Mid-Ohio Regional
Planning Commission

A REGIONAL DATA SITE FOR CENTRAL OHIO

Outcomes of the Regional Data Task Force

OCTOBER 2015

TABLE OF CONTENTS

EXECUTIVE SUMMARY 03
THE TASK FORCE 04
FINDINGS 05
NEXT STEPS 16
GLOSSARY 20
REFERENCES AND LINKS 21
APPENDICES 22

EXECUTIVE SUMMARY

INTRODUCTION

We look to data to reveal issues and help us find solutions. With so much data available, how can we be certain we are using legitimate information to inform decision makers, or even make our own decisions? Danger lies in using inaccurate data, and easing access to sound, reliable information can assist analysts, academics, business and government in their work.

FORMATION OF A REGIONAL DATA SITE TASK FORCE

The Mid-Ohio Regional Planning Commission (MORPC) formed its Regional Data Task Force to make recommendations to MORPC leadership and members about creating a regional data site. The task force met in 2015 from February through October. It consisted of members from a broad set of disciplines representing business, government, social services and academia.

THE WORK OF THE REGIONAL DATA SITE TASK FORCE

The task force determined that for the site to succeed, there must be collaboration for engaging the data community to identify and find solutions to problems that face the Central Ohio region. Moreover, the site should become a branded single point of entry, shared and supported by key stakeholders with MORPC functioning as the coordinating agency. The task force was challenged with four tasks:

Task 1: Affirm or deny the demands and benefits of a regional data site.
Findings: There is a need and a benefit to create a regional data site in Central Ohio.
Easy access to information can assist analysts, academics, business and government in their work. Increased awareness of ongoing analyses can lead others to value-added opportunities. The site could be a catalyst for entrepreneurialism and addressing social issues.

Task 2: Outline a high-level structure and layout for an online data site.
Findings: The site should include a variety of tools to encourage collaboration and coordination.
Include a data catalog, opportunities for public interaction, ready-made maps and tables, and a platform for sharing analyses.

Task 3: Recommend a governance structure for continued maintenance and operation of the site.
Findings: Governance must have clear responsibilities, broad representation, and be well organized.
MORPC should act as the lead agency for coordination. A governance group should be multi-disciplinary and charged with overseeing the management and funding of the site. Technical support and guidance should include input from a variety of analysts and users.

Task 4: Recommend a business strategy and financial plan to ensure sustained success of the site.
Findings: Funding is needed for both building and supporting the maintenance of the site.
Solicit initial funding from a philanthropic foundation to construct the site. Continue support and expansion of the site through key stakeholders' participation.

ACTION STEPS FOR MOVING FORWARD

- Engage with the data community to build momentum
- Create a governance group
- Build a data catalog
- Develop the site and identify ongoing support
- Determine feedback and success measures

THE TASK FORCE

COMPOSITION

The task force was composed of members from a broad set of disciplines including representatives from city government, health, economic development, social service, academia, research, local business and public utilities. The group was chaired by Anthony Jones, Director of Planning and Development for the City of Gahanna and MORPC Commission member, and Jung Kim, Managing Director of Research and Business Intelligence for Columbus 2020. The full list of task force members is included as Appendix A.

TASKS

The task force had four main tasks:

Affirm or deny the need for a data site.
The goal of this task was for the group to educate itself about current data needs and data offerings in Central Ohio and determine whether a regional data site would provide value to the region.

Outline a high-level structure and layout for an online data site.
This includes identifying tools that the site might contain and recommending general data categories and other content. Tools could include search and download, ready-made visualizations and community engagement features.

Recommend a governance structure for continued maintenance and operation of the site.
Governance means establishing an agreed-upon organizational structure for overseeing the operations of the data site, setting priorities and securing funding.

Recommend a business strategy and financial plan to ensure sustained success and operation of the site.
For this task the goal was to identify both short- and long-term funding strategies – short-term for the initial creation of the site and long-term strategies to ensure sustained success of the site. This included identifying specific cost ranges based upon similar efforts and other research.

The task force met for six months between February and July in 2015 to address these tasks.

DELIVERABLES

The deliverable from the task force is this document – an action plan based upon the findings from the tasks above.

FINDINGS FROM THE TASK FORCE

TASK 1: AFFIRM OR DENY THE NEED FOR A REGIONAL DATA SITE

The task force members felt it was important to first reach consensus that a regional data site would indeed provide value to Central Ohio before working through the details of how to manage a data site and what it may contain. Defining the demand for the site and questions about its purpose and clarification of the audience for whom it should be designed influenced discussions about site features, governance and funding.

Some of these challenges, as identified below, were easy to answer while others are considerations that will need to be considered as the site evolves.

Defining a data site.
A data site is a public website designed specifically for finding and retrieving data. A data site can exist as a standalone website (e.g. data.gov) or as part of an existing website (e.g. the Mid-Ohio Regional Planning Commission's Data, Maps and Tools section found within the MORPC website). The data site as discussed in this paper is a standalone website for finding and retrieving data for Central Ohio.

