

The Data Maturity Index (DMI)

Produced by Powered by Publics' Data Integration Cluster (C13)

March 2023

**Suggested Citation: Powered by Publics. The Data Maturity Index. March 2023. Association of Public and Land-grant Universities (APLU). <https://www.aplu.org/news-and-media/blog/powered-by-publics-learning-memo-the-data-integration-cluster>.*

The DMI or DMM is an assessment of the institutional data capacity by users of the data to support a learning community to improve the use data at institutions. Data maturity is the ability of an institution to improve the organizational, procedural, and technical levels of its data ecosystem to leverage institutional data to produce insights that inform strategic initiatives that meet institutional goals and support student success. Users report their perceptions of data capacity.

The guiding questions to assess institutional data capacity:

- Does the institution have a data governance structure?
- Does the institution have a data-informed culture?
- How are data collected and managed?
- How is data quality maintained?
- How are the data distributed?
- How do we ensure that the reports and analysis add value?
- Are units using the data in their strategic or operational discussions?
- Are the data leading to insights and change?
- Is the overall effectiveness of our data strategy and interventions evaluated?

Definitions

Many institutions have a Chief Data Officer who leads the effort to improve data use and data literacy. Titles may include: Chief Data Officer, AVP or Director of Institutional Research, Planning, Institutional Effectiveness, Business Intelligence, or Data Analytics. The position takes primary responsibility for the management and dissemination of data, including data quality and the user experience. The Chief Data Officer may head the Data Governance Committee.

Data governance is the structure, policies, and practices at an institution that support an evidence-based culture, where data are treated like a vital resource to support priorities, strategies, and operations. It is a collaboration across the campus including the executive leadership, student advising, academic leaders, faculty, staff, IR, IT, legal, and budget resources.

Phase 1: To be completed by data consumers from various offices across the university.

What is your role at your institution? _____

What is your institution? _____

Instructions: Select the level that describes your perception of your institution’s data maturity for each row.

Data Maturity Index						
Indicator description	Question	Level 1	Level 2	Level 3	Level 4	Level 5
1. Data Governance – Systemic and intentional practices and policies to support the effective use of data						
Identify the key stakeholders who will help to create a new culture.	Who are the key stakeholders that will transform the culture?	The institution has not identified key stakeholders for data culture.	The institution has an informal set of stakeholders who may occasionally interact with one another, but none are part of the leadership.	The institution has a set of stakeholders that does not include executive leadership and little engagement comes from areas of the institutional community.	The institution has a set of stakeholders that includes executive leadership with limited engagement.	The institution has a committee of key stakeholders to include executive leadership team, institutional research, information technology, academic affairs, student services, admissions, registrar, facilities management, enrollment management, public affairs, and marketing. They meet regularly and are engaged.
Transparent policies and procedures on how to use	Are policies and practices around the use of data	There are no standard policies or procedures on	There are some practices that are expected, but nothing formal.	There is an informal group that meets to discuss standard	There is a formal group, but no standard policies or practices are	There is a formal group and standard policies and practices

the data are communicated to the users.	communicated to the campus community?	the use of data.		acceptable practices, but it is not official.	communicated to users.	are communicated to users.
Data Governance Committee	Is there a data governance committee that meets regularly to guide the institutional use of data?	The institution has no data governance committee. Governance is adhoc or decentralized.	The institution has an informal group that meets to discuss governance issues.	The institution is forming a data governance committee; beginning to review institutional policies and examine the use of data at the institution.	Data governance structures are implemented and fully operational with a diverse group of officers from IT, IR, Academic Affairs, Legal, Student Services, etc. and an identified Chief Data Officer.	The institution has a data governance committee that include officers from IT, IR, Academic Affairs, Legal, Student Services, etc. and an identified Chief Data Officer. The DGC meets regularly, develops policy, and makes recommendations for improving data practices.
2. Data Collection – Comprehensive processes that integrate and transform data from disparate sources						
Include all data sources that relate to student engagement and success	Are all relevant data sources available and accessible?	Data are not used for planning purposes. Limited data sources may available to limited users and	Limited data sources relevant to student success are available to users but not into a central data repository	Some data sources relevant to student success are incorporated into a central data repository or some data are available, but not in a central repository	All data sources relevant to student success are incorporated into a central data repository and available to a limited number of users	All data sources relevant to student success are incorporated into a central data repository and available to authorized users
Data should be stored in a central repository that is shared across	Are data stored in a central repository?	Data are not stored in a central repository. A central repository may	Most data are stored in a central repository. Accessing the central repository may involve	All data are stored in a central repository. There are clear ways to access the central repository.	All data are stored in a user-friendly central repository. There are clear ways to access the central	All data are stored in a user-friendly central repository with robust capabilities. There are clear ways to access the central repository.

