morpc

Mid-Ohio Regional
Planning Commission

IDENTIFYING CENTRAL OHIO DATA USER AUDIENCES

DECEMBER 2017

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Mid-Ohio Regional Planning Commission (MORPC) does work around data and mapping that is guided by a Regional Data Advisory Committee (RDAC)—a group of MORPC Commission members and other community leaders whose work is integrally linked to regional issues, and enhanced and supported by the development of cooperative strategies around regional data. The RDAC members have diverse data expertise—they represent different sectors (public, private, nonprofit, and academic) and geographic areas. The RDAC has posed the following question:

> *As data becomes more numerous and complex, more open, and easier (through technology advances) to deliver to more people in more formats, how do organizations with key roles around data deliver tools and resources that are practical, relevant and accessible to the people who need them?*

User personas are characterizations of key audiences or market segments that can be used to refine the development and marketing of a product or service. User personas are one component among a broader set of principles of User Experience (UX) Design. Other public organizations with roles around data are using UX Design tools to improve their effectiveness and efficiency in their work.

In late 2017, MORPC undertook a three-part research effort in order to create user personas that represent groups of Central Ohio data users. The three components of this research are:

1. **Peer Organization Interviews** We spoke with seven organizations with roles around public data about how they identify, engage, and build resources for their audiences. These conversations revealed some consistent challenges and aspirations, regardless of the size or location of the organization. The highlights (found on pages 3-5) emphasized a need for engagement, education and outreach as essential for building data and mapping tools with a meaningful impact.

2. **Focus Groups** We held three focus groups that helped us better-articulate the data landscape in Central Ohio, from an internal MORPC staff and committee perspective. These conversations (summarized on pages 6-9) helped us map the dynamics between our roles, work activities and audiences around data. Furthermore, our focus groups generated some useful ideas about how different audiences relate to data differently—a key component of the final user personas.

3. **Online Survey** We broadly distributed an online survey to gather information to help us segment and describe groups of data users. From our 445 respondents, we identified six general groups, based on each of their similar behaviors, needs and challenges, and preferences when interacting with data. The general characteristics of the survey respondents are on pages 10-15. One group, the civic tech community, was not visible in our results. For that group, we created a persona based on our experiential knowledge of that group.

The end result are user personas, found in Appendix E (pages 29-35). Savvy Sonja, Manager Marco, Engaged Elaine, Decisive Delaney, Hopeful Hadiya, Specialist Samir, and Coding Corey are the seven characters. As representatives of their respective communities of data users, MORPC and our partners around data can use these personas to aid in getting the right data to the right people in the right formats.

BACKGROUND

The Mid-Ohio Regional Planning Commission (MORPC) is a voluntary association of Central Ohio governments and regional organizations that envisions and embraces innovative directions in transportation, energy, housing, land use, the environment and economic prosperity. Our makeup is representative of the rural, urban and suburban communities that comprise our burgeoning region. In addition, our programming and public policy throughout the 15 counties we serve are supported by our role as a consumer and provider of data.

Our work around data and mapping is guided by a Regional Data Advisory Committee (RDAC)—a group of MORPC Commission members and other community leaders whose work is integrally linked to regional issues, and enhanced and supported by the development of cooperative strategies around regional data. The RDAC members have diverse data expertise—they represent different sectors (public, private, nonprofit, and academic) and geographic areas.

Several years ago, MORPC partnered with Thoughtwell (formerly Community Research Partners) to develop a regional data hub. The result of that effort was a tool called DataSource, which was a detailed, robust, and high-maintenance web-based data hub that proved difficult to sustain. Furthermore, it was unclear who used the resource and whether it met their needs effectively.

With the recent retirement of DataSource, MORPC perceives a continued need for a regional data resource, but one that is built with specific purposes and audiences in mind. As data becomes more numerous and complex, more open, and easier (through technology advances) to deliver to more people in more formats, how do organizations with key roles around data deliver tools and resources that are practical, relevant and accessible to the people who need them?

In support of the express priority to define and understand our data audiences, we looked to the principles of user experience (UX) design, which have been cultivated and refined with great success in the software development industry over the past several decades. The main idea of UX design is to build products that meet a specific need for a specific group or audience, instead of building products, then searching for an audience to use them—a shift in focus that improves both the success and efficiency of product development.

The principles of UX design can be integrated into any kind of work that involves creating tools and resources that could be configured in many ways or serve many purposes. Many public and nonprofit organizations that work to provide diverse data and analysis to multiple audiences are turning to UX design for tools to help them create more impactful data and mapping resources.

MORPC recognizes that we—alongside our many partners—play important roles around the complex and varied data that impact decisions and planning throughout Central Ohio. As such, the agency has conducted this user persona research effort to support a central commitment to getting the right data to the right people in the right formats.

RESEARCH APPROACH

In order to learn more about Central Ohio data users and their needs, MORPC conducted three research activities: (1) interviews with peer organizations, (2) focus groups, and (3) an online survey.

The objectives of these three efforts were to (1) learn how other similar organizations around the country work to understand their audiences as they build data and mapping resources; (2) assess known and presumed audiences and needs from discussions with staff and data experts in the region; and (3) understand data needs, motivations, and comfort levels across the region through a broad survey of users. This document will provide a summary of the results from each of these components, along with a discussion about how the analysis informed the creation of Central Ohio data user personas.

PEER ORGANIZATION INTERVIEWS

To enhance our understanding of other audience segmentation efforts, we interviewed staff from peer organizations around the country. Four of the organizations are part of the National Neighborhood Indicators Partnership (NNIP), a peer network of locally-focused organizations that work to connect people with neighborhood data. The other two interviewed organizations were selected based on a review of numerous websites, which provide publically-available data and mapping resources. Websites that appeared to successfully provide resources in a variety of formats and subject-areas were selected. Additionally, some of the organizations were known to have a UX design program or interest. Appendix A is the list of interview questions that were asked.

The interview questions were designed to gather insights from other organizations about their knowledge of key audiences—whether researched or intuitive—and the way that knowledge shapes the data and mapping resources they provide. While the organizations vary in type, size and scope, many commonalities exist among them. Along with some overarching themes, the responses yielded some important details that can guide other organizations seeking a shift to more user-centered design. Several interviewees cited resources used in their own work in understanding and engaging with audiences. These sources, along with details about the organizations interviewed, are referenced in Appendix B.

Avoid the Multi-Tool

In an age of more breadth and depth of data and advancing technology for delivering that data to audiences, it is easy for public-focused organizations to fall into building tools for the sake of building tools. While the tools may serve a purpose with an intention to "do good", packing in too many features often leads to a tool that is overly-complicated for the audiences that might have otherwise used them. In fact, even the most data- and technology-savvy audiences are deterred by resources with no clear or elegant purpose. Building single-purpose tools in response to a specific need garners greater success than tools that try to meet too many needs at once.

"[The most effective data tools were] the ones that did one thing well and that was it."

-Steven Spiker, Urban Strategies Council

Know your Audience, Stretch your Resources

All of the respondents agreed, even the best tool has no value without an audience to use it. With limited staff hours and budgets, the organizations we interviewed all spoke about audience awareness as a mechanism for deliberately spending those finite resources. While the approaches taken by each entity varied, they all had used some approach (whether informal or formal) to help them think through the audience—or, user experience—when creating something new.

"We have some funding right now to do [user experience research] about partners around how they consume and absorb the data that we produce and…rethink what it is that we produce that meets their needs more directly"

-Steve Spiker, Urban Strategies Council

"[We're] getting people to think about how…data is going to be used. Do you really need a dashboard, or is that just spending time on something that may not really have an impact?"

-Bob Graedeck, University of Pittsburgh

Iterative Design

Building smaller, simpler tools may be enough to meet the needs the designer wants to fill, and it gives users something to react to. A simple tool that meets one need really well today can be repurposed or expanded if a user or an astute analyst makes a connection to some other audience's need. Furthermore, if the tool misses the mark, fewer resources were spent, making it easier for resource-limited organizations to change course. These smaller, more deliberate investments have made the data-focused organizations we spoke to feel more nimble and effective.

"Typically the people who come to us are not the audience, [and may not know] what to ask for. I think an iterative approach is best. We tend to build something and then modify it later, based on user requests."

-Michelle Riordan-Nold, Connecticut Data Collaborative

"Send people out to community meetings…and just show off the stuff…[G]et regular feedback from people…that are using it. You can't just release it, wash your hands and move on. You've got to stay engaged so you can refine it and learn from what you've done"

-Bob Graedeck, University of Pittsburgh

Relationships are Key

Organizations with a role in local, public data are often responding to customers—whether hired as a consultant or directed by elected officials to address a particular need. In many cases, the audiences with the need, however, are not the same people as those who bring it to light--this means that building relationships through engagement in the community is imperative. In some instances, hosting listening sessions or user groups was said to generate ideas for content or formats for the resources provided. In other cases, user testing (which might be as simple as providing a paper version of a resource and asking for feedback) was used to refine an approach before investing more substantial resources.

"Before we publish a report, we run some preliminary thoughts either through a focus group or a full meeting of a group in the neighborhood that we analyze…because they just know it so much better."

-Erica Raleigh, Data Driven Detroit

Keep the Audiences Learning

While data impacts all of our lives, not all of us make it our living. All of the interviewees acknowledged a need to educate audiences, in an ever-expanding data age. Every organization described their role as keeping audiences engaged in effectively using resources, defining the problems or needs that can be addressed in communities through data, or even participating in the analysis itself. Data education and outreach were viewed by most as critical components to reaping the greatest benefit from public, open data. Some of the suggestions included:

- 'Digital storytelling', or simple pre-made data visualizations to demonstrate to audiences what can be done with data
- Data workshops and presentations for varying skill levels and interests of users
- Providing a small amount of free technical assistance to anyone who requests it, as an opportunity to build relationships and educate new audiences
- Use libraries for workshops to improve access to the general public
- Engage community leaders through participatory analysis
- Find opportunities to show the civic tech community the data sets that are available

"I think data at its core is still not fulfilling its destiny. ... [I]f open data is all about transparency and empowerment, we need to do a much better job at reaching [all] populations."

