The State of Ohio: unlocking enterprise data and building an analytics program

Derek Bridges
Ohio Department of Administrative Services
Office of Information Technology
Agenda

State of Ohio Challenges

The Ohio Data Analytics Initiative

Addressing the Challenges: Silos, Expertise, Technology

Applied Analytics: Addressing Infant Mortality & Opioid Use Disorder

Additional Analytics In Progress
Ohio Data Analytics Initiative

ODA is a collaborative initiative spanning State agencies and dedicated to using data and analytical methods to solve problems affecting Ohioans.
## Analytical Maturity

**From Traditional BI to Analytics**

<table>
<thead>
<tr>
<th>Analytics 1.0</th>
<th>Analytics 2.0</th>
<th>Analytics 3.0</th>
<th>Analytics 4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>Diagnostic</td>
<td>Predictive</td>
<td>Prescriptive</td>
</tr>
<tr>
<td><strong>“What has happened?”</strong></td>
<td><strong>“Why did it happen?”</strong></td>
<td><strong>“What will happen?”</strong></td>
<td><strong>“What should I do?”</strong></td>
</tr>
<tr>
<td>- Data is stored in underlying databases and warehouses</td>
<td>- Explains why something happened</td>
<td>- All kinds of “raw” data is stored in Data Lakes</td>
<td>- All kinds of “raw” data is stored in Data Lakes with data relationships captured in graph databases</td>
</tr>
<tr>
<td>- Focus is on reports and dashboards</td>
<td>- Data discovery and correlations</td>
<td>- Focus is on algorithms and models that are able to predict business scenarios</td>
<td>- Focus is on guiding actions to be taken based on an insight</td>
</tr>
<tr>
<td>- Information is retrospective</td>
<td>- Drilling down to the root-cause</td>
<td>- Business managers act on optimizing performance based on probabilities</td>
<td>- Information points to the future</td>
</tr>
<tr>
<td>- Information portrays the state of business as “it has happened”</td>
<td>- Ability to isolate all confounding information</td>
<td>- Information portrays the state of business “as it could happen”</td>
<td>- Information points to what actions to take</td>
</tr>
<tr>
<td>- Business managers act on trends from the past</td>
<td>- Understand causes</td>
<td>- Business managers act on optimizing performance based on probabilities</td>
<td>- Information portrays the business steps to take to get to an outcome</td>
</tr>
<tr>
<td></td>
<td>- A look at past performance</td>
<td></td>
<td>- Business management functions can be automated for execution</td>
</tr>
<tr>
<td></td>
<td>- End result is commonly an analytic dashboard</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Analytics Challenges

- **Silos**: Organizations often face many barriers to data sharing; some very real and some part of organizational lore.

- **Expertise**: It is not easy to attract and retain individuals with the right skills and experience in data analytics and big data technologies.

- **Technology**: Traditional technologies don’t always keep pace with analytical needs and emerging technologies are moving quickly.
The Ohio Data Analytics Initiative

Taking Ohio to where it needs to be in the 21st Century by embracing technology *(H.B. #49 Sec. 125.32)*

- **Compute & Storage**
  - Hybrid deployment options

- **Data Sharing & Analytics Platform**
  - State of the art analytical platform and tooling available to State agencies

- **Data Analytics Support Services**
  - Technology, security & privacy, strategy, and procurement support

- **Pre-Qualified Analytics Firms**
  - 70 prequalified firms with analytical expertise across 14 domains

- **Visual Analytics**
  - Widely adopted visual analytics and data discovery capabilities
Collaborative Approach to Applied Analytics

Ohio’s unique and innovative approach drives value in months rather than years.

- Innovative Procurement
- Communication
- Education and Training
- Organized Data Sharing
- Ideation and Support
- Analytics Industry Engagement
- Multi-Disciplinary Teams

360° Entity Resolution

Commercial Cloud
- Hybrid Platform
- Visual Analytics
- BYO Tool Options
- Technical Support

Private Cloud

Insights & Outcomes

Ohio’s Data Analytics
Self-Service and Data Management

Big data capabilities help agencies unlock the value in their data.

