MiGreatDataLake:
Partnership in Support of Educational Innovation in Michigan Public Schools

Diane Talo
Consolidation Grant Director
MAISA

Bruce Umpstead
Consolidation Grant Manager
MAISA

Jason Kronemeyer
Data Science Consultant, MAISA

Executive Summary

	Feasibility of Creating the Michigan Educational Data Lake
	MiGreat Data Lake Feasibility Study Executive Summary

Consolidation of Service Opportunity

Evaluating the Feasibility of creating an Educational Data Lake for Michigan's 56 Intermediate School Districts (ISDs), 537 school districts, and 293 public school academies (Districts) at the statewide level.

Feasibility Summary

By moving to a modern, API-driven data architecture, MAISA-sponsored and other statewide applications that today source data from the Michigan Data Hub backend database can instead s**ource data from the data warehouse that lives downstream from the MiDataHub** and leverages the Ed-Fi API.

This strategy has three key advantages:
1. It **reduces overhead** by decoupling downstream use cases from changes to the Ed-Fi Operational Data Store data schema. This strategy takes advantage of the stability of the Ed-Fi API.

2. This strategy provides a **common source of truth for key metrics**. In the current state, where each application sources raw data from the MiDataHub and transforms those data into metrics, every downstream application that uses the same metrics (e.g., attendance rates) is duplicating the same work repeatedly, and it is likely that each application is making slightly different decisions so that metric calculations look different in each application. Every application that sources a particular indicator from the MAISA data warehouse will be using the same information.

3. It establishes a source of truth for security and user permissions. In a strategy of applications consuming data from the Ed-Fi API (without a data warehouse layer), the Ed-FI API can authorize what data the application

has access to, but the application must then decide how to show each user the appropriate data that only that user should see. In the proposed strategy, the data warehouse can provide helper services to consuming applications to **make the security straightforward to implement, which should have the dual advantages of reducing costs and reducing the possibility of security issues**.

Recommendation

Development of a statewide Data Lake for educational entities in Michigan and their partners. With its Michigan Data Hub offering and program, Michigan is uniquely poised to lead the nation in using data to lead and support decisions that can change the way data is used to impact instruction.

By interconnecting multiple, disparate systems, districts will benefit from increased efficiency in data handling, which can lead to savings in both time and resources. Additionally, the scalability of the data lake ensures that it can accommodate future growth without requiring major infrastructure overhauls, further enhancing long-term sustainability.

Michigan and MAISA boast the most longstanding and robust K-12 data sourcing and data storage infrastructure in the country. Looking forward to the next step in data modernization, via a three-year, $15 to $20 million dollar investment, Michigan will become the nation's leader in a fully modernized data system. This initiative realizes the potential of Michigan's Data Hub by leveraging a data infrastructure downstream of the MiDataHub that turns data into insights with impact from the classroom to the boardroom.

Table of Contents

MAISA — MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS — SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

Introduction

In today's rapidly evolving educational landscape, data is increasingly becoming a critical asset for driving student success, enhancing instructional practices, and improving district-wide decision-making. However, many Michigan school districts face challenges in managing the vast amounts of data required to deliver high quality instruction, develop interventions for at-risk students, rely on effective early warning systems, and comply with state and federal regulations.

Current data systems are fragmented and often not available to critical users. These restrictions limit districts' ability to harness the full potential of their data. The MiGreat Data Lake, a unified, scalable, and secure platform for managing educational data across the state, aims to address these challenges.

The MiGreat Data Lake feasibility study explores the proposition to build on to the exemplary work of the Michigan Data Hub (MiDataHub), a successful statewide data collaboration by creating a shared data lake. Because of the success of MiDataHub, Michigan is uniquely poised to explore the logical evolution of the way the Data Hub securely moves student and school data. A shared data lake would offer a statewide, comprehensive data infrastructure for centralizing and standardizing data from various district-level systems using an integrated data architecture that facilitates real-time data sharing and advanced analytics.

By building a single collaborative data platform, districts can break down silos, gain deeper insights into student performance, and make better informed decisions to drive educational outcomes. This data infrastructure will also improve districts' ability to personalize learning, anticipate challenges, and design more effective interventions, ensuring that every student receives the support they need to succeed.