-Sari Laden, DataLA

Open Data and Find Partners

Many of the organizations spoke to the power of simply publishing high-quality raw data sets via a data portal, while developing relationships with data partners. There are people everywhere who are doing or want to do useful things with data, so it is imperative to stay tuned-in and find opportunities to build on existing efforts, instead of always starting from scratch. All of the organizations played some role in data policy and coordination. The shared challenges about pervasive needs to (1) develop more effective strategies for privacy protection with locally-generated data, (2) build trust with smaller data-collecting entities, (3) better-connect existing data portals, (4) keep the focus simple and targeted in an increasingly-complex environment, (5) leverage partnerships and funding, and (6) keep audiences learning and engaged.

"...[P]robably the biggest lesson is to leverage your institutional partners. Communicate as much as you can about what's going on...Let people know what you're working on or if you're having problems...Don't reinvent the wheel—work with what's out there already...Use those relationships and those institutions as a way to improve your work."

-Bob Graedeck, University of Pittsburgh

FOCUS GROUPS

Three focus groups were held to gather information from staff at MORPC, as well as from data experts in the region. Focus Group A was attended by 11 staff on the Agency's Communications Team--with representation from each of MORPC's departments. This group included both technical and non-technical staff. Focus Group B was attended by 7 staff, suggested by each of the department Directors. These were staff who regularly work with and communicate about data, generally in technical roles. Focus Group C was attended by 6 members of the Regional Data Advisory Committee, Communications Subcommittee. These participants are high-level committee members representing entities with key roles around data in the Central Ohio Region.

Aside from a few minor revisions, the questions asked in each focus group were the same (the full list of questions can be found in Appendix A). Each session lasted approximately one hour, and was documented via audio recording and visible notetaking.

While each of the Focus Groups had some unique perspectives, there were many recurrent themes in the responses. Overwhelmingly, the responses served as an orientation tool—a 'you are here' mark on the roadmap of MORPC's evolving role with data about Central Ohio, as well as MORPC's relationships with other people and organizations in the regional data ecosystem. Additionally, the conversations summarized some important hypotheses about various audiences' needs. Finally, the discussions surfaced a number of unmet current and hypothetical future needs, as changes around how people use data are anticipated.

MORPC's Data Roles and Audiences

In defining MORPC's roles around Central Ohio data, the three groups mentioned a range of functions pertaining to data, including collection, coordination, creation, analysis, processing, knowledge, provision, resources, and guidance. In Figure 1, these roles are more broadly categorized as gathering, working with, and providing data. Figure 1 highlights MORPC's key work in each of these roles (what we gather, what we do with it, and how we provide it to our audiences). The lists displayed are not comprehensive, but rather representative of the type of work we do in each of the three generalized roles.

In each segment of our work around data, MORPC engages, broadly, with data providers, partners and collaborators, and consumers. Figure 2 summarizes the audiences named by the focus groups, and defines how each group intersects with these broad categories. The focus groups identified a clear distinction about the data that are continually kept in an inventory, and those that are 'pulled in' and analyzed for specific purposes, but not updated or maintained on a regular basis. In parallel to this delineation about our work, the summary of audiences required that categories like 'State Agencies' be subdivided to clarify the difference between agencies we work with directly, as part of our core activities (e.g. Ohio Department of Transportation) and those we may reach out to only periodically for data or best practices (e.g. Ohio Department of Public Safety). State Agencies, Federal Agencies, and Non-Profits were all divided into *primary* and *secondary* types, to differentiate between agencies with core versus peripheral relationships.

FIGURE 1 MORPC DATA ROLES, ACTIVITIES & RELATIONSHIPS
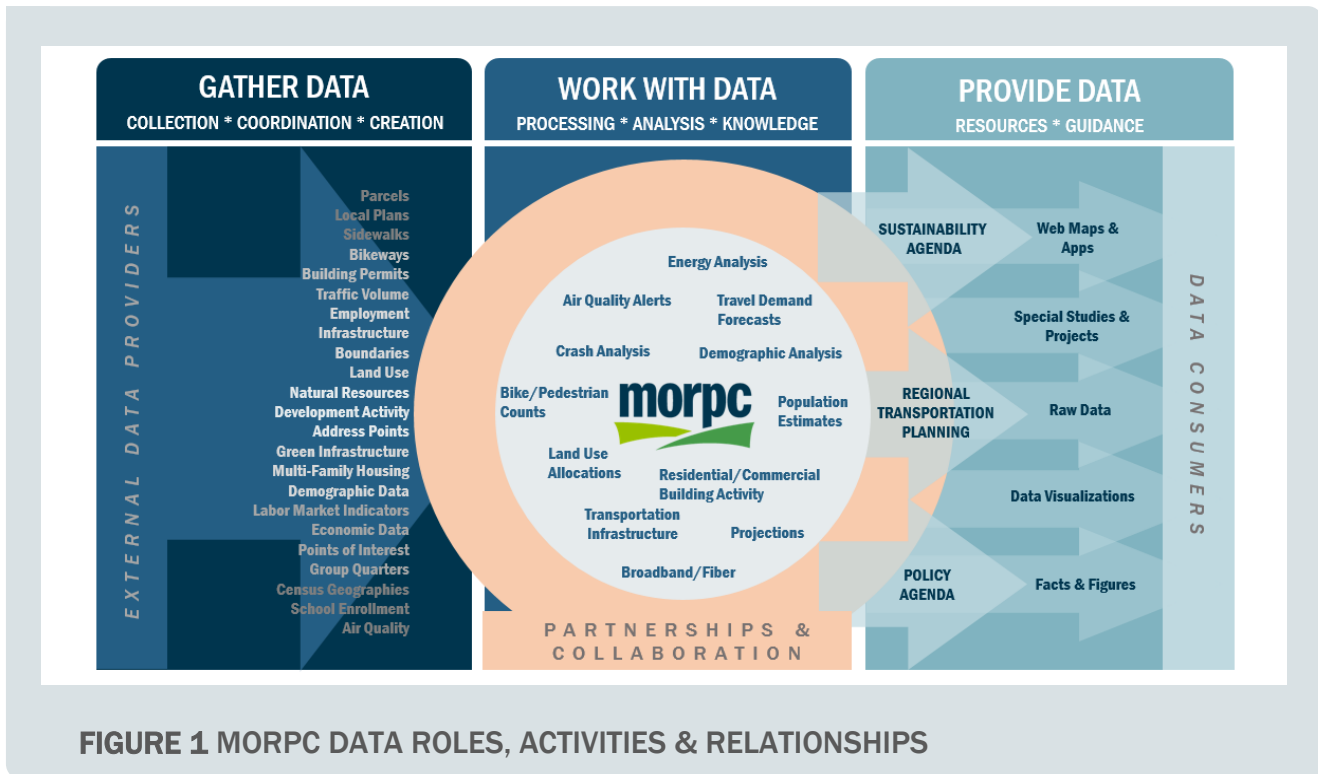
The determination of how audiences relate to MORPC's roles was inferred from respondents' answers to the question 'does MORPC have clearly-defined roles around data?' The responses naturally tended to assign certain audiences as viewing MORPC's roles as 'clearly defined', and others as more abstract or unknown altogether. Other data audiences were identified and described, as they relate to MORPC data roles, with the questions 'who are our audiences?' and 'what are their needs?'

The resulting categorization of our data audiences presented in Figure 2, provides a framework of understanding. Audiences are subdivided into three tiers. Tier One audiences are those that intersect with our work in all three roles—they are providers, partners and consumers. One clear example of a Tier One audience is the Ohio Department of Transportation—they provide some data, direct the focus of our work, and consume the data we generate for the region to prioritize project funding.

Tier Two audiences are partners or collaborators—individuals or entities who participate in shaping and directing our work. These audiences are also interested in consuming the data that are generated from our work efforts. An example is our Commission, which sets the big picture goals that prioritize much of our work. The data and analysis generated is then consumed for purposes such as local decision-making or communicating with constituents.

Tier Three audiences neither provide us data, nor do their goals directly influence the focus of our work. These groups, however, may still have needs that our data can serve, even though it wasn't created expressly for that purpose. A final group of audiences are only sources of data. We may pull data in from these providers for a specific purpose. An organization could move into a different tier at any time, as new initiatives can change our focus and lead to new relationships and audiences around data.

| MORPC Role	GATHER DATA	WORK WITH DATA	PROVIDE DATA
Audience Relationship	Data Providers	Partners/ Collaborators	Data Consumers
TIER ONE			
Primary State Agencies	X	X	X
Local Government Staff (Technical)	X	X	X
Primary Federal Agencies	X	X	X
Consultants	X	X	X
Primary Non-Profit Agencies	X	X	X
Transit Authorities	X	X	X
Universities	X	X	X
Utility Companies	X	X	X
Metro Parks	X	X	X
Other MPOs	X	X	X
TIER TWO			
Commission		X	X
Public Service Administrators		X	X
Elected Officials		X	X
Local Government Staff (Non-Technical)		X	X
TIER THREE			
News Media			X
Community/Civic Groups			X
General Public			X
Secondary Non-Profit Agencies			X
Unaffiliated Data Community			X
Students			X
Public Service Staff			X
Real Estate Developers			X
Businesses			X
SOURCES			
Secondary State Agencies	X		
Secondary Federal Agencies	X		
Data Brokers	X		

FIGURE 2 MORPC DATA AUDIENCES

Data Needs versus Data Resources

Each of the Focus Groups considered what different audiences need when seeking data. Ideas about different audiences' data needs were also generated from a discussion about existing data resources that do a good job meeting those needs. Examples cited varied, depending on the type of audience. For example, some audiences would need more granular, detailed data; whereas others prefer less detail, but require data that answers a specific question (about a certain place and topic). It was also suggested that some audiences may need data that are pre-processed into user-friendly formats, whereas others would prefer data in machine-readable formats.

Figure 3 illustrates five categories of audience data needs that are summarized from the focus group responses. The 5 types of needs are considered to be spectrums; oriented to correspond to a range of user types that goes from audiences with low data resource capacity to users with high data resource capacity. Audiences on the low data resource end of the spectrum may have some limitations in their ability to collect, process or analyze data independently. They may have technical skills, but limited access to necessary software or hardware; or perhaps they lack time due to other responsibilities; or they may be uncomfortable working with raw and diverse data sets, but are comfortable in their topic of expertise to interpret data on their own.

It should be noted here that many Focus Group respondents referred to data guidance, data literacy and data education in relationship to all types of data audiences. All audiences are assumed to need information about the source of the data, how the data should or should not be used and what the data represents. The 'Expert Guidance' need, while arguably a spectrum, includes a minimum requirement of data documentation and/or guidance, even for audiences on the highest end of the data resource capacity spectrum.

