



The Massachusetts Data Hub

A one-stop connected and secure ecosystem

Over the last decade, awareness and concerns over student data privacy issues gained attention across the United States. With 2020 came a newly heightened awareness of privacy, as nearly all students were thrown into online learning, and families had greater visibility to the technology tools their children were using. At the same time, the need for more robust interoperability has spiked as more and more technology tools are needed to provide online, hybrid, and in-person education. In Massachusetts, a partnership between the Executive Office of Education (EOE), the Student Data Privacy Consortium (SDPC), Cambridge Public Schools (CPS), and Cedar Labs was formed to address these twin needs.

The Massachusetts Data Hub is a project to create a connected and secure ecosystem, balancing the needs of privacy and interoperability in one place. The main goals of the project are to:

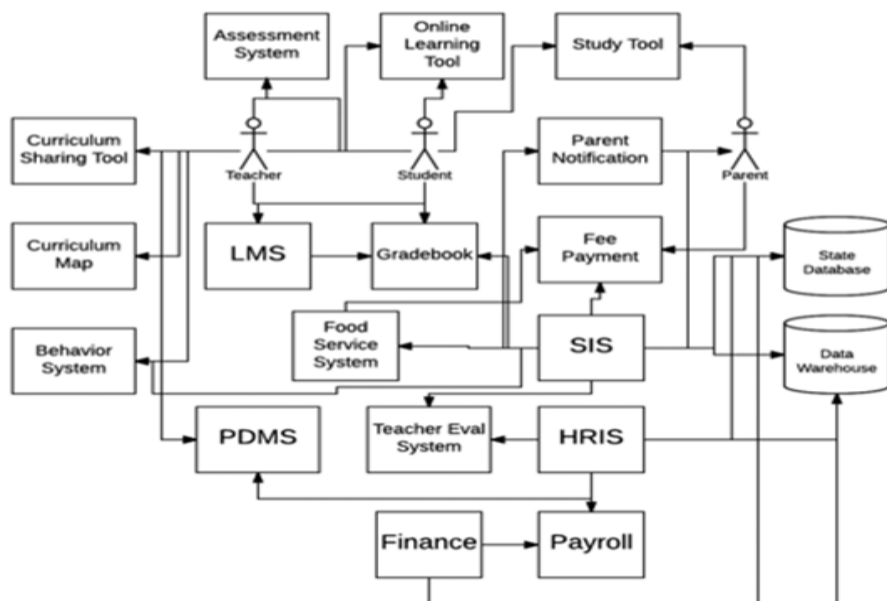
- Align to open data standards
- Provide an easy-to-manage tool-

set for districts to handle connecting their many software systems

- Leverage existing state data connections
- Meet vendors “where they are” to avoid external dependencies for project success
- Handle privacy and interoperability at the same time, in one system

Like most school districts across the country, CPS uses dozens of software systems. Like all districts in the state of Massachusetts, CPS is also required to report various student and staff data to the state’s department of education regularly throughout the year. In Massachusetts, this process of state reporting is made more efficient through a cloud-based data integration system provided by Cedar Labs. However, ex-

Figure 1: A rough diagram of typical data exchange architecture in schools



changing data with the state is just one part of the tangled mess of data connections between software systems, which school districts are responsible for maintaining. *Figure 1* is a diagram of the “old way” of exchanging data.

MA EOE, CPS, and Cedar Labs began to explore a way to leverage the existing data connection between the district and the state to branch the flow of data not just for state reporting, but to all of the many systems that CPS needs to keep updated with current data. Because state reporting in Massachusetts is done using the *open* Schools Interoperability Framework (SIF) Standard, this creates an opportunity for all vendors in the ecosystem to speak a shared, freely-available language for data, meaning no key data elements will be left out of the picture. Similarly,

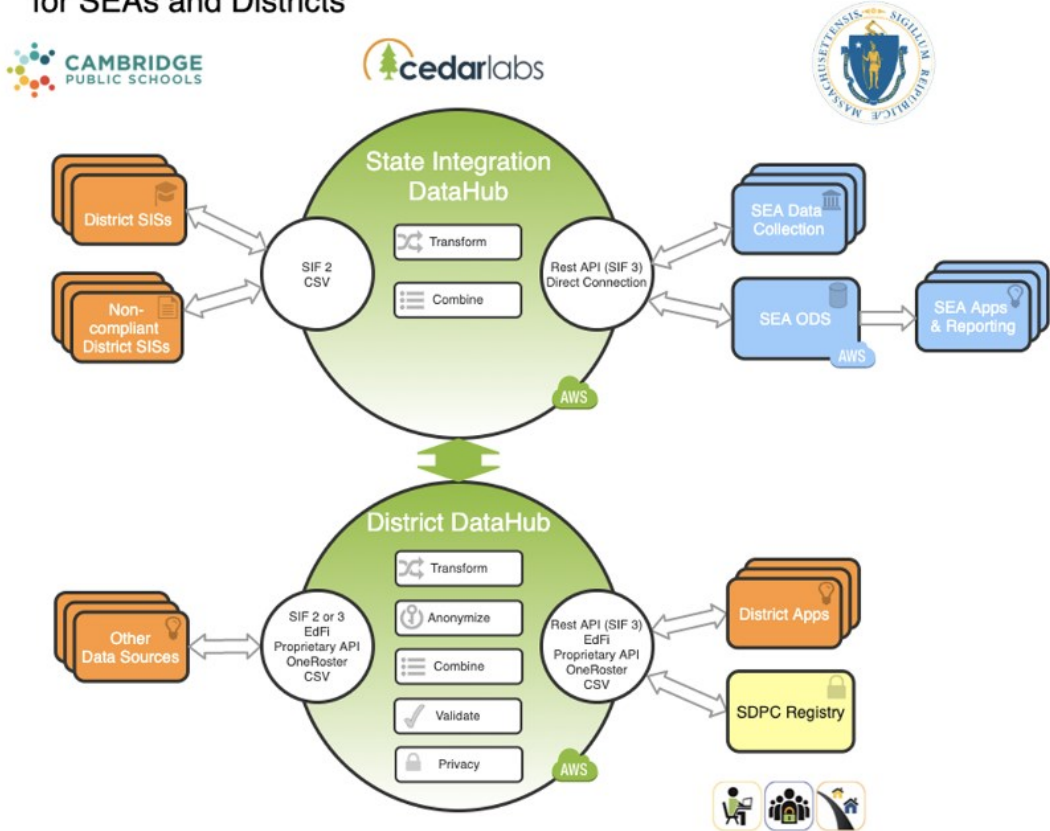
because the Access 4 Learning Community, the organization that drives Community development of the SIF Standard, is also the parent organization for the fast-growing Student Data Privacy Consortium (SDPC), this was a perfect opportunity to realize the SDPC vision of an ecosystem that was both connected *and* secure, by including both data and privacy “over the wire.”

“Interoperability, without embedded privacy, actually increases risk.” reports Steve Smith, Chief Information Officer at CPS and Founder of the SDPC, “School districts need an easier way to take control of their data, both to make sure that software vendors only get access to the data they need, and to make sure those vendors acknowledge the privacy and security obligations they agreed to in their contract.”

A pilot project was designed, which took advantage of Cedar Labs’ “hub and spoke” architecture to branch this state reporting data connection and create an easy-to-manage data integration platform for school districts. CPS began approaching vendors to participate in the pilot project. However, they quickly ran into roadblocks - vendors either were strapped for time and resources and unwilling to deviate to adopt an open standard for importing data, or vendors reported that they followed an open standard but in actuality the vendors required significant adaptations to that standard in order to exchange data. Yet again, the bulk of the effort for data integration was being *left to the school district to manage*.

However, Cedar Labs technology includes the ability to transform da-

**The MA DataHub:
Real-time statewide data integration and privacy management
for SEAs and Districts**



ta, in real-time, between different and otherwise incompatible data formats. So, Cedar Labs employed this technology to create a data hub service that districts can easily manage. The Data Hub leverages the open standards that districts are using to communicate data for state reporting, but doesn't mandate that vendors consuming data also subscribe to these standards. This "meet them where they are" approach reduces the risk to CPS, because the district is no longer dependent on vendors being willing to make changes to their systems in order to participate. Data are translated in real time, as they are exchanged between systems, and the "hub and spoke" approach still means that districts connect to each system **only once**, rather than having to manage multiple connections to each software system that consumes or shares data with multiple other systems. *Figure 2* is a diagram showing how the "district hub" and the "state hub" work in tandem to provide data to all the various software systems that both the district and the state need to exchange data with.

Because this centralized system can flexibly manage all the district's data connections, it can also provide a **privacy layer** that blankets all of the district's handling of confidential student information. Cedar Labs, in partnership with A4L and the SDPC, built privacy directly into the Data Hub. Districts can continue to manage their contracts using the SDPC Resource Registry, and the Registry's open API provides this information to the Data Hub. In addition to giving the district a single place to go to manage both data interoperability and privacy, the Data Hub also allows privacy-focused vendors to participate in the exchange of the Privacy Obligation Document (POD). The POD is a separate set of metadata, that communicates the specifics of the privacy obligations each vendor is contractually bound to, and ensures acknowledgement of those obligations by the vendor before they receive data. This split-second exchange and acknowledgement provides school districts, students, and parents with the assurance that their data privacy will be protected.

PODs contain all the required privacy metadata in a standard, industry accepted format. These obligations are driven by national and state laws as well as local requirements. The obligations are driven by Data Privacy Agreements (DPAs) executed by and between the LEA and Provider. The DPAs contain references to all applicable state and federal laws, technical obligations and security requirements where applicable. This will ensure that privacy *and* interoperability can be managed at the same time, bringing much-needed support to districts trying to keep up with the demands of exchanging data across their growing software infrastructure, while protecting the privacy of sensitive student, parent, and staff information.

For more information, contact:
Benjamin Silberglitt, Ph.D., CETL
CSO, Cedar Labs
Ben.Silberglitt@cedarlabs.com
www.cedarlabs.com



"Interoperability, without embedded privacy, actually increases risk. School districts need an easier way to take control of their data, both to make sure that software vendors only get access to the data they need, and to make sure those vendors acknowledge the privacy and security obligations they agreed to in their contract."



Steve Smith
Chief Information Officer, Cambridge Public Schools
Co-Founder of the Student Data Privacy Consortium (SDPC)

A4L Community

PO Box 1024

New Albany

OH 43054

202.621.0547

<https://www.A4L.org>
<https://privacy.A4L.org>