**source**

- **Data Ingestion**
  - sourcing data from diverse systems and formats without coding

- **Discovery & Prep**
  - reviewing data, discovering and repairing inconsistencies

**share**

- **Data Lineage**
  - traceability of the sources of data, lifecycle and curation

- **Governance & Audit**
  - ensuring privacy, security and authorization of use

- **Data Sharing**
  - enabling collaboration

**understand**

- **Enhanced Data Science**
  - common tools for exploration and insight

- **Visual Analytics**
  - driving insights and understanding
Addressing the Challenges

- Silos
- Expertise
- Technology

Policies and design that enable data-sharing
O.R.C. 125.32

“...measure the outcomes of state-funded programs, develop policies to promote the effective, efficient, and best use of state resources, and identify, prevent, or eliminate the fraudulent use of state funds, resources, or state programs.”

- Data is a Strategic Resource
- Data is a State Shared Asset
- Data is for Analytical Purposes
- Data is Private and Confidential
- Data is Governed by Accountable Entities
- Data is Reusable
- Data is Interoperable
Data Sharing Protocol

Laying the foundation to enable data-sharing and addressing overall enterprise challenges

**Platform Design**
Built considering HIPAA, FERPA, PHI, 1075, CJIS, and PCI requirements and regulations

**Security**
Partnered with state CISO to leverage agency security and IT / CIO representation to develop comprehensive security plan

**Privacy**
Continuous partnership with privacy office and legal counsel to establish standard processes for data-sharing agreements

**Data Ownership**
Provide clarity, process and responsibilities for data ownership to remain with agency sources
Addressing the Challenges

Silos

Expertise

Technology

Engaging experts across domains and growing an analytical culture
Procuring Analytical Expertise

In 2017, the State issued an RFP that encapsulates and generalizes the needs of the State of Ohio’s 120 Agencies, Boards and Commissions across more than 1,600 operational systems.

Analytical Expertise

Ohio & Nationally Eminent Firms

14 categories of specialization firm capabilities, methods, & tools

Demonstrated innovations and outcomes

Team, expertise and resumes

Open and proprietary tools (generalized and domain specific)

69 Firms Offered Contracts

117 Business Entities

26 Ohio Companies

57 New State Vendors
Launching a Project

Ohio Data Analytics supports State agencies through the entire project lifecycle.

Project Ideation
- Subject Area
- Problem Statement
- Analysis Dimensions
- Desired Outcomes
- State Agency Sponsor
- State Participants
- Pragmatics & Timing

Assemble State Team
- Sponsorship and Oversight
- Exploratory Project Manager
- Program / Policy Experts
- Intervention Specialists
- Data Analytics Leaders
- Agency Systems/Data SMEs
- Agency System Providers (as required)

Assemble Project Data
- State Agencies and Systems
- Data Taxonomies and Volumes
- Data Anonymization & Handling
- Data Sharing Agreements

Solicit Expert Firms
- Develop Exploratory Project Work Solicitation
- Let to Pre-Qualified Contractors
- Evaluate / Award Proposals
- Commission Work

Execute Project
- Progress Reporting & Course Corrections
- Evaluate Ancillary & Additional Analysis
- Validate Findings, Recommendations
- Design, Implement and Measure Interventions
Addressing the Challenges

Silos
Expertise
➤ Technology

Bringing the right tools to solve problems without being locked into technical limitations
Ohio Data Analytics Big Data Platform

A hybrid data analytics ecosystem is available to State agencies, providing leading tools and the flexibility to run analytical workloads in the State’s private data lake or in the cloud.

