



Better Management for Data Maturity



Introduction

Government agencies have been talking about treating data as a strategic asset for some time, but with the advent of the latest generation of artificial intelligence-based solutions — which are only as good as the data feeding them — those efforts have a new urgency.

This first installment of our three-part data guide series will focus on the strategies and tactics that can help agencies elevate their data management game and increase their data maturity. One challenge is to understand what constitutes “maturity.” So, we begin by highlighting several frameworks that describe the stages of maturity as well as some of the key dimensions, from visualization and storytelling to data governance in the AI era.

Although data management itself is a well-established discipline, the practices, techniques and tools continue to evolve. We look at the latest developments in four key areas: literacy, interoperability, privacy and analytics.

Finally, an ongoing initiative in Fairfax County, Virginia, illustrates how strong policies and practices can pay off in concrete ways. The county, not far outside the nation’s capital, is taking a data-centric approach to strategic planning, using dashboards and digital maps to plan, manage and track progress on a multiyear effort to improve quality of life across the county.

These articles will lay the groundwork for the next two parts of this series.

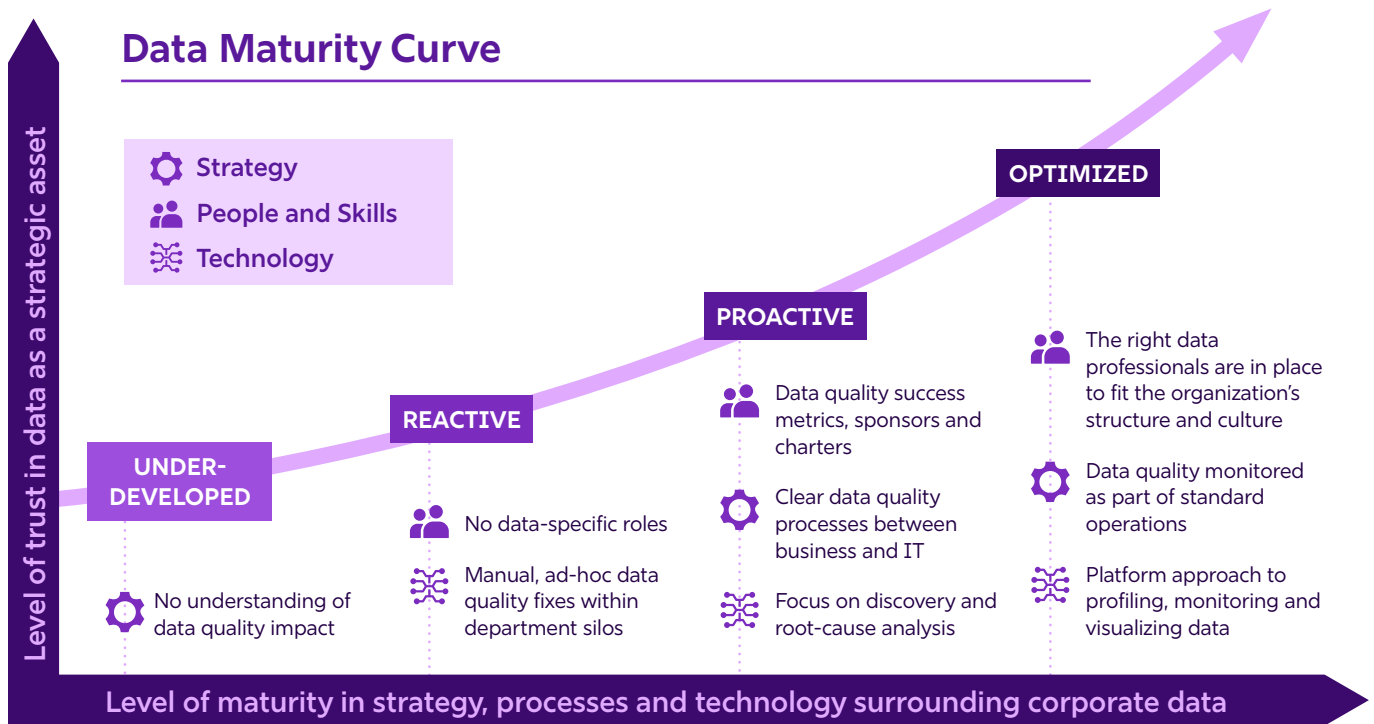
Contents

- 3 Why the Big Deal About Data Maturity?
- 5 How to Use Cross-Skilling to Address the Data Skills Gap
- 6 4 Tips for Effective Data Management**
 - 7 Cultivate Data Literacy
 - 8 Break Down Silos
 - 9 Implement Privacy-Enhancing Technologies
 - 10 Use Advanced Data Tools
- 11 How to Take the Next Step Toward Data Resilience
- 12 GIS Reshapes Strategic Planning in Fairfax County, Virginia
- 15 Conclusion and Acknowledgments

Why the Big Deal About Data Maturity?

