider organization and the actual direct care provider staff is often difficult.
Granularity: The usefulness of the service data will depend on the detail reported. When providers bundle multiple services into one bill, information on the individual services provided is lost; it is difficult to determine who provided an individual service, who received an individual service, when a service was provided, and what each individual service was.

Enrolled providers. Not all service providers will be registered providers recognized by Medicaid or through a department’s licensing function. (That is, not all providers will have an identifier. This might be because the service provider is a neighbor providing a one-time service or the provider is associated with a registered group but is not registered individually.)

Enrolled clients. Not all clients to whom services are provided are enrolled, or recognized as a legitimate client of the department. Some services might be provided to groups of consumers (e.g., a group prevention service) and no individual information is recorded. Or identifying information was not collected (e.g., in a crisis or emergency situation).

Burden on providers and political will. Submitting accurate and thorough claims information for individual service encounters increases provider workload, without necessarily increasing provider reimbursement, at least for those programs not reimbursing based on individual services. Enforcing complete service reporting will be a matter of each department’s willingness to address the concerns of providers.

Financial. Ideally the financial domain will capture budget information at the program level, allowing the State to analyze programmatic cost per person. This analysis will have to be a loose coupling of the cost with the number of people served; a rigorous accounting of program costs will not be possible. It is not the mission of this project or the integrated database to perform the role of a financial accounting system. This system will not perform complex cost allocation or an accounting of services provided to various funding sources.

Outcomes. Issues relating to outcome data will depend on the outcomes measured. Because waiting lists are of major importance for this project, the discussion below focuses on the issues related to waiting lists.

Types of waiting lists. There are at least two types of waiting lists:

Waiting to get into a program. Some services are provided through entitlement programs (e.g., Medicaid and special education) and the State cannot have waiting lists to get into these programs if the eligibility criteria are met. Waiting lists are possible, however, for non-entitlement programs. For example, people might need to wait for a slot to open up under a Medicaid waiver program if all slots are filled. People might also be on a waiting list for state-funded programs, which are often capped by available resources.
• Waiting to get a service. A person might be on a waiting list to receive a service, regardless of funding source. The primary reason for these waiting lists is the shortage of providers available.

Point in process. It is important for the State to know where in the process a person is waiting for a service. For the purposes of this exercise, there are three points in the service delivery process where a person may be on a waiting list:

• Request for service. A person may have requested a service but the need for that service has not been confirmed.

• Need determined, no provider identified. The need for a service may be identified, but there is no provider identified to provide the service.

• Provider identified, wait for service. A person is on a list to receive a service but must wait for an appointment. Presumably, a case manager can keep track of whether a person with an identified need has received those services. Or if a person does not have a case manager, the provider of services should also be able to keep track of who is waiting for those services.

Ideally, the departments would have access to data for all three types of waiting lists.

Point in time. Ideally, providers (or other waiting list keepers) will routinely update who is waiting for a service. For the purposes of this exercise, we recommend that the record reflect the date of the last update on waiting list status. We also recommend that an automatic reminder to update wait list status be built into the care management function or any other function that monitors waiting lists.

Using the Data

Who is the client? As discussed above, the integrated database will include information on any person for whom a record has been created. It will be up to the end user of the data to establish the rules for determining when or whether a person is a “client.” For example, sometimes person-level information is recorded when the only “service” provided is a referral to another resource. In some cases, information is recorded for family members closely involved in the treatment of the primary client. The end-user will need to develop rules for determining when a person has achieved status as a client.

In addition, services might be provided across a broad spectrum of time, depending on the individual. It will be up to the end-user to establish rules for determining whether a person is a client, based on when services are received, etc.

Communicating business rules. The integrated system will need to be able to communicate the “business rules” for collecting the data to the end-user. Since knowing how to use the data will
depend on why and how it was collected, the system should have some infrastructure (which will need to be kept up to date) for communicating these “business rules” to the end-user.

Privacy. Until otherwise justified, only non-identifying information will be available through the integrated data system. Even if all collection and use issues are resolved, the issue of privacy must still be addressed. The integrated system must encrypt identifying information, control and monitor access to information, and otherwise protect the data as required under the strictest of any departmental confidentiality and privacy policy.
Data Elements

This section identifies the data elements that would be needed to support the preliminary, interdepartmental “Olmstead measures” for services for people with disabilities. These data elements are a first pass, to be confirmed during the next stage of development. Because the proposed design for an integrated system is flexible and adaptable to changing needs, the “core data set” can evolve with a better understanding of the value and possible uses of integrated data (including beyond the “Olmstead measures”).