**Data Sources**
- Structured
- Semi-structured
- Unstructured

**Data Acquisition**
- Batch Processing
- Near-real time
- Streaming

**Data Storage & Processing**
- Data Lake Platform
  - Processing Layer Zone
  - Agency Restricted Access
  - Shared Data Zone
  - Plug & Play BI Data
  - Refined Datasets
  - Project Zone

**Data Access**
- General Business Users
- Data Analyst Users

**Visual Analytics**
- Tableau

**Data Insights**
- Cloudera Data Science Workbench
- Alteryx

**Data Security & Governance**
De-Identification Strategy

Data Owner(s)

Personally Identifiable Information
Any data that could potentially identify a specific individual and can be used to re-identify pseudonymized data is considered PII.

Data Owner(s) & Business Associates

Pseudonymized Information
Substitutes the identity of the data subject in such a way that additional information is required to re-identify the data subject.

Public Users

Anonymized Information
Irreversibly destroys any way of identifying the data subject.

Pseudonymization Tactics
- SHA Hashing (Key Linking Fields)
- Obfuscation (Other PII)

Anonymization Tactics
- Aggregation
- Obfuscation

Agency Source Systems

Original Form Datasets
De-Identification Process

Data Lake

Exploratory Project

Project Deliverables
- Data Results
- Analytical Models
- Reports & Findings

Accessible By: Internal Users & Approved Agents of the State
Governed By: Business Associates Agreements, Data Sharing Agreements, State Data Handling / Access Standards (Supp N)

Required for agency use and operational analytics for intervention purposes, e.g. at risk subject identification.

Supports the ability to link datasets between agencies for analytical purposes.

Accessible By: Public Users
Governed By: State Data Handling Standards

Allowed results and findings from analytical studies to be published via a report or open datasets for public consumption.
Applied Analytics: Addressing Infant Mortality

Designing Targeted Interventions to impact Ohio’s Infant Mortality Rate
“Expand and enhance predictive models and profiling models to determine those at-risk for infant mortality in Ohio and design targeted interventions.”
Using Analytics to Help Reduce Infant Mortality

KEY QUESTIONS

Which mothers and infants are most at risk of infant death?
Which families are most likely to benefit from targeted interventions?
Which families are most likely to participate in targeted interventions?
Which intervention programs yield the best return on investment?

KEY MODELS

Evaluating Efficacy of State Intervention Programs
Identifying Mothers at High Risk of Infant Mortality and Constructing Their Profiles
Predicting the Characteristics of Mothers Most Likely To Benefit From An Intervention Program
Predicting Which Intervention Program(s) At-Risk Mothers Should be Enrolled In
Identifying Mothers Most At-Risk of Having a Baby that will Require a NICU Admission
Keys to Success

+++

MOVING BEYOND THE EXPLORATORY AND THE ACADEMIC

+++

INDUCTIVE REASONING AND EVIDENCE-BASED INTERVENTIONS

+++

STATE PLANNING DRIVING LOCAL IMPACT
360° Analytical Model

De-Identification
Secure Record Linkage

Medicaid

Job & Family Services

Mental Health & Addiction Services

ID: xe8rcfowh2

Health

Social Determinants of Health

Risk Models
Red Flags
Alerts
Interventions
Data-driven Interventions

Piloting, Tracking and Monitoring Targeted interventions to reduce IMR and Preterm Birth in key Ohio metros, building on the data insights in the analytics, targeted to the highest risk mothers

- Targeted Interventions delivered via Home Visiting
- Referrals to programs, e.g. WIC, SNAP
- Navigating to care e.g. progesterone therapy or antenatal steroid if indicated
- Reduction of risk behaviors e.g. smoking, drinking, drugs
- Evidence based reduction of other risk factors e.g. Domestic Violence or Postpartum Depression
- Introduction of Enhanced Screening in prenatal care settings birth hospitals
- Results are Tracked and Monitored by data analytics assets and approach is refined in tight iterations
Measuring Provider Efficacy & Outcomes

Executive Summary

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Households Served</th>
<th>Total Children Visited</th>
<th>Avg Time with Households (min)</th>
<th>Total Households on Waitlist</th>
<th>Provider Avg Missed Visit (All Model)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2017</td>
<td>8,590</td>
<td>9,738</td>
<td>72.96</td>
<td>154</td>
<td>51.76%</td>
</tr>
</tbody>
</table>

Map of Home Visiting Regions

Filter by selecting a year and model below.