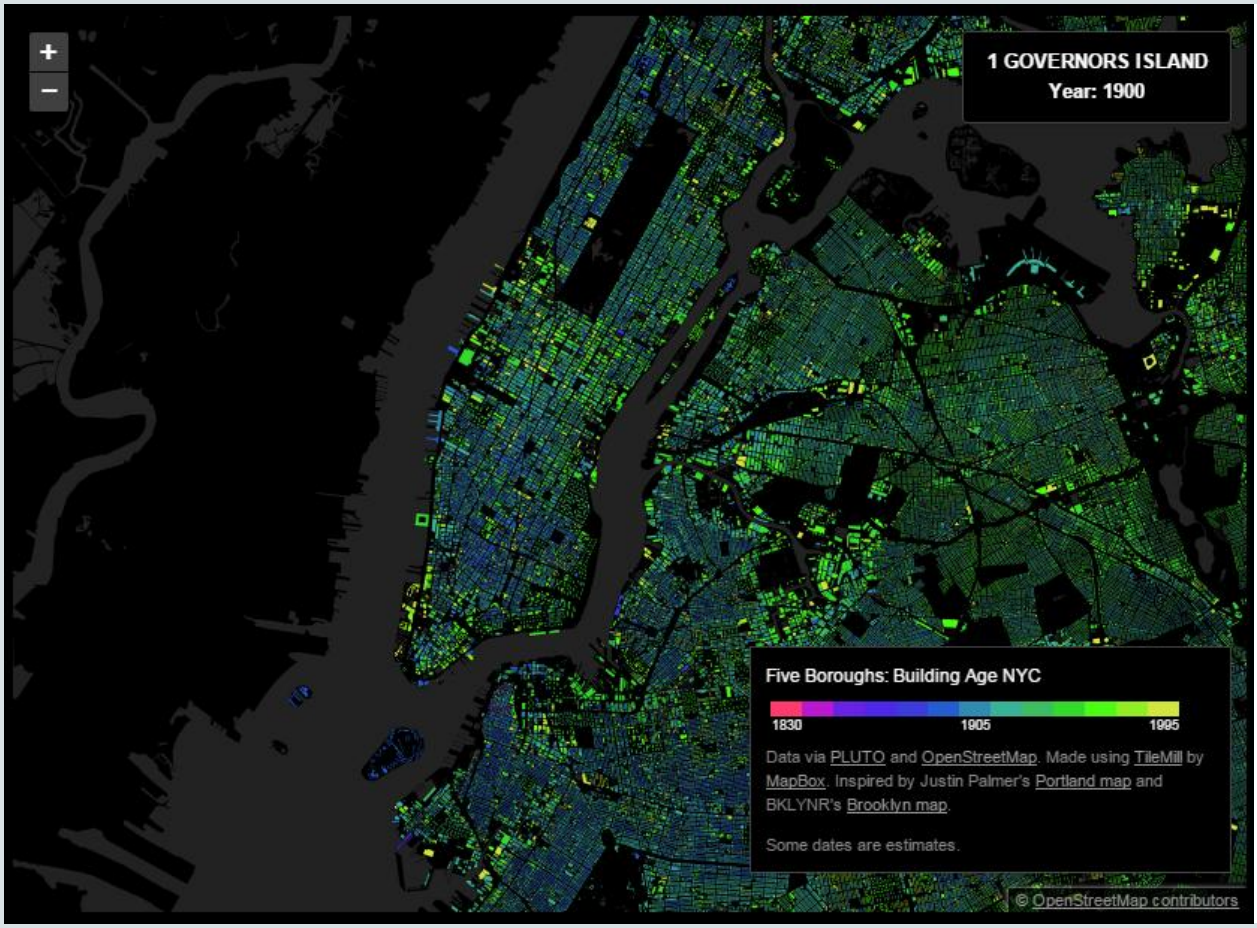
What is Central Ohio?
There are a few definitions of Central Ohio, but Central Ohio in this case refers to a 15-county region that includes the following counties: Delaware, Fairfield, Fayette, Franklin, Hocking, Knox, Licking, Logan, Madison, Marion, Morrow, Perry, Pickaway, Ross and Union.

Proving value.
How do stakeholders know that a data site is valuable to them? Though it may be a challenge, measures should be defined that prove the tangible value of the site.

Who is the audience for the data site?
This question came up repeatedly. Some task force members felt that the site should be developed for solving specific problems for a specific audience, and this approach is being recommended by the task force as a way to focus momentum and to support an application for grant funding for initial development. However, as potential users could be anybody, the task force generally agreed that the site should be as inclusive as possible. Specific audiences may self-identify over time once they use and realize the value of the site.

Maintaining relevance and vibrancy.
Maintaining a data site, from website development through data processing, requires a lot of work. A sustainable approach would involve finding ways of sharing the load between data providers and site maintainers. A single agency should not be solely responsible for all aspects of a large-scale data site.

