

Connecting the people and processes
involved with data to promote
communication between, and
integration of, formerly siloed data
across executive branch agencies

Connecticut State Data Plan

Submitted pursuant to C.G.S.
Sec. 4-67p

Office of Policy and Management, November 1, 2024

Introduction2

 Mission3

 Implementation.....3

 About the Process.....4

Principles.....4

Progress and Lessons Learned5

Goals for 2025 – 20267

 Access to data7

 Building agency capacity.....9

 Using data for decision-making.....10

 Artificial intelligence.....13

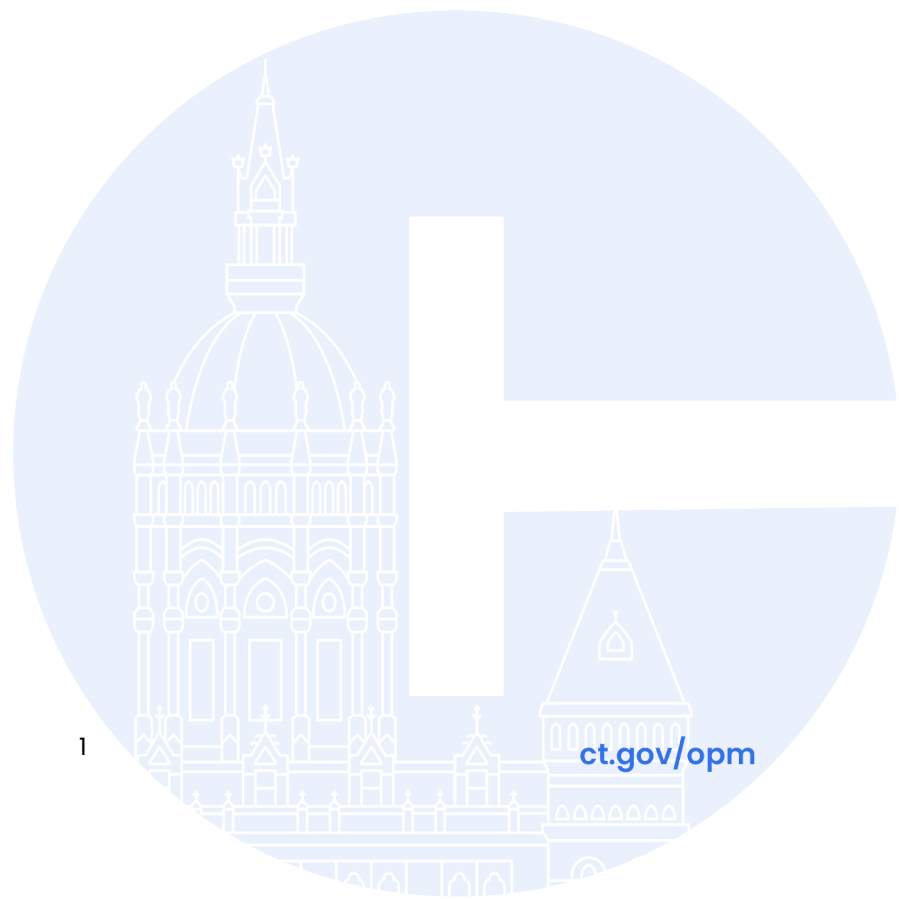
Metrics13

Appendices.....14

 Appendix A: Glossary15

 Appendix B: Results of Survey to Agency Leadership17

 Appendix C: State Agency Leadership Meeting Summary Report19



Introduction

The Connecticut State Data Plan serves as a framework for the state's executive branch agencies to engage in a consistent approach to data stewardship, use, and access. The State Data Plan is applicable to all data in the custody and control of executive branch agencies and covers open data, data sharing and data analysis, in accordance with [C.G.S. § 4-67p](#). Specifically, the state data plan shall:

- establish **management and data analysis standards** across executive branch agencies;
- include **specific, achievable goals** within the two years following the adoption of such plan, as well as longer term goals;
- make **recommendations to enhance standardization and integration of data systems and data management practices** across executive branch agencies;
- provide a timeline for the **review of any state or federal legal concerns or other obstacles to the internal sharing of data among agencies**, including security and privacy concerns; and
- set goals for improving the **open data repository**.