The goal of this feasibility study is to:

- evaluate the potential benefits, operational improvements, and cost savings associated with the implementation of the MiGreat Data Lake.

This study assesses the technical requirements, infrastructure upgrades, and governance structures needed to support the data lake, while also exploring its impact on data security, compliance, and long-term sustainability.

MAISA — MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS — SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

Why Michigan Needs A MiGreatDataLake

Michigan's education sector can benefit greatly from a data lake due to several key reasons:

1. Centralized Data Storage and Management

A data lake centralizes data from various educational institutions, such as K-12 schools, community colleges, and universities, into a unified repository, making data management and access more efficient. With its scalable design, it can handle large volumes of diverse data types, including student records, academic performance, formative assessments, behavioral indicators, social-emotional surveys, attendance, and extracurricular activities.

2. Enhanced Data Analytics and Insights

Data lakes support the ability to make advanced analytics available to practitioners in a usable, natural language format, available at a teacher's fingertips. Teachers would be better able to tailor educational experiences to meet individual student needs by combining data from multiple sources that have not previously been available. Data lakes also allow educators to analyze data trends, identify students at risk of falling behind, and implement early intervention strategies to improve student outcomes through increased personalized learning not currently available.

3. Improved Decision Making

A data lake provides policymakers with insights to develop data-driven education policies and allocate resources more effectively. Schools and districts can use the data to optimize resource allocation, ensuring that funds and materials are directed to where they are most needed.

4. Research and Development

Data lakes provide universities and research institutions with access to extensive secure, anonymized data for academic research, fostering innovation in teaching methods, curriculum development, and educational technologies. They also enable researchers to conduct longitudinal studies by analyzing historical data to understand long-term trends and outcomes in education.

5. Performance Monitoring and Reporting

Data lakes enable educators and administrators to use real-time analytics to monitor school performance, track student progress, and make responsive

adjustments to teaching strategies. They also enhance transparency by providing clear, accessible data on school performance and student achievements.

6. Collaboration and Sharing

Data lakes facilitate easier data sharing between educational institutions, enabling collaboration between schools, districts, and higher education. By analyzing data from multiple sources, educators can identify and share best practices, promoting continuous improvement across the state.

7. Support for Educational Technologies

A data lake supports the integration of educational technologies, such as adaptive learning platforms and virtual classrooms, by providing the necessary data infrastructure. It also drives innovation in EdTech by offering comprehensive data that helps develop new tools and applications to enhance teaching and learning.

8. Student and Teacher Support

Data insights from a data lake can identify areas where teachers may need additional training, promoting targeted professional development, as well as identify bright spots where educators are beating the odds for student success. Schools can also use data to enhance student services, such as counseling, tutoring, and extracurricular programs, by helping educators to better understand student needs and preferences.

Conclusion

A data lake can revolutionize Michigan's education sector by providing a robust infrastructure for data storage, management, and analysis. By enabling personalized learning, data-driven decision-making, enhanced research capabilities, and improved collaboration, a data lake can help create a more effective, efficient, and equitable education system in Michigan.

Logic Model, Guiding Tenets, and Context

A. Logic Model

> **Goal:**
> To create a unified, statewide data lake infrastructure that enhances data equity, accessibility, and real-time analytics across Michigan's Intermediate School Districts (ISDs) and school districts, facilitating

MAISA | MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

> improved educational outcomes and decision-making

- The complete logic model can be found in [Appendix A](#).

.

B. Guiding Tenets

1. Educational Data Equity

Provide all Michigan Intermediate School Districts (ISDs), Public School Districts, and Academies (districts) equal access to high-quality data infrastructure. This tenet ensures that all districts, regardless of geographic location or financial resources, can leverage advanced data systems for informed decision-making, promoting educational equity across the state.

2. Data Accessibility and Interoperability

Centralizing data into a unified, cloud-based platform will streamline data management and reporting processes. This will reduce redundant systems, automate data updates, and improve educational data management efficiency. Districts will save time and resources previously spent on manual data handling and siloed systems.

3. Improve Data Security and Governance

Implement robust security measures that comply with local, state, and federal regulations to protect sensitive student, staff, and district data. Standardized security protocols across the state will reduce vulnerabilities and ensure that all districts maintain consistent data privacy practices.