different stakeholders		not exist or there have not been scaled efforts to move to a central repository.	multiple steps/parties. Only some parties on campus can easily access the repository.	Most parties on campus are familiar with how to use the repository.	repository. Most parties on campus are familiar with how to use the repository.	Most parties on campus are familiar with how to use the repository.
The data managers are clearly identified.	Are data stewards identified?	There are no official data stewards. Everyone independently collects/uses data.	There are no official data stewards; however, IT or similar group may act as unofficial data stewards. There are no official guidelines, but there is some understanding of how to work with stewards.	There are official data stewards at operational levels. There are some guidelines in place for working with data stewards.	There are official data stewards at senior- or executive levels.	There are official data stewards across all levels of the institution. All institutional groups (faculty, researchers, staff, and admin) have clear guidelines and understand how to work with data stewards and their delegates.
Maintain a data dictionary for all data elements used in reporting	Is there a resource that provides users with data definitions?	Data definitions are not available	Data definitions are available to a limited number of users	Text documents are available but not in a central location for user	Text documents are available centrally located for all users to access and are updated regularly	Data definitions are available and regularly updated on the internet and searchable by users; training may be available to users on demand.
3. Data Quality – Transparent quality management techniques to ensure usable data						
Data are evaluated for quality and reliability	Are there procedures to ensure quality data?	No process to evaluate the data quality	There is a process to clean the data and it is not used widely	There is a process to clean the data and it is widely used	There is a process to clean the data; there is a process to test the data for accuracy and reliability;	There is a process to clean the data; there is a process to test the data for accuracy and reliability; predictive models are validated

Students are asked to update their information on regular basis	Do students know how to update their information?	Student information is updated only if the student initiates the process and they must update their information in different offices/systems.	The process for students to update their information, requires that the student complete a form	There is an online process for students to update their information, but that information is not propagated to other enterprise systems, and they are not prompted to update	There is a student-friendly online process for students to update their information and that information is propagated to other enterprise systems;	There is a student-friendly online process for students to update their information and that information is propagated to other enterprise systems; students are asked to update their demographic information each time they register
Self-reported data should be identified in the data dictionary and in the metadata (% of missing cases)	Do users know the sources and definitions of the data such as: are these data self-reported?	No	A limited number of users are aware of which data elements are self-reported	The data dictionary identifies which data elements are self-reported	The data dictionary identifies which data elements and some reports identify self-reported metrics	The data dictionary identifies which data elements and the enterprise reporting tool will identify metrics that are based on self-reported data
4. Data dissemination – Distribution of data, analyses, and reports to the people who need them						
The institution has an enterprise reporting system for communicating information	What tools are used for distributing data and communicating information?	There is no centralized tool for easy access to data. Data may be distributed on an adhoc basis.	The institution has a centralized tool, but the information is not disseminated or there is no access to users.	The institution has a tool that disseminates information, but the use of information is not tracked.	The institution has an interactive tool to help users engage with the data with interpretive language to understand its meaning	The institution has an interactive tool to help users engage with the data with interpretive language to understand its meaning allowing them to extract insights with access to data stewards and managers
Different departments/units can create, or	Are there processes in place for units to develop	There is no centralized tool for accessing unit level data.	There is a centralized tool that is available to	Reports are created and available for most users to drill	The business units meet with IR to request reports that may be	The institution's chief data officers meet regularly with the institution's business

access custom reports designed for their needs in a timely manner	custom reports that unique to their department?	Data may be distributed on an adhoc basis.	a limited number of users.	down (analyze) without additional support	developed and become useful.	units to assure that reports are useful and relevant, and the data stewards understand business needs in order to prepare data for those needs. A shared model is used to develop individual reports.
The institution provides training to faculty and staff	Is there training on how to use data?	Formal training is not available. Users must either learn on their own or from a colleague.	Training is available on adhoc basis and is voluntary. The training is not updated to reflect the changing needs of institutional data.	Training is available on an infrequent and regular basis and is voluntary with incentives.	Training is available on a regular basis and is mandatory.	The institution requires training for all faculty and staff who have access to the data. Trainings are easily accessible and feedback on the trainings are used to improve training sessions. Faculty and staff learn how to access and use the data and how to follow up if there are issues or insights.