The statutes also require *Information technology-related actions and initiatives of all executive branch agencies, including, but not limited to, the acquisition of hardware and software and the development of software* to be consistent with the plan. The plan is required to be updated every two years. The draft plan is organized as follows:

- **Principles** represent a framework for state agencies to organize their use of data;
- **Goals** describe the desired outcomes of plan implementation and the concrete implementation steps to achieve those outcomes;

The 2025 - 2026 plan includes the following specific, achievable goals for the next two years. The goals are based directly on feedback from agency data officers, agency leadership and data users within and outside government and echo the priorities in Appendix C:

Goal 1: Develop a centralized hub to provide access to existing online data-related resources, including tools, software, and training materials.

Goal 2: Improve metadata and documentation for accessibility and interoperability.

Goal 3: Conduct a needs assessment to identify additional data-related training and resource requirements.

Goal 4: Offer guidance and support to agencies in hiring qualified data analysts.

Goal 5: Establish clear and consistent statewide guidelines for data sharing across agencies.

Goal 6: Develop data governance and data quality policies to support responsible use of AI

The goals reinforce each other and are covered in more detail below.

Mission

The mission of this plan is to connect people, processes and technology to ensure the safe, ethical and secure management, analysis and use of data by executive branch agencies. The plan describes a general framework for management and use of data through the Mission and Principles, with concrete, achievable Goals, clear Metrics and steps for Implementation.

State agencies collect and manage data to operate programs and services and have used this information for reporting for decades. Only recently have state officials begun to harness administrative records for analytical purposes to achieve desired goals or outcomes though.

Implementation

The plan supports agency efforts and brings those efforts into alignment to provide increased efficiency in the use of limited resources for data stewardship, use, and access. The plan does not require agencies to dedicate additional resources to its implementation, nor does it generally require agencies to begin collecting or creating new data. Implementation of the plan relies on the Chief Data Officer, the efforts of individual Agency Data Officers, and support from the Office of Policy and Management and leadership of executive branch agencies.

In addition to establishing the framework for the State Data Plan, [C.G.S. § 4-67p](#) establishes the position of Chief Data Officer within the Office of Policy and Management, and requires each executive branch agency to designate an agency data officer, with roles and responsibilities outlined below. Further, the statute requires that each agency conduct an inventory of its “high-value data” annually and formulate plans to enhance the availability of open data, known as “open data access plans.” The following are the roles and responsibilities of the Chief Data Officer and Agency Data Officer that are established by [C.G.S. § 4-67p](#):

Chief Data Officer:

- Directing executive branch agencies on the use and management of data to enhance the efficiency and effectiveness of state programs and policies;
- Facilitating the sharing and use of executive branch agency data (A) between executive branch agencies, and (B) with the public;
- Coordinating data analytics and transparency master planning for executive branch agencies;
- Creating the state data plan; and
- Providing a procedure for each agency head to report regarding the agency's progress toward achieving the plan's goals.

Agency Data Officer:

- Coordinating and submitting agency high value data inventories annually;
- Coordinating and submitting agency open data access plans;
- Serving as the main contact person for inquiries, requests or concerns regarding access to the data of such agency; and
- Establishing procedures to ensure that requests for data that the agency receives are complied with in an appropriate and prompt manner in consultation with the Chief Data Officer.

[C.G.S. § 4-67p](#) requires Agency Data Officers to establish procedures to ensure that requests for data that the agency receives are complied with in an appropriate and prompt manner, as well as establish open data access plans, but does not create a separate entitlement or an alternative to the Freedom of Information Act process.