Data maturity refers to how effectively an organization uses data to advance its mission. At peak maturity, an organization’s data management processes and abilities are well-defined and in harmony with the entity’s overall purpose.

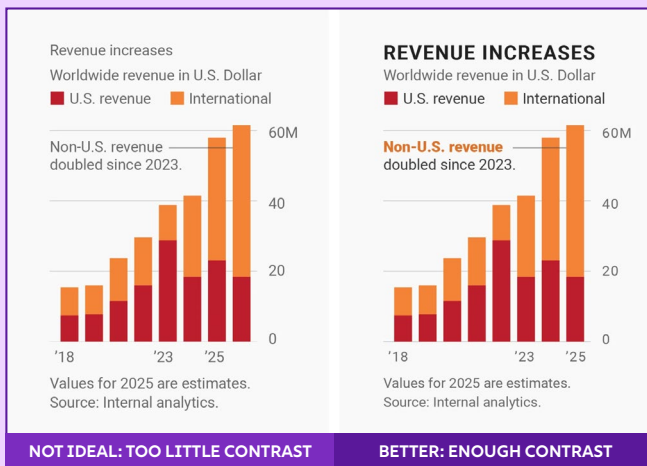
But that requires careful strategy and thoughtful deployment of people and technology, and it doesn’t happen overnight.



Source: Vanguard X

To use data in a sophisticated way, agencies need to find connections among various datasets — drawing richer, more holistic insights. That’s called **interoperability**, and it offers six key benefits to the public sector:

- 1** Improves resident experiences
- 2** Increases administrative efficiency
- 3** Enables data-driven policymaking
- 4** Maximizes the potential of open data
- 5** Enhances data protection and privacy
- 6** Reduces fraud, waste and abuse



See It — Don't Just Say It

The human brain [processes information](#) faster when it's in a visual format, so rather than let raw data speak for itself as black-and-white marks on a page, use visualization tools — e.g., charts, graphs, diagrams — to help tell the data's story.

Here's how, [according to](#) a visualization specialist at Harvard Library.

- **Choose your message.** You probably can tell different stories with the same data.
- **Select the right type of visual.** There are many options, from simple graphs to complex heat maps.
- **Draw the eye to the main story.** Use simplified labels, among other techniques.
- **Let the data dictate palette colors.** And remember: More colors can be worse.
- **Add descriptions directly on a graph.** Put any contextual language on the visual itself, rather than as a side note.
- **Include enough data to tell your story.** Too little data, and your visual can be misleading.
- **Follow C.R.A.P., the four principles of graphic design.** Contrast (make certain elements pop), repetition (repeat visual elements to create a theme), alignment (beware of centering) and proximity (place related items close together).

And then there's AI...

AI has highlighted the acute need for strong data management — and not all agencies are up to the task. In fact, according to [Gartner](#), 63% of organizations either do not have, or are unsure if they have, the data management practices they need for AI. And, through 2026, organizations will abandon 60% of AI projects that aren't supported by AI-ready data, says Gartner.

9 Best Practices for Data Governance — in the AI Era

How can you ensure that your AI applications use high-quality, secure data? You need strong data governance practices.

1. **Define your data governance objectives.** AI doesn't govern itself, after all.
2. **Build an AI data governance team.** Don't just add AI responsibilities to an existing IT group.
3. **Implement data quality controls.** You've heard it before: Garbage in, garbage out.
4. **Lock down your data security.** Protect AI systems from malicious actors and data corruption.
5. **Control and monitor who accesses data.** That includes the information that AI algorithms access.
6. **Implement data retention and deletion policies.** You don't want AI algorithms drawing on outdated data.
7. **Monitor compliance.** Continuously monitor the activity of employees, data systems and AI applications.
8. **Continuously adapt.** AI isn't "set it and forget it," and its governance policies shouldn't be either.
9. **Communicate, train and reinforce the policies.** AI technology doesn't enforce itself.