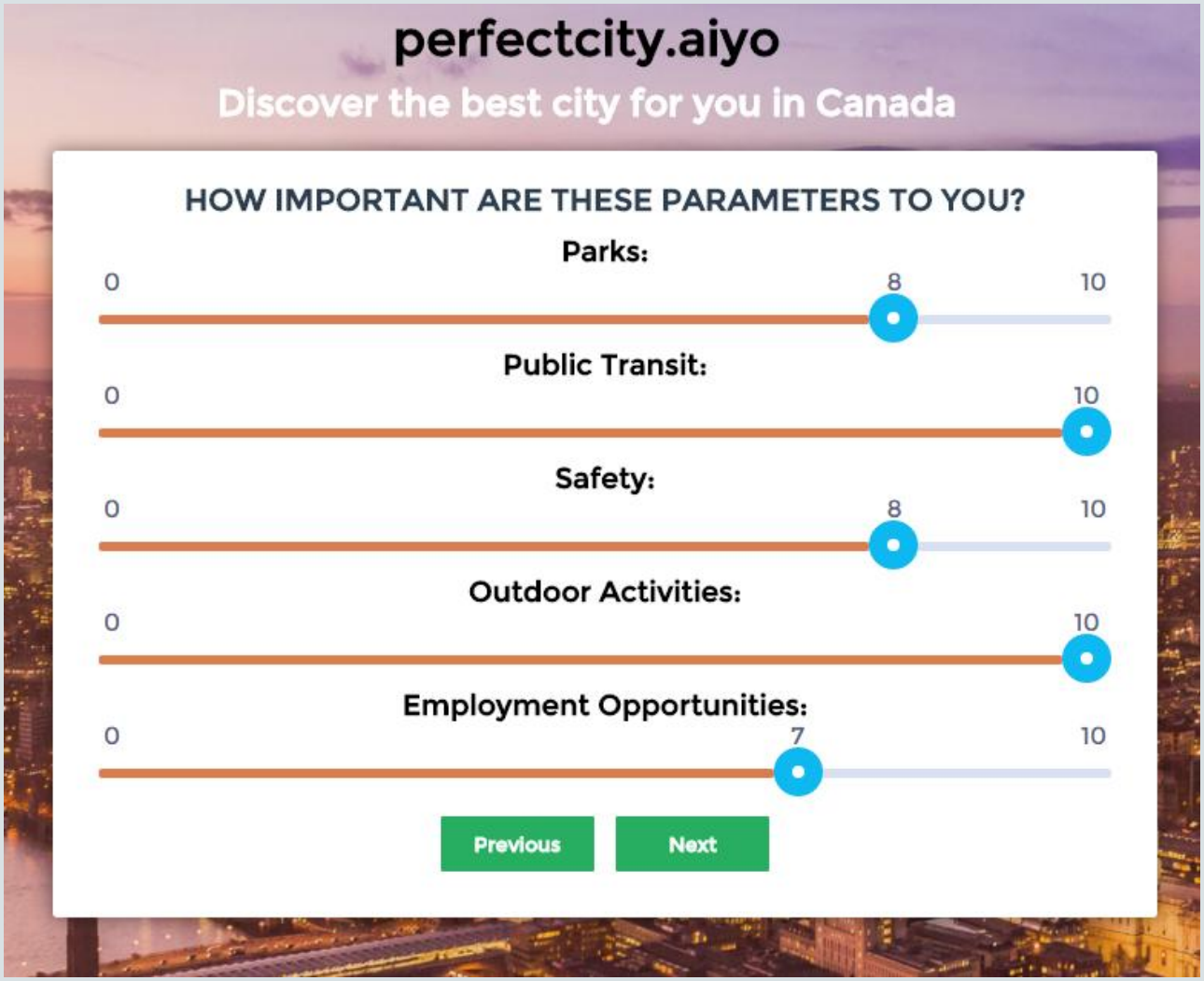
Making data attractive.
How do you get the general public excited about data and drawn to a data site? One idea is to show what others have done for inspiration and to show what is possible with the data. The following are a couple of examples from other regions:



EXAMPLE 1 - BUILDING AGE NYC

Building Age NYC visualizes New York City building age using a dataset that was made available to the public in 2013.

http://pureinformation.net/building-age-nyc



EXAMPLE 2 – PERFECTCITY.AIYO

This is an example where entrepreneurs used a number of open datasets to create a way for people relocating to Canada to find which cities may suit them best.

http://aiyotech.github.io/perfectcity.aiyo/index.html

Learning from the past.
There are few examples of regional data sites. Most data sites reviewed through the task force were from a single entity. While there were several instances including multiple departments, the sites still fell within a single jurisdiction. That is not to say regional sites should not exist, but they require commitment to be successful. Perhaps the most notable local example is DataSource2.0, made in partnership between MORPC and Community Research Partners (CRP). DataSource2.0 was built around a powerful web-based data visualization tool developed by an academic institution with funding from a consortium of which MORPC and CRP were part. The development and maintenance of DataSource2.0 provides insight for subsequent initiatives. While the visualization tool was powerful for technical applications, a more user-friendly interface was not delivered and required MORPC and CRP to create a custom solution. It required an extraordinary and unexpected amount of resources that placed a heavy burden on both organizations. Success was also hindered because it was developed without a clear understanding of desired features and data.

Impact on data providers.
There are numerous impacts on data providers to consider when planning and designing the data site. Data collection and processing, education and outreach efforts, meeting data sharing standards, maintaining data sharing agreements and keeping data up-to-date all place demands on data providers. On the other hand, participating in an organized data-sharing effort may help identify efficiencies and opportunities for streamlining data workflows. Coordination among data providers is critical to assure that impacts are kept to a minimum. Perceived unfunded mandates on participants could result in reduced participation.

Background on Research

Initial group discussion covered a broad range of ideas, including the possibility of collaboration with data initiatives at The Ohio State University, how to provide value to site stakeholders and ways of involving the community to find and document sources of data.

A number of non-local data websites were reviewed at the first meeting, with some discussion about how they were managed and funded and what motivated their creation. The group asked MORPC to further investigate some sites in order to learn more about the origins of the sites and the costs involved.

Additionally, after the first meeting, MORPC staff interviewed Columbus and Franklin County Metro Parks, the Columbus Metropolitan Housing Authority and the Columbus Metropolitan Libraries to explore the data needs of organizations in the region in more depth.

Findings
The group generally agreed that a regional data site could provide value to Central Ohio.

There is a demand for data.
This sentiment is reflected in the plans and documents from many other data site initiatives. MORPC's Data, Maps and Tools section on its website is the most popular portion of the website, but it is not comprehensive.