About the Process

The State Data Plan is developed through an iterative process, with opportunities for both public and agency input as follows. The timeline and methods for data collection are as follows:

Draft Plan – A draft of the State Data Plan must be presented to the Data Analysis Technology Advisory (DATA) Board pursuant to [C.G.S. § 4-67p](#) and [C.G.S. § 2-79e](#), by November 1, with a public hearing held within 30 days.

Stakeholder Engagement – Development of the plan involved outreach to executive branch agency staff, the DATA Board and data users within state agencies. We sought broad stakeholder engagement and distributed information about the plan and surveys to education, nonprofits, private sector and local government. Agency Data Officers completed surveys and met to discuss agency priorities in August 2024. Agency leadership completed the same survey and met to discuss priorities for data in September 2024. A summary of the agency priorities and survey results is included in the appendices. The DATA Board met in October 2024 to review results from agency outreach. Survey respondents were asked if they were interested in a follow-up conversation and staff from the OPM Data and Policy Analytics division coordinated to meet with anyone that expressed interest. Copies of the survey responses and summary results are included in the appendices.

Final Plan – The final plan must be issued by December 31, 2024. The final plan will incorporate feedback from the public hearing, scheduled for November 15, 2024 and additional outreach through the agency and external data users surveys.

Principles

The Principles below are a framework for agencies to mature and enhance their management, use, sharing, and analysis of data. The Principles remain the same as in prior State Data Plans to serve as guidelines for continuous improvement and inform decision-making on an ongoing basis.

1. Appreciate that the data we collect require a considerable investment of resources and have value beyond the purpose for which they are collected.
2. Perform continuous data quality and analytics improvement to ensure the value of data is protected and maximized.
3. Ensure standardized data governance in order to protect data and improve its quality and utility.
4. Create, acquire, use, and disseminate data deliberately and thoughtfully; in compliance with federal law and state statute, and considering, quality, consistency, privacy, equity, client data dignity, value, reuse, and interoperability from the start.
5. Coordinate and prioritize data needs and uses, utilize data from multiple sources, and acquire new data only when necessary.
6. Protect individual privacy and maintain confidentiality using effective data stewardship and governance, and by maintaining modern data security practices and technology.
7. Employ ethical standards in the use, analysis, sharing, and integration of data to avoid intrusion into the lives of Connecticut residents, and disparate impact.
8. Improve data sharing and access with ongoing input from users and other stakeholders, including those whose personal and protected data are collected in state agency systems.
9. Create clear and predictable pathways for data sharing which are necessary for effective data use and sharing.
10. Manage a data asset one time and use it for multiple purposes, to the extent possible, within legal and regulatory constraints.
11. Promote a culture of continuous and collaborative learning, with data and about data.
12. Embrace openness, transparency, and accountability.

Progress and Lessons Learned

Agency leadership reported successes from 2023 – 2024 in the following areas:

Survey Results Summary: Reported Successes

Successes building staff capacity to use data themes	Successes improving access to data	Successes using data to make decisions
<ul style="list-style-type: none"> • Hiring qualified staff into key roles • Developing skills of junior staff 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Managing call wait times, real-time data to the 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Reviewing call wait times to identify areas for improvement

<ul style="list-style-type: none"> • Acquiring additional resources to grow area • Access to Fellows to increase support • Investing in trainings and technical assistance for staff • Investing in new tools 	<ul style="list-style-type: none"> public via maps, dashboards • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication 	<ul style="list-style-type: none"> • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication
---	---	--

The OPM Data and Policy Analytics (DAPA) division worked to support the 2023 – 2024 plan implementation and to improve policies and governance for data, by collaborating with state IT and policy leadership to launch the [responsible AI framework](#) for Connecticut and organizing the state response to [federal changes in race / ethnicity reporting](#). We developed a new [high-value data inventory process](#) to catalog agency data usage and integrated the [CT Geodata Portal](#) with the [Open Data Portal](#). We conducted user surveys through both portals to understand how people interact with data and the challenges they experience and, based on survey results, updated metadata, reducing datasets missing metadata from 94% to under 20% today.

