Smart Columbus
Powered by Integrated Data Exchange (IDE)
Smart Columbus
Confluence of Civic, Research & Startup Cultures
• Winner $40M DOT Smart City Challenge
• $367M investment local businesses & partners
• $93M for research (largely from The Ohio State University)
• $10 M Vulcan Inc.

“Smart Columbus is just the start. There is a lot to this program no one has ever done before.”
Mayor Andrew Ginther, 2016
Confluence of 3 Disruptive Data Trends

Anyone can access, understand, use and share Open Data

IoT: Devices embedded in the physical world—e.g., bodies, busses, and businesses

Smart cities actively collect data and adapt policies and services to make citizenry successful

Smart Columbus
Data Flows From and Through Cities

- Public-Private Partnerships: COTA is funded by 0.5% sales tax through 2019
- Starting 2017, On-board Wi-Fi will produce 4.7 TB per month
- How does the city ingest, analyze and use all of this data?

Central Ohio Transit Authority Ridership
(Millions of Rides Per Year)

- 2006
- 2008
- 2010
- 2012
- 2014
Smart City Applications Abound

- CoGo opens trip data
  - https://www.cogobikeshare.com/system-data
  - Discover new bike paths
  - Over-provisioned and under-provisioned paths
  - Activity levels of citizens
Who rises to the challenge? Who makes Smart Columbus?

Challenges: Align For-Profit and Not-For-Profit Incentives, Find people with the right skills, Politics

Any problem in computer science can be solved with another level of indirection. David Wheeler, 1993
Who rises to the challenge? Who makes Smart Columbus?

• Solution: The city needs a system to (1) manage data resources from multiple sources and (2) provide a common interface for application developers & researchers using data

• Integrated Data Exchange (IDE)
  – Collect data from multiple sources (IoT), govern access, ensure privacy and enable analysis (data API)
  – Moving data is a new city service
Integrated Data Exchange Benefits/Challenges

IDE makes application developers and entrepreneurs the bottleneck, **not data access**
IDE makes researchers and IoT builders the bottleneck, **not integration**

• Challenges
  • Provide security
  • Provide fair access to the public
  • Protect the public from dangerous analysis
  • Protect the city from legal problems
  • Scale to support city-scale data ingest
  • Scale to support city-scale IoT
Smart Columbus Partnerships
Powered by Integrated Data Exchange

Smart Columbus Sandbox 🔄 Open Data 🔄 Translational Data Analytics & Research

Smart City 🔄 Internet Of Things 🔄 Columbus Integrated Data Exchange

Powered by Integrated Data Exchange
Giving developers access to Smart City datasets

- Smart Columbus Sandbox brings together entrepreneurs, developers to use existing open data sets
Parking Startup: Easton Mall Pitch

Request Reply Semantics

```json
"result": "success", "trafficCounts": [
  {
    "reading_time": "2017-01-01 11:00:00",
    "occupancy": "2.36", "volume": "65", "median_speed": "0"
  },
  {
    "reading_time": "2017-01-01 12:00:00",
    "occupancy": "3.38", "volume": "93", "median_speed": "0"
  }
]
```
Early Success: A Hack-A-Thon

Bonsai Data
VR Data Visualization in 3D with Trees
Smart Columbus Partnerships

Smart Columbus Sandbox → Open Data → Translational Data Analytics & Research

Smart City → Internet Of Things

Columbus Integrated Data Exchange
Translational Data Analytics Institute (TDAI) is the foundational element of the University’s $500 million investment into Discovery Themes.

- **University-wide Educational Investments**
- **Research across 30+ disciplines**
- **105 faculty & faculty-in-residents**
- **Communities of Practice**
- **Strategic business partnerships**
(1) Collect GPS and sensor data from IoT embedded in transit vehicles (e.g., phones & wi-fi access points)

(2) Model and manage demand, using OSU COTS testbed for future deployment in COTA

Prof. Anish Arora
Battery powered radar mesh networks for pedestrian and bicycle identification & wireless, sensor networks to expand Internet access passively & cost effectively

Andre Carrel – Public Transportation

Harmony Bench
What modes, networks, and types of transportation link performers within a city

Smart Campus Testbed - Prof. Rabi Mishalani
Route planning and fleet management using video & other sensing and tracking technologies.
Smart Columbus Partnerships

Smart Columbus Sandbox

Open Data

Translational Data Analytics & Research

Smart City

Columbus Integrated Data Exchange

Internet Of Things
“The big play and leverage point for us is that internet of things space. But the only way we get there is making sure that we have a secure, sound, accessible integrated data exchange that’s going to stand the test of time.”
Mayor Andrew Ginther, 2017
OUR VISION

ACCESS TO JOBS | SMART LOGISTICS | CONNECTED RESIDENTS | CONNECTED VISITORS | SUSTAINABLE TRANSPORTATION

The City of Columbus “Smart Columbus” vision won the U.S. Department of Transportation $40 million Smart City Challenge in June, 2016 after competing against 77 cities nationwide to implement a holistic vision for how technology can help all residents to move easily and to access opportunity. Columbus was also awarded an additional $10 million grant from Paul G. Allen’s Vulcan Inc. to reduce greenhouse gas emissions through the decarbonization of the electric supply and transport sectors.