Data sites exist – and for a reason.
Data site initiatives are numerous, both in the nation and worldwide. This larger movement around data has been underway for at least a decade, driven by at least a few different motivations, such as transparency, better customer service and solving societal problems.

Data sharing is already a common practice.
Many agencies already share data, and while some datasets are sensitive and need special treatment, many standard datasets could be shared in a more streamlined way. A tool such as the data site could make data sharing easier.

Data is everywhere, but knowing what is reliable is difficult.
Data providers likely duplicate efforts around data, but how do they know what each other is doing? Data on some websites lack citations, and data in newspaper or reports may be just a snapshot in time and become quickly outdated. The casual browser may need to do extra work to find reliable sources.

Regional Data Sites are rare.
To be successful there must be buy-in from a variety of users. One single entity cannot bear the burden alone and meet the expectations of the community at large.

Regional data initiatives should be coordinated.
The site can benefit value-added analysis by reducing time and resources spent on data collection.

Engage citizens.
Citizen engagement was identified as an important element of informing which features and data are available on the data site. Community participation features as recommended later in the document include public feedback tools as a way of incorporating community members' ideas into the development of the data site. Additionally, efforts should be made to market the site and to educate the public about the data movements happening around it, to explain concepts like "open data," "big data" and "data sites," and to help it realize the value of a data site.

Engage the technical and entrepreneurial community.
The regional data site could benefit entrepreneurial efforts of technologists by having curated, accurate and current datasets available. Successful ventures using public datasets already exist. Examples include companies using climate data to help farmers with crop production, creating value-added educational data to help students choose educational paths, and augmenting real-estate data with school, walkability, crime rates and other public datasets to provide better real-estate decision-making.

TASK 2: OUTLINE A HIGH-LEVEL STRUCTURE AND LAYOUT FOR AN ONLINE DATA SITE

This task was intended to help organize thoughts about what a regional data site might contain in terms of functionality and general information.

Background on research
The task force reviewed a list of national, regional and local sites and rated them on a number of characteristics, such as ease of use, comprehensiveness and searchability. A summary of the review is available in Appendix B.

This exercise informed the group about the types of tools and data being made available by other entities providing data sites. Ratings and comments helped to capture what items were most important from the group's perspective.

In another exercise, group members explored data needs and problems in breakout sessions. Specifically, they discussed the kinds of problems they would like to solve with data, trouble they had finding data and datasets they would like to have available.

Both of these exercises guided the following recommendations regarding site features and datasets:

Findings
Based on the exercises and information presented, the task force came to the following conclusion regarding the site features and content:

Recommended site features

- **Ready-made data and simple visualizations.** Charts, tables, maps.

- **Data catalog.** With robust search capabilities and export options.

  The Task Force learned that the Ohio Geographically Referenced Information Program (OGRIP), a geographic data program for the entire state of Ohio, has committed to using the Esri software platform – specifically the Esri Open Data component of the ArcGIS Online product, to provide a data catalog of public geographic data. Using this product allows opportunity for coordination because many public agencies already have access to the Esri platform.

- **Community participation elements.** This includes mechanisms for users to recommend datasets and to showcase how the community, organizations and 'civic hackers' are making good use of data from the site.

- **Help forum.** A value-added approach to data, this component would allow site users to approach data experts and other data users to get help with deeper data analysis.

Recommended datasets

- Place-based economic data
- Regional infrastructure
- Health
- Weather
- Socioeconomic
- Environmental
- Geographic data (GIS)
- Crime and safety
- Economic

TASK 3: RECOMMEND A GOVERNANCE STRUCTURE FOR CONTINUED MAINTENANCE

Good governance of the regional data site will ensure that it is trustworthy, its operations and processes are well-organized, and that it is funded properly.

What does a governance group do?

- **Create guidance.** To enable easy participation in the data site, the group recommends data standards and clearly outlines expectations for involvement with the site.
- **Incorporate public feedback.** To keep the site relevant and useful, the governance group includes feedback from site users to help the site evolve.
- **Set timelines.** The group sets appropriately ambitious timelines for site development and maintenance activities.
- **Create data quality processes.** As best as possible, the group decides and evaluates minimum quality standards for data being published to the site.
- **Explore partnerships.** The group explores public/private partnerships and other collaboration opportunities to bolster funding and create community around using data.
- **Set and review site success measures.** The group determines clear and obtainable measures that help it know if the site is meeting expectations.
- **Identify incentives for participation in the site.** Participating in the data site either as a maintainer or data publisher requires resources, and the group will help make the case for the return on investment for those who participate.
- **Ensure adequate funding.** The group identifies and explores funding opportunities and strategies for creating and continued operation of the site.
- **Prioritize datasets.** In order to provide the most relevant data, the group will evaluate data needs in the region to prioritize the release of datasets on the site.
- **Solicit and encourage new data for the site.** Some datasets may not yet be created or available, so the group will lobby for new and useful data from the public and potential data providers in the region.

Background on research

Recommendations for governance were informed by a few sources.

The Sunlight Foundation offers a comprehensive guide for many aspects of open data initiatives, and its recommendations are largely relevant to the current initiative. In its Open Data Guidelines document[1], the Sunlight Foundation suggests in guideline 24, *Create or Appoint Oversight Authority*:

> "Specifying an authority, review board or similar body is an important step to making sure that an open data policy can be executed and provides a resource to address unforeseen hurdles in implementation. Oversight bodies should conduct their work independently and publicly …"

The Sunlight Foundation guide was referenced upon a recommendation of Ohio State University Professor David Landsbergen during his presentation to the task force about his open data research.