There are metrics on the number of times a report is used.	Are there analytics to track how often a report is accessed and by whom?	No	The Chief Data officer maintains analytics on how often the data is used, but access is limited.	The Chief Data Officer has information on the number of times the report is used, but it may not be shared with users on a regular basis and feedback is not collected about how to improve reports.	The Chief Data Officer monitors the use of individual reports and shares that information with business units on a regular basis through an automated feedback loop.	The Chief Data Officer monitors the use of individual reports and shares that information with business units on a regular basis through an automated feedback loop and collects information on the usefulness of the data. The CDO will make adjustments to the reporting portfolio based on this feedback.
--	--	----	--	--	--	--

5. Data Analysis – Connections between the data and the questions

The data are useful and complete	Do the data address key questions that help the institution meet its strategic goals?	There are no data to support strategic goals or initiatives.	Data may be available that can address strategic goals but may not be accessible.	Data are available to address strategic goals but are not distributed widely or are incomplete.	Data are available to address strategic goals and are distributed widely but may not be complete.	Data are available to address strategic goals, are distributed widely, and feedback is sought to assure that the data are complete and relevant.
The institution has a set derived metrics that inform strategic initiatives and decision-making.	What insights are we deriving from the data?	There are no derived metrics for informed decision making.	Metrics are developed on an ad hoc basis.	Metrics are available, but not used or disseminated when needed.	Metrics are available and disseminated to the right users but may not be incorporated into decision making.	Metrics are available, disseminated, and users incorporate their insights into decision making, daily operations, and strategic initiatives.

6. Culture & Application – Creation of a new social practice around using data						
The reports are incorporated into the unit's workflow processes	Are reports the from the central reporting platform incorporated into business processes for individual departments?	Reports are not available at the department level to include in business processes. If a department wants a report, they will have to create it themselves.	Reports are available at the department level but must be requested by the department to use in their business processes.	Reports are available at the department level and are disseminated or used in a consistent manner.	Reports are available at the department level, are disseminated in a user-friendly format and reflect department priorities and align with strategic institutional priorities and used in a consistent manner.	Reports are available at the department level units demonstrate how the data are used in their business processes and decision-making. The reports are evaluated for their value and usability. This information is provided back to the DGC or the CDO for continuous improvement.
Data, insight, innovation, interventions are part of the institutional culture	Are insights derived from the data and included in discussions and decision-making?	Data may not be available, so most units make decisions without metrics.	Data are available to some units on campus either through ad hoc reports or an enterprise system, but most decisions are made without consistent metrics.	Leadership and academic and business units have access to data on a consistent basis through an enterprise reporting system. But no training is offered to users and no data are collected to evaluate the use of the data.	Institutional leaders and academic and business units have access to the data through an enterprise reporting system, participate in data training, and use the data. Data are available on the frequency of use, but these data are not used for improvement.	Senior leaders and academic and business units use the data from an ERP, training is required, insights are extracted, interventions are created, and metrics are tracked to evaluate progress on initiatives. The evaluation on the use of data is used to enhance users' data experience.
Policies are modified based on data insights.	Do changes in practices or policies take place as a result of data-	Data are not used in policy discussion.	Data are available but not used to inform policy decisions.	Data may inform policy reviews but not directly supporting policy decisions	Data are used for policy discussions and are directly linked to policy decisions.	Data on effective (or ineffective) interventions lead to evaluation and changes in policies.