Connecticut was awarded honorable mention in 2024 for building and using evidence and data to improve residents' lives from the [Results for America State Standard of Excellence](#). We developed a draft evaluation policy with Results for America and J-PAL North America and created a new [Research and Evaluation unit](#) within DAPA. New reports and analysis included a [housing and segregation study](#), a [Housing Data Hub](#), and a [report on streamlining affirmative action plans](#).

Connecticut's work integrating data was recognized as an [Advanced/Leading state for work on data sharing in the Strada Foundation State Opportunity Index](#). Data from P20 WIN, Connecticut's longitudinal data system, was highlighted in a [Postsecondary Labor and Earnings dashboard](#), research on the [Roberta Willis state financial aid program](#), reporting on [Disconnected and At-Risk Youth in Connecticut](#), and analyses of college and career readiness and outcomes for special education students. P20 WIN was supported with new federal grants for [Workforce Data Quality Initiative](#) and the [2023 State Longitudinal Data System](#) grant to establish a research collaborative to conduct research related to agency policy priorities.

The GIS Office completed the first [GIS Strategic Plan](#) in May 2024, and managed development of high-resolution imagery, elevation data, and additional products from statewide flights for publication in 2024. We collected and processed municipal digital parcel and computer-aided mass appraisal (CAMA) files into a single [statewide parcel dataset](#) now available on the [CT Geodata Portal](#). Additional GIS resources include updated [broadband maps](#), collaboration with DEEP's Bureau

of Energy and Technology Policy on the [Broadband Equity, Access, and Deployment \(BEAD\) Program](#); with the Department of Public Health on a [web mapping application](#) to visualize local health departments; and updated [municipality](#), [county](#), [Council of Governments](#), and [planning region boundaries](#).

In March 2024, Connecticut celebrated [Open Data Day and 10 years of the Open Data Portal!](#) To commemorate the anniversary, we convened agency data users and developed [open data screencasts](#) to encourage data use, and guidance on accessible data visualizations for the [Open Data Handbook](#). Agency capacity-building included [workshops](#) on data literacy, data storytelling, evaluation, and data analysis; an R Users Group with over 100 state employees; and training on data wrangling and storytelling to 57 agency staff from 16 agencies.

New and updated Open Data resources included an [ARPA Funding and Project Inventory Dashboard](#), with analyses of state spending and visualizations by focus area and agency and data stories for [Juvenile Justice Policy Oversight Council Equity Dashboard](#), [supplier diversity data](#), [HVAC indoor air quality grants for public schools](#), [sales and use tax data](#), [personal income tax data](#), [highway use fees](#), [cannabis prices](#), [amusement park rides registered](#), [CT Prescription Drug Monitoring Program](#), [CT Business Registrations](#), [State of CT Executive Branch Workforce](#), [Municipal Fiscal Indicators](#), [Census Tracts Identified for PA 23-205](#), and [Disproportionately-Impact Areas for Cannabis Legalization](#).

A theme from the surveys was to increase public awareness of existing resources – one respondent recommended “Increase public awareness of the state’s data portals so residents and people working within state government know they can access the data available.” To raise awareness, we have begun GIS and DAPA newsletters, which cover recent activities and new data and reports.

Goals for 2025 – 2026

The goals for 2025 – 2026 are based on the stakeholder outreach and survey results. The surveys and agency leadership feedback identified “low-hanging fruit” – goals that are achievable because they involve raising awareness of existing resources or consolidating fragmented resources.

To track progress on the goals, we will continue to hold quarterly meetings with all agency data officers and regular meetings with agency leadership and the DATA Board.

The goals are grouped into the three categories used in the previous 2023 – 2024 plan:

Access to data

Goal 1: Develop a centralized hub to provide access to existing online data-related resources, including tools, software, and training materials.

Both agency staff and external data users highlighted the desire to make it easy to find the right data. Almost 70% of respondents to the external survey said finding the correct dataset was a barrier to using state data and over 60% said knowing where to find data was a barrier.