Lastly, the primary developer and maintainer of OpenDataPhilly, Robert Cheetham, offered information about his experience and likewise recommended appointing a cross-disciplinary and perhaps nonprofit group to oversee a regional data site.

Findings

The task force concluded that the data site should be governed by a multidisciplinary group that can represent the interests of multiple sectors of the region, but have one lead agency leading the charge.

- MORPC should function as the lead agency. It should initiate the formation of a new standing Regional Data Advisory Committee organized under the administrative by-laws of the agency. Representation should be from key interests that include health, government, business, academia and socially oriented non-profits.

- There is a need to engage technical experts as well. One possible forum would be an existing group of experts that meet regularly as part of the Central Ohio GIS User Group that is coordinated by MORPC. This forum could be expanded to include other data experts. This forum would provide technical guidance to the governance group.

- The Regional Data Advisory Committee should be responsible for driving funding decisions and soliciting financial support.

1 *http://sunlightfoundation.com/opendataguidelines*

TASK 4: RECOMMEND A BUSINESS STRATEGY AND FINANCIAL PLAN TO ENSURE SUSTAINED SUCCESS AND OPERATION OF THE SITE

There are short-term costs associated with building the site and long-term costs associated with its maintenance. Building the site can be considered the easy part, but keeping it going requires buy-in from stakeholders. There will be demands on personnel to keep the site current, maintain strong coordination and make necessary upgrades to software supporting it.

The task force framed discussions of costs to build and maintain the site into three scenarios:

1. Subscribe to a data site development and maintenance service.
2. Seek foundation support to hire a consultant to build the site (possibly using open source software such as CKAN) and then maintain it in-house.
3. Build and maintain the site completely in-house.

Background on research

Most of the websites reviewed while exploring site features and content were built using pre-existing software with some level of customization. Three common products used by sites reviewed are:

- **Socrata.** Used by a few of the sites we reviewed. NYC Open Data is a prominent example.
- **Esri Open Data.** Used by Charlotte Open Data and the State of Ohio's Geographically Referenced Information Program (OGRIP).
- **CKAN.** Used by OpenDataPhilly.

These all aim to streamline the process of creating data sites, particularly around a data catalog concept. Besides making data available and searchable, they provide features for data providers and usually offer basic data preview and visualization tools.

Socrata and Esri Open Data allow limited customization, while the free and open source CKAN platform is highly customizable, albeit most likely with a significant cost in development time. Some pros and cons of each approach are listed in the following table:

Development approach	Pros	Cons
Software subscription (Socrata, Esri)	• Company support • Some features already built • Updates, fixes and upgrades done by the company	• Less flexibility for custom features or site design • Costs can be very high • May require all participating organizations to have their own subscriptions – high cost
Consultant created or customized site (e.g. customized CKAN solution)	• Consultant has existing expertise • Customizable	• New features and customizations (scope changes) could be costly • Eventually must be owned by a sponsoring organization
In-house development (e.g. customized CKAN solution, or entirely custom web application)	• Complete control of design and features	• Unpredictable costs • Must have development expertise in-house

Other communities that have data sites were consulted to get estimates of costs they incurred for their data site activities. Minneapolis' regional planning group reported that it spent $30-40k on a customizable map tool alone. The primary developer whose design firm created the first OpenDataPhilly site had more insight to offer. Their approach was to customize the open source solution CKAN to meet their needs. They made an initial version of the site for approximately $50k, and then found that it didn't easily accommodate new needs identified by users who provided feedback after its release. This required a reworking of the site, with the total cost (not including pro bono work by the design agency) estimated to be $150k.

The subscription-based Socrata platform was also explored, given its popularity. The subscription cost at the time of the conversation was $3,000 - $5,000 per month for site development and subsequent maintenance. The Socrata process involves a startup phase during which time they create a customized home page for the site, but additional customizations are not supported. An additional complicating factor for a service like Socrata is that they do not support a multi-organizational approach that allows multiple data providers to use the same subscription, but rather require each to have their own subscription.

Esri is a mainstream software used by most GIS offices. The Esri Open Data tool is included in the licensing cost of ESRI products. At this point, Esri Open Data will not meet all of the expectations of the identified features on the complete site. For example, functionality regarding open dialogue among the data community is not part of the off-the-shelf software. However it will accommodate the data catalog and sharing visualizations elements on the site. While communities will not be required to purchase an Esri license to use the site, those without it will require assistance from an organization that does have an Esri license to post their information into the data catalog. Yearly licensing costs for the full Esri product vary widely, ranging anywhere from less than $3,000 to $50.000 and higher, depending on the size and type of the organization buying the software and whether there is a need for advanced features.

Findings

Scenario 1 was thought to be too expensive and didn't allow for flexibility. Scenario 3 was dismissed based on previous experience. For example, the implementation of DataSource by MORPC and Community Research Partners revealed that it is not feasible for one, or even two, organizations to build a site of this scale within the confines of existing workloads and budgets. Scenario 2 suggests an approach that will allow us to modify the site to meet future expectations without overburdening available resources.

- It was generally agreed to adopt Scenario 2. While Esri Open Data will be used for building the data catalog, the cost of developing the site is expected to be a minimum of several hundred thousand dollars. This estimate is based on conversations with other communities. OpenDataPhilly's example set the cost of initial development at $150k, not including pro-bono development contribution.
- The preferred short term strategy to build the site includes making application to a foundation around a specific data-driven problem that the governance group recommends.
- Long term strategies to maintain the site should include:
  - Pay-to-play by getting line-items in stakeholder budgets
  - Allowing sponsorships from experts advertising their services via the site

NEXT STEPS

The following are next steps to move forward with the creation of a regional data site. Generally, these are aimed at establishing governance, securing funding, determining success measures and digging deeper into data needs for the region. It is understood that MORPC is recommended to coordinate these efforts.