For the purposes of this exercise, we are not worrying about whether this data can be collected from existing resources. This “ideal” data set is proposed within the scope of this effort, to provide a roadmap for both exploring existing resources and for mapping out future data integration efforts. Deeper inquiry into existing data systems will help us whittle this “ideal” down to what is possible or doable, based on existing data.

<table>
<thead>
<tr>
<th>Domain/Element</th>
<th>Proposed Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>People</strong></td>
<td></td>
</tr>
<tr>
<td>Identifiers</td>
<td>Identifier assigned by payer (department, school, etc.) and provider</td>
</tr>
<tr>
<td>Name</td>
<td></td>
</tr>
<tr>
<td>Social Security Number</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Male, female</td>
</tr>
<tr>
<td>Date of Birth</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>U.S. Census categories</td>
</tr>
<tr>
<td>Hispanic Origin</td>
<td>U.S. Census categories</td>
</tr>
<tr>
<td>Primary language spoken</td>
<td>Language most frequently used, including sign</td>
</tr>
<tr>
<td>English proficiency</td>
<td>Ability to communicate in English</td>
</tr>
<tr>
<td>Address/Residence</td>
<td>Including zipcode</td>
</tr>
<tr>
<td>Living Arrangement</td>
<td>Usual living arrangement (<em>e.g.</em>, alone, with family members, with unrelated individuals</td>
</tr>
<tr>
<td>Residential Arrangement</td>
<td>Usual residential arrangement (<em>e.g.</em>, home, institutional setting, etc.)</td>
</tr>
<tr>
<td>DOMAIN/ELEMENT</td>
<td>Proposed Values</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Accommodation Needed</td>
<td><em>e.g.</em>, wheelchair accessible, interpreter</td>
</tr>
<tr>
<td>Guardianship/legal status</td>
<td>Whether under guardianship (public, private, other) or in the correctional system; whether voluntary or involuntary admission</td>
</tr>
<tr>
<td>Employment status</td>
<td>U.S. census categories</td>
</tr>
<tr>
<td>Annual earnings</td>
<td>Amount earned</td>
</tr>
<tr>
<td>Annual income</td>
<td>All income received</td>
</tr>
<tr>
<td>Years of schooling</td>
<td>Years of schooling at time of report</td>
</tr>
<tr>
<td>Marital status</td>
<td>U.S. Census categories</td>
</tr>
<tr>
<td>Custodial parent</td>
<td>Identifier, name, address</td>
</tr>
<tr>
<td>Program enrollee</td>
<td>Enrolled in department program</td>
</tr>
<tr>
<td>Program eligibility</td>
<td>Eligibility status for relevant programs, including Medicaid and state-funded programs</td>
</tr>
</tbody>
</table>

**Provider**

| Provider identifier: organization | Unique provider identification number                                      |
| Provider location                 | Provider address or addresses                                               |
| Type of facility                 | Type of facility code for provider type (range of facility and agency provider categories) |
| Licensed beds                    | For facility the number of licensed beds                                    |
| Medicare-certified beds          | For facility, the number of Medicare-certified beds                        |
| Staffed beds                     | For facility, the number of beds set up and staffed                        |

