



Visualizing a Better Ohio through the Power of Data

Category: Data management, analytics & visualization
State: Ohio
Contact: Derek Bridges, Ohio Department of Administrative Services;
derek.j.bridges@das.ohio.gov
Project initiation
and end dates: Initiation date: July 2016; End date: November 2017

Executive Summary

In January 2017, Ohio Governor John Kasich's office issued a report entitled "Embracing Technology: Taking Ohio to Where We Need to be in the 21st Century." The Governor's office outlined the importance of unlocking the power of the state government's data resources. The report addressed taking "the benefits of advanced data analytics to the next level by requiring all state agencies, boards and commissions to share the data they now store in more than 1,600 databases." Sharing data and making data actionable through better data and insights positions the State to "tackle complex problems with solutions that improve Ohioans' health, security and well-being."

In parallel, the Ohio Department of Administrative Services, Office of Information Technology (OIT) had begun planning work to improve Ohio's capabilities to create an analytics environment to facilitate agency data sharing and execute analytical workloads. It was with this combination of OIT's planning efforts and the Governor's report that the Ohio Data Analytics initiative (ODA) began to expand toward its current form. ODA is a comprehensive program, including a big data platform, which allows for the collection, sharing, and analysis of Ohio's data across State agencies. The big data platform includes a hybrid approach allowing agencies to choose whether to run their workload in the cloud or on the State's on premises Hadoop data lake. ODA exists to make the State's data actionable, increasing efficiencies and improving citizen services. Ohio's state agencies generate a massive amount of data – 4 petabytes spread across 1,600 databases and 120 different agencies. Ohio is one of only five states to earn an "A" from industry experts for its work in advanced technology.

Until the Ohio Data Analytics program and platform, there was not an efficient and secure way for state agencies to share their data. Data existed in silos within each agency; varying in source, structure, and actionability. In addition, the data was not able to be shared securely. In short, data was not being used to its fullest extent.

To address this problem, the Ohio Department of Administrative Services (DAS) was tasked with putting in place an innovative platform that would pool the vast amount of data managed by the state's many different agencies. For the first time, data could be shared seamlessly across agencies and analyzed to measure and improve important state programs for constituents.

This essentially formalized the work of the Governor's office and OIT and thus ODA was born, which has already begun to meet and exceed the role it plays in these ambitious goals. Today, through ODA, the state is realizing its goal to turn data into actionable information. Multiple agency programs are analyzing critical societal problems for Ohio and the country, including issues like Infant Mortality Data, Opioid use, transparency, and program efficiency. Operationally, data storage costs have come down dramatically and ODA is quickly eliminating data silos across the state through an enterprise roll out of a single big data platform and supporting services program (ODA).

Concept

The Birth of ODA

What would become ODA began with an RFP, issued on January 5, 2017, that called for outside companies and experts to submit proposals to help the state identify and act upon trends and insight, utilizing the state's big data platform.

The state sought greater collaboration among state agencies to solve the state's most pressing problems. State leaders made it clear that greater sharing of information, and more insight into the data being shared, would be instrumental in solving the state's most critical problems. The RFP specifically defined 14 domains of expertise for firms to pre-qualify as vendors to the State of Ohio, including education, public health and safety, transportation, and the identification of fraud, waste, and abuse.

The issuance of Ohio House Bill #49 (OHB49) in July of 2017 accelerated ODA's development.

OHB49 called for the Ohio Department of Administrative Services (DAS) to "establish an enterprise data management and analytics program to gather, combine, and analyze unspecified types of data provided under the program by state agencies that participate in the program." OHB49 specifically noted the need to "measure outcomes of state-funded programs, to develop policies to promote effective, efficient, and best use of state resources, and to identify, prevent, or eliminate fraudulent use of state funds, resources, or programs."

The State of Ohio, for years, experienced many of the same problems plaguing most large corporations. Data use was typically limited to traditional reporting; there were no advanced analytical capabilities and very limited data sharing across agencies. When data was shared, it was often restricted due to privacy concerns.

Forced to use infrastructure not built for advanced analytical capabilities, the process of collaborating across agencies inhibited the usefulness of the state's data. Agencies were unable to take on more reports, use cases, and users. Data silos created inefficiencies and prevented agencies from accessing information that could be useful. The traditional infrastructure was designed for curated reports, not iterative, self-service analytics. It was never intended to be elastic, and therefore was not built for modern, flexible storage practices, such as object storage.