FIGURE 3 SPECTRA OF AUDIENCE DATA RESOURCE CAPACITY
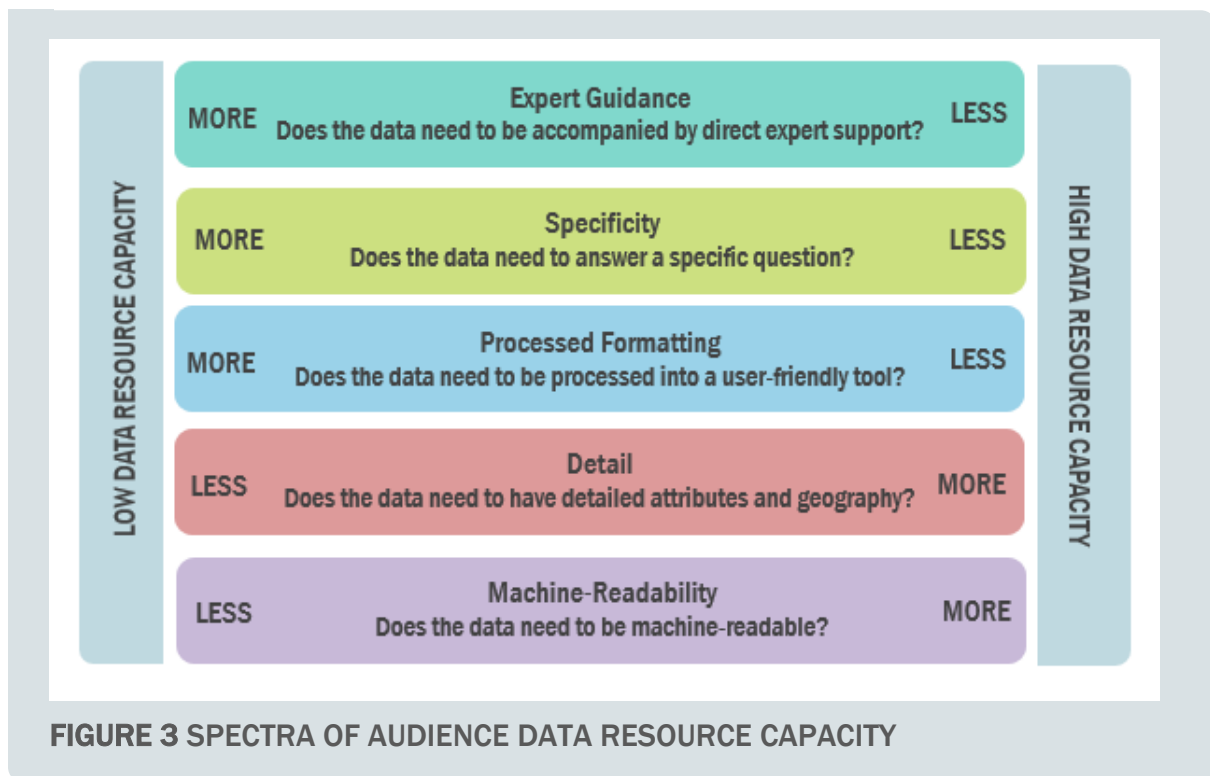
Unmet and Future Needs

The focus groups discussed 'gaps' in data in Central Ohio, as well as emerging trends that may require attention in the coming years. Those conversations are generalized into three main points, as follows:

1. *Everyone* needs data education
   Regardless of skill level or purpose with data, every single data user needs some degree of data education--whether that's basic data literacy, guidance about when and how to appropriately use a particular dataset, or creating better, richer datasets.

2. Bigger. Better. Faster. *More!*
   The importance of the role of data in our everyday lives is expanding, as we, as a society, generate more data with greater complexity every day. To keep up, organizations with roles around data have a critical need to build infrastructure and operations that can support the increased availability and demand for current, clean and detailed data.

3. Data in the "Big Picture"
   The public needs education. The analysts need organization. The organizational leaders and policy-makers need to set the stage. Conversations about data privacy, standards, security and governance are already buzzing. The need is to ensure that actions taken are able to adapt to the fast-changing circumstances.

ONLINE SURVEY

An online survey was conducted to obtain data directly from Central Ohio data users. The complete survey questionnaire can be found in Appendix D. The survey was disseminated using a 'snowball sample' approach. It was distributed to MORPC mailing lists with a general request to complete the survey. Also, it was promoted on multiple social media sites, in an attempt to reach a broader audience. Finally, specific contacts were invited directly to pass the survey on to their constituents and/or mailing lists, again, in an effort to reach audiences beyond our direct contacts.

During the two and a half week survey window, the Central Ohio Data User Survey was completed by 445 respondents—of those, about 390 answered the majority of the questions. For most of the questions, we used 445 as the total number. For those that used weighted scores or averages (see Figure 9 and 10) we used the number of total responses. To gain some insight about the participants, the survey asked for some basic descriptive information, such as their affiliation with MORPC, the organization they represent, and their occupation. From these characteristics, key groupings of responses were analyzed in greater detail for the purpose of generating user personas. The following discussion summarizes the more general analysis that was used to *identify* and *describe* user groups, and then outlines the key groups of Central Ohio data users that resulted.

Characteristics for Identifying & Describing User Groups

Based on the survey analysis, occupation and MORPC affiliation of respondents were the key characteristics for grouping users. Other characteristics, such as purpose when seeking data and preferred data formats were helpful in identifying subdivisions between larger occupational, or affiliation-based groups. Finally, some characteristics contributed to describing these groups for the user personas themselves. These include questions related to comfort-level when working with various types of data, which issues are most important to the respondents, educational attainment, what geographic scale of data is of greatest interest, and which sources respondents go to for data. Furthermore, several open-ended questions were included to further explore the goals, needs, and traits of the user groups.Figures 4-10 on the following pages highlight some of the overall characteristics of this survey's respondents.

Occupation

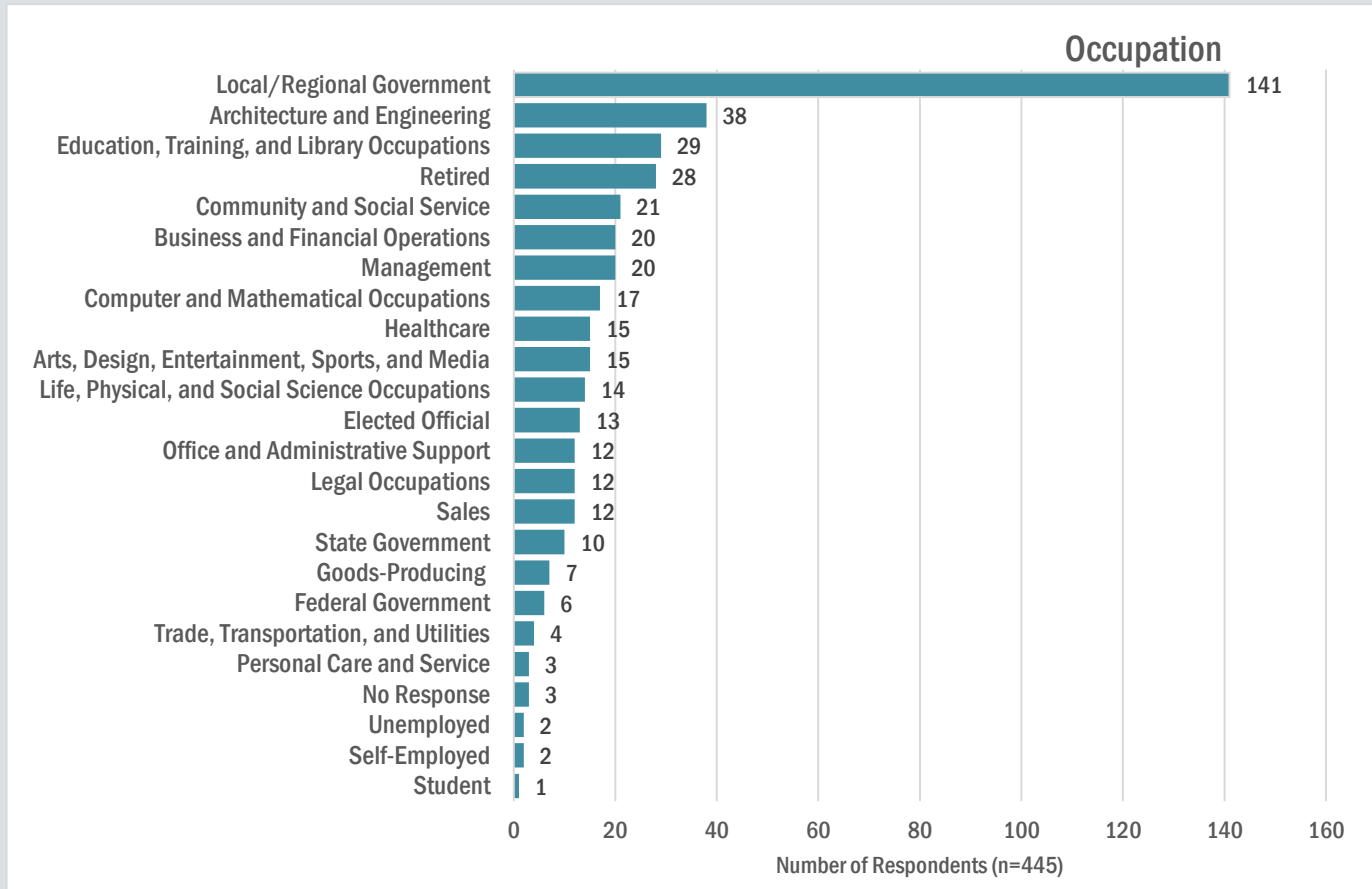| Occupation	Number of Respondents
Local/Regional Government	141
Architecture and Engineering	38
Education, Training, and Library Occupations	29
Retired	28
Community and Social Service	21
Business and Financial Operations	20
Management	20
Computer and Mathematical Occupations	17
Healthcare	15
Arts, Design, Entertainment, Sports, and Media	15
Life, Physical, and Social Science Occupations	14
Elected Official	13
Office and Administrative Support	12
Legal Occupations	12
Sales	12
State Government	10
Goods-Producing	7
Federal Government	6
Trade, Transportation, and Utilities	4
Personal Care and Service	3
No Response	3
Unemployed	2
Self-Employed	2
Student	1