Source: [Project Management Institute](#)

How to Use Cross-Skilling to Address the Data Skills Gap



Tony Holmes

Practice Lead for Solutions Architects for Public Sector, Pluralsight

WATCH VIDEO

As state and local agencies have increased their reliance on data in all aspects of their operations, from policymaking to cybersecurity and service delivery, their need for data experts has grown substantially. Unfortunately, the same data experts are also in high demand in the private sector, where companies can offer bigger salaries and more flexibility and hire more quickly. And the emergence of new artificial intelligence (AI) solutions is only widening that skills gap and making it harder still to hire and retain talent.

Instead of relying on external talent, agencies need to invest more in their existing staff, expanding their data skills to address current and emerging mission priorities, said Tony Holmes, Pluralsight's Practice Lead for Solutions Architects for Public Sector. This approach is known as cross-skilling.

In this [video interview](#), Holmes discusses how agencies can use cross-skilling to address their data skills gaps. Topics include:

- Supporting more secure and responsible use of AI and generative AI
- Building teams of expert generalists who can tackle problems across domains
- Providing employees with greater internal career mobility

“This is beyond filling jobs. It’s about future-proofing the IT team by creating cross-skilled expert generalists who can think across disciplines and adapt to new technologies as they arrive.”

— Tony Holmes, Practice Lead for Solutions Architects for Public Sector, Pluralsight

About Pluralsight

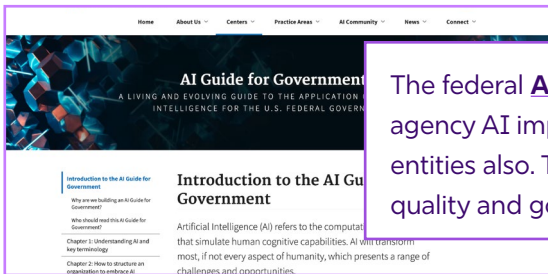
Pluralsight is the leading technology workforce development company that helps government agencies develop critical skills, improve processes and gain insights through data, and provides strategic skills consulting. Pluralsight helps build technology skills at scale with expert-authored courses on today's most important technologies, including cybersecurity, cloud, AI, data science, and more. Its platform includes tools to align skill development with agency objectives, virtual instructor-led training, hands-on labs, skill assessments and one-of-a-kind analytics.

[Learn more about Pluralsight](#)

Data to AI Maturity

Source: [Medium](#)

Manual Data Drudgery	Death by Dashboards	Data Tells a Story	Emerging Intelligence	Transformed Organization
<p>Manual reports</p> <p>Spreadsheets & PowerPoints communicate status</p> <p>Disagreements on how data was processed</p>	<p>Shadow data teams</p> <p>Only privileged employees can create reports</p> <p>Big spend on reporting, dashboarding or business intelligence systems</p> <p>Employees flooded with irrelevant data</p> <p>Multiple, inconsistent sources of truth</p>	<p>Glance-able answers start to simplify employee processes</p> <p>Multisource data merging</p> <p>Consistent view of info up & down the organization</p> <p>IT & business leadership coordinate work</p> <p>Measurable results emerge</p>	<p>Consistent measurable results</p> <p>Proactive information supports employees</p> <p>Experience tuned for each customer and employee</p> <p>Smart systems know what to look for</p> <p>Data crosses silos</p>	<p>AI/Machine learning is real</p> <p>New ways of working</p> <p>Employees focused on high-value work, all low-value work eliminated</p> <p>Recommendations are right for the employee</p> <p>New business models emerge</p>



The federal **AI Guide for Government** is a wide-ranging blueprint for agency AI implementation, with important insights for state and local entities also. The resource addresses structural, workforce, ethics, and data quality and governance concerns.

4 Tips for Effective Data Management

Data-mature organizations know how to manage their data well. They have practices, techniques and tools in place so that employees can consistently and securely access and deliver data across various subject areas and departments. The data consumption needs of all agency applications and business processes are met.