**Provider staff**

<p>| Provider identifier: staff/individual | Unique provider identification number for department, school, or organizational provider with whom affiliated or by whom employed |
| Affiliation/employer                | Unique provider identification number for department, school, or organizational provider with whom affiliated or by whom employed |</p>
<table>
<thead>
<tr>
<th>Domain/Element</th>
<th>Proposed Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provider credentials</td>
<td>Education level, type of degree, field of degree, year of receipt, etc.</td>
</tr>
<tr>
<td>Provider code</td>
<td>Code identifying type of provider</td>
</tr>
<tr>
<td>Provider specialties</td>
<td></td>
</tr>
<tr>
<td>License/certification</td>
<td>Provider license or certification number, status</td>
</tr>
<tr>
<td>Employment/affiliation status</td>
<td>Employed by or contracted to department, school, or organizational provider</td>
</tr>
<tr>
<td>Languages other than English</td>
<td>Identify languages spoken, other than English, including sign</td>
</tr>
<tr>
<td><strong>Assessment</strong></td>
<td></td>
</tr>
<tr>
<td>Assessment Type</td>
<td>E.g., MDS, LOCUS, CALOCUS, BERS, CAFAS, etc.</td>
</tr>
<tr>
<td>Assessment Description</td>
<td></td>
</tr>
<tr>
<td>Assessment Type ID</td>
<td>Unique identifier for assessment type</td>
</tr>
<tr>
<td>Assessment Type Version</td>
<td>Version of assessment tool</td>
</tr>
<tr>
<td>Assessment Date</td>
<td>Date assessment conducted</td>
</tr>
<tr>
<td>Location at time of assessment</td>
<td>Location of client at time of assessment (e.g., in hospital, at home)</td>
</tr>
<tr>
<td>Client assessed</td>
<td>Identifier for client for whom assessment conducted</td>
</tr>
<tr>
<td>Provider performing assessment</td>
<td>Identifier for provider performing assessment</td>
</tr>
<tr>
<td>Results</td>
<td>Substantive content of assessment, organized into domains; see DOMAINS discussion in text.</td>
</tr>
<tr>
<td><strong>Care Plan</strong></td>
<td></td>
</tr>
<tr>
<td>Plan type</td>
<td>E.g., Individualized Support Plan, Individualized Education Plan.</td>
</tr>
<tr>
<td>Plan description</td>
<td></td>
</tr>
<tr>
<td>Plan type ID</td>
<td>Unique identifier for plan type</td>
</tr>
<tr>
<td><strong>DOMAIN/ELEMENT</strong></td>
<td><strong>Proposed Values</strong></td>
</tr>
<tr>
<td>--------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Plan type version</td>
<td>Version of plan type</td>
</tr>
<tr>
<td>Plan date</td>
<td>Date plan developed</td>
</tr>
<tr>
<td>Client for whom plan developed</td>
<td>Identifier for client for whom plan developed</td>
</tr>
<tr>
<td>Provider developing plan</td>
<td></td>
</tr>
<tr>
<td>Plan development team</td>
<td>Participants in plan development process</td>
</tr>
<tr>
<td>Results</td>
<td>Substantive content of plan, organized into domains. See DOMAINS discussion in text.</td>
</tr>
<tr>
<td><strong>Services</strong></td>
<td></td>
</tr>
<tr>
<td>Claim identifier</td>
<td>Number assigned by processor to identify claim</td>
</tr>
<tr>
<td>Claim service period</td>
<td>Beginning and end dates for the service period covered</td>
</tr>
<tr>
<td>Procedure code and type of service</td>
<td>Code identifying the procedure, medications, product or service, or category of services</td>
</tr>
<tr>
<td>Procedure date</td>
<td>Date when the procedure was performed</td>
</tr>
<tr>
<td>Service type code</td>
<td>Code identifying the classification of services</td>
</tr>
<tr>
<td>Billing provider identifier and information</td>
<td>Unique identification number for billing provider, name and address</td>
</tr>
<tr>
<td>Service provider identifier and information</td>
<td>Unique identification number, name and address for provider of service</td>
</tr>
<tr>
<td>Place of service code</td>
<td>Code that identified where the service was performed</td>
</tr>
<tr>
<td>Units &amp; frequency</td>
<td>Units, dosage, frequency of services and procedures</td>
</tr>
<tr>
<td>Person receiving service</td>
<td>Identifier for person receiving services</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>An ICD-9-CM diagnosis code for diagnosis associated with service</td>
</tr>
<tr>
<td>Diagnosis date</td>
<td>Date the diagnosis was established or recorded</td>
</tr>
<tr>
<td>Domain/Element</td>
<td>Proposed Values</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Financial</strong></td>
<td></td>
</tr>
<tr>
<td>Service cost</td>
<td></td>
</tr>
<tr>
<td>Total charge</td>
<td>Total charges for encounter</td>
</tr>
<tr>
<td>Consumer co-pay</td>
<td>Amount of out-of-pocket copayment made or charged at time of visit, paid by consumer</td>
</tr>
<tr>
<td>Consumer payment toward deductible</td>
<td>Amount of copayment counted toward deductible.</td>
</tr>
<tr>
<td>Primary payer identifier and information</td>
<td>Identifying information for primary payer</td>
</tr>
<tr>
<td>Primary payer paid amount</td>
<td>Payment made by primary payer</td>
</tr>
<tr>
<td>Other payer identifier</td>
<td>Identifying information for other payer.</td>
</tr>
<tr>
<td>Other payer covered amount</td>
<td>Amount determined by other payer to be covered for the claim for coordination of benefits</td>
</tr>
<tr>
<td>Billing provider</td>
<td>Identifying information for provider receiving payment</td>
</tr>
<tr>
<td>Client identification</td>
<td>Identifying information for client for whom payment made</td>
</tr>
<tr>
<td>Date of payment</td>
<td>Dates on which claim or prospective payment is made</td>
</tr>
<tr>
<td><strong>Program/activity costs</strong></td>
<td></td>
</tr>
<tr>
<td>Original budget</td>
<td>Program budget</td>
</tr>
<tr>
<td>Amended budget</td>
<td>Amended program budget</td>
</tr>
<tr>
<td>YTD expenditures</td>
<td>Year to date program expenditures</td>
</tr>
<tr>
<td>YTD available</td>
<td>Year to date available program funds</td>
</tr>
<tr>
<td>YTD encumbered</td>
<td>Year to date encumbrances on program funds</td>
</tr>
<tr>
<td><strong>Human Resources</strong></td>
<td></td>
</tr>
<tr>
<td>Staff identifier</td>
<td>Unique provider identification number</td>
</tr>
<tr>
<td>Staff position or title</td>
<td>Provider address or addresses</td>
</tr>
</tbody>
</table>