4. Leverage Shared Infrastructure and Expertise

Pool resources, expertise, and data storage capabilities using a shared, cloud-based infrastructure. This will reduce individual district costs and increase access to cutting-edge data analytics tools and AI-driven insights, empowering districts to make data-driven decisions without the financial burden of maintaining separate systems.

5. Increase Efficiency and Cost-Effectiveness

Centralizing data into a unified, cloud-based platform will streamline data management and reporting processes. This will reduce redundant systems, automate data updates, and improve educational data management efficiency. Districts will save time and resources previously spent on manual data handling and siloed systems.

MAISA | MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

6. Support Professional Development and Capacity Building

Provide comprehensive training and professional development opportunities to build capacity for effective data management across all districts. By equipping IT staff, administrators, and educators with the skills needed to operate within a centralized data platform, districts will be better prepared to utilize data for improving educational outcomes.

7. Facilitate Adoption and Support:

Recognize the varying capabilities and resource constraints of different districts, especially those that may be slower to adopt new technologies. The initiative will provide a clear pathway for late adopters, ensuring that all districts can transition smoothly to the new system, regardless of their starting point.

C. Context

Michigan's ISDs and districts face significant challenges in effectively managing and leveraging student and administrative data. The current landscape is fragmented, with many districts relying on various disconnected data systems, leading to inconsistencies in data quality, delays in reporting, and difficulties in making data-driven decisions. This lack of a unified data infrastructure hampers districts' ability to monitor student progress accurately and limits their capacity to implement timely interventions based on real-time insights.

Many districts, particularly those in rural or under-resourced areas, struggle with managing large datasets and accessing the advanced analytics tools needed to meet modern educational demands. Disparities in data access across districts exacerbate the gap between well-resourced and under-resourced schools, contributing to inequities in academic outcomes. As Michigan's schools increasingly rely on data to inform teaching and learning, a centralized, scalable solution is more critical than ever.

The Michigan Data Hub, discussed in detail in the next section, a vital component of the state's educational data infrastructure, has made significant strides in centralizing and standardizing data across districts. While the Hub has improved data-sharing capabilities for many schools, additional efforts are needed to expand these capabilities into a comprehensive, cloud-based data lake that can handle a wider array of data types and provide more advanced analytics and reporting. The MiGreatDataLake initiative aims to build on the foundation set by MiDataHub, addressing these broader challenges and ensuring that all districts, regardless of location or resources, have equitable access to the data tools they need to succeed.

MAISA | MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

D. Comparable Programs

Several state-level programs provide evidence that other states are moving forward with establishing an equitable data infrastructure. Examples include:

1. South Carolina

The South Carolina Department of Education, in partnership with the state's District Data Governance Group, has been supported by a technical partner since 2018. This partnership began by providing student growth measures based on annual state assessments, enhancing accountability, and offering insights into school performance statewide. The next phase involved implementing a statewide data framework powered by a shared Operational Data Store (ODS), enabling districts to push real-time student information, such as rosters, attendance, and behavioral incidents, from their local student information systems to the statewide system.

The work in South Carolina includes the ongoing rollout of tools for real-time data integration (ODS/API), data warehousing, administrator dashboards, and teacher-focused data applications. Future efforts aim to expand the range of assessments and other data types incorporated into the system to provide a more comprehensive view of student performance.

2. Texas Education Exchange

Since early 2023, the Texas Education Exchange has been supported in its Ed-Fi data initiative. The project focuses on implementing tools for real-time data management (ODS/API), a centralized data warehouse, administrator-facing dashboards, and teacher-oriented data applications. The Exchange is currently engaging districts across Texas to integrate their data sources and utilize these tools and other Ed-Fi-based applications.

In addition to technical implementation, the Texas Education Exchange receives high-level advising and management support, helping navigate the complex ecosystem of state stakeholders and vendors. This support includes managing subcontractors developing applications within the Exchange and facilitating collaboration between the Exchange and the Texas Education Agency as the state builds a parallel data system.