That's a lot to unpack. So here are tips for successful data management that address the challenges of:



Data Literacy



Data Silos



Data Privacy



Data Analytics



Cultivate Data Literacy

Why It Matters:

Think of raw data as a two-dimensional resource. In a data literate agency, employees can read, understand and communicate that data in context — that is, in whatever way is appropriate based on someone’s job function — in order to derive useful meaning and bring the data to life. That leads to better and faster decision-making. And as AI gains prominence, data literate employees know how to find and interpret what’s available — and what to ignore.

Key Elements

Data literacy is a gradual process that evolves in five phases.

LEVEL 1

Level 1: Foundational/Data Awareness

At first, you simply appreciate what data is, where to find it and why it’s important, plus you might know about basic data-related technology and software.

LEVEL 2

Level 2: Data Competency

At this stage, you can boast simple data skills and complete lower-lift tasks, such as sorting, filtering and making basic calculations. You also might be acquainted with data visualization tools, such as charts and dashboards.

LEVEL 3

Level 3: Data Proficiency

The next step is performing more advanced data manipulation and analysis work (i.e., regression analysis, hypothesis testing and data modeling) and having a solid understanding of data analysis methods and tech. You also know how to use statistical software and programming languages.

LEVEL 4

Level 4: Data Fluency

In the fourth stage, your data skills are advanced, and you’re well-versed in data visualization and storytelling for various audiences. You can work easily with complicated datasets; incorporating multiple sources and performing advanced statistical analysis is no problem.

LEVEL 5

Level 5: Data Expertise

Ultimately, you develop expertise in specific areas of data science, such as machine learning (ML), and contribute to cutting-edge research and the creation of innovative solutions. You also provide strategic know-how for data-related initiatives in your organization.

8 Ways to Create a Data Literacy Plan

1. Distinguish between data literacy and technical literacy.
2. Start with a baseline of employee skills.
3. Use common language.
4. Build a culture of learning and reward curiosity.
5. Remember that everyone learns differently.
6. Define success.
7. Make sure leaders are involved.
8. Know that data literacy alone isn’t enough.

Source: *MIT Sloan School of Management*

Break Down Silos

Why It Matters:

There are many reasons for siloesd data: decentralized business or IT operations, legacy technology, workforce culture, weak data management, etc. But such barriers to data sharing lead to inaccurate or incomplete information and, ultimately, to faulty decision-making, subpar customer experience and AI inaccuracies. There are financial ramifications — for example, poor data quality costs organizations an average of \$12.9 million annually, according to Gartner. And data silos increase cybersecurity risk and the likelihood of duplicative technology.

Key Elements:

Some commonsense tools and techniques can break down silos.

Treat Data as an Agencywide Asset — The goal is to establish coherency, and that includes identifying all the data an agency holds, cataloguing it, and adopting a centralized governance and security framework. A uniform, agencywide naming convention provides clarity and makes data analysis more meaningful.

Embrace Cloud Technology — Cloud-based data platforms offer a scalable, secure and centralized way of storing and using data from multiple data sources in real time. This improves efficiency and helps create a single source of truth, but it also fosters a collaborative environment that can improve workplace culture.

Adopt Centralized Data Tools — Data warehouses consolidate and identify relationships among data from various agency sources. Because they find connections between datasets, the technology is particularly useful for data analysis. Another option is data lakes, which store vast amounts of structured, unstructured and semi-structured data waiting to be processed.

Maximize Employee Skills — Create cross-functional teams that identify and close gaps in how the agency integrates and uses its data. If the positions don't yet exist, hire and empower both a chief data officer and a chief AI officer who each report to the CEO, so the agency's data and AI strategies are aligned.

Dismantling Data Silos With AI

AI technology can unify business operations and improve data sharing through:

Data Integration and Automation

- ▷ collects, scrubs and incorporates data from many sources
- ▷ creates a single dataset for users to analyze

Advanced Analytics and ML

- ▷ scours data and finds patterns that humans cannot
- ▷ provides more holistic insight

Natural Language Processing (NLP)

- ▷ searches data using common language
- ▷ removes technical barriers that inhibit data sharing

Real-Time Data Sharing and Collaboration

- ▷ allows for synchronized data access and collaboration
- ▷ leads to better decision-making

Data Governance and Security

- ▷ provides centralized data management
- ▷ makes enforcement of data policies and standards easier

Source: AIThORITY

Implement Privacy-Enhancing Technologies

Why It Matters:

Data privacy policies always have been vulnerable both to the malicious intent to ignore or work around approved processes and to simple mistakes or misunderstandings. In recent years, however, several privacy-enhancing technologies (PETs) and techniques have emerged that make it easier to enforce privacy policies without overly limiting data's value.