	informed insights?					Evaluation of the policy occurs on a regular basis.
7. Evaluation – Consistent assessment of the effectiveness of the data in affecting outcomes						
Interventions are evaluated and measured to determine their success	Are data available on how effective interventions are?	Interventions are not identified	Interventions are not measured, or their effectiveness is based on anecdotal evidence of “hunches”	Data are collected but are not linked to activities related to the effectiveness of the intervention.	Data are collected but are linked to activities related to the effectiveness of the intervention, but not linked to future activities.	The data show the effectiveness of selected interventions and impact future activities. Evaluation of interventions occur on a regular basis.
The institution is moving the needle on key metrics	Do institutional leaders evaluate progress on key metrics and communicate their progress to the community?	Data on key metrics may not be collected or reported.	Data on key metrics are collected and available but are not routinely shared with leadership.	The data are collected, available, and shared with leadership but the data are not directly tied to strategic initiatives.	The data are collected, available, and shared with leadership but the data are tied to strategic initiatives.	Data are collected and shared with leadership. Metrics are tied to the strategic plan and actively guide activities and resources.
8. Overall Rating						
Overall Rating	How would you rate your institution’s overall data maturity?	We don’t have the components for data maturity	We have some components of a mature data infrastructure.	We have some components of a mature data infrastructure, but it is not embedded in the culture.	We have many components of a mature data infrastructure, and it is embedded in the culture.	We have all the components of a mature data infrastructure, it is embedded in the culture, and it is improved regularly.
The institution’s members practice all aspects of data	Are all aspects of data literacy practiced by the institution’s	Faculty and staff do not have sufficient data literacy to identify, collect,	Faculty and staff have sufficient data literacy to identify needed data, but do not have	Faculty and staff have sufficient data literacy to identify and collect needed data, but do not	Faculty and staff have sufficient data literacy to identify, collect, and analyze needed data, but do	Faculty and staff have sufficient data literacy to identify, collect, analyze, and apply institutional data.

literacy: Identify, collect, analyze, and apply institutional data.	faculty and staff?	analyze, and apply institutional data.	sufficient data literacy to collect, organize, analyze, communicate, and apply institutional data.	have sufficient data literacy to analyze, and apply institutional data.	not have sufficient data literacy to apply institutional data.	
---	-----------------------	---	---	--	---	--

Phase 2: To be completed by the Senior Institutional Researcher or Chief Data Officer (Data Manager).

What is your role at your institution? _____

What is your institution? _____

Instructions. For each of the questions below, write a few sentences to describe your institution’s practices on the topics and subtopics listed. Describe successes and good practices. Identify challenges and indicate changes you’d like to see take place. Responses will be aggregated and analyzed to identify common focus areas and best practices. The results will be shared with a broader community of learners to engage conversations designed to help institutions improve their data capacity. Individual responses will be confidential.

General questions about your institutions data infrastructure:

1. Please describe your institution’s student information system.
2. Does your institution have a customer relations management system (CRM)? If so which one?
3. Does your institution have a centralized data repository or data warehouse? If so, what platform is used for housing data?
4. What reporting system(s) does your institution use?

Open-ended responses – for each of the areas below, describe strengths in your organization and challenges your organization faces. (Note these areas correspond to the items in Phase 1.)

1. Please describe your Data Governance practice - systemic and intentional practices and policies to support the effective use of data
 - a. Identify the key stakeholders who will help to create a new culture.
 - b. Transparent policies and procedures on how to use the data are communicated to the users
 - c. A data governance committee oversees data collection and use

2. Data Collection – Comprehensive processes that integrate and transform data from disparate sources
 - a. Include all data sources that relate to student engagement and success
 - b. Data should be stored in a central repository that is shared across different stakeholders
 - c. The data managers are clearly identified.
 - d. Maintain a data dictionary for all data elements used in reporting

3. Data Quality – Transparent quality management techniques to ensure usable data
 - a. Data are evaluated for quality and reliability
 - b. Students are asked to update their information on regular basis
 - c. Self-reported data should be identified in the data dictionary and in the metadata (% of missing cases)

4. Data dissemination – Distribution of data, analyses, and reports to the people who need them
 - a. The institution has an enterprise reporting system for communicating information
 - b. Different departments/units can create, or access custom reports designed for their needs in a timely manner
 - c. The institution provides training to faculty and staff
 - d. There are metrics on the number of times a report is used.

5. Data Analysis – Connections between the data and the questions
 - a. The data are useful and complete
 - b. The institution has a set derived metrics that inform strategic initiatives and decision-making.

6. Culture & Application – Creation of a new social practice around using data
 - a. The reports are incorporated into the unit's workflow processes
 - b. Data, insight, innovation, interventions are part of the institutional culture
 - c. Policies are modified based on data insights.

7. Evaluation – Consistent assessment of the effectiveness of the data in affecting outcomes
 - a. Interventions are evaluated and measured to determine their success
 - b. The institution is moving the needle on key metrics