These steps are considered only a first phase of data site development. The Ohio Department of Administrative Services Open Data report[1] recommends an iterative and incremental approach to providing public data, and this initiative adopts the same principle. Future datasets, tools and revisions should happen in stages based upon reviewing past experience, success metrics and public feedback.

A timeline for next steps follows at the end of this section.

ENGAGE WITH THE DATA COMMUNITY TO BUILD MOMENTUM AND STRENGTHEN THE APPLICATION FOR DEVELOPMENT FUNDING

Formalize a data subgroup at the Central Ohio Geographic Information Systems (GIS) User Group (COGUG). COGUG is a quarterly meeting of regional geographic information systems professionals, mostly from local governments. This group deals primarily with geographic data, but many of those attending are experienced data users who have insights into providing, presenting and processing data.

COGUG meetings include subgroup breakout sessions, where the attendees speak in-depth about particular topics in smaller groups. Attendees are free to move between subgroups from meeting to meeting, so formalizing this group involves getting attendees to commit to regular participation in the data subgroup.

Additionally, in order to represent perspectives from other interests in the region, other regional members should be identified and invited to participate in the data subgroup.

Coordinate with other regional data groups. This includes groups at The Ohio State University, such as the Battelle Center and Kirwan Institute, and the State of Ohio's Ohio Geographically Referenced Information Program (OGRIP). MORPC has also worked with Community Research Partners on past data initiatives and will continue that relationship through the current initiative.

These groups should be contacted and discussions should be initiated to identify data sharing and other collaboration opportunities. Opportunities should be incorporated into development plans as appropriate.

Connect with civic hackers. This includes the local tech community and perhaps a national group like Code for America. Key contacts should be identified and contacted to initiate discussion about the involvement of these groups in the development of the data site.

1 *http://das.ohio.gov/Portals/0/DASDivisions/InformationTechnology/IS/Optimization/Open Data Report March 2015.pdf*

CREATE A STANDING REGIONAL DATA ADVISORY COMMITTEE AT MORPC

The creation of the Regional Data Advisory Committee should be facilitated by MORPC, and the committee will serve as the governance group. Key stakeholders from the region should be identified and invited to participate. Once in place, this group will have actions to take, including:

- Providing guidance for obtaining funding for initial development
- Establishing strategies for long-term funding
- Administering a survey of potential data providers and users in the region to further clarify the data and tools they need and what benefit they could get from a regional data site
- Defining branding and marketing strategies

BUILD A DATA CATALOG

A recommendation from the Ohio Department of Administrative Services Open Data report[1] that is also a task force recommendation is to create an initial data inventory and data catalog as a starting point for providing data to the public. A data catalog was identified as one of the desired data site features in Task 1. Specific actions for building a data catalog, to be undertaken by the governance group, include:

- Creating a data inventory and a roadmap and timeline for release of datasets
- Recommending minimum data standards for data providers

MORPC has secured funding to contribute to the creation of a data catalog. Existing commonly used datasets from MORPC (and possibly others) could be added to the data catalog as a first release. In the short term, this catalog can be on MORPC's website and linked to other stakeholders sites to assist in branding and building awareness.

DEVELOP THE SITE

To secure funding, a high-value data project should be explored that will clearly show the worth of a regional data site. This exemplary project will be used to apply for a foundation grant for initial development.

FEEDBACK AND SUCCESS MEASURES

As mentioned in the introduction to this section, success metrics and public feedback should guide the development of the data site. For this first phase:

- Feedback mechanisms for users should be included in the initial site design
- The governance group should establish additional communication strategies
- The governance group should define success measures for the site

1 *http://das.ohio.gov/Portals/0/DASDivisions/InformationTechnology/IS/Optimization/Open Data Report March 2015.pdf*

ACTIONS AND TIMELINE

Recommendation	Actions/Notes	Who Does it	Time Frame
Accept MORPC as the Lead Agency	• Present Task Force Findings and Recommendations • Endorse creation of Regional Data Advisory Committee	MORPC Executive Committee MORPC Board of Commissioners	November-December 2015
Engage with Data Community	• Several other regional data efforts were identified during this process. Reach out to those managing those initiatives. Identify ways of consolidating efforts. Examples: OGRIP; OSU — Battelle Center, Kirwan Institute, Community Research Partners • Build relationships with local entrepreneurial groups • Explore partnership with Code for America. • Add firmer structure to the Data Subgroup of Central Ohio GIS User Group	MORPC Staff	Now
Implement Data Catalog for Central Ohio	• Use Data Subgroup of Central Ohio GIS User Group • Coordinate with OGRIP • Pull together data experts to establish minimum metadata requirements	MORPC Staff	• Open discussion at the November 2015 User Group Meeting • Framework established by end of 2015 • Quarterly communication
Establish Regional Data Advisory Committee	• Invite constituents from private business, academia, nonprofit, healthcare, local government, tech community • Select chair	MORPC Leadership	January- March, 2016