*Edmund S. Muskie School of Public Service*
<table>
<thead>
<tr>
<th><strong>DOMAIN/ELEMENT</strong></th>
<th><strong>Proposed Values</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff credentials</td>
<td>Education level, type of degree, field of degree, year of receipt, etc.</td>
</tr>
<tr>
<td>Affiliation/employer</td>
<td>Unique provider identification number for department, school, or organizational provider with whom affiliated or by whom employed</td>
</tr>
<tr>
<td>Employment/affiliation status</td>
<td>Employed by or contracted to department, school, or organizational provider</td>
</tr>
<tr>
<td>Languages other than English</td>
<td>Identify languages spoken, other than English, including sign</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td></td>
</tr>
<tr>
<td>Waiting lists</td>
<td>All data collected by departments, providers on waiting lists</td>
</tr>
<tr>
<td>Unmet need</td>
<td>All department measures of unmet need for the people it serves.</td>
</tr>
<tr>
<td>Barriers to access</td>
<td>All department measures of barriers including distance to providers, access to transportation services, language barriers, etc.</td>
</tr>
<tr>
<td>Most integrated setting</td>
<td>All department measures of “most integrated setting.”</td>
</tr>
<tr>
<td>Other outcome measures to support identified measures</td>
<td>placeholder for other outcome measures that might be identified to be collected.</td>
</tr>
</tbody>
</table>

**Applying the Measures**

The next phase in developing an integrated system will be to develop the “detail design” for the “high-level” design described in this document. The detail design will paint the software screens that an end-user would like to see when querying integrated data. (With funding through CMS’ Real Choices initiative, the detail design will be developed over the course of the next year.) This section sketches what an end-user might want to see when measuring the State’s performance against the “Olmstead measures” identified, while recognizing the likely shortcomings of the underlying data.

*Who the State serves.* An end-user of the integrated system will want an unduplicated count of how many people with disabilities are served by the State. With identifying information, the integrated data set can produce a linked and unduplicated count (within some margin of error) of the people served by the State. An end-user will also want some basic information about the characteristics of the people served, including in which region of the State they reside, their
gender, their age, etc. The integrated system will produce basic demographic data from the multiple sources to provide this information.

