

MARYLAND LONGITUDINAL DATA SYSTEM (MLDS)
550 West Baltimore Street
Baltimore, MD 21201

GOVERNING BOARD MEETING

December 16, 2014

MINUTES

The meeting of the Maryland Longitudinal Data System Governing Board was held on December 16, 2014 in the University of Maryland, Francis King Carey School of Law, Krongard Board Room. Vice Chair Catherine Shultz called the meeting to order at 10:05 a.m. and noted that a quorum was present. Ms. Shultz asked if there were any additions to the agenda and noted that the discussion initiated by Dr. Wilson from the last meeting regarding achievement gap would be continued under old business.

The following Governing Board members were in attendance:

Ms. Catherine Shultz, Acting Secretary of Higher Education
Mr. Brian Roberts, Change Management Specialist, Montgomery County Government
Mr. Leonard Howie, Secretary of Labor, Licensing, and Regulation
Ms. Tina Bjarekull, Maryland Independent Colleges and University Association
Mr. Steven D. Rizzi, Vice President, PAR Government
Ms. Jennifer Mullinix, Math Teacher, Wilde Lake Middle School, Columbia, Maryland
Dr. Ben Passmore, University System of Maryland (Designee for Dr. Kirwan)
Mr. John White, Maryland State Department of Education, (Designee for Dr. Lowery)
Mr. Brad Phillips, Maryland Association of Community Colleges (Designee for Dr. Sadusky)
Mr. Jason Perkins-Cohen, Jobs Opportunities Task Force

The following staff members were in attendance:

Mr. Ross Goldstein, Executive Director, MLDS Center
Dr. Jon Enriquez, Director of Reporting Services, MLDS Center
Ms. Laia Tiderman, Director of Data Management, MLDS Center
Dr. Laura Stapleton, Associate Director of the Research Services Branch, MLDS Center
Ms. Dawn O’Croinin, Assistant Attorney General for the Governing Board and MLDS Center
Ms. Jamese Dixon-Bobbitt, Executive Associate, MLDS Center
Ms. Tejal Cherry, Director of System Management Branch, MLDS Center
Mr. Chuck Shelton, Senior System Architect, MLDS Center
Mr. Michael Chen, ETL Developer, MLDS Center

MLDS Center Report

Staffing

Mr. Goldstein began by informing the Board of the Center’s recent success in hiring two new employees: Michael Chen, who was hired to serve as an ETL Developer and will be responsible for loading data and working through any issues involving data quality or matching; and Bob Murphy, who was hired to serve as an OBIEE Analyst and will be developing applications that are necessary to support operations including dashboard development and data management.

The Center is also close to hiring a WebCenter Developer. The position is not filled yet, but there is currently an individual working via a contract. That individual also applied for and has been offered the

State position. The official hiring is pending various bureaucratic challenges - including H1-B Visa application process and a reclassification with the Department of Budget and Management.

Mr. Goldstein noted that with the recent hires, there are only 6 remaining vacant MLDS Center positions, which include:

- a. The three positions shared with MSDE, MHEC, and DLLR.
 - o The MSDE shared position was held by Chandra Haislet, who left for a new position within MSDE. Laia Tideman, who has been with MSDE for eight years and has an in-depth background with education data, has been selected to serve as Ms. Haislet's replacement. Ms. Tideman's appointment is pending final approval by the State Board of Education (meeting 12/16).
 - o The MHEC shared position was held by Jon Enriquez, who left for a new position within MHEC. There was no-one at MHEC who could take Dr. Enriquez's place. MHEC is actively recruiting to fill the position. In the meantime Dr. Enriquez continues to provide support to the MLDS Center.
 - o Finally the DLLR shared position has never been filled. DLLR is in the process of recruiting to fill the position.
- b. The fourth vacant position is being filled by Chuck Shelton who is the senior system architect. Mr. Shelton successfully built a similar system in Washington State. The Center plans to keep him in the contractual position for the foreseeable future.
- c. The fifth vacant position is for a Database Administrator. The Center is actively recruiting for this position. Interviews have been conducted, but so far the right candidate has not been identified. Currently the Center is utilizing the services of a MSDE contractor.
- d. Finally, the sixth vacant position is for an IT Technology Support Specialist. This position is being held until the Center determines the exact needs to be addressed.