Key Elements

Here is a sampling of technologies and techniques that have shown the most promise. Synthetic data, described in the next section, is another one.

Tokenization — Sensitive data, such as Social Security numbers, is replaced with a token that represents that data but has no inherent value, while the original data is stored in a separate, secure token vault. Someone might hack the database and steal the tokens, but without the tokenization system, they have no way to map the tokens to the data.

Trusted Execution Environment — When dealing with especially sensitive data, a computer's hardware and software might not provide adequate protections. In such instances, encrypted data can be stored in a secure area on the computer processor that cannot be accessed via the computer's main operating system.

Homomorphic Encryption (HE) — The idea of encrypting sensitive data has always had a weakness: The data typically must be decrypted before it can be processed or analyzed, at which point it is exposed. HE makes it possible to compute data while it is still encrypted. This technique provides another layer of defense in a zero-trust security environment.

Federated Learning — This technique can help specifically when different offices or organizations want to use their own data to train an AI model. Rather than integrating those datasets, which might not be permissible with sensitive data, each organization can train its own model, then the resulting algorithms can be combined to create a smarter global model. The same approach can be used to refine the model over time.

Overview of Privacy-Enhancing Technologies

Here are four general categories of PETs, with examples of each.

Data obfuscation tools

- Anonymization/
Pseudonymization
- Synthetic data
- Differential privacy
- Zero-knowledge proofs

Encrypted data processing

- Homomorphic encryption
- Multiparty computation
- Trusted execution environments

Federated and distributed analytics

- Federated learning
- Distributed analytics

Data accountability tools

- Accountable systems
- Threshold secret sharing
- Personal data stores/personal information management systems

Use Advanced Data Tools

Why It Matters:

In a [September 2024 report](#), the BCG Henderson Institute described what it called “The Immediate Aptitude-Expansion Effect,” that is, knowledge workers’ ability to use generative AI (GenAI) tools to complete tasks outside their current skillsets. BCG and other experts see data analytics as an especially promising area for aptitude expansion.

Key Elements

Here are some of the common use cases that BCG, Gartner and other experts cite.

Data Exploration — Because GenAI can translate everyday language into a technical data query, people who aren’t data experts can generate insights without waiting for a data scientist to conduct analysis. This is sometimes referred to a [conversational user interface](#).

Data Summarization — Just as many people ask GenAI to summarize long documents, data users can do the same with [large datasets](#). For example, the technology can describe the overall structure, identify the types of data involved and highlight key statistics.

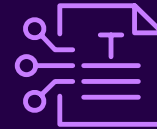
Report Generation — For many people, translating analysis into an engaging report can be a slog. GenAI can [draft a report](#) that provides basic analysis and relevant charts and graphs, freeing up the analyst or subject-matter expert to focus on interpreting the data and its ramifications.

Synthetic Data Generation — Government agencies that work with sensitive or personally identifiable information struggle to find data to train AI models without compromising security or privacy. In some cases, they can use GenAI to anonymize that data or [create artificial data](#) that resembles but does not mirror existing datasets.

Code Generation — GenAI has proven adept at [writing code](#) based on natural language prompts. Besides empowering users who have no programming expertise, this can save data engineers time in writing code and designing processes.

GenAI Use Cases in Analytics

A [Gartner study](#) predicts that these will be the most popular data-related GenAI use cases in the next two years:



Data narrative generation
(41%)



Conversational interfaces for analytics
(39%)



Data exploration
(39%)



Productivity (e.g., generating summaries)
(35%)



Code generation
(33%)

How to Take the Next Step Toward Data Resilience



Greg Carl

Principal Technologist for U.S. Federal at Pure Storage

WATCH VIDEO

Data security continues to evolve. In recent years, many government agencies have shifted from perimeter-based security to a zero-trust network architecture, with a focus on identity-based access management. While important, this strategy does not address the full scope of data security requirements.

The next step is to evolve from security to resilience by developing a data-centric architecture that aligns with the full data life cycle, said Greg Carl, Principal Technologist for U.S. Federal at Pure Storage.