Demonstration of Concept: Michigan Data Hub

The Michigan Data Hub (MiDataHub) is a statewide initiative to provide a centralized data management platform for Michigan's educational institutions. It was developed to help school districts efficiently manage and share student information, enhance

data-driven decision-making, and reduce the burden of managing multiple data systems. MiDataHub integrates various district-level data systems, such as student information systems (SIS), special education platforms, and assessment tools, enabling seamless data exchange across districts and between schools and the state. This ensures that accurate, up-to-date information is readily available to educators, administrators, and policymakers.

MiDataHub is designed to create a standardized data architecture, enabling schools to adopt best practices in data management. Through this shared platform, schools can eliminate silos and the need for duplicate data entries, leading to improved data quality and consistency. The platform supports real-time data updates, allowing school districts to gain insights into student performance, attendance, behavior, and other critical decision-making areas. Data centralization also supports compliance with federal and state reporting requirements, making it easier for schools to meet regulatory standards.

Beyond data integration, MiDataHub fosters collaboration across districts by providing a common platform for communication and problem-solving. By sharing data securely, school districts can identify trends, address challenges collectively, and leverage data analytics to improve teaching and learning. The platform also supports the growing need for interoperability, where districts can integrate new applications and tools without the complexity of overhauling existing systems. This adaptability makes MiDataHub a vital resource for districts across Michigan, enabling them to harness data to improve educational outcomes.

A. Michigan Data Hub Adoption

According to the 2023-2024 Legislative Report, MiDataHub enjoys:

- **100% District Adoption**

Districts have overwhelmingly embraced the initiatives, nearing full adoption. This showcases the widespread acceptance and integration of funded projects. 100% of districts have implemented at least one MiDataHub feature.

Annual Investment: $3,500,000 or about **$2.45** per student.

MiDataHub currently provides 10,533 active connections (integrations) between school data systems, a 21% increase from one year ago. The total annual value of these integrations is over $77 million annually.

MAISA — MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS — SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

- An **integration** refers to the automated connection between a school district's data systems and a centralized platform, such as MiDataHub. It allows data from various local systems, like student information systems (SIS) and assessment tools, to be transferred and standardized in real time. This process ensures that data can be aggregated, analyzed, and shared efficiently across districts, fostering interoperability between systems and providing a unified source of truth for school data.

B. MiDataHub Direct Return on Investment (ROI)

The MiDataHub routinely releases ROI studies to inform the Michigan Legislature and its other stakeholders of the benefits of collectively addressing data interoperability. The 2024 study found:

Average Cost per Integration		Savings	
Non-MiDataHub Integration	$7,532		
Average of all Integrations	$6,055	$1,477	19.6%
MiDataHub Integration	$3,711	$3,821	50.7%

ROI is calculated based on the time and material savings this investment generates for districts:

861 Districts	9.7 MiDataHub Integrations (avg.)	861 x 9.7 x $3,821 = **$31,911,846** in savings

> Annual ROI = $31,911,846 ÷ $4,959,000 = **643.5%**

C. Additional Savings, Funding, and ROI from Data Management:

Across 8 specific data management tasks, MiDataHub helped districts save over $4.4 million by streamlining processes that would otherwise require significant time and resources. MiDataHub also helped districts avoid tasks valued at over $26.8 million, such as manual data entry or building and repairing integrations (coding).

- By adding the savings from data management tasks ($4.4M) and the value-added work avoided ($26.8M), the total combined savings amount to $31.2 million.

In addition to added saving, MiDataHub also assisted districts in accessing $9.0 million in additional state funds, which they could obtain with almost no effort due to the platform's capabilities. This brings the total savings to between $41.2 million and $41.8 million.

MAISA
MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

- The total of added saving/revenue totals between $41.2M and $41.8M in savings or $29 per student.
- The Net ROI for this added savings/revenue is between $36.2 million and $36.9 million annually, or $25 to $26/student..

> The added savings/revenue ROI percentage ranges from 830% to 843%.
> 8 times that value of every dollar invested.