**Strengths, preferences, needs.** The assessment data collected across departments will be grouped by domain (e.g., medical, mental health, housing), bringing together whatever information each department or program has collected on individual strengths, preferences, and needs. Each type of assessment will contribute different types of information, with some areas of overlap. An end-user will be able to query the integrated system to find out how many people have a housing need across all five departments, or the need for a particular service. Ideally, the end-user would also be able to identify the preferences of the people served (depending on whether this information is collected), including when a person prefers one residential setting rather than another, or one service rather than another.

**Services received.** The integrated data system would provide data on the services received across five departments. The end-user might query the system to compare the services consumed by persons with a particular diagnosis or some other set of characteristics. Data integrated across five departments will provide a more comprehensive picture of the services consumed and when there are gaps in services. Combined with demographics, needs, and other data, it will be easier to more reliably determine when a preventive measure or a better service combination might improve services or when a gap in services may be causing harm.

**Which departments are serving whom.** The integrated system will need to develop rules for determining when an individual is served by any one department. The following pieces of data, depending on the business rules of the individual departments and the purposes of the end-user, might be used to indicate which department is serving whom:

- A person record exists within a department.
- A department has identified a person as enrolled in a program.
- A service has been recorded by a department;
- The existence of a plan reported by a department (demonstrates that someone provided a service)

An end-user might want to get an unduplicated count of all people served by the Bureau of Child and Family Services, the Department of Behavioral and Developmental Services, the Department of Education and the Department of Corrections. An end-user authorized to use identifying information might use the integrated system to determine whether there is a need to coordinate case management for children served by these multiple departments.

5. **Which providers are serving whom.** Claims information, the assessment data, and the plan can all be used to identify the provider providing the services. Having provider data across departments serves multiple purposes. A consumer end-user accessing the data through the internet can use the integrated provider data to meet specific needs. Or an end-user can use the data to measure network adequacy, in meeting the Olmstead objectives.

6. **Which of the identified needs and preferences are unmet.** The ability of an end-user to measure unmet need will have to evolve over time. Unmet need can be measured by the service
or services for which the waiting list is kept (e.g., case management services or personal care assistance). It is expected that the accuracy of this data will be highly suspect, inconsistent and out-dated. As an alternative, but more complicated, the State might measure unmet need by matching an assessment of need with the service plan and then with the delivery of services. Or if a person does not have a plan, tracking the referral and the lack of services. This level of analysis requires rules and standardization. It is unlikely that this measure will be easily attained in the short term.

7. Waiting lists. The integrated data will not address underlying shortcomings in the data collected by individual departments. The waiting list data is not expected to provide accurate, up-to-date, information or identify the people (as opposed to the number of people) on waiting lists. Because the fair administration of waiting lists is critical to the State’s ability to defend itself against an *Olmstead*-type lawsuit, it is in the State’s interest to develop standard business rules for collecting and the type of data collected for waiting lists. To provide an end-user with an accurate picture of waiting list status, the state programs should collect:

- **Type of wait.** Whether for entry into a program or for a particular services.
- **Where in wait.** Whether need has been assessed, and a provider identified
- **Type of service.** What type of services are being waited for?
- **Provider.** Which provider is being waited for?
- **Start date on wait list.** What is the person’s start date on the waiting list?
- **Date of last update.** What is the date of the last update for the waiting list?

Until accurate and current waiting list data is kept across departments, an inter-departmental assessment of waiting list data is not possible through an integrated system.

8. Setting. An end-user should be able to use the integrated data to identify the residential setting of people served by the State. Given the wide variation in types of settings within licensed categories, work will need to be done in reconciling each department’s definition of “setting” to make this determination as well as in developing a standardized approach to collecting residential data. Data acquired through the integrated system can be tailored to the needs of end-users to make its data meaningful for individual programs.

9. *Most integrated setting.* Under the *Olmstead* decision, a state must provide services to people with disabilities in the most integrated setting appropriate to their needs. In *Olmstead,* the court said that the state’s designated health professional could determine whether a setting is the most integrated setting appropriate to a particular individual. For the purposes of this data set, if that determination is made it will be collected. In addition, information on the setting in which a person resides will be collected. Matched with the data collected on an individual’s needs, the State will have some means of evaluating the appropriateness of a setting. Overtime, with either better efforts at collecting this information or the development of rules for deriving it, an end-user may be able to measure the State’s compliance with this *Olmstead* requirement.