In this [video interview](#), Carl discusses how agencies can take a more holistic approach to protecting their data. Topics include:

- The risks created by the proliferation of data management tools
- The benefits of adopting a tiered resilience model
- Key metrics for your assessing business continuity/disaster recovery plans

About Pure Storage

Pure Storage® delivers the industry's most advanced data storage platform to store, manage, and protect government data at any scale. With Pure Storage, agencies gain ultimate simplicity and flexibility, enabling them to save time, reduce costs, and increase operational efficiency. From AI to archive, Pure Storage provides a modern cloud experience through one unified Storage as-a-Service platform across on-premises, cloud, and hosted environments. Built on Evergreen architecture, the platform evolves with an agency's mission — continuously improving with zero planned downtime, guaranteed. Federal agencies are enhancing their capacity and processing power while significantly reducing their energy consumption and carbon footprint. Trusted across civilian and defense agencies, Pure Storage is proud to prioritize customer success, as reflected in its industry-leading Net Promoter Score.

[Learn more about Pure Storage](#)

“It’s not about perimeter-based security, like putting a moat around the castle. It’s about protecting every single aspect of our data beginning from inception all the way to decommission.”

— Greg Carl, Principal Technologist, US Federal, Pure Storage



GIS Reshapes Strategic Planning in Fairfax County, Virginia

By many measures, Fairfax County, Virginia, is big and diverse. Its population of close to 1.2 million (49% white, 20% Asian, 16.5% Hispanic and 10% Black) is spread across 406 square miles of suburban, urban and rural areas, according to [county data](#). The median household income is \$150,113, as per [Census data](#), with slightly more than 6% living at or below the poverty line.

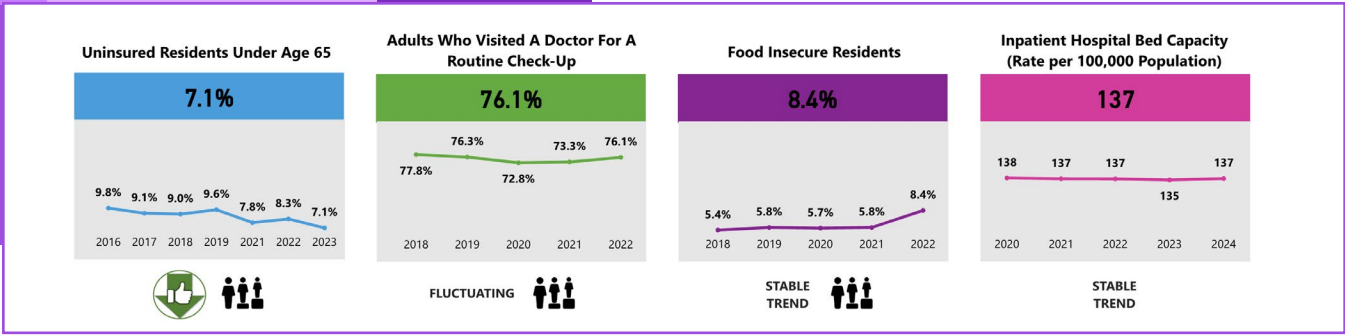
When officials began developing the first [Countywide Strategic Plan](#) in 2019, those data points loomed large. Given the county's size and diversity, how do you assess the outcome of any given program?

The strategic plan aligns with [One Fairfax](#), a policy adopted in 2017 that directs agencies to identify and remove barriers to racial and social equity, said Aimee Brobst, Director for the Countywide Strategic Plan in the Department of Management and Budget (DMB).

"It's not just, 'How is a program faring?'" Brobst said, "but 'Is it doing what we need it to do to serve the community — not just for the county on average, but in these specific places and populations that we're looking to target as part of our One Fairfax policy?'"

This is why GIS plays a critical role in the strategy. GIS enables county officials, constituents and other stakeholders to explore how key metrics play out across different regions of the community, said Michelle Gregory, Division Director for Countywide Data Analytics in the department.

"We're trying to make sure that wherever possible, we have maps that are interactive," Gregory said. "Yes, we can have a general map to provide some insights, but where possible, providing that ability for a tool to be interactive allows the user [to] capture the information that they're looking for in the way that's most helpful for them."