Comprehensive Data Infrastructure

For all the effectiveness of MiDataHub, it does not provide a comprehensive data infrastructure as several key elements are missing. The following chart provides a visual representation:
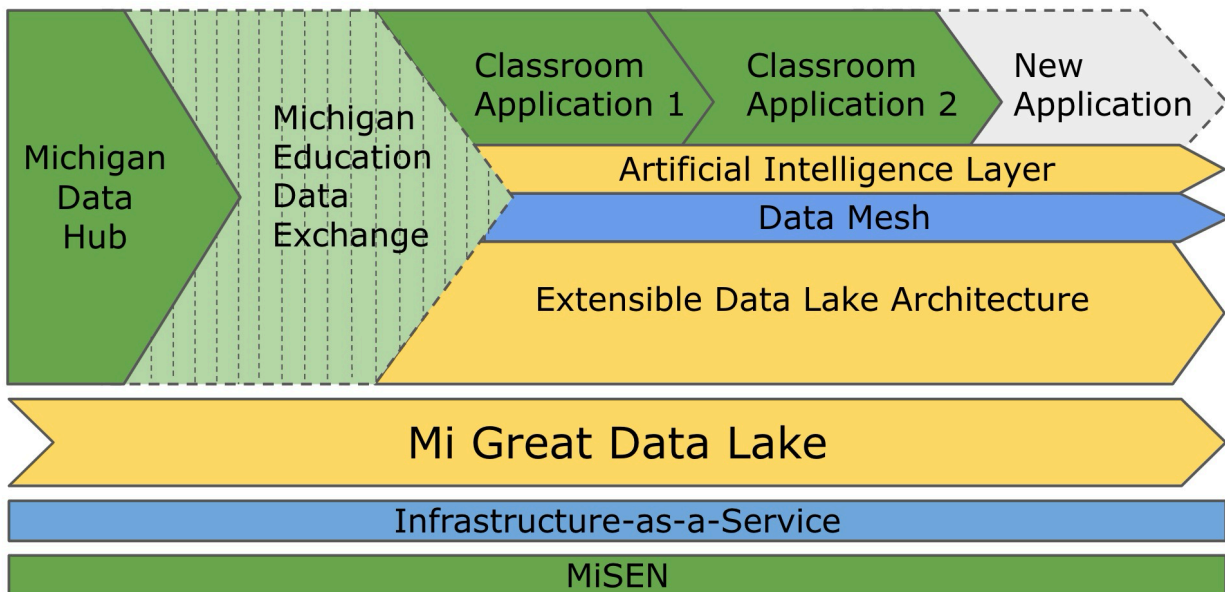
A. Data Infrastructure Elements

1. Michigan Data Hub (Existing, Requires Upgrades)

A statewide initiative that enables the secure, automated sharing and management of data across Michigan's educational entities. It integrates various district-level data systems, such as Student Information Systems (SIS), and centralizes that information for reporting, analysis, and compliance purposes. The goal is to standardize data collection and ensure data accuracy, providing a critical foundation for data-driven decision-making across all districts in Michigan.

2. Michigan Education Data Exchange (Existing, Requires Upgrades)

A data warehouse that is becoming rapidly outdated, originally designed to serve as a central repository for data from across Michigan's districts. It has facilitated the storage and analysis of large datasets, and enabled efficient data sharing among districts, the state, and other authorized parties. By aggregating data at a state level, the Data Exchange has supported compliance with state and federal reporting requirements and enables broader insights into statewide educational trends.

3. Classroom Applications 1 & 2 (Existing, Requires Upgrades)

These represent specific classroom-focused software tools integrated into the MiGreatDataLake architecture to pull real-time data, ensuring that educators can access the most up-to-date information for making instructional decisions. These include:

- **MiCIP (Michigan Integrated Continuous Improvement Process)** is a platform that supports school districts in planning, implementing, and evaluating continuous improvement strategies. It focuses on aligning district goals, data, and resources to promote student success. By streamlining the improvement process, MiCIP integrates various plans into a cohesive approach, making it easier for districts to manage and track their progress toward educational goals.

- **MiECC (Michigan Early Childhood Collaborative)** aims to improve early childhood education by fostering collaboration among providers and educators. The system is focused on supporting young learners and ensuring their readiness for school. MiECC integrates early learning data to enhance the effectiveness of programs, ultimately improving outcomes for children in their formative years.

- **MiEWIMS (Michigan Early Warning Intervention and Monitoring System)** helps schools identify students at risk of academic failure by monitoring key indicators such as attendance, behavior, and course performance. The system provides early warning signals that prompt, timely interventions, allowing schools to support at-risk students before they fall behind. MiEWIMS supports data-driven decision-making, helping to improve student success rates and reduce dropout rates.