Mr. Goldstein next turned to the organizational changes that have been implemented. Primarily the changes affect supervision. Mr. Goldstein stated that he did not feel that the MSDE and MHEC shared employees made effective supervisors since they are only engaged in MLDS Center work part-time and because they, as non-IT professionals would be required to supervise all IT staff.

To address this concern Tejal Cherry has assumed the duties of Systems Management Director. The MSDE shared position will serve as the Data Management director and the MHEC position will continue to serve as the Reporting Services director. The MHEC and MSDE positions continue to play an important role, just not a supervisory role.

Mr. Goldstein addressed questions regarding the efficacy of having shared positions by stating that the shared positions are worthwhile and serve an important purpose as a liaison and coordinator between agencies. Mr. Howie asked if the shared positions were statutorily required. Mr. Goldstein responded that it was not a statutory requirement, but it was a requirement under the inter-agency agreements. The shared employees are MLDS employees and are fully funded by the MLDS, but fifty percent of their duties are for the partner agency. Ms. Shultz and Dr. Passmore noted their support for continuing the shared staffing arrangement.

Finally, a Research Coordinator has been hired. Dr. Angela Henneberger has been appointed to a Research Associate Professor position with the School of Social Work. The position will be funded by the MLDS Center. Dr. Henneberger is currently finishing a two-year postdoc training program in educational research at the Pennsylvania State University. Dr. Henneberger will join the team in January.

Legislative Audit

Mr. Goldstein informed the Board that the Office of Legislative Audits is currently auditing MSDE and has decided to include MLDS Center in that audit. There are two components of a Legislative Audit: fiscal and information technology. The fiscal audit makes sense since the MLDS finances are managed by MSDE. The IT audit is arguably too soon and would be more productive once the Center has had time to become fully operational. Nonetheless, the auditors are proceeding and it could prove to be a useful review.

Legislative Briefing

Mr. Goldstein informed the Board that the Senate Education, Health and Environmental Affairs Committee has asked him to provide a 10-15 minute presentation on January 21st about the Center and its progress to date. The presentation is part of an overall hearing on higher education topics, which will include presentations by USM, Morgan State, St. Mary's College, MACC, MICUA, and a discussion on textbooks.

System Management and Development

Ms. Cherry began by providing the Board with a system development timeline and highlighted the accomplishments and next steps in the following areas:

- Security - Staff has created a data breach plan as required under the *Data Security and Safeguarding Plan*. The data center installation at MSDE is complete and the servers have been configured using VM Ware. Staff has also completed background checks and the necessary security training requirements. Work is ongoing in establishing user accounts and user groups. The user accounts will ensure access is limited to only the portions of the system necessary authorized for that user.
- Development Environment - The IT staff has been working closely with the research team and has installed software and analytic tools for use by the research team. In addition, the website has been redesigned consistent with State standards and naming convention (maryland.gov). WebCenter and the business intelligence tools have been integrated, which will allow for dynamic dashboards. Upcoming tasks include installation of various Oracle products on the new data center servers and continued building, testing and publishing of data dashboards.
- Test and Production - Staff is building test and production environments, installing software, EMC storage, VLANs, and building firewalls, with the goal of having the MSDE data center fully operational by late January. Staff is also exploring offsite backup options.
- Data Load - While data has been loaded, the process stopped while the Center was working to find an ETL Developer. The loading of wage data is behind schedule. However, Ms. Cherry notes with the addition of Michael Chen, the data load process should pick-up. Staff is also working to build and test an automated ETL process.
- Standard Operating Procedures - Procedures will be developed starting in March of 2015. The procedures include change control, log aggregation, audit reporting, disaster recovery, and vulnerability testing.

Ms. Cherry also spoke about some of the procurement issues that have impacted system development. Specifically, network cards had to be procured in order to install the servers at MSDE. The procurement took months longer than expected which slowed the data center installation progress. Those cards have been received and installed.

In response to a question from Mr. Rizzi, Ms. Cherry clarified that the data center installation at MSDE only represents a change of location - not a change in system design. New servers and equipment needed to be installed to host the full system. Currently the system is running on the legacy servers at DPSCS, but once the data center at MSDE is complete, the system will be moved over to that location.

Dr. Stapleton clarified that the Research Team has not had the access to conduct research using the system. To date, the engagement with the system has been for planning and learning purposes.

Data Matching Overview

Ms. Cherry provided a brief overview of the data matching process. Providers send data to MLDS Center. Center staff determines whether the data is usable. If it is, they remove duplicate records, standardize the data, and match the data with any existing records in the system. If no record exists, a new record with a unique party identifier is created. In response to a question from Mr. Rizzi, Mr. Shelton clarified that data is determined to be unusable if it is the incorrect file or does not contain the correct information necessary for identity resolution (i.e. a file that does not contain Social Security Numbers would be rejected). Mr. Shelton also explained that the deterministic data matching requires an exact match between identifiers.

Data Matching Challenges

Mr. Shelton explained that the challenge for data matching is that the different sectors are not using the same identifiers. Early Childhood uses SSN (Social Security Number) as the identification number. MSDE uses the SASID (State Assigned Student ID); MHEC uses SSN (but has also begun collecting SASID); and workforce uses SSN. This creates a challenge when trying to match an individual throughout time and across sectors. There are strategies for improving the matching capabilities, but they will take time.

Mr. Shelton provided information on identity matching in the Master Data Management (MDM) system.

1. There are 393,921 organizations identified in the MDM database.
2. There are 2,785,528 people identified in the MDM database.
3. There are 2,788,703 Social Security Numbers in the MDM database. Of these SSN's:
 - a. 2,766,903 SSN's are associated with 1 person.
 - b. 21,710 SSN's are shared by 2 people.
 - c. 90 SSN's are shared by 3 or more people.
4. There are approximately 340,404 people with invalid social security numbers. Of these:
 - a. 322,217 people have an SSN of 111-11-1111.
 - b. 18,125 people have a value of "N/A" for the SSN.
 - c. 49 people have an SSN of 555-55-5555.
 - d. 8 people have an SSN of 999-99-9999.
 - e. 5 people have an SSN of 123-45-6789.

Mr. Rizzi, noted that approximately 10% of the population do not have a valid social security number. This is a high number of inaccurate SSNs and a discussion ensued regarding the impact of these inaccuracies, its cause, and steps that can be taken to resolve it.

Due to the importance of the Center's ability to match data, Mr. Rizzi asked to make these matching statistics a standing agenda item to continually update the Board at future meetings.

Mr. Shelton next turned to the merged data in the Operational Data Store, which currently contains 1,039,381 people (including their transactional data). This includes 293,108 Maryland high school graduates for the years 2009-2013 and 1.6 million postsecondary fall enrollment records covering the same time period. Based on initial, draft data analysis, there are 75,566 of the 293,108 Maryland high school graduates enrolled in Maryland postsecondary schools during this time period. Dr. Passmore noted that this enrollment figure is well below known findings on Maryland high school students going into a Maryland higher education institution.

Dashboard

In response to Dr. Passmore, Dr. Enriquez displayed the data dashboard that staff has been working on for the website. The dashboard shows the percentage of Maryland high school graduates who enroll in Maryland postsecondary education. The dashboard indicates that only one-third of high school graduates enrolled in Maryland postsecondary education. This is a low percentage due to the matching process used by the Center and the fact that certain key identifiers were absent. For example, one county had virtually no students enrolled in Maryland postsecondary education because the data from that county did not include SSN, resulting in the Center not being able to match students from that county to the postsecondary records. There is a need to improve data quality and that is an effort that the Center will require the assistance of the partner agencies. There are also steps the Center can explore that will improve matching rates - such as exploring the use of a probabilistic matching as opposed to deterministic matching and using SSN validation or including National Student Clearinghouse data. Mr. Shelton noted that in Washington State, drivers license data is used to improve data quality and matching.

The challenge for the Center right now is determining whether, given the limitations of the data set, the data can still provide meaningful information and how best to convey that information (and the limitations) to the public and policy makers.

In response to a question by Mr. Rizzi, Ms. Tiderman noted that there are values that can be trusted and can be used as a check on Center output. For example, with the current subject of high school graduation rates and postsecondary enrollment, both MHEC and MSDE have data that can be used to assess the accuracy of the MLDS output. There will also be instances when there will not be expected results to verify against. In those cases the Center will rely on the research team to provide an assessment of the accuracy and reliability of the Center's findings. Mr. Rizzi noted that the Governing Board needs to know the accuracy and data quality so the Board can provide good governance.

Data Inventory

Ms. Tiderman presented the data inventory to the Board for approval. A copy of the data inventory was provided to all of the members. Mr. Goldstein noted that unlike in past meetings, this data inventory included all of the elements in the system, not just new elements. New elements to be added are listed as "proposed." Ms. Tiderman also explained that the inventory is broken out by source. The status of the data and whether it is loaded and active in the system is also indicated.

- Workforce (DLLR) - The new proposed data elements are adult education, including GED and NEDP data;
- PK-12 (MSDE) - The new proposed data elements are postsecondary student information from the National Clearinghouse (Ms. Tiderman explained that these data elements are listed under MSDE since MSDE has a contract with the National Clearinghouse and provides the data to MLDS);
- Higher Education (MHEC) - There are no new proposed elements at this time; and
- External Sources - The proposed data elements include IPEDS (Integrated Postsecondary Education Data System) and Census data.

In response to a question from Mr. Rizzi, Ms. Tiderman explained that "EL Staff" was early learning staff. This data is necessary to answer questions required under the federal SLDS grant. Similar teacher data will be available from K-12. The status of higher education teacher data has not been determined.

Ms. Shultz asked for a motion to approve the data collection schedule. Mr. Rizzi moved approval, which was seconded by Dr. Passmore. The motion was unanimously approved.

Data Collection Schedule

Ms. Tiderman presented the data collection schedule to the Board for approval. A copy of the data collection schedule was provided to all of the members. Ms. Tiderman explained the data collection

process, which includes a collection window, data transfer deadline, cross agency reconciliation window, and a required sign off from a representative of the agency. The schedule being presented includes dates for the 2014-2015 school year. The schedule is organized around the agency calendar (MSDE and MHEC - academic calendar). In response to a question by Ms. Shultz, Ms. Tideman stated that the agencies have reviewed the data collection schedule.

Ms. Shultz asked for a motion to approve the data collection schedule. Mr. Roberts moved approval, which was seconded by Mr. Perkins-Cohen. The motion was unanimously approved.

Regulations

Mr. Goldstein first noted that all of the items on the agenda for Governing Board approval, including the regulations, have been presented to either the Data Governance Advisory Board or the Research and Policy Governing Board prior to the meeting to get feedback and input from various partners.

Mr. Goldstein presented two sets of regulations to the Governing Board for approval. One set of regulations are new proposed regulations and the other set are regulations for final adoption.

The proposed regulations are COMAR 14.36.05 and .06. Chapter 05 establishes procedures for Data Collection that generally codify the current practices of the Center, including:

1. All data must be in the Data Inventory approved by Board before it can be collected or used;
2. Data collection must be conducted pursuant to the schedule approved by Board;
3. Data must be transmitted as specified by MLDS Center; and
4. Clarifying the source from which the Center collects data, as follows:
 - a. While State law authorizes the Center to collect data directly from LEAs and institutions of higher education, the regulations specify that if an agency is collecting data the MLDS Center must get the data from that agency;
 - b. In the event that a private secondary school chooses to provide data to the Center, the Center may designate MSDE to collect that data;
 - c. The regulations authorize the Center to designate MHEC as its agent to collect data on its behalf from for-profit and private nonprofit institutions of higher education (this is necessary since MHEC's authority to independently collect that data is not clearly established under State law); and
 - d. The collection of third party data is authorized provided the data meets the definition of student and workforce data, is determined to be accurate, and is approved by the Board.

The proposed regulations also provide guidelines for appointing authorized staff to the Center. There was a concern at the last meeting that the process was too open ended and might create an expectation that appointments were readily available to the public. Chapter 06 clarifies that authorized staff includes State employees and individuals authorized by the executive director pursuant to the following limitations. First, the appointment must be necessary to carry out the mission of the Center. Second, the number of appointments must be restricted for the purpose of maintaining control over access to the system. Pursuant to those limitations, the executive director may designate researchers, information technology experts and State agency employees. The regulations also reiterate the required compliance with security requirements and criminal history background investigations.

Ms. Shultz asked for a motion to approve the proposed regulations COMAR 14.36.05 and .06. Mr. Howie moved approval, which was seconded by Mr. Perkins-Cohen. The motion was unanimously approved.

Next Mr. Goldstein presented COMAR 14.36.01 - .04 for final adoption and noted that the regulations had been published in the Maryland Register and no comments or questions were received. The

regulations establish procedures for compliance with the Maryland Public Information Act (Chapter 01), compliance with Open Meetings Act (Chapter 02), addressing requests to correct public records created and maintained by the Center (Chapter 03), and managing requests for longitudinal data requests (Chapter 04).

Ms. Shultz asked for a motion to approve the proposed regulations, COMAR 14.36.01 and .04, with any necessary nonsubstantive changes related to citation changes. Mr. Rizzi moved approval, which was seconded by Ms. Bjarekull. The motion was unanimously approved.

Annual Report to the Governor and General Assembly

Mr. Goldstein presented a draft of the Annual Report to the Board. Mr. Goldstein noted that State law (Ed. Art. §24-705, Annotated Code of Maryland) lists five reporting requirements and provided a brief overview of how each of those requirements are being addressed in the report.

(1) An update on the implementation of the MLDS.

This section discusses both the implementation of the system and the Center. The system update includes a statement that the development of the system is complete and an acknowledgement that, due to a variety of factors, the system development took longer than expected and as a result the amount of research and web content is less than what was planned. The system update also provides a detailed technical explanation of the system architecture, the decision to change data hosting locations, and a listing of the dashboards to be completed in the next three months. The Center implementation update includes an overview of the advisory boards, the development of the data collection schedule, a summary of the proposed and final regulations, staffing overview and challenges, a listing of the completed interagency agreements, a discussion of the implementation of the Data Security and Safeguarding Plan, and the research series conducted during the past year.

(2) A listing of studies conducted.

This section provides an overview of the work done by the research team and a discussion on the changes made to the Research Agenda over the past year.

(3) A list of data determined to no longer be necessary.

There is currently no data determined not to be necessary. However, the report does note that certain data items have been flagged for further review and possible removal during the upcoming year. In addition, the report discusses the disposition of the P20W system and whether the data in that system will be purged or will continue to be maintained in a separate database or incorporated into the MLDS.

(4) A list of data in the system.

This requirement is addressed by providing the complete Data Inventory as an attachment to the report.

(5) Any recommendations of the Governing Board to the General Assembly and Governor.

This section of the report is not complete. Mr. Goldstein stated that neither he nor staff had any suggested recommendations at this time and was looking to the Board for input.

The Board had a lengthy discussion regarding the inclusion of a recommendation in the Annual Report for the establishment of procedures and standards for determining when data is sufficiently accurate and complete to allow the Center to publicly report that data in a dashboard or other report.

During this discussion Dr. Passmore stated that there was a need for the Center to develop a step-by-step plan for improving data quality. The data quality issues need to be communicated to the agencies so that they can take steps to address the data quality issues. Dr. Passmore also noted that there are other reports and analyses that the Center can use as a benchmark to determine the accuracy of its information.

Mr. Howie noted that data always has limitations. There will always be caveats and if the Center holds out for perfection then there will be no MLDS. Mr. Goldstein agreed by noting that there are several built in limitations: for example K-12 data does not include private schools and Unemployment Insurance data does not include federal employees. The goal, therefore, is to define what is acceptably complete and develop standards for communicating those limitations.

Ultimately there was a consensus to include a statement in the recommendation section of the Annual Report addressing data quality and reporting standards. Accordingly, Ms. Shultz directed staff to develop and circulate within the next two days, a statement for Board review and approval. Mr. Perkins-Cohen recommended removing the statement that the list of proposed new dashboards would be completed within three months.

Ms. Shultz asked for a motion to approve the Annual Report with the change noted by Mr. Perkins-Cohen and pending the approval of the recommendation section to be drafted and circulated by staff. A motion was made by Ms. Bjarekull and seconded by Dr. Passmore. The motion was unanimously approved.