Smart Columbus aspires to:
- Improve access to jobs through expanded mobility options in major job centers
- Compete globally through smart logistics
- Connect Columbus residents to safe, reliable transportation that can be accessed by all
- Better connect our visitors to transportation options
- Develop a more environmentally sustainable transportation system

THE COLUMBUS WAY

Columbus has built an unprecedented culture of collaboration. By knocking down silos and building partnerships, Columbus has quickly become one of the fastest growing cities in the country, leading the Midwest in job and wage growth. The City of Columbus is matching the USDOT and Vulcan grants with more than $360 million in pledges from public and private sector partners.

OUTCOME: A SAFER, MORE MOBILE AND SUSTAINABLE CITY

Columbus will become the nation’s epicenter for intelligent transportation systems (ITS) research to improve safety, enhance mobility, create ladders of opportunity for those who may have been left behind in the past, and reduce emissions.
Common Payment System

Multi-Modal Trip Planning Application

Smart Street Lighting

Pedestrian Collision Avoidance

Smart Mobility Hubs

Mobility Assistance for People with Cognitive Disabilities
Connected Electric Autonomous Vehicle

- Electric Autonomous Transit Shuttle
- Automated Routes
  - COTA to logistics
  - Transit center to corporate campus
  - Remote parking to retail center
## 15 Projects (USDOT Grant)

### ENABLING SYSTEMS & APPLICATIONS

<table>
<thead>
<tr>
<th>CCTN</th>
<th>IDE</th>
<th>DISTRICTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Smart Street Lighting</td>
<td></td>
<td>6. Multi-Modal Trip Planning Application</td>
</tr>
<tr>
<td>3. Transit Pedestrian Collision Avoidance System</td>
<td></td>
<td>7. Smart Mobility Hubs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8. Mobility Assistance for People with Cognitive Disabilities</td>
</tr>
</tbody>
</table>

### Smart Columbus Electrification Plan’s (Vulcan Foundation) 5 Elements

- **RESIDENTIAL**
  - 5. Common Payment System
  - 6. Multi-Modal Trip Planning Application
  - 7. Smart Mobility Hubs
  - 8. Mobility Assistance for People with Cognitive Disabilities

- **COMMERCIAL**

- **DOWNTOWN**
  - 10. Delivery Zone Availability
  - 11. Enhanced Permit Parking
  - 12. Event Parking Management

- **LOGISTICS**
  - 13. Truck Platooning
  - 14. Oversize Vehicle Routing
  - 15. Interstate Truck Parking Availability
The IDE is a web based dynamic governed platform at the heart of the Smart Columbus data environment that integrates data and data services from multiple sources and tenants, including the planned Smart Columbus technologies and traditional transportation data. The IDE embodies an open-data approach using best of breed technologies, including both open-source and commercial off the shelf components to enable better decision-making and problem solving for all users to support a replicable, extensible, sustainable platform for data ingestion and dissemination. The IDE drives performance metrics for program monitoring and evaluation.
The IDE will serve the needs of internal and external users.

**Internal**
- City Wide Collaboration & Sharing
- Visualization
- Analytics
- API Access
- Data Files
- 3rd Party

**External**
- City Public Engagement & Value Creation
- Citizens
- Journalists
- Researchers
- Developers
- NGO’s
- Vendors
- Governments
- IoT Sensor Fabric
- Data Files
- Applications Search
- Governments

**Categories**
- IoT Sensor Fabric
- Data Exchange
- Visualization
- Analytics
- API Access
- Data Files
- 3rd Party

**Domains**
- Public Safety
- Public Service
- Building & Zoning
- Development
- Utilities
- Culture & Events
- Finance
- Health
- Government
- Neighborhoods
- Mapping
- Recreation & Parks
- Culture & Events

**Logos**
- American Electric Power
- Nationwide
- Columbus Collaboratory
- Battelle
- Battelle
- Columbia University
Integrated Data Exchange: Nascent to Exemplar?

Starting Premise: We don’t know what we don’t know
   **BLUF:** Other governments are ahead of us, we need to learn from them

Understanding Data Management: Are There Reference Architectures?
   **BLUF:** It isn’t just about a place to store data

Seeking Exemplars: What Can We Learn From Other Governments?
   **BLUF:** Data quality and governance matter

Open-Source: What Can We Learn From Proprietary Solutions?
   **BLUF:** Nobody has all the answers, but architectures are similar
Challenge: Current Data Environment

- No Catalog of Data Sources
- Point to Point Integration is Messy
- Data is Not Shared
- No Universal Data Quality Processes
- Data Tethered to Applications
- Change is Risky, Expensive and Time Consuming
Cities As Data Owners

• Information Technology in organizations has evolved primarily as a means to:
  • Automate business processes
  • Enable better decisions
• For most “users” of IT systems, the data is the “application”
• Cities have massive amounts of data within their information technology systems
• City of Columbus has approximately 500 software applications and Terabytes of data
Lessons Learned: Data Management Architecture

INTEGRATED DATA
- Master/Reference Data (meta)
- Enterprise DW
- Operational data stores

DATA LAKE
- Structured & Unstructured Data

BUSINESS BRAINS & ANALYTICS
- Query Engine
- Semantic Modeling
- Security
- APIs

DATA VISUALIZATION
- ENGAGEMENT LAYER

ARCHIVAL

DATA VISUALIZATION

ENGAGEMENT LAYER

APPLICATIONS

RESEARCH
Lessons Learned: Treat Data as an Asset (Governance)

The City’s data is an asset. Like any physical asset it must be properly managed and maintained by people.
Road Map: IDE Initial Development Plan
Thematic Targets (2017)

Product Vision
• January

Platform Framework
Theme 2
• August

Governance & Conceptual Architecture
Theme 1
• March

Data Exchange
Theme 3
• December 2017

Retrospective and Year 2 Plan
• January 2018
IDE Development (as of June 2017):

Theme 1: Governance & Conceptual Architecture
2-Day Workshops, User Stories: 14, Chores: 4, Artifacts: 32

Theme 2: Platform Framework
2-Day Workshops, User Stories: 16, Chores: 10, Artifacts: >20

Theme 3: Digital Exchange
Theme 1 Artifacts

- Architecture
- Evaluation Criteria
- Data Publishing

Process

The publishing process is initiated by a user (or tenant) who has access to the publishing process API's or forms.

010 – Form or API will have requirements that ensure a dataset package has the proper “ingredients” for processing. This includes metadata, ownership information, well-formed schemas and structure, sample data, quality parameters, etc.

020 – The dataset package is processed and elements are recorded in the IDE database.

030 – The IDE connects to the data source and begins to harvest the data into the IDE. This could be an API endpoint, or an uploaded static file for example. At this point the data is in its original format from the source.

040 – Metadata is processed and updated in the IDE register.

050 – Data enrichment and improvement processing will improve the quality of the data where possible. This could include cleaning, reformatting, or aggregating data.

060 – The data is then stored and/or buffered in the IDE. For example pass-through data would be buffered from the source with its IDE enrichments, this will improve performance of data access to users.

070 – API Access to the data can expose the data to other platforms without the need to put the data on the public data portal. This data pipe is particularly useful when routing between machines and applications where human readable parsing isn’t required.

080 – Data to be made publicly available through the portal is then released onto the IDE Public Portal.

* It is yet to be determined how periodic data set releases will be handled. There are both technical and policy questions involved. There are several basic approaches, including:
  - Replacement
  - Appending
  - Versioning
  - Updating in place
Theme 1 Artifacts
- Data Stewardship
- API Framework
- Data Classification
- Data Usability Rating
- Reference Architecture
Theme 2 Artifacts:

Three Tracks:
- Open Data Portal
- Subscribed API’s
- Data Management Platform

- Deployed Open Data Portal Framework
Theme 2 Artifacts

- Open Data Portal (CKAN)
  - Data available in different formats
  - Multiple tenants
### Extending Engagement: Work Group Matrix

<table>
<thead>
<tr>
<th>Plan</th>
<th>Build</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
<tr>
<td>City</td>
<td>Vendor</td>
<td>Partners</td>
</tr>
</tbody>
</table>

- **Data Governance**
- **Data Architecture**
- **Data Asset Development**
- **Data Quality**
- **Data Analytics**
- **Cyber Security & Privacy**
- **Infrastructure**
- **Business Management**
- **Resource Dev & Education**
- **Community Engagement**
- **Data Sharing Policy & Laws**
- **User Experience**
If you want to change outcomes, you need to realize that outcomes are the result of systems. Not the computer systems, but the way people work and interact. And these systems are the product of how people think and behave. So, if you want to change outcomes, you have to change your systems, and to do that, you have to change your thinking.

*John Morgan*