- **MiRead** is a literacy-focused platform designed to help educators and students improve their reading proficiency. It provides tools for educators to monitor student reading progress and identify areas needing intervention. MiRead aligns with Michigan's Read by Grade Three law, helping to ensure students meet critical literacy milestones early in their educational journey.

MAISA
MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

- **MiStrategyBank** is a repository of evidence-based strategies that schools can use to improve student outcomes. It allows districts to document and share successful strategies with other schools, fostering a collaborative approach to educational improvement. By offering access to various proven practices, MiStrategyBank supports continuous improvement efforts across the state.

4. New Application (New)

A placeholder for future or yet-to-be-developed applications that will eventually integrate with the MiGreatDataLake. This ensures that the architecture is flexible and scalable, able to accommodate evolving educational technologies and tools that meet emerging needs in classrooms and administration.

5. Data Visualizations (Existing, Requires Upgrades)

Data visualization tools within the MiGreatDataLake architecture will enable educators, administrators, and policymakers to easily interpret complex datasets through interactive charts, graphs, and dashboards. These tools will transform raw data into visual insights, making it easier to track student performance, attendance, behavior trends, and other critical metrics in real-time. By providing intuitive, user-friendly interfaces, the data visualization component will help stakeholders identify patterns, make informed decisions, and quickly respond to emerging challenges within the educational landscape. Additionally, these tools will support custom reporting and analysis, allowing districts to tailor data insights to their specific needs.

6. Artificial Intelligence Layer (New)

A crucial component designed to leverage advanced AI algorithms to analyze the vast amount of data within the data lake. This layer provides predictive analytics, identifying trends and potential issues before they arise. It also offers automation for tasks such as data entry, reporting, and analysis, enabling educators and administrators to make proactive, data-informed decisions based on real-time insights.

7. Data Mesh (New)

A decentralized data architecture approach that allows various parts of the data system to operate independently while maintaining secure connectivity. This model enables different applications and data sources to share and process information while ensuring that data ownership, governance, and security protocols remain localized. Incorporating governance, security, and digital identity, the data mesh establishes robust frameworks for managing data access, compliance, and protection. Adding a secure digital identity management system will ensure proper authentication and authorization of users, safeguarding sensitive data and enhancing scalability, flexibility, and real-time access across the state.

8. Extensible Data Lake Architecture (New)

The core infrastructure that serves as the backbone of MiGreatDataLake. It provides scalable storage for structured and unstructured data, ensuring that new data sources and applications can seamlessly integrate. This architecture is designed to grow with the system, allowing future applications, datasets, and services to be added without overhauling the existing infrastructure.

9. Mi Great Data Lake (New)

The overarching platform that unifies all data systems and applications across Michigan's education system. This centralized data lake enables real-time access to critical data, allowing districts to share, analyze, and act on information efficiently. It facilitates better data-driven decisions at both the classroom and administrative levels and supports the long-term goal of educational equity through accessible and actionable data insights.

10. Infrastructure-as-a-Service (IaaS) (Existing, Needs Upgrades)

The cloud computing infrastructure that underpins the MiGreatDataLake. This service provides essential computing, storage, and networking resources to efficiently host and run the data lake. IaaS allows districts to leverage high-performance computing power without the need to maintain their physical servers, reducing the cost and complexity of managing such a large data environment.

11. MiSEN (Michigan State Education Network) (Existing, Upgrade)

The state's high-speed education network provides the foundational internet connectivity for all districts. MiSEN ensures that schools have reliable, fast internet connections, essential for supporting the real-time data exchanges, cloud computing, and communications required by the MiGreatDataLake platform. It serves as the backbone for the networked education services in Michigan.

Estimates of Cost

The following are order of magnitude estimates for the costs associated with upgrading and building this shared, statewide data infrastructure:

Element	Description	Cost
Michigan Data Hub	Upgrade: Expand its capacity to integrate with and feed the other MiGreatDataLake elements by adding support for a broader range of data types and formats. This would also include addressing outstanding data architecture and data security needs and adding digital identity and rostering management features.	$1,545,450
Michigan Education Data Exchange	Upgrade: Develop a more robust data storage and retreival system capable of scaling with the large amounts of structured and unstructured data that will be funneled into the data lake. This would involve enhancing its storage, indexing, and retrieval capabilities to support advanced analytics and reporting tools.	$3,399,990
Classroom Applications	Upgrade: Modify the applications to directly interface with the data lake for real-time data sharing. This may require adjustments to the application's API or integration logic to ensure it can both push and pull data from the central data lake.	$479,090
New Applications	New: To ensure that future applications are designed with cloud compatibility and seamless integration into the MiGreatDataLake, build a Software Development Kit (SDK) with robust APIs for efficient data exchange and ensuring that all applications meet security and compliance standards.	$324,550

MAISA | MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

Element	Description	Cost
Data Visualizations	Upgrade: Develop and integrate advanced data visualization tools that can directly interface with the MiGreatDataLake to provide real-time, interactive visual insights. This would involve creating dynamic dashboards, charts, and graphs that pull data from the lake to enable stakeholders to monitor student performance, attendance, and other key metrics.	$1,530,000
Artificial Intelligence Layer	New: Implement advanced AI models that can process large, complex datasets within the data lake. These models would need to be trained to analyze data in real-time, identify patterns, and offer predictive insights. Additionally, the AI layer would need to integrate with other applications to offer insights directly within operational dashboards.	$2,704,540
Data Mesh (Governance & Security)	New: Build out the mesh architecture to ensure that each data source and application can operate independently while maintaining secure data connectivity. This would involve establishing decentralized governance structures that define clear roles and responsibilities for managing data access, compliance, and security across different districts.	$1,622,730
Extensible Data Lake Architecture	New: Expand the existing architecture to accommodate new data sources, applications, and analytical tools. This may include upgrading storage capacity, improving scalability, and building more flexible integration points to allow for future innovations and developments in data processing.	$2,395,450

Element	Description	Cost
Mi Great Data Lake	New: Build out the infrastructure to manage large-scale data ingestion, storage, and processing capabilities. This would involve upgrading the data pipeline to ensure that data can flow from various systems into the lake in real-time, while ensuring data quality, security, and compliance.	$1,777,270
Infrastructure-as-a-Service	New: Load exsiting data sets and scale the cloud infrastructure to meet the growing demands of a data lake, ensuring sufficient computing, storage, and networking resources.	$927,270
MiSEN	Upgrade: Implementing mirroring or pairing (redundancy or failover) over the fiber network to connect with Clouding hosting service	$154,550
MAISA Program Management	New: Implement a structured approach to oversee the coordination and execution of multiple interrelated projects within the MiGreatDataLake initiative. This involves establishing a centralized management framework to ensure that timelines, budgets, and objectives are aligned across all components of the project.	$1,700,000
Total Projected Program Cost		$18,560,890

- Total Projected Three-Year Program Cost: **$12.99 per student**.

A. Third Party Cost Comparison

MAISA engaged Education Analytics (EA), the national consulting firm that delivered the South Carolina and Texas Education Exchange solutions, to do a comprehensive needs analysis for MiGreatDataLake. EA's order of magnitude estimate:

- Base Price: $7,300,000 + High End Analytics $ 8,750,000 + MAISA Program Management Costs $1,700,000 = $17,750,000

Details of the EA proposal are provided in **Appendix A.**

MAISA | MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS
SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

B. Total Cost of Operation

After three years of implementation, we estimate the comparable annual total cost of operation of the current MiDataHub annual costs, understanding that it requires the MiDataHub and doesn't replace the MiDataHub:

Annual Investment: $3,500,000 or **$2.45** per-student (100% of students)

C. Return on Investment (ROI)

Return on Investment for the MiGreatDataLake is challenging because the results will be measured in gains in classroom outcomes that are no easily articulated in dollars and cents.