Number of Respondents (n=445)

FIGURE 4 OCCUPATION

About one-third of respondents work in local or regional government occupations. Architecture and engineering; and education, training and library occupations had the next highest representation in the survey—each comprising about 7% of the respondents. Interestingly, retirees accounted for 6% of all respondents. Other occupations included healthcare, business and finance, computer science and math, law, and community and social services. About 3% of respondents described themselves as elected officials.
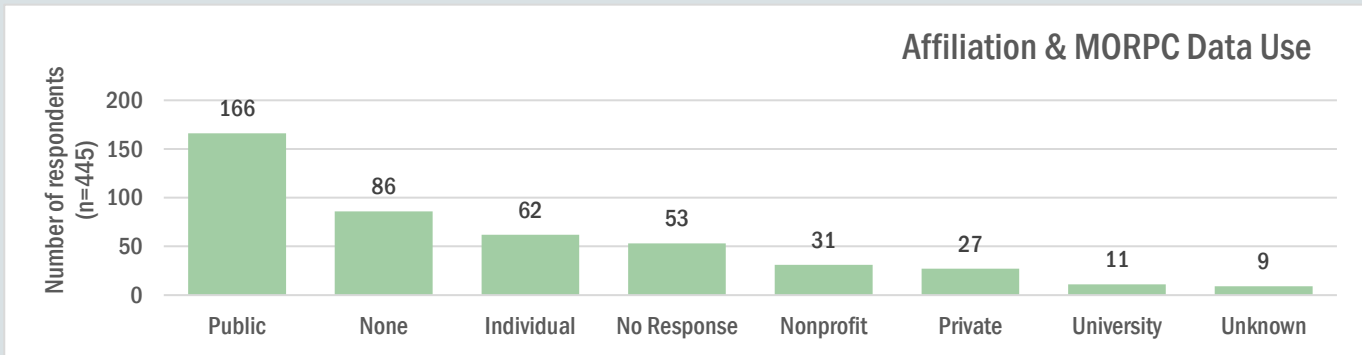
Affiliation & MORPC Data Use

Number of respondents (n=445)

Public	166
None	86
Individual	62
No Response	53
Nonprofit	31
Private	27
University	11
Unknown	9

FIGURE 5 MORPC AFFILIATION

MORPC Affiliation Survey respondents were asked 'Do you work for or represent an organization that provides data to or uses data from MORPC?' If they did, they were asked to name up to three organizations through which they are affiliated with MORPC data. Respondents were also asked 'As an individual (outside of your work or professional life) do you use MORPC data?' Responses to these questions were coded to generate groups of respondents based on whether or how they are affiliated with MORPC data. About 37% of respondents either provide data to or use data from MORPC in their professional roles in the public sector. 19% reported no use of MORPC data, either in their professional roles or as individuals. 14% use MORPC data as individuals, but not in their professional roles. 7% are affiliated via roles in non-profit organizations, 6% via the private sector, and 3% via a role in an academic institution. 14% either did not respond to this question, or did not list any specific organizational affiliation.
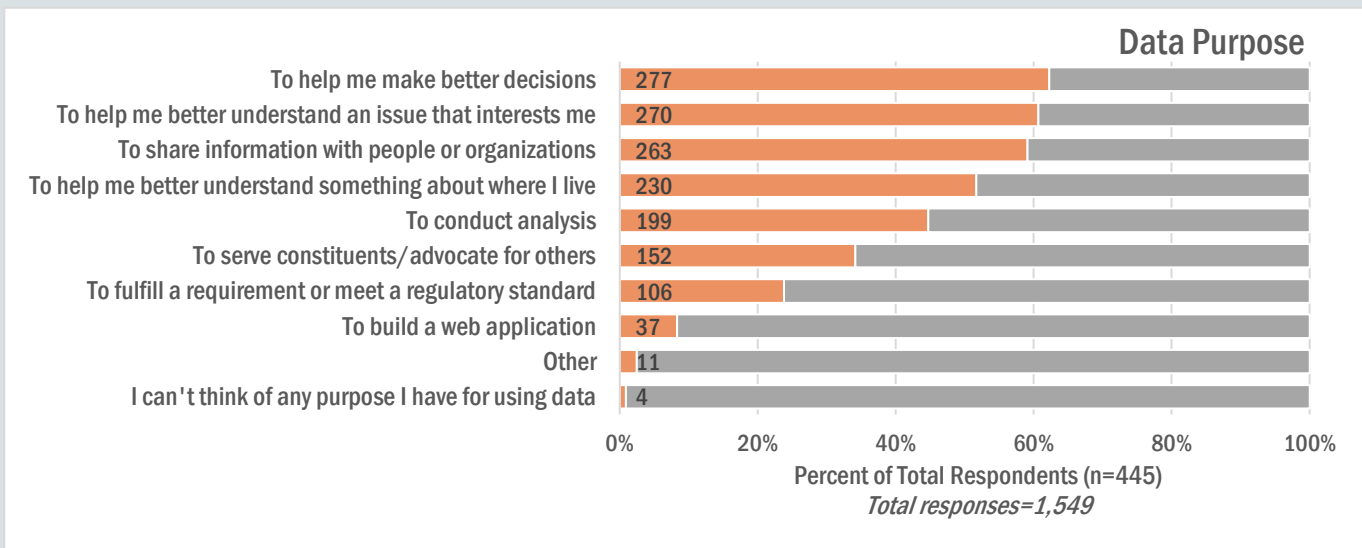
Data Purpose

Purpose	Respondents
To help me make better decisions	277
To help me better understand an issue that interests me	270
To share information with people or organizations	263
To help me better understand something about where I live	230
To conduct analysis	199
To serve constituents/advocate for others	152
To fulfill a requirement or meet a regulatory standard	106
To build a web application	37
Other	11
I can't think of any purpose I have for using data	4

Percent of Total Respondents (n=445)
Total responses=1,549

FIGURE 6 DATA PURPOSE

Survey respondents were asked to select (from a list of options) for which purposes they use data. More than half of the respondents included 'to help me make better decisions', 'to help me better understand an issue that interests me', 'to share information with people or organizations', and/or 'to help me better understand something about where I live' in their selections. Between one-quarter and one-half of respondents included 'to conduct analysis' and/or 'to serve constituents/advocate for others' in their answers. Less than a quarter included 'to fulfill a requirement or meet a regulatory standard' or 'to build a web application'.
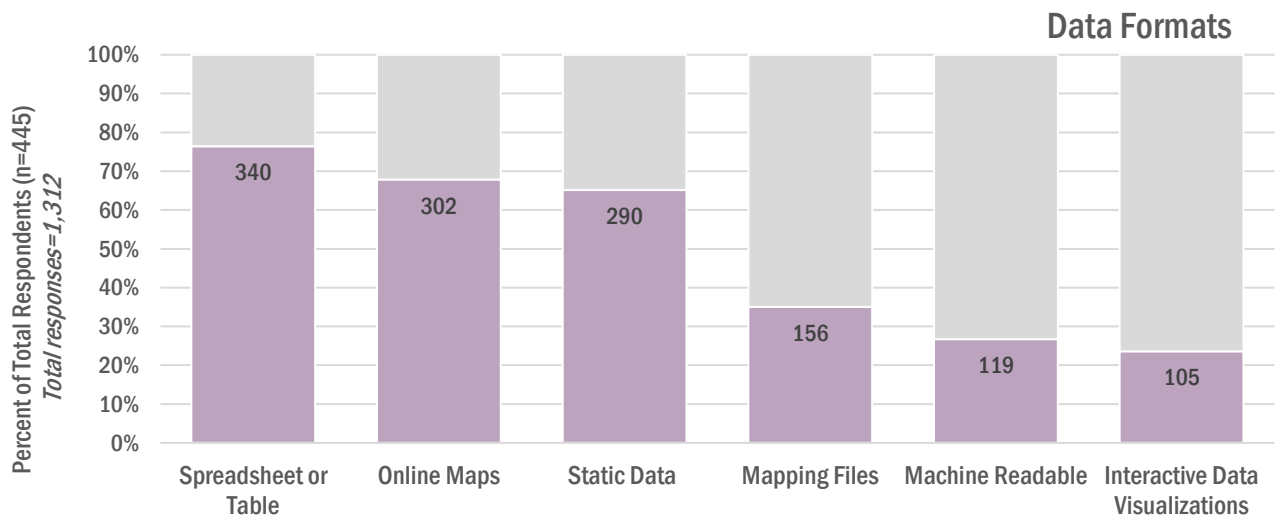
Data Formats

FIGURE 7 DATA FORMATS

Respondents were asked to select formats from a list of options in response to the question 'which of the following data formats do you use?' About 75% reported using spreadsheets or tables (e.g. Excel), and around 65% of respondents selected online maps and/or static data (e.g. charts, graphs, reports). Smaller segments of the survey respondents reported use of more technical file types–mapping files (e.g. KML, shapefiles, GEOJSON), machine readable (e.g. TXT, JSON, XML, CSV), or Interactive Data Visualizations (e.g. Tableau, PowerBI, Insights).
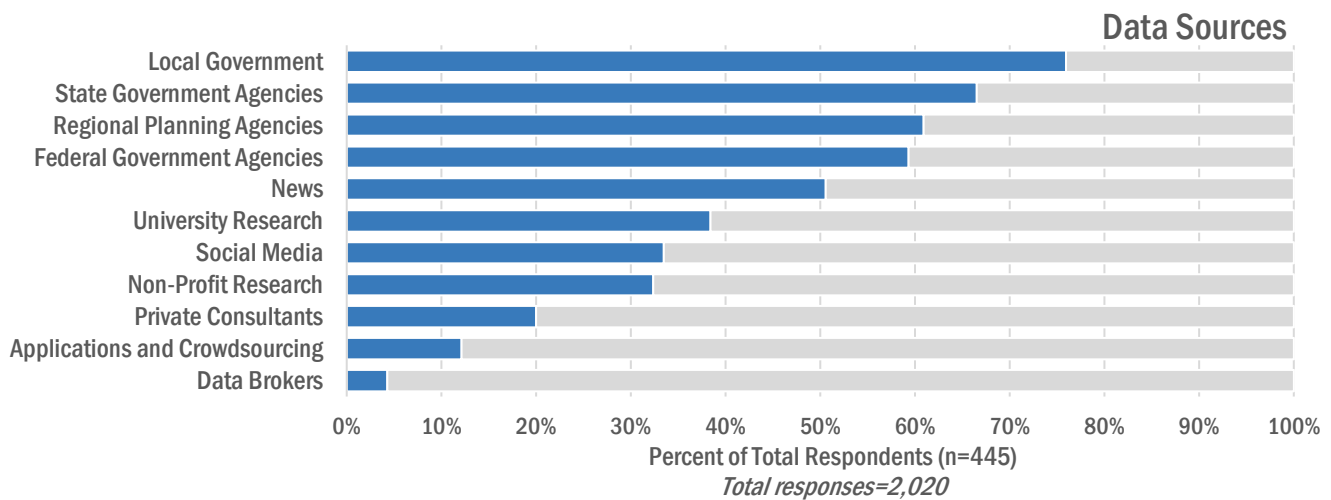
Data Sources