10. Resources. The combination of provider data and human resources data will provide the State with the means to measure the State’s capacity to meet the needs of people with disabilities and identify opportunities for improvement.

---

Institute for Health Policy
11. **Cost.** The state budget information matched with the people served should provide a ballpark estimate of programmatic costs per person. Using other elements, like age categories, region, level of need, etc., an end-user could have a much more comprehensive view of the costs associated with the services provided and for whom.

12. **Barriers to access.** The assessment and plan data will identify elements that can be used to develop indicators measuring barriers to access. For example, the assessment data might be used to determine:

- whether the lack of transportation services is a barrier to access;
- travel distance between a client and a client’s provider; or
- language barriers (spoken and sign).

Querying tools would need to be developed to allow the end-user to quickly derive this information across programs.
Next Steps for Data System Integration

An integrated system needs to be developed in overlapping stages. Overarching any data integration initiative must be an inter-departmental effort at educating and building buy-in for integrated data by those most in control of the quality and access to data. Information system staff, program managers, quality management staff, and front line workers all need to have a shared understanding of the State's commitment to collecting and integrating data, and using integrated data to manage and improve the effectiveness of services. To the degree that state leadership recognizes the value of integrated data, it needs to create the "cultural shift" critical to making it happen.

Reinforced by the commitment of the state leadership, the following immediate next steps are recommended:

**Detail design.** With the guidance of some key stakeholders, the State needs to design solutions to a number of the core challenges identified in this document (e.g., linking and duplication, security, query capabilities). The detail design will identify the requirements for each function of the integrated system. For example, for search functions, the search capabilities will be identified (e.g., search provider by type, region, etc., with specified logic and security restrictions) as well as the process flows for that function (e.g., the series of screens associated with that function). The input of key stakeholders will assist in identifying the desired functions and requirements. Key stakeholders should include likely end-users, including both state staff and consumers likely to access the integrated data over the Internet. In addition, it will be important to have the input of those with concerns about the security and access, so that their concerns can be addressed in the design phase.

The product of the design phase will be a detailed description of the requirements of the integrated system - what the system must be able to do - providing sufficient granularity for the implementation phase to begin. The detailed descriptions will include a mock-up or prototype of the integrated system. This prototype will consist of screen layouts and screen navigation. Screens developed will include the meta-data screens used to link to the data sources, for data queries, business rules, policies, help facilities, data transformation rules, crosswalks, etc.

Because the design detail will depict the end-use screens that will ultimately be created, the design detail can convey the potential uses of integrated data in a very concrete way.

The detail design phase will be funded under CMS' Real Choices initiative.

**Inter-Departmental Agreement.** The State also needs to formalize inter-departmental commitment and investment in launching an integrated data initiative. Here, on a commissioner-level, and with the Governor's endorsement, the departments should:

- agree on the goals of an integrated data initiative;
- agree on a strategy for data integration;
- adopt some guiding principles to govern departmental ownership and control over access to data (with the understanding that a fully developed inter-departmental
strategy for confidentiality and privacy will be developed before full implementation of the integrated system);

- develop a strategy for financing the development of the integrated system;
- agree on a strategy for ongoing maintenance of an integrated system.

Formalizing inter-departmental agreement in a memorandum of understanding will be helpful in making sure all parties have a mutual understanding of the agreement reached. Inter-departmental agreement on this initiative will further not only inter-departmental interests triggered by Olmstead, but also inter-departmental efforts at coordination prompted by the Children’s Cabinet and the children’s legislative oversight committee.

*Develop Implementation Workplan and Activities.* A detailed work plan will also need to be developed. This work plan will delineate success criteria, risk assessments, cost estimations, funding proposals, roles and responsibilities, definitions, project and steering management structure, and project timelines.
EDMUND S. MUSKIE SCHOOL OF PUBLIC SERVICE educates leaders, informs public policy, and broadens civic participation. The School links scholarship with practice to improve the lives of people of all ages, in every county of Maine, and in every state in the nation.

EDMUND S. MUSKIE SCHOOL OF PUBLIC SERVICE
96 Falmouth Street
PO Box 9300
Portland, ME 04104-9300

TELEPHONE (207) 780-4430
TTY (207) 780-5646
FAX (207) 780-4417
www.muskie.usm.maine.edu

A member of the University of Maine system