We believe a majority of ISDs and districts would not be capable of building an effective data infrastructure articulated above. Perhaps 5 ISDs with the technical expertise and capacity to attempt the work as a group could do so:

	Districts	Students
Wayne RESA	33	280,000
Macomb ISD	21	110,000
Oakland Schools	28	220,000
Kent ISD	20	106,000
Ottawa ISD	11	50,000
	113	766,000

53.6% of Michigan Students

We then calculated the group's Total Program Cost, assuming a smaller scale. If the groups costs were:
- 53.6% of the stateside estimate (~$10.2M), the statewide ROI would be 0.0%.
- 60% of the stateside estimate (~$11.4M), statewide the ROI would be 7.4%.
- 70% of the stateside estimate (~$13.3M), statewide the ROI would be 20.7%.

The statewide scale of the project lowers would lower Total Program Cost per-student for the 5 ISDs considered by providing equitable data infrastructure for the entire state with potential **7.4% to 20.7% ROI** over a regional or group approach in addition to serving **+663,000 students** with a majority in rural and remote areas of Michigan.

Recommendation

Based on the analysis conducted throughout this feasibility study, it is clear that implementing the MiGreat Data Lake offers significant potential for return on investment (ROI) for Michigan's school districts. The centralization and standardization of data infrastructure will streamline operations, reduce redundancy, and ultimately lead to considerable cost savings across districts.

By interconnecting multiple, disparate systems, districts will benefit from increased efficiency in data handling, which can lead to savings in both time and resources. Additionally, the scalability of the data lake ensures that it can accommodate future growth without requiring major infrastructure overhauls, further enhancing long-term sustainability.

The MiGreat Data Lake will improve the quality of educational outcomes by enabling more precise data analytics and real-time insights. These capabilities will allow districts to implement personalized learning strategies, anticipate challenges, and design targeted interventions. The integration of advanced data tools, such as artificial intelligence and predictive analytics, will help districts identify early warning signs and proactively address student needs, contributing to measurable improvements in performance. This will generate long-term value by enhancing both instructional decision-making and resource allocation.

Given the projected ROI—driven by both operational cost savings and improvements in student outcomes—it is recommended that the MiGreat Data Lake be pursued as a high-priority investment. The potential savings, combined with the increased capacity for data-driven decision-making, provide compelling justification for state and district-level stakeholders to move forward with this initiative.

Appendix A: Logic Model

	Description
Goal	To create a unified, statewide data lake infrastructure that enhances data equity, accessibility, and real-time analytics across Michigan's ISDs.

Inputs	Description
Stakeholders	Michigan Department of Education, ISDs, local school districts, data governance bodies, and technology vendors.
Resources	MiDataHub, Michigan Education Data Exchange, MiRead, MICIP, MiCloud IaaS (AWS), funding (State School Aid Act), IT staff, AI, and analytics tools.
Partners	External vendors, state IT services, cybersecurity teams, and educational analytics experts.

Activities	Description
Data Integration	Integrate student and administrative data from various sources (MiDataHub, MiRead, etc.) into the unified data lake.
Infrastructure Development	Build and scale cloud-based data storage (AWS or MiCloud), enabling real-time access, storage, and management.
Training and Capacity Building	Provide training for district IT staff, administrators, and educators on effectively using and navigating the data lake system.
Governance and Security	Establish data governance protocols, ensure compliance with state and federal laws (FERPA), and secure access to sensitive data.
Data Accessibility and Reporting	Develop dashboards and reporting tools to provide real-time insights into student performance, attendance, and other key areas.

MAISA — MICHIGAN ASSOCIATION OF INTERMEDIATE SCHOOL ADMINISTRATORS — SUPPORTING REGIONAL EDUCATION SERVICE AGENCIES

Outputs	Description
Data Lake Infrastructure	A fully integrated data lake accessible by all ISDs and districts.
Real-Time Data Access	Enable real-time data and analytics for educators and administrators.
Streamlined Reporting	Simplify reporting processes for state and federal requirements.
User-Friendly Dashboards	Dashboards that regularly update on student performance and attendance.
Training Resources	Comprehensive training modules for district IT staff and educators.

Outcomes	Description
Short-Term	Increased data accuracy, real-time access, and data-driven decision-making.
Medium-Term	Improved educational outcomes, stronger collaboration across districts, and increased staff capacity to manage data.
Long-Term	A standardized statewide data infrastructure providing equitable access and sustainable data management across all districts.

Appendix B: Education Analytics High Level Prospectus

High Level Budget Guidance for MAISA's Data Modernization Efforts