FIGURE 8 DATA SOURCES

Survey respondents were asked 'what sources do you go to for data?' The respondents showed the greatest use of data from one or more government sources, ranging from local to federal entities. News was selected by half of the respondents. Fewer than a quarter of respondents selected private consultants, applications and crowdsourcing or data brokers as sources they use.
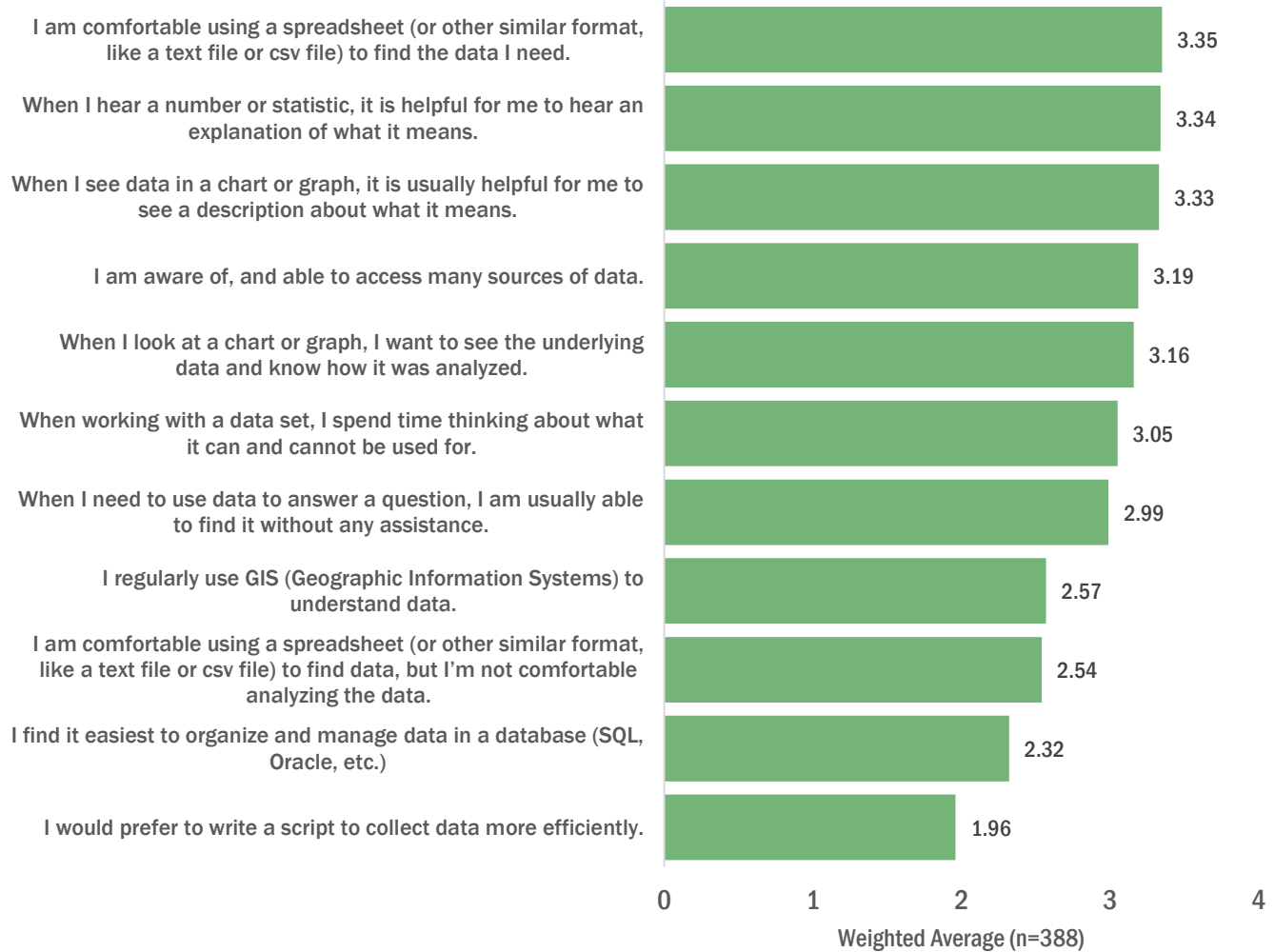
Comfort with Data

Statement	Weighted Average
I am comfortable using a spreadsheet (or other similar format, like a text file or csv file) to find the data I need.	3.35
When I hear a number or statistic, it is helpful for me to hear an explanation of what it means.	3.34
When I see data in a chart or graph, it is usually helpful for me to see a description about what it means.	3.33
I am aware of, and able to access many sources of data.	3.19
When I look at a chart or graph, I want to see the underlying data and know how it was analyzed.	3.16
When working with a data set, I spend time thinking about what it can and cannot be used for.	3.05
When I need to use data to answer a question, I am usually able to find it without any assistance.	2.99
I regularly use GIS (Geographic Information Systems) to understand data.	2.57
I am comfortable using a spreadsheet (or other similar format, like a text file or csv file) to find data, but I'm not comfortable analyzing the data.	2.54
I find it easiest to organize and manage data in a database (SQL, Oracle, etc.)	2.32
I would prefer to write a script to collect data more efficiently.	1.96

Weighted Average (n=388)

FIGURE 9 COMFORT WITH DATA

Respondents were asked to rate a series of questions related to comfort working with data on a 4-point scale from 'strongly disagree' to 'strongly agree'. Figure 10 shows the weighted average score for each of the 11 statements presented. Some trends in responses to these questions were used to validate or sub-divide user groups. Many respondents (>75%) strongly or somewhat agree that they want to hear or read an explanation of data that is presented to them—whether it's a statistic, chart or graph—pointing to an overwhelming need among audiences for clear communication about data.

At least 70% of respondents strongly or somewhat agree that they are comfortable using spreadsheets to find data. Over 70% of respondents also feel like they have access to lots of data. Also over 70% of people who took the survey think beyond the chart or graph presented to them, and consider the data and analysis method behind it. At least 65% reported comfort finding answers using data on their own, and also thinking about the strengths or limitations of a particular dataset.

Other statements which describe more technical skillsets had a higher divide between respondents who agreed and disagreed—less than 50% use GIS to understand data, are comfortable analyzing data, or use databases to organize data. Just over a quarter of respondents prefer to write a script to collect data more efficiently.
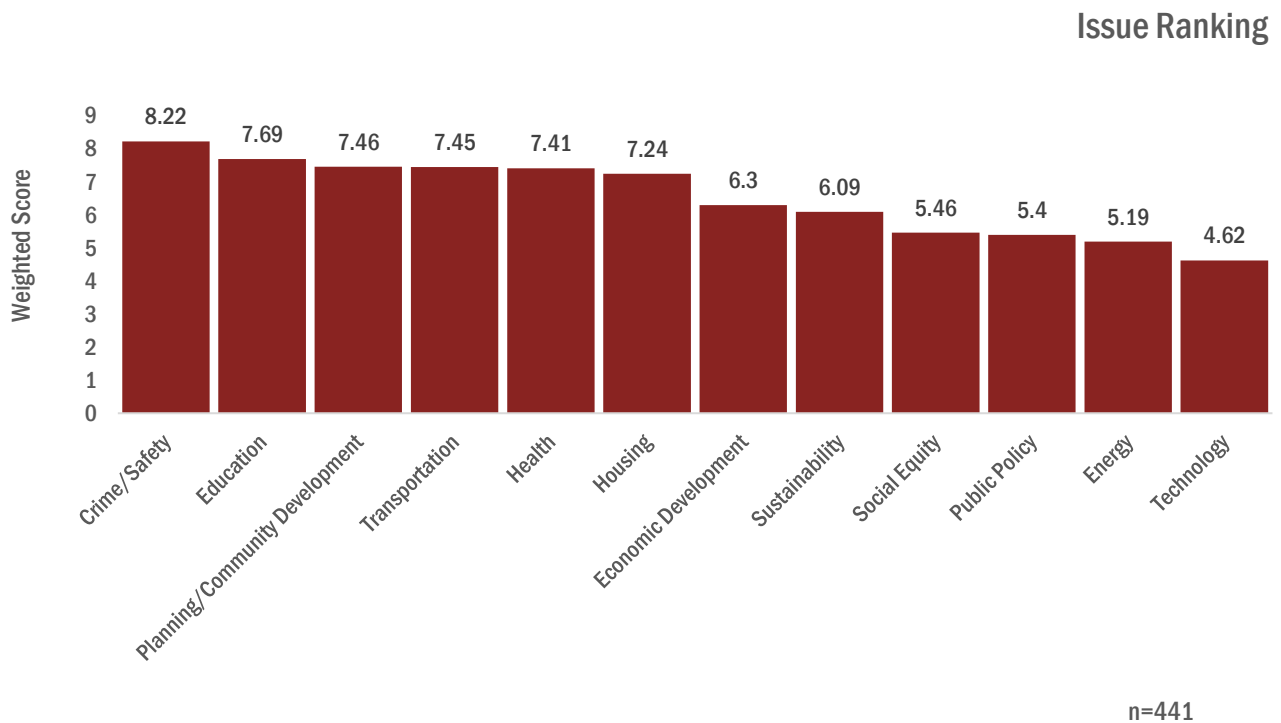
morpc

Issue Ranking

Issue	Weighted Score
Crime/Safety	8.22
Education	7.69
Planning/Community Development	7.46
Transportation	7.45
Health	7.41
Housing	7.24
Economic Development	6.3
Sustainability	6.09
Social Equity	5.46
Public Policy	5.4
Energy	5.19
Technology	4.62

n=441

FIGURE 10 ISSUE RANKING

The survey included a question that asked respondents to rank 12 broad issues in order of importance to the individual. This question was intended to aid in describing user groups in more detail, as an estimate of the type of data in which user groups may be interested. Figure 9 shows the weighted average scores for each issue. 'Crime/safety' stands out as the most important issue, with 'technology' ranked least important, on average.

morpc

USER PERSONA DEVELOPMENT

The three components of this research undertaking each contributed to the development of *user personas* as a tool for guiding our work around data. The survey results contained the actual data about our audiences, which most directly informed how to group and describe users. The 'audience tiers' (Figure 2), defined by the focus groups were used to help assign broad user groups as well. The 'data user needs' chart (Figure 3), also generated by our focus groups, was incorporated into the user personas as a quick-reference when thinking about how an audience thinks about and needs data.