Note: On December 17, 2014, staff emailed to the Board the following recommendation for inclusion in the Annual Report. The Governing Board reviewed and emailed their approval on December 18, 2014.

The accuracy of information reported by the Maryland Longitudinal Data System is of the utmost importance to the Governing Board and the Center. Accuracy is affected by the quality and completeness of the data received, the ability of the Center to match that data across sectors (early childhood, PK-12, higher education, and workforce), and the manner in which the information is presented to the public.

The Governing Board therefore recommends that the Center develop, for Board review and approval, a set of standards and protocols for assessing the accuracy of information reported to the public. The standards should include:

- 1. An assessment of whether the data relied upon for a report is sufficiently complete to support the information reported;*
- 2. An assessment of whether the information presented can be reconciled against other sources;*
- 3. Criteria for determining whether information based on incomplete data is appropriate to be reported; and*
- 4. Methods for informing the public regarding the information published by the Center.*

The Governing Board recommends that these standards and protocols be established prior to the release of public information from the MLDS.

Dual Enrollment Report

Jon Enriquez presented the *Dual Enrollment Report*, which is another report that the Center is statutorily required to submit to the Governor and General Assembly. The report finds that, despite the fact that the College and Career Readiness and College Completion Act of 2013 (CCRCCA) instituted a new funding method (effective July 1, 2013) for students in public high schools pursuing dual enrollment, there is no significant change in the number of students dually enrolled in Fall of 2013 as compared to Fall of 2012.

Dr. Enriquez did note that there was a slight increase in the number of credits taken. Some of the other findings included the fact that 97% of dual enrollees go to Community Colleges; females make up 60% of the dual enrollees, and whites are statistically overrepresented in the dual enrollment population.

Dr. Enriquez also discussed the fact that the data for the report still relies on data provided by MHEC. However, both MHEC and MSDE are in the process of improving data collections that will benefit this report. MSDE will begin collecting additional course information as well as information on course payment. MHEC is also developing improvements on course information.

In response to a question about online course data, Dr. Stapleton responded that the online course data was a separate project by the Research Team in fulfillment of one of the SLDS Grant requirements. The research team produced a white paper that discusses the need to include online course data in the MLDS data collection. Ms. Shultz requested a copy of the report.

Assistant Attorney General's Report

Ms. O'Croinin informed the Board that all of the administrative and data sharing MOU's between the MLDS Center and MSDE, DLLR & MHEC are complete. The MOU between the Center and the University has been in place since February 2014. The MOU regarding the relocation of the Center's data center from DPSCS to MSDE should be completed today pending the completion of technical specifications.

At the last meeting there was a question about member liability and the applicability of the Maryland Tort Claims Act. Ms. O'Croinin advised that the authorizing statute for the MLDS does not specify any protections from liability for members of the Governing Board for actions falling within the scope of the duties performed on behalf of the Board. However, the Maryland Tort Claims Act, provides those protections. Members of the MLDS Governing Board are defined as "State personnel" within the statute and state personnel have the immunity from liability described under §5-522(b) of the Courts and Judicial Proceedings Article. That section provides, "State personnel, as defined in §12-101 of the State Government Article, are immune from suit in courts of the State and from liability in tort for a tortious act or omission that is within the scope of the public duties of the State personnel and is made without malice or gross negligence, and for which the State or its units have waived immunity under Title 12, Subtitle 1 of the State Government Article, even if the damages exceed the limits of that waiver."

Old Business Item

Ms. Shultz tabled the pending old business item regarding the achievement gap and the role of the MLDS in making policy recommendations. Ms. Shultz noted that this topic should be held until a subsequent meeting when Dr. Wilson, who began the discussion, is present. Mr. Rizzi stated that he wanted to have a more in depth discussion about the Research Agenda, which would likely be a part of the overall discussion of the policy role of the MLDS.

New Business

No new business was discussed.

Adjournment

Ms. Shultz asked for a motion to adjourn the meeting. A motion was made by Mr. Rizzi and seconded by Mr. Pscherer. The meeting adjourned at 12:15 p.m.

Respectfully submitted,
Ross Goldstein
Executive Director

Approved: January 22, 2015