A Data-Centric Plan

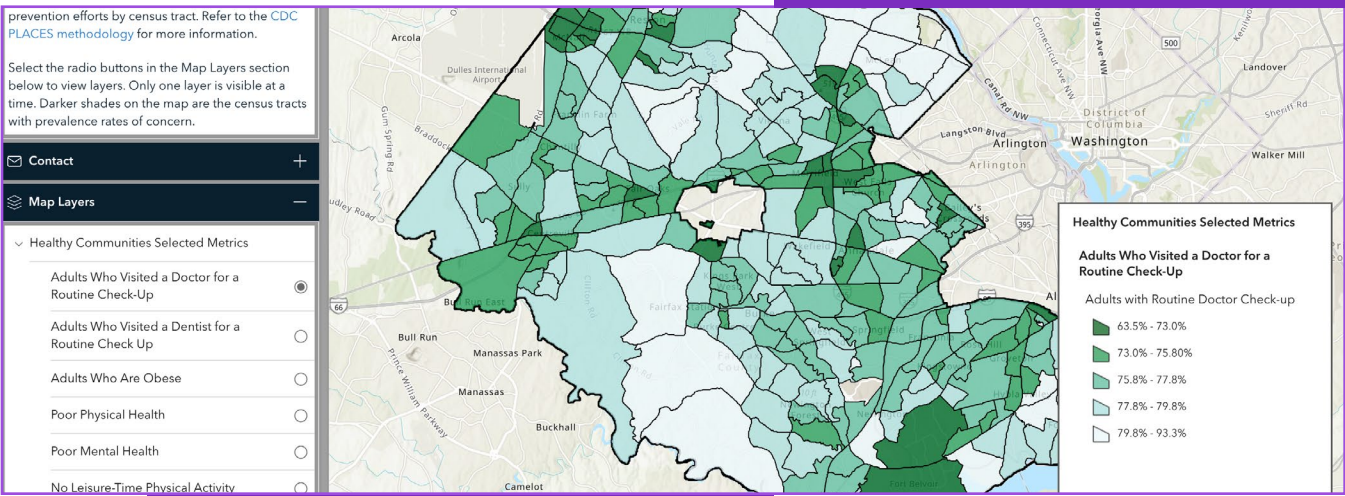
The plan, published in 2021, identifies goals and metrics around 10 “community outcome areas” that reflect county residents’ top concerns, such as cultural and recreational opportunities, economic opportunities, environment, health, and others.

Metrics are key. In carrying out the plan, the county is shifting to “a data-informed, strategic process that allows us to prioritize actions that will have a direct and positive impact on the people and neighborhoods who need it most,” writes County Executive Bryan Hill in the introduction to the [2024 Countywide Strategic Plan Annual Report](#).

In practice, the county is developing [dashboards](#) to track basic “headline” metrics around each of the 10 outcome areas. For example, the [Healthy Communities](#) dashboard shows 15 headline metrics, including the ratio of residents to primary care physicians, dentists and mental health providers; the percentage of uninsured residents; and the percentage of residents who experience food insecurity.

In many cases, the charts include an equity icon, showing three people of various heights. By clicking on it, users can see how the data breaks down by race, ethnicity and geography. Additional pages provide more details on the headline metrics, including narratives that put the data in context, plus interactive maps that enable users to drill down further by geography. As of February 2025, the county provided dashboards for five of the 10 outcomes, with the others in the works.

Additionally, the county has developed a [Communities of Opportunity Dashboard](#), which looks at different regions of the county (based on Census tracts) through two lenses: “Access to Opportunity” indicators, such as home ownership, employment, education and access to public transportation, and “Wellbeing” indicators, such as heat vulnerability, health insurance coverage and median household income (see next page).



Data for the People

As these dashboards take shape, they are playing a critical role in helping county officials carry out the strategic plan, said Brobst. Often, when officials meet to discuss possible actions, they will invite Gregory and her team to share relevant data and explain its context.

“If we’re talking about health, if we’re talking about public safety, we want to have a data case — data analysis, data insights — to be a key part of that,” Brobst said. “And that’s the expectation that our leadership has.”