Establish Committee's purpose, role and responsibility	Define mission, goals, and responsibilities • Rules of conduct and by-laws • Short- and long-term funding needs to construct and maintain the site • Establish policy for participation • Branding and marketing strategies • Build ladders for analysis • Funding solicitation plan • Contingency plan should foundation-based funding solicitation not succeed • Define success measures	Data Advisory Committee MORPC Staff	January-March 2016
Wireframe site design	Basis for developing an RFP to build the site • Drawings of site structure • Drawings to represent user interaction with the site	MORPC Staff Data Subgroup	March 2016
Conduct a survey to identify audiences and answer the question: "What's in it for me?"	• Survey potential stakeholders throughout the region about their data needs, what data they provide and how a data site might benefit them • Discuss findings	Data Advisory Committee MORPC Staff Consultant	March-June 2016
Identify a 'hook' project to support building the site	Use the survey and other means to choose a topic to build the foundation request around	Data Advisory Committee	June 2016
Make application to funder for funding to build site	MORPC staff makes application for funding based on recommendation from the Data Advisory Committee	MORPC Staff	July 2016

GLOSSARY

Activity stream	In the context of this document and the proposed data site, an activity stream is an actively updated listing of activities around data, including new visualizations, software, data releases and events such as hackathons.
Big Data *Source: Adapted from the Open Knowledge Foundation Open Data Handbook*	A collection of data so large that it cannot be stored, transmitted or processed by traditional means. The increasing availability of and need to process such datasets (for example, huge collections of weather or other scientific data) has led to the development of specialized computer technologies, architectures and programming languages.
Central Ohio	There are a few definitions of Central Ohio, but Central Ohio in this case refers to a 15-county region that includes the following counties: Delaware, Fairfield, Fayette, Franklin, Hocking, Knox, Licking, Logan, Madison, Marion, Morrow, Perry, Pickaway, Ross and Union.
Civic Hacking *Source: Adapted from the Open Knowledge Foundation Open Data Handbook*	Building tools and communities, usually online, that address particular civic or social problems. Examples could be tools that help users meet like-minded people locally based on particular interests, report broken infrastructure to their local council, or collaborate to clear litter from their neighborhood. Community-level open data is particularly useful for civic hacking projects.
CKAN *Source: Open Knowledge Foundation Open Data Handbook*	An open-source software platform for creating data portals, built and maintained by Open Knowledge. CKAN is used as the official data-publishing platform of around 20 national governments and powers many more local, community, scientific and other data portals.
Data site or data portal	A website for finding and retrieving data. Common features include browsing and searching a data catalog, data preview and visualization, publishing tools for data providers and tools for other computerized data tools to access data (Application Programming Interfaces, or APIs).
Hackathon *Source: Adapted from Wikipedia*	Also known as a **hack day**, **hackfest** or **codefest**, is an event in which computer programmers and others involved in software development and hardware development, including graphic designers, interface designers and project managers, collaborate intensively on software projects in competition with other teams. Some hackathons are intended simply for educational or social purposes, although in many cases the goal is to create usable software. Hackathons related to data often focus on creating software such as mobile phone apps using publicly available data.
Metadata	Data that describes and provides information about other data, including such things as the types of data, where the data originated and how often it is updated. Metadata is used to make finding and understanding data easier.
OGRIP	The Ohio Geographically Referenced Information Program
Open Data *Source: Adapted from the Open Knowledge Foundation Open Data Handbook*	Data is open if it can be freely accessed, used, modified and shared by anyone for any purpose - subject only, at most, to requirements to provide attribution and/or share-alike. Specifically, open data is defined by the Open Definition and requires that the data be **A. Legally open**: that is, available under an open (data) license that permits anyone freely to access, reuse and redistribute **B. Technically open**: that is, that the data be available for no more than the cost of reproduction and in machine-readable and bulk form.

REFERENCES

What	Where	What's in it
Ohio Department of Administrative Services open data report	http://das.ohio.gov/Portals/0/DASDivisions/InformationTechnology/IS/Optimization/Open Data Report March 2015.pdf	Recommendations for open data implementations in Ohio.
The Sunlight Foundation Open Data Policy Guidelines	http://sunlightfoundation.com/opendataguidelines/	Comprehensive open data implementation guidelines; also aimed at government.
Govloop Open Data Playbook	https://www.govloop.com/resources/the-open-data-playbook-for-government/	High-level guidelines for implementing open data, stories and interviews to provide examples
The Open Knowledge Foundation; Open Data Handbook	https://okfn.org/ http://opendatahandbook.org/	Another comprehensive guide; Value Stories - make the case for open data sites; open data resource library; describes various business models from other data site implementations.
The Open Data Institute	http://theodi.org	Guides, case studies, open data stories, UK-based.
OpenColorado Data Publishing Guide	http://opencolorado.org/publishing-guide/	Example of data standards for data collaboration.
Data Science for Social Good	http://dssg.io/	Examples of using data to solve specific societal problems.

APPENDICES

A. Task Force members
B. List of websites reviewed
C. Website review summary sheet
D. Presentation slides

APPENDIX A: TASK FORCE MEMBERS

Name	Email	Affiliation
Anthony Jones, Co-Chair	anthony.jones@gahanna.gov	City of Gahanna Development
Jung Kim, Co-Chair	jung_kim@columbus.org	Columbus 2020
William Murdock	wmurdock@morpc.org	MORPC
Andrew Brush	abrush@delawareohio.net	City of Delaware / City Council
Emily Buser	ebuser@columbuslibrary.com	Columbus Metropolitan Library
Lynnette Cook	lcook@researchpartners.org	Community Research Partners
Shoreh Elhami	shelhami@columbus.gov	City of Columbus / GIS
Megan Johanson	mjohanson@researchpartners.org	Community Research Partners
Devin Keithley	dkeithley@researchpartners.org	Community Research Partners
Bill LaFayette	bill.lafayette@att.net	Regionomics, LLC
David Landsbergen	landsbergen.1@osu.edu	Kirwan Institute, OSU
Ty McBee	tmcbee@adena.org	Adena Health System
Ernest Perry	eperry@handsoncentralohio.org	HandsOn Central Ohio
Sarah Poe	spoe@nisource.com	Columbia Gas of Ohio
Erin Prosser	eprosser@campuspartners.org	Campus Partners Community Relations
Jay Ramanathan	jayram@cse.ohio-state.edu	John Glenn School of Public Affairs (OSU)
Scott Sanders	ssanders@co.delaware.oh.us	Delaware County Regional Planning Commission
Scott Tourville	stourville@pickerington.net	City of Pickerington / City Engineer
Michael Wilkos	mwilkos@columbusfoundation.org	The Columbus Foundation / Community Research and Grants Management
Rob Powell	rpowell@morpc.org	MORPC, Task Force Coordinator
Nancy Reger	nreger@morpc.org	MORPC

APPENDIX B: LIST OF WEBSITES REVIEWED

Local/Regional/State Examples

MORPC Data and Mapping	http://morpc.org/our-region/data-maps-tools
Columbus 2020	http://columbusregion.com/Data-Reports.aspx
Columbus Dispatch Data	http://www.dispatch.com/content/sections/news/databases.html
DataSource	http://www.datasourcecolumbus.org
HandsOn Central Ohio	http://www.handsoncentralohio.org
Transportation Information Mapping System (TIMS)	http://tims.dot.state.oh.us/tims
Ohio Labor Market Information	http://ohiolmi.com/data.htm
Dublin GIS	http://dublinohiousa.gov/gis-maps
City of Columbus GIS	http://www.columbus.gov/Templates/Detail.aspx?id=68551
City of Columbus MyNeighborhood	http://myneighborhood.columbus.gov
Delaware County Regional Planning Commission	http://www.dcrpc.org
Delaware County Auditor	http://www.delco-gis.org/auditor
Franklin County Auditor	http://www.franklincountyauditor.com
Ohio State University Kirwan Institute Opportunity Mapping	http://kirwaninstitute.osu.edu/community-development-collaborative/

What Other Cities and Regions Have Done

Hartford Info	http://hartfordinfo.org/
Metropolitan Council (Minneapolis)	http://www.metrocouncil.org/Data-and-Maps.aspx
Minnesota Compass	http://www.mncompass.org/
Minnesota OpenData	http://www.state.mn.us/opendata/
MAPAS - Greater Grand Rapids area, MI	http://gis.cridata.org/maps/mapas
Denver Regional Council of Governments (DRCOG)	https://drcog.org/services-and-resources/data-maps-and-modeling
Baltimore Neighborhood Indicators Alliance	http://bniajfi.org
NYC Open Data	https://nycopendata.socrata.com/
California data	http://Data.CA.Gov
NEO CANDO (Cleveland area)	http://neocando.case.edu
Connecticut Data Collaborative	http://ctdata.org
Philadelphia Open Data	http://opendataphilly.org
Charlotte Open Data	http://clt.charlotte.opendata.arcgis.com
Austin Open Data	https://data.austintexas.gov
Seattle Open Data	https://data.seattle.gov
Portland (Oregon) Maps, GIS & Open Data	https://www.portlandoregon.gov/28130

National/Other

RAIDS Online	http://raidsonline.com/
American FactFinder	http://factfinder.census.gov/
Bureau of Labor Statistics	http://www.bls.gov/data/
Federal Reserve Economic Data	http://research.stlouisfed.org/fred2/
Bureau of Economic Analysis	http://www.bea.gov/itable/index.cfm

APPENDIX C: WEBSITE REVIEW RESULTS SUMMARY

MORPC Regional Data Task Force
Homework #1 Results Summary

Top 10 sites by score			
Rank	**Site**	**URL**	**Notable comments**
1	Seattle Open Data	https://data.seattle.gov	Mapping, search filters, some design
2	Austin Open Data	https://data.austintexas.gov	Simplicity, Data of the Month
3	Charlotte Open Data	http://clt.charlotte.opendata.arcgis.com	Download options
4	Philadelphia Open Data	http://opendataphilly.org	Nice search options, activity stream
5	NYC Open Data	https://nycopendata.socrata.com/	Nice categorization
6	Metropolitan Council (Minneapolis)	http://www.metrocouncil.org/Data-and-Maps.aspx	Make-A-Map
7	Connecticut Data Collaborative	http://ctdata.org	Datasets easily accessed, data gallery
8	Denver Regional Council of Governments (DRCOG)	https://drcog.org/services-and-resources/data-maps-and-modeling	Data catalog
9	American FactFinder (Census)	http://factfinder.census.gov/	Comprehensive, but needs more documentation
10	Minnesota OpenData	http://www.state.mn.us/opendata/	Map-based search, neighborhood-based option

Likes	Dislikes
Multiple download options / multiple formats of data	Stale data or links
Searchable, ability to filter	Choices too narrow
Comprehensiveness of datasets	Busy designs / clutter / too much text
Modern design / simple and clean design / intuitive interfaces	Interactive maps too busy or dysfunctional
Freshness / activity – seeing that data are being used – engagement	Too slow or site down completely
Help easily available	Login required

Mid-Ohio Regional
Planning Commission

111 Liberty Street, Suite 100
Columbus, Ohio 43215

T 614.228.2663

www.morpc.org