Bounce Rate

<table>
<thead>
<tr>
<th>CareStar</th>
<th>Franklin</th>
<th>Marion</th>
<th>Delaware</th>
<th>Guernsey</th>
<th>Muskingum</th>
<th>Miami</th>
<th>Montgomery</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CATHOLIC CHARITIES CORPORATION</th>
<th>Cuyahoga</th>
<th>Lorain</th>
<th>Geauga</th>
<th>Lake</th>
<th>Summit</th>
<th>Medina</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FRS Counseling</th>
<th>Ross</th>
<th>Clinton</th>
<th>Adams</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OhioGuidestone</th>
<th>Cuyahoga</th>
<th>Lorain</th>
<th>Tuscarawas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Topics Covered

Referrals
Applied Analytics: Opioid Use Disorder in the Criminal Justice System

Designing Targeted Interventions to improve outcomes for individuals with OUD
Understanding the profiles of individuals within the criminal justice system to prioritize resources and interventions and more effectively target those at-risk of opioid use disorder to avoid recidivism, dependencies, and death.
Bringing Data Together

**OPIOID SURVEILLANCE DATA**
- Prescription rates and pharmaceutical data
- Bad actors
- Treatment program data
- Historical interventions and successes

**CRIMINAL JUSTICE DATA**
- Incarceration, clinical, release, and recidivism data
- Locations of Residential Reentry Centers

**HEALTH OUTCOMES DATA**
- Opioid relapse and death rates
- Mental health prevalence
- Medical utilization rates, such as ER visits and claims data

**PROGRAMS / SOCIAL DATA**
- Demographics & education levels
- Employment rates
- Child welfare data & food assistance data
- Community support systems (i.e., school, church, and non-profit data)

**Individuals in Ohio aged 12 or older in 2014-2015 had used heroin in the past year: 37,000**

**Individuals seen for OUD in the publically funded behavioral health system according to MHAS Claims Data in 2017: 59,120**

**Of substance abuse treatment referrals are made by criminal justice services: 34.4%**
Using Analytics to Help Reduce Opioid Use Disorder in the Criminal Justice System

Who should we target for early intervention?

What service needs are we overlooking?

Where should we be focusing our efforts?

When is the best time to intervene?

How do we improve re-entry programs?

A focus of the project is determining gaps in diagnosis, treatment, social supports, and care management for individuals with OUD.

This project will use a larger breadth and depth of rich datasets than what has been used in prior research.
Opioid Risk View (ORV)

12+ publicly available data sources related to the opioid epidemic in the ORV to garner state and county level insights.
Additional Analytics in Progress
<table>
<thead>
<tr>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Health</td>
</tr>
<tr>
<td>Business 360</td>
</tr>
<tr>
<td>Cyber Security</td>
</tr>
<tr>
<td>Program Efficiency and Optimization</td>
</tr>
<tr>
<td>Fraud and Risk</td>
</tr>
<tr>
<td>Workforce &amp; Economic Development</td>
</tr>
<tr>
<td>Education</td>
</tr>
</tbody>
</table>

Using Data for Public Good

Additional opportunities under development
**Silos**

- Statewide laws, policies, and protocols are a valuable foundation.
- Collaborate to address the needs and requirements of individual programs.
- Sponsorship is essential.

**Expertise**

- It's difficult to attract and retain expertise in this market.
- Innovative approaches to procurement can help tap into the analytics market and reach expert firms that may not have traditionally pursued government contracts.

**Technology**

- This technology is evolving rapidly.
- There is a tool for every job – probably several.
- Don’t limit your options aligning to a single set of technical tools.
Thank you!

ODA Ohio Data Analytics

analytics.ohio.gov