The following discussion lists the seven personas, with some description of how they were identified in the survey results, and what other audiences might be treated similarly. The actual user personas are included in Appendix E.

1. **Savvy Sonja**

   Sonja represents audiences, generally working in the public sector, who are generalists, and usually quite technical, especially in their use of GIS in their work. In the survey results, this group was identified as people with a public sector affiliation with MORPC data, who use mapping file formats (n=82). We think some more generally focused nonprofit and university types could fall in this category too.

   These users tend to be knowledgeable about data sources, and comfortable working with data. Analyzing data to communicate and inform decision-making is an integral part of their work. They are more likely than other groups to use data in their non-work lives as well as in their work lives. They seem to know what they need, naming very specific data, tools and functionality when asked what other types or formats they'd like to have.

2. **Manager Marco**

   Marco represents audiences from the public sector, with non-technical roles. This group was isolated in the survey results as people with a public sector affiliation with MORPC data, and no use of either mapping files, machine readable formats or interactive data visualization software (n=51).

   These users mention many data sources they use, including news sources and libraries. They're most likely to use data to communicate, but not as likely as Sonja to analyze data or use it to inform decisions. This group is less comfortable working with different kinds of data, and unlikely to view themselves as 'data users' outside of their work lives. When asked what data or formats they wish they could find, they either mentioned processed data like statistics, charts or graphs or mentioned not being entirely sure what could be out there that would help with their jobs.

3. **Engaged Elaine**

Elaine represents the diverse community of engaged citizens in the region. While Elaine is portrayed as a retiree (since they represent 6% of survey respondents), this persona could work in any occupation at any stage of their career. It might also include 'front line' public service workers (e.g. librarians, educators or social workers). This group was qualified as people with individual affiliations with MORPC data, with potentially more technical users culled out by excluding people who use mapping file formats (n=41).

These users are most focused on using data to understand an issue or something about where they live. While they prefer more processed data sources, they tend to be pretty keen analytical thinkers, based on their responses to the data comfort questions. They use sources like local groups, news and reports more than raw data. They name practical data, property data, and special issue data when asked what's missing that they'd like to use.

4. **Decisive Delaney**

Delaney represents elected officials. We reasonably assume this could stretch to include people in high-level public service administration roles. These were people in the survey who marked 'elected official' as their occupation (n=13).

These users are most focused on using data to inform decisions and engage with residents, about the issues and places that matter to them. While they may have some data skills, they'd really like quick information that answers a specific question. Like Marco, Delaney is less likely to have a data wish list than other more technical types.

5. **Hopeful Hadiya**

Hadiya represents the issue-focused nonprofit people in the region. This category is assumed to also encompass more issue-focused public or university staff as well. These people were identified in the survey responses as those with a nonprofit-affiliation to MORPC data (n=31).

Not surprisingly, their top data purpose is to understand an issue. They're also likely to want to use data about their focus issue to analyze, communicate and inform decisions. The data they want is more detailed, neighborhood-level numbers that can help them get funding and shine a light on the issues they're passionate about.

6. **Specialist Samir**

Samir is a specialist, like Hadiya, but his interests are defined by his industry rather than an issue. These private sector data users might do consulting (e.g. planning, architecture and engineering), land development, logistics, or journalism. These users were isolated in the survey responses as those with a private sector affiliation to MORPC data (n=27).

These users are most likely to analyze and inform decisions using data. Some of them are focused on using data to meet regulatory requirements. While they certainly use more processed data formats (e.g. reports and online maps), they tend to be comfortable using raw data and mapping files to get what they need. Like Sonja, Samir can easily name specific datasets or formats he'd like to have access to for his work.

7. **Coding Corey**

While the above groupings only capture 245 of the survey respondents, the other responses didn't reveal any noticeable groupings or trends. We intentionally looked in the data to see if we could find any civic tech people, but weren't able to locate this group from the response data available.

Knowing this could be an important data user group to consider, we created a *proto persona* (a user persona based on educated guesses, rather than user data) to represent these socially-minded, tech savvy people in the community. These users are characterized by their interest in using diverse datasets to build practical tools to address social challenges. They have programming skills that they want to put to use, making machine-ready data their preferred format.

APPENDIX A

INTRODUCTION
Describe who we are, the user persona project, and the data and mapping resources we currently produce/maintain.

GOAL
Describe your process for developing your current web-based data and mapping resources, and especially how you consider the user experience in that process.

QUESTIONS
Role Do you believe your organization has a clearly-defined role related to data in your area? If so, what is the role? If not, are there any roles your organization plays by 'default'?

Audiences As you develop your data resources, do you have a strategy for identifying and targeting certain audiences with certain resources?

If yes:
- What is your strategy?
- Who are your main audiences?
- What kinds of data resources do you provide them, and why?
- Do you use analytics to understand who uses your resources? If so, please describe.
- How do you communicate with your audiences about data?
- Are there any audiences you find more difficult to reach than others, and why?
- Are there any needs that are not being met, or audiences that are underserved? If so, please describe.

If no:
- Who do you believe are your audiences (partners, users, providers)?
- Do you use analytics to understand who uses your resources? If so, please describe.
- What needs to your audiences have that you're working to meet with your data resources?
- How do you communicate with your audiences about data?
- Which of your resources seem to meet one or more of your audiences' needs well, and why?
- Are there any needs that are not being met, or audiences that are underserved? If so, please describe.

Reflection Looking back at your process developing data resources so far, what lessons have you learned?

Future How do you see the data needs of your audiences changing in the future?

- What opportunities, tools or technology are available/on the horizon that can serve the data needs of the audiences?
- How might changes in the way we think about data (or what constitutes "data") create new audiences or needs?

APPENDIX B

PEER ORGANIZATION LIST/RESOURCES CITED BY INTERVIEWEES

Organizations Interviewed

ORGANIZATION	DESCRIPTION
Connecticut Data Collaborative	Statewide data initiative
Data LA	Mayor-initiated centralized city data program
Urban Strategies Council	Non-profit social research organization (NNIP)
Western Pennsylvania Regional Data Center	University-based regional data hub (NNIP)
Data Driven Detroit	L3C social research organization (NNIP)
Atlanta Regional Commission	Regional planning and intergovernmental coordination agency (NNIP)

Resources Cited

Association of Public Data Users --APDU

Bloomberg Intelligence – Bloomberg Terminal

Civic User Testing Group – The CUTGroup Book

MIT– Media Lab

MIT – Civic Design Data Lab

National Neighborhood Indicators Partnership –
Monitoring Impact: Performance Management for Local Data Intermediaries

National Neighborhood Indicators Partnership – NNIP

APPENDIX C

FOCUS GROUP SCRIPT

WELCOME: Thank you. Quick overview of the project purpose. We will be recording.

1. MORPC's Role **[WHAT IS THE SYSTEM?]**
    a. Does MORPC have a clearly-defined role related to data in the Central Ohio Region?
        i. (If yes) What is/are the role(s)? What roles does MORPC play in data throughout the Region?
        ii. (If no) Do we default into particular roles, even if they are not clearly defined?
        iii. (If prompts are needed) Are we a communicator? Are we an educator? Are we a provider? Are we a trustworthy source of data?

TRANSITION: Thinking about MORPC's current role around data in Central Ohio, let's shift to thinking about other individuals and organizations that are involved.

2. Audiences **[WHO USES THE SYSTEM?]**
    a. From whom does MORPC currently collect data?
    b. To whom does MORPC currently provide data?
    c. Whom else do we partner with around data? (Whom do we work alongside in collecting, creating and/or analyzing data?)
    d. Looking back at the roles, are we missing any audiences? What about people who don't consider themselves "data users"?

TRANISITION: We've generated some ideas about who else is involved in MORPC's data role in the Region. Now I'd like us to think about those people or organizations, and try to imagine what their needs or objectives may be. This is whether or not those needs are met.

3. Needs & Objectives **[WHY & HOW DO THEY USE THE SYSTEM?]**
    a. What needs do each of these audiences have?
    b. If met, what will those needs help them to achieve?
    c. (If MORPC itself is not mentioned as an audience) What needs do we have? If met, what will those needs help us to achieve?
    d. What's being done well to meet those needs?
        i. Can you think of any tools, provided by MORPC or someone else that meet individuals' needs well?
    e. Are all needs/objectives being met by us or by someone else?
    f. Are there any audiences whose needs are underserved, compared to others?

TRANSITION: Whether or not current needs are being met, we know that audiences, systems, and needs may change in the future, which will likely impact our role. Let's shift into talking about the future.

4. Future **[HOW IS THE SYSTEM CHANGING?]**
    a. How do you see the data needs of these audiences changing in the future?
    b. What opportunities, tools or technology are available/on the horizon that can serve the data needs of the audiences?
    c. How might changes in the way we think about data (or what constitutes "data") create new audiences or needs?

APPENDIX D

ONLINE SURVEY QUESTIONS

The Mid-Ohio Regional Planning Commission (MORPC) is conducting research to find out why and how people use data about the Central Ohio region. We need feedback from current and potential data users to help us understand how to provide data in the ways that are most useful to the people and organizations who need it.

What do you mean by 'data'?
Data means different things to different people, which is part of what we'd like to better understand. When completing this survey, we are interested in any kind of data (from any source) you use that is related to the Central Ohio region, and all of the communities and neighborhoods within it. We know data can be used in many ways and in many formats—we want to know what this looks like to you.

Should I take this survey?
You may be wondering if you are the 'right' kind of data user to take this survey—the answer is almost certainly 'yes'. We know that data is everywhere these days--whether it's a big or small part of your life, chances are, you use data in some way. Regardless, we'd like to know more about that. Participation in this survey will provide MORPC, and its partner organizations, with valuable information about how to collect, maintain, analyze and provide data to the Central Ohio region.

All survey responses will be kept confidential. Only aggregate data will be included in any reported results, and all potentially identifying information will be excluded.

Haven't heard of MORPC?
MORPC is a voluntary association of Central Ohio governments and regional organizations that envisions and embraces innovative directions in transportation, energy, housing, land use, the environment and economic prosperity. Find out more by visiting our website.

We will be collecting survey responses through September 15th. If you have questions about the survey, please email Liz Whelan at lwhelan@morpc.org.

1. Which of the following best describes your current occupation?

[dropdown ▲▼]

2. What is the highest degree or level of school you have completed? (If you're currently enrolled in school, please indicate the highest degree you have received.) 💬1

◯ Less than a high school diploma

◯ High school degree or equivalent (e.g. GED)

◯ Some college, no degree

◯ Associate degree (e.g. AA, AS)

◯ Bachelor's degree (e.g. BA, BS)

◯ Master's degree (e.g. MA, MS, MEd)

◯ Professional degree (e.g. MD, DDS, DVM)

◯ Doctorate (e.g. PhD, EdD)

3. In what community (city, village or township) do you live?

4. Please rank the following issues in order of importance to you, with 1 being "Most Important" and 12 being "Least Important": (*click on the six dots to the left of each issue to drag them into the order you choose*)

⬍ Crime/Safety

⬍ Housing

⬍ Transportation

⬍ Energy

⬍ Public Policy

⬍ Planning/Community Development

⬍ Economic Development

⬍ Social Equity

⬍ Technology

⬍ Health

⬍ Education

⬍ Sustainability

5. At what scale(s) are you interested in data about Central Ohio? (*check all that apply*)

☐ Neighborhood

☐ City/Village/Township

☐ County

☐ Region

☐ Other (please specify)

6. Which types of technology do you use to access data? (*check all that apply*)

☐ Mobile Device

☐ Tablet

☐ Desktop Computer

☐ Laptop Computer

☐ Printed Documents

☐ Other (please specify)

☐ _____

7. Which of the following data formats do you use? (*check all that apply*) 💬

☐ Spreadsheet or Table (e.g. Excel)   ☐ Machine Readable (e.g. TXT, JSON, XML, CSV)

☐ Static Data (e.g. charts, graphs, reports)   ☐ Mapping Files (e.g. KML, Shapefiles, GEOJSON)

☐ Interactive Data Visualizations (e.g. Tableau, Power BI, Insights)   ☐ Online Maps

☐ Other (please specify)

_____

8. What sources do you go to for data? (*check all that apply*) 💬

☐ Local Government

☐ Regional Planning Agencies

☐ State Government Agencies

☐ Federal Government Agencies

☐ University Research

☐ Private Consultants

☐ Data Brokers

☐ Non-Profit Research

☐ News

☐ Social Media

☐ Applications and Crowdsourcing

☐ Other (please specify)

_____

9. For each of the following statements, please select the answer that most accurately describes your comfort working with data: 💬 3

	Strongly Disagree	Somewhat Disagree	Somewhat Agree	Strongly Agree
When I look at a chart or graph, I want to see the underlying data and know how it was analyzed.	○	○	○	○
I find it easiest to organize and manage data in a database (SQL, Oracle, etc.)	○	○	○	○
When I need to use data to answer a question, I am usually able to find it without any assistance.	○	○	○	○
When I hear a number or statistic, it is helpful for me to hear an explanation of what it means.	○	○	○	○
When I see data in a chart or graph, it is usually helpful for me to see a description about what it means.	○	○	○	○

I am comfortable using a spreadsheet (or other similar format, like a text file or csv file) to find data, but I'm not comfortable analyzing the data.

I am comfortable using a spreadsheet (or other similar format, like a text file or csv file) to find the data I need.

I regularly use GIS (Geographic Information Systems) to understand data.

When working with a data set, I spend time thinking about what it can and cannot be used for.

I would prefer to write a script to collect data more efficiently.

I am aware of, and able to access many sources of data.

10. Select the statement with the description that best completes the following sentence: "I most often encounter Central Ohio Data when I am_____." 💬 1

- ○ watching or listening to news
- ○ completing a task or assignment for work
- ○ completing a task or assignment for school
- ○ attending a presentation or conference
- ○ doing something related to a hobby or personal interest
- ○ Other (please specify)

- ○ in a meeting at work
- ○ talking with friends
- ○ doing something related to volunteer work
- ○ in a public meeting
- ○ I don't remember the last time I encountered data

11. For what purpose(s) do you use data? (*check all that apply*) 💬 2

☐ To fulfill a requirement or meet a regulatory standard.

☐ I can't think of any purpose I have for using data.

☐ To conduct analysis.

☐ To help me better understand an issue that interests me.

☐ To share information with people or organizations.

☐ To build a web application.

☐ To serve constituents/advocate for others.

☐ To help me better understand something about where I live.

☐ To help me make better decisions.

☐ Other (please specify)

_____

12. Do you work for or represent an organization that provides data to or uses data from MORPC? 💬

○ Yes

○ No

○ I don't know

13. If applicable, please list the name of the organization(s) that you represent: (*you can enter up to three organizations*) 💬

1 _____

2 _____

3 _____

14. As an individual (outside of your work or professional life) do you use MORPC data? 💬

○ Yes

○ No

15. Are there any organizations or sources of data that work very well for your purposes? Please list up to five:

1 _____

2 _____

3 _____

4 _____

5 _____

16. Are there types or formats of data that you would like access to, but that are not available? Please list up to five:

1 _____

2 _____

3 _____

4 _____

5 _____

17. Please list any emerging technologies, innovations or policies that you know of that might change how you use data: (*list up to five*)

1 _____

2 _____

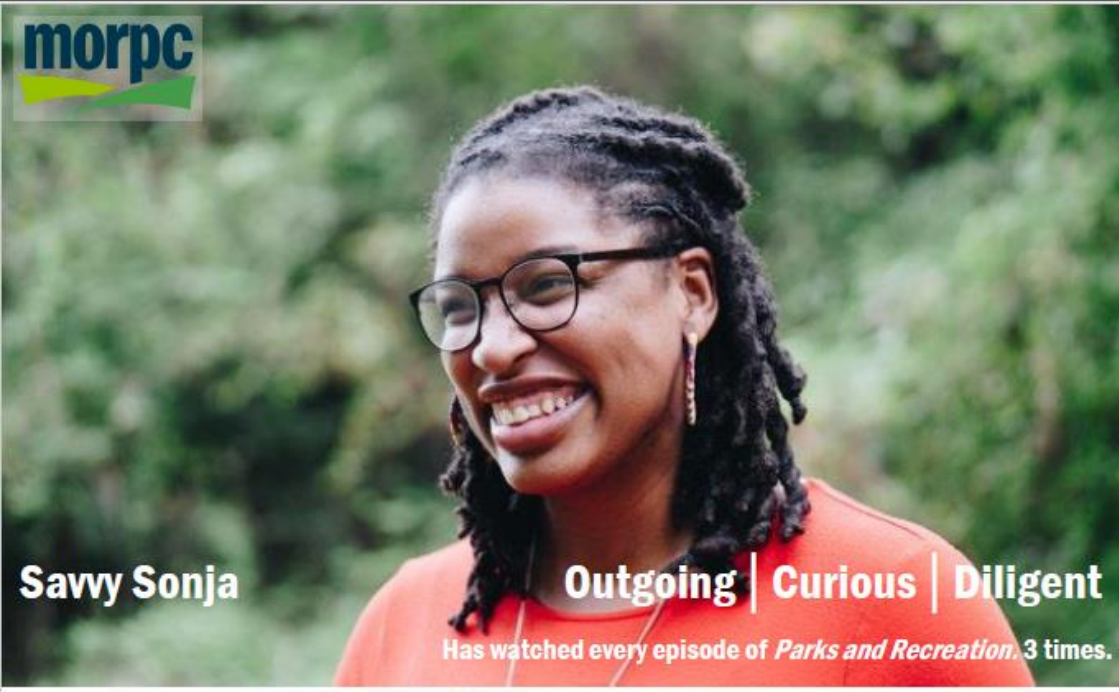3 _____

4 _____

5 _____

18. If you have additional ideas about how data might be affected in the near future, please use this space for your comments:

APPENDIX E

USER PERSONAS

morpc

Savvy Sonja

Outgoing | Curious | Diligent

Has watched every episode of *Parks and Recreation*. 3 times.

Planner for a mid-sized city government. Analyzes data to find answers to questions and share information with stakeholders.
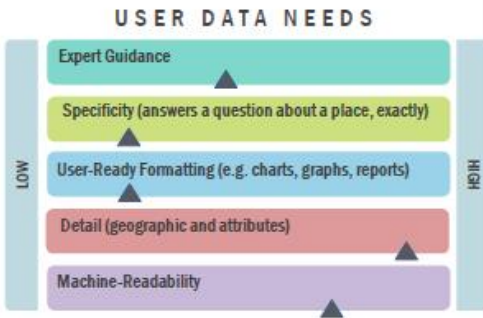
Behaviors

- Uses many local, regional, state and federal open data sources.

- Keeps up with issues and skills by attending local conferences and user groups.

- Uses intermediate GIS skills regularly.

- Creates maps and charts to communicate with internal staff, stakeholders and the public.

Needs & Challenges

- When she finds an interesting online map or other resource, wants to have easy access to the underlying data to use in her own work.

- A 'generalist', she needs data with different subjects.

- A local planner, she needs geographic detail.

- Even though she tries to stay tuned-in, frequently discovers data sets that have been available, but that she wasn't aware of before.

USER DATA NEEDS

Expert Guidance

Specificity (answers a question about a place, exactly)

User-Ready Formatting (e.g. charts, graphs, reports)

Detail (geographic and attributes)

Machine-Readability

LOW — HIGH

Manager Marco — Steady | Resilient | Compassionate

Soccer dad. Spends weekends coaching his kids' teams.

Administrator for a township. Uses data to share information and understand practical issues.

Behaviors

- Focused on key functions of the township, such as permit applications, street maintenance and asset management.

- Works in a small team in a building with all other township government staff.

- Works closely with the adjacent city, and the county to deliver services to residents.

- Stays busy just keeping up with the workload.

USER DATA NEEDS

- Expert Guidance
- Specificity (answers a question about a place, exactly)
- User-Ready Formatting (e.g. charts, graphs, reports)
- Detail (geographic and attributes)
- Machine-Readability

LOW — HIGH

Needs & Challenges

- Comfortable with his 'go-to' sources of data needed to keep up with work.

- Would like to see some updates in software and workflow, but it's hard to imagine how a small team like his would implement changes.

- Points people to entities with more staff resources when they ask for data or analysis.

- Needs assistance and support in pursuing funding for operations and infrastructure projects.
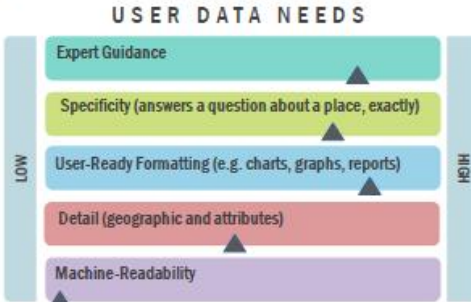
morpc
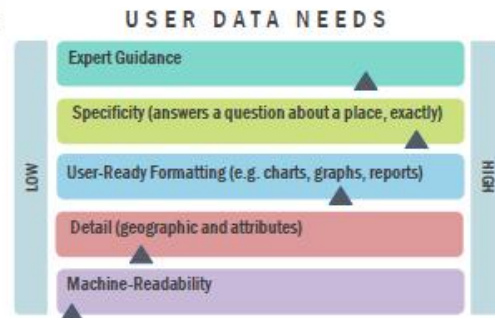
Engaged Elaine — Insightful | Determined | Energetic

World Traveler. Can handle much spicier food than you.

Retiree, neighborhood advocate, watershed protection group board member. Seeks to engage in participatory analysis about areas of expertise.
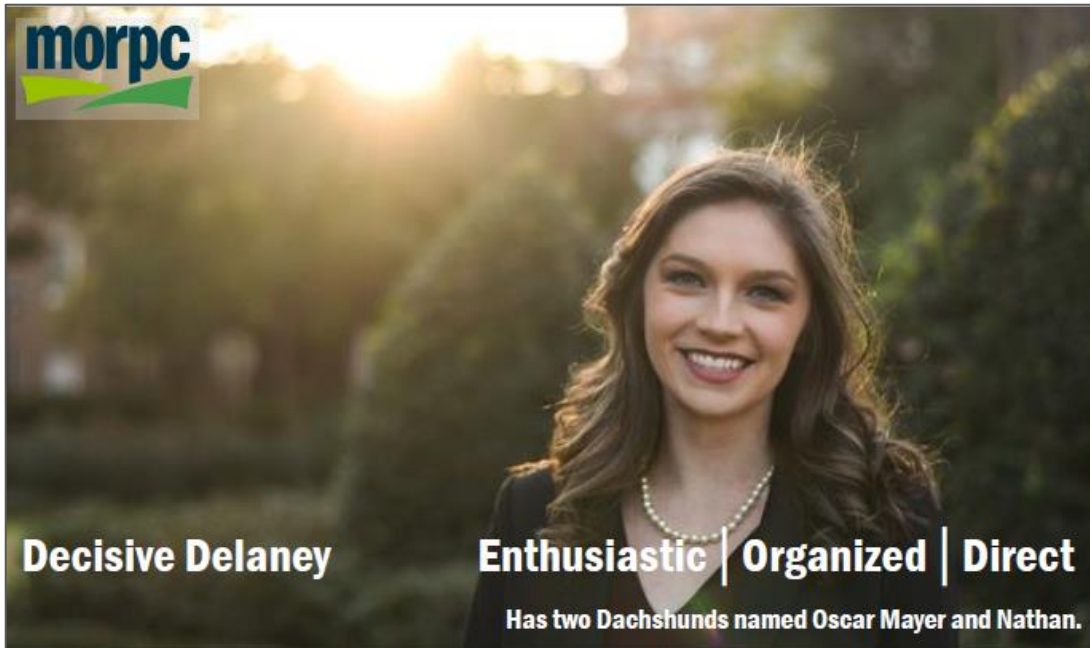
Behaviors

- Follows local news to keep up with happenings in the community.

- Uses apps to navigate when getting around town.

- Engaged in neighborhood civic group.

- Frequently attends public meetings.

- Uses social media to discuss local and national politics.

USER DATA NEEDS

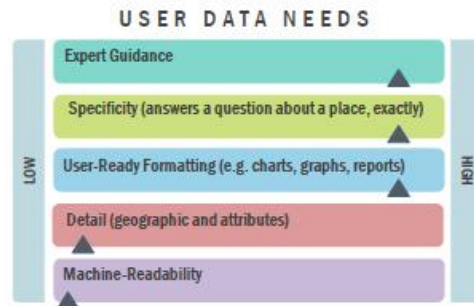|	LOW	HIGH
Expert Guidance	▲	
Specificity (answers a question about a place, exactly)	▲	
User-Ready Formatting (e.g. charts, graphs, reports)	▲	
Detail (geographic and attributes)	▲	
Machine-Readability	▲	

Needs & Challenges

- Looking for new ways to use technology to simplify daily life.

- Would like to see data about her community—things like property sales and development plans—in an accessible format.

- Likes to share data from the news and reports, but wants to feel confident in her sources.

Decisive Delaney

Enthusiastic | Organized | Direct

Has two Dachshunds named Oscar Mayer and Nathan.

Elected village council member. Looks for accessible data to inform decisions and communicate with residents.

Behaviors

- Reaches out to regional data providers to find information that addresses a particular issue in her village.

- Talks regularly to constituents and stakeholders with an interest in development in her community.

- Reads news about the region regularly.

- Attends presentations about regional planning and development issues and reads reports.

USER DATA NEEDS

Expert Guidance

Specificity (answers a question about a place, exactly)

User-Ready Formatting (e.g. charts, graphs, reports)

Detail (geographic and attributes)

Machine-Readability

LOW — HIGH

Needs & Challenges

- Analysis to aid in choosing and communicating the impact of public projects.

- Would like an easy way to find stats, charts and graphs specific to her village.

- Understanding projects, initiatives and trends in the region and how they impact her community.

morpc

Hopeful Hadiya

Sincere | Loyal | Eloquent

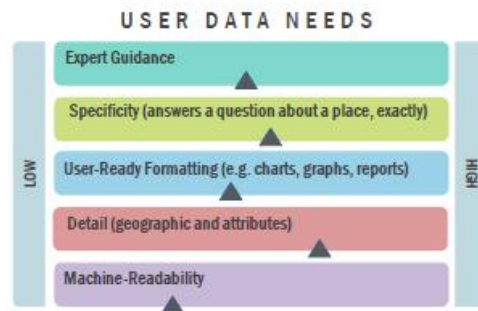Debate champion. Can and will beat your argument.

Affordable housing nonprofit employee. Wants to understand and communicate housing needs, and demonstrate program impacts.

Behaviors

- Passionate about her work, advocates for affordable housing on and off the clock.

- Writes grants to secure funding.

- Attends regional events to network and stay current.

- Uses government and nonprofit-produced reports and resources in her work.

Needs & Challenges

- Needs data about neighborhood housing.

- Always seeking better ways to identify and justify areas with funding needs.

- Looks for data to help tell stories about health and safety outcomes related to housing

- Wishes she had time and resources to do her own neighborhood surveys, instead of always relying on small-sample Federal demographic data.

USER DATA NEEDS

LOW		HIGH
Expert Guidance	▲	
Specificity (answers a question about a place, exactly)	▲	
User-Ready Formatting (e.g. charts, graphs, reports)	▲	
Detail (geographic and attributes)	▲	
Machine-Readability	▲	

Specialist Samir

Ambitious | Systematic | Affable

Biggest. Movie nerd. Ever.

Project manager for a planning and engineering consulting firm. Needs project-specific data to meet contracted goals and regulatory requirements.

Behaviors

- Attends industry conferences to keep up with best practices in his field.

- Travels for work often, sometimes internationally.

- Enjoys solving one type of problem really well.

- Sometimes works long hours, and just wants to relax whenever he gets a chance.

- Creates specific kinds of data, when the project requires.

USER DATA NEEDS

| Expert Guidance |
| Specificity (answers a question about a place, exactly) |
| User-Ready Formatting (e.g. charts, graphs, reports) |
| Detail (geographic and attributes) |
| Machine-Readability |

LOW — HIGH

Needs & Challenges

- Knows what he's looking for, but sometimes needs help from local experts to know where to find it.

- When working in a new place, uses reports, graphs and charts to get familiar.

- When it's time to get the work done, needs clean, high-quality, detailed, machine-readable data for the people on his team.

morpc

Coding Corey

Creative | Socially-Minded | Geek*
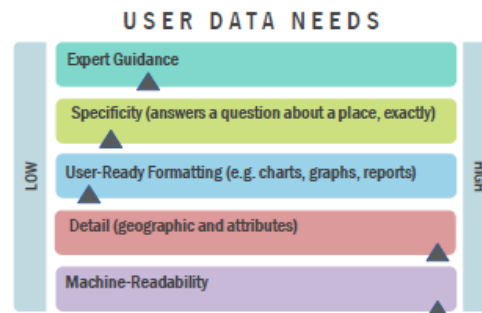
*Self-proclaimed and unapologetic

Civic tech enthusiast. Wants to make cool solutions to real community problems with any and all data they can get their hands on.

Behaviors

- Sometimes stays up all night programming...just for fun.

- Engaged in local arts, politics and social-service organizations.

- Believes in the power of technology to do good.

- Attends tech user groups, meetups and hackathons, whenever possible.

- Publishes data visualizations on their blog.

Needs & Challenges

- Indiscriminately consumes data.

- Frustrated when the data are out there, but aren't open.

- In spite of their ravenous data appetite, they 'get it' that data can be misleading, and do their best to use it responsibly.

USER DATA NEEDS

	LOW		HIGH
Expert Guidance	▲		
Specificity (answers a question about a place, exactly)	▲		
User-Ready Formatting (e.g. charts, graphs, reports)	▲		
Detail (geographic and attributes)			▲
Machine-Readability			▲

morpc

**Mid-Ohio Regional
Planning Commission**

111 Liberty Street, Suite 100
Columbus, Ohio 43215

T   614.228.2663
TTY 614.228.2663

www.morpc.org