Most data users said they access data directly from state websites, compared to the Open Data or Geodata Portals. Several comments from data users reinforced interest in a centralized portal:

*Have a **directory that provides links** to all CT data sources including DOT, DEEP and CT ECO. The several locations and database differences can be confusing for casual users and frequent users wont need several bookmarks.*

*A description on the CT State website of information that can be located or searched and from what web site or source where it can be found. A **directory of sorts would be helpful**. Especially when searches are started.*

Centralize data and make it easy to download with minimal jumping between websites.

Have **pages for all data sources on specific populations**, such as children, veterans, etc.

A **centralized list of databases with links** (including all state-funded or -populated datasets); databases which synthesize disparate datasets (i.e HMIS homeless population figures and locations X rates/concentration of cost-burdened renters).

*I do not use CT Open Data Portal because it is **not very well organized or easily navigable**. I find it easier to find the necessary data on the agency (DPH, DMHAS, etc) data portals.*

The goal is not to put all data (or all resources) into one place. Creating copies of data or resources goes against the Principles in the plan. A hub or directory would provide links to existing resources but would not replace or duplicate those resources.

As a related improvement, we plan to increase use of enterprise data platforms, which provide increased access but also greater security and control over management of shared data resources.

Goal 2: Improve metadata and documentation for accessibility and interoperability.

A second set of comments reflected a need to improve documentation and metadata for online data resources. Improved metadata supports data use and interoperability. Sample user feedback includes:

*Sometimes we **need better description of how the data was collected and what was done to it to ensure that data is credible** (lots of skeptical audiences out there!). Also - somewhere there should be a better **breakdown of terms** used and what they mean*

Data standards are needed across the state

*It would be **helpful to have and utilize metadata standards** that allow the data user to understand what the data set is, where it comes from, and who manages it in case there are follow up questions. I have found that this information is difficult to find.*

A metadata audit in 2023 focused on completing missing metadata on the open data portal, but there is still room for improvement to create consistency across different data resources. Providing additional data on how data was collected can help identify limitations or potential bias in data for data users. At the same time, agencies often adopt specific metadata standards to meet their own operational needs and there is not currently a 'one size fits all' solution for the state.

Along with creation of the hub, we will revisit and refine the "open data access plans," required under C.G.S. Sec. 4-67p. Open data access plans are intended to provide transparency about agency plans to publish data. A new approach to the open data access plans can leverage improvements made to the high-value data inventory and encourage agencies to take a strategic, rather than compliance, approach to publishing data. We will also continue related work to improve the open data and geodata portals, including regular reviews of data freshness, audits of accounts and access for data users and further alignment of standards and data publication guidelines for spatial and non-spatial data.

Building agency capacity

Two related themes from stakeholder engagement focused on building agency capacity - one to provide training for agency staff, and the other to provide agency support to attract and retain qualified staff.

Goal 3: Conduct a needs assessment to identify additional data-related training and resource requirements.

A needs assessment will focus on agency priorities for training and capacity building, from agency data officers, analysts, managers and agency leadership to develop a concentrated program for agency staff. The assessment will review resource requirements, onboarding processes and the preferred frequency and modality for training. External data users have also requested more training and communications about data, how to access different resources and clarity on handling requests for data.

Goal 4: Offer guidance and support to agencies in hiring qualified data analysts.

In addition to training and capacity-building, agencies expressed that attracting and retaining qualified staff is a critical challenge. Job classifications and job specifications need to be kept up to date with current trends and career pathways and compensation established to remain competitive in the job market. Fellowship and intern programs, from data science or public sector-focused programs have proven to be a good option and increasing awareness and building new relationships is an additional strategy to augment agency capacity.

Using data for decision-making

Goal 5: Establish clear and consistent statewide guidelines for data sharing across agencies.

Data sharing increases the value of data assets by allowing agencies to draw richer insights from integrated data. Data sharing supports interdisciplinary efforts, like supporting environmental justice or social determinants of health, by bringing together disparate datasets, including spatial data. At the same time, it is critical to put in place legal and technical guardrails to ensure safe, ethical and secure data sharing.