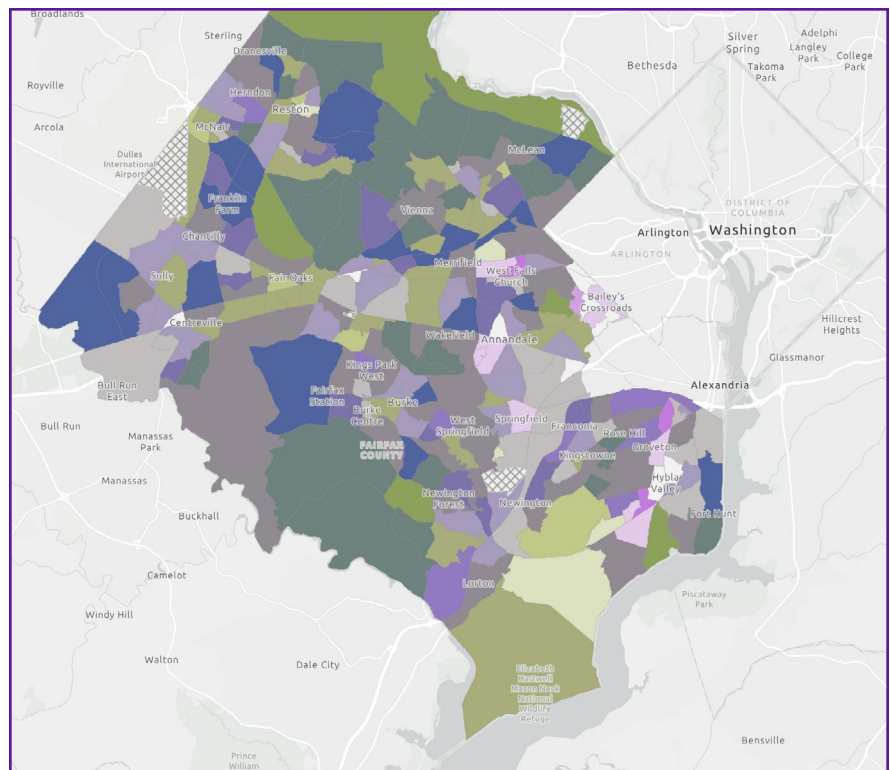
In the same vein, GIS makes this data readily accessible to constituents who otherwise might lack data skills. The dashboards summarize and contextualize data in a way that makes it easy to understand and navigate, said Gregory.

“We’re in this age where people want information, and they want it to be quick and be able to dive into it and digest it quickly,” she said. “The goal is to try to make sure that these dashboards allow for that, so that people can consume information at the level that they see fit for them and their needs.”

More to Come

Gregory said her team continues to look for new ways to incorporate GIS data into strategic planning and to get more value out of it. “We’re not only looking at maps, but at applications that give us a better ability to manipulate that data from a variety of [angles].”

But the real story is not about the technology. It’s the way in which data has become so integral to the planning process, said Brobst. “It’s great to have tools like this, but I think what’s more exciting for us is that we are actively and continuously using them.”



Communities of Opportunity Dashboard

Colors reflect performance against key indicators of opportunity and well-being. For example, white regions rate low against both sets of indicators, while green indicates areas of low opportunity but high well-being.

Conclusion

This guide is Part 1 of our 2025 data guide series. But be sure to check out our 2024 guides. Here are some highlights:

Part 1: A Fresh Look at Data

- ▷ Guiding principles and strategies for maximizing your data's potential
- ▷ Proven tactics for using data to facilitate policymaking

Part 2: Leveling Up With Data

- ▷ Case studies in using data to improve delivery in health care, procurement, emergency response and other fields
- ▷ Data governance best practices from an in-the-weeds government official

Part 3: Strength in Numbers

- ▷ Case studies in making key government functions more data-driven
- ▷ Using data for the public good

Preregister for the next guides in our 2025 Data Series >

About GovLoop

GovLoop's mission is to inspire public-sector professionals by serving as the knowledge network for government. Govloop connects more than 300,000 members, fostering cross-government collaboration, solving common problems and advancing government careers. GovLoop is headquartered in Washington, D.C., with a team of dedicated professionals who share a commitment to the public sector.

For more information about this report, please reach out to info@govloop.com.

govloop.com | [@govloop](https://twitter.com/govloop)

Thank You

Thank you to Pluralsight and Pure Storage for their support of this valuable resource for public-sector professionals.

Authors

Candace Thorson, Managing Editor

John Monroe, Director of Content

Lauren Walker, Senior Staff Writer

Designer

Kaitlyn Baker, Senior Creative Manager

Kelly Boyer, Motion Graphics Team Lead