ODA was created to address these shortcomings. The ultimate goal was to tap into the enormous potential of the state's enormous data pools – making it shareable across the enterprise and applying visualization and analysis tools to make it more operational. Far from simply existing to feed various reports, ODA turns raw data into actionable insights that

- Data is a **Strategic Resource**
- Data is a **State Shared Asset**
- Data is for **Analytical & Research Purposes**
- The State will **Preserve Confidentiality & Privacy**
- Data is **Governed by Accountable Entities**
- Data is **Reusable**
- Data is **Interoperable**

agencies can use to identify important trends and address real-life scenarios and issues that impact the lives of Ohioans. Additionally the platform was designed to take action in agency transactional systems and programmatic touch points with constituents by providing analytical answers from the “black box” algorithms running within the platform.

The Development of ODA

In creating ODA, DAS sought to clear several hurdles that posed significant challenges that inhibited better data analysis for the state of Ohio:



Here is how the DAS team addressed each of these:

Silos

To eliminate data silos and share information across agencies, the ODA team partnered security and privacy both from the state and from agencies to create a data sharing protocol. The protocol helped to break down barriers to data sharing and address other overall enterprise challenges. They designed the platform with state and federal regulations in mind, including HIPAA, FERPA, PHI, 1075, CGIS, and PCI.

Data privacy considerations were at the forefront from the very beginning, and the team engaged in a continuous partnership with the state of Ohio’s privacy office and legal counsel to establish standard processes for data sharing agreements. To address security concerns, ODA partnered with the state CISO to leverage agency security and have IT and CIO representation. Finally, the team made sure to provide clarity, process, and responsibilities for data ownership, stipulating that ownership would reside with agency sources.

Expertise

The January 2017 the RFP was released encapsulating and generalizing the needs of the state of Ohio’s agencies, boards, and commissions across the state’s more than 1,600 operational systems and featured two supplements.

An RFP was released which centered on the expertise of the various firms. This RFP was uniquely designed to allow vendor firms to bring tools and resources on premise as a competitive component of their bid, allowing the State to effectively manage costs of



analytical projects. The RFP focused on 14 domain areas across Ohio and throughout the rest of the country (see illustration above).

Technology

A core mission of the State of Ohio is to embrace the use of technology to improve citizen services, as specified in OHB49. There are a number of different types of solutions and tools available in the world of big data analytics; including cloud compute and storage, including hybrid deployment options, data sharing and analytics, visual analytics; and data analytics support services. The platform contains many advanced and leading technologies but also allows outside experts to bring their own tools to execute workloads against the platform. The team responsible for supporting adoption of the technology is an additional component of the ODA program. The team engages and supports agencies through their data analytics exploration, from idea generation to ongoing support once their data is transferred to the big data platform.

In addition, the data analytics platform leverages Tableau to provide visual analytics for agency users to data shared and enriched on the platform. The on premises platform exists in the state's data center and is available to authenticated users who desire to share data or execute an analytic use case. Flexibility was an objective when standing up the platform. The State wanted to easily integrate state agency systems with the ODA platform, and provides a number of tools to accomplish this objective.

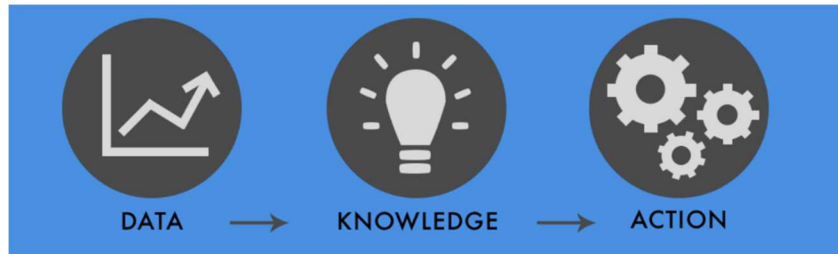
Significance

Insights alone do not benefit constituents. Constituents benefit when their state agencies *do something with those insights*. ODA helps state agencies compare data on performance, evaluate outcomes throughout the state, use prescriptive and predictive analytics to anticipate trends, and more. Together, these are helping to drive the creation and augmentation of new programs and initiatives that directly benefit the lives of Ohioans.

ODA has allowed agencies, including the Ohio Department of Health (ODH), to glean actionable and operationalized data analytics from the massive amounts of information at their disposal. With the introduction of this platform, the state of Ohio has encouraged its agencies to move beyond simple research to using their data to truly make a difference for public good.

ODA also adheres to several of NASCIO's State CIO Top Ten Priorities for 2018. These include Data Management and Analytics, Consolidation/Optimization, Digital Government, Shared Services, and Enterprise IT Governance. The platform incorporates big data, data governance, data architecture, strategy, and predictive analytics in an ambitious effort to make the state's data repositories work to improve the lives of Ohioans.

Impact

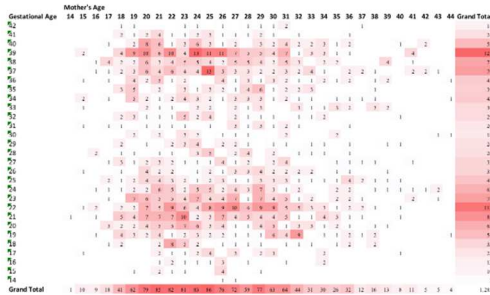


In a short time frame, ODA is already making a great difference within agencies and the lives of Ohioans.

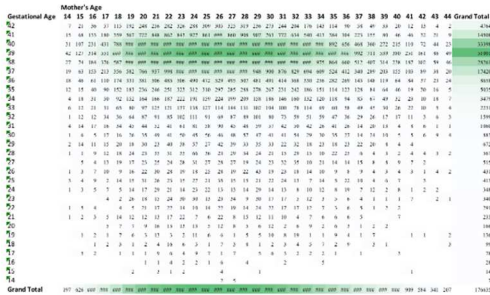
For example, Ohio has consistently ranked among the highest in the country for infant mortality rates. It is a problem the State has taken very seriously and infant mortality was the first use case that ODA was challenged with addressing. ODH is utilizing the ODA program to expand and enhance predictive and profiling models to determine those at risk for infant mortality and design targeted interventions based on this data.

When ODH reviewed the infant mortality data, it began to uncover interesting trends in how certain counties report causes of death. Some of the information was incomplete or inaccurate; there were blind spots in the data. Identifying these blind spots and the inconsistencies in reporting can help enhance analysis and data collection, and ultimately change the way programs are administered from the State. Below, is work in progress from phase I of the infant mortality project, providing ODH insights into the gestational age ODH can target intervention, and therefore make the biggest impact with premature babies.

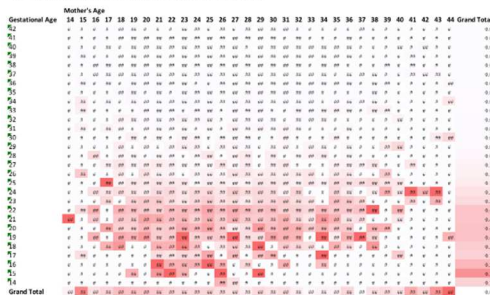
1. Infant Mortalities, OEI Counties 2013-16



2. Live Births, OEI Counties 2013-16



3. Infant Mortality Rates (%) OEI Counties 2013-16



DAS has undertaken a large-scale educational and promotional initiative to get additional agencies on board with the platform. DAS' implementation team routinely engages with

agency representatives to identify their challenges and goals and develop manageable statements of work (SOWs). They guide the agencies through procurement, train them on the technology and data analytics, and stage data for them. Promotion is handled through a variety of channels to reach the widest audience, including newsletters, conferences, and one-on-one conversations.

To date, DAS has partnered with other state agencies to release three additional analytic project SOWs. Multiple agencies have on-boarded to the platform for executing workloads to drive additional value for their agencies and constituents. The platform has succeeded in reducing administrative overhead and streamlining access to data for users across the state's agencies. Users are able to get the information they need and data warehouse storage costs have been reduced.

The opportunities to improve performance and policymaking based on actionable data are only going to increase as the data from state agencies itself increases. ODA was designed to be dynamic; its flexible architecture allows the platform to grow as enterprise needs grow, demand increases, and policy questions become more complex.

Once the data is staged, it becomes highly iterative in nature, since it does not have to be spun up again. Whereas the first leg of the infant mortality project will take six to nine months to deliver actionable information, that time is already being significantly shortened; thus state agencies are getting more accurate, insightful information, more quickly.

Each question that is input will derive more answers. As more data is shared and the system is used, advanced models can be created which will drive better outcomes. Ultimately, agencies throughout the State will benefit from this intelligence. More importantly, so will their constituents.