We are at a critical moment to be able to expand state analytic capacity and have a mandate, under C.G.S. Sec. 4-67p, to “enhance standardization and integration of data systems and data management practices across all executive branch agencies,” as part of the State Data Plan. Agency leaders have asked clearly for “clear and consistent statewide guidelines for data sharing across agencies” in the next two years.

We have the authority to use data “to enhance the efficiency and effectiveness of state programs and policies” and “coordinating data analytics and transparency master planning for executive branch agencies” through the responsibilities of the Chief Data Officer, under C.G.S. Sec. 4-67p. Data sharing and increasing data interoperability supports both priorities.

Through the Preschool through 20 Workforce Information Network (P20 WIN), we have an existing “state data system for the purpose of matching and linking longitudinally data of state agencies and other organizations to inform policy and practice for education, workforce and supportive service efforts,” established under C.G.S. Sec. 10a-57g. While there are gaps, this system covers cradle-to-career data sharing and using that data to inform decision-making.

Consistent with these goals and responsibilities, we plan to review executive branch agency opportunities for data sharing and to propose solutions that address multiple related, but distinct, high-profile use cases.

Concrete and specific use cases for data sharing span the cradle-to-career continuum. These use cases include, but are not limited to:

- Linking unique child and family information within early childhood, health and family support data systems in real time
- Development of an integrated workforce data system to enable access to up-to-date, high-quality workforce training and outcomes data to drive policy, funding, and programming
- Residents will have an easier time accessing and using state service through a future, centralized Service Center experience. The Service Center will be a streamlined, one-stop digital hub for residents to easily interact with the State for services: renew licenses and make payments, set up appointments, get notifications, and stay informed.

- Developing a secure online portal to facilitate cross-agency data sharing for basic critical information and to improve service delivery for individuals with intellectual and developmental disabilities
- Measure and improve understanding of the number and nature of at-risk and disconnected youth across state and local levels and build state-level data infrastructure, including a data platform and intermediary to support secure data and information sharing across agencies, Youth Service Bureaus, and community providers
- Develop a population health panel of people who are incarcerated to track outcomes post-release, to support a Medicaid demonstration waiver to better serve both youth adults who are incarcerated as part of a multi-state agency initiative

These use cases may not have a common solution, but approaching each piecemeal ensures a fragmented response. Expanded analytical capacity and data sharing can build upon each of the following:

- Development of a hybrid cloud-based solution, based on recommendations from the Data Integration Support Center (DISC), that centralizes the platform for matching, analysis and data movement for the state longitudinal data system, P20 WIN, using a CT-built environment and the Coleridge Initiative Administrative Data Research Facility (ADRF).
- Current and proposed technology modernization projects for agencies, potentially including updates to the Comprehensive Child Welfare Information System (CCWIS), the CT Medicaid Enterprise Technology System (CT METS) Enterprise Data Warehouse and Analytics platform (EDW&A), the unemployment insurance (UI) system and others.

Expanding data sharing will require resources to build on existing investments and expand analytical capacity and data sharing efforts, in the following categories:

- People: the staff required, directly and indirectly, to expand analytical capacity and data sharing, including roles for analytics, technology, legal, communications and related functions. This can and should also include efforts to build partnerships with external research and analytical capacity, including the potential for applied data analytics training and a research collaborative involving local institutions of higher education.
- Process: the related policies and procedures, whether they are enterprise data or technology policies, or specific to individual initiatives, including results from the annual review of legal obstacles to interagency data sharing, completed by OPM in accordance with C.G.S. Sec. 4-67z and a forthcoming review of consent management requirements for P20 WIN agencies
- Technology: the technical infrastructure required, including any documentation or training required for staff to implement the solutions

Sustained funding for data sharing can leverage the mechanisms identified in the Field Guide for Financing Integrated Data Systems, developed by the National Association of Public Administrators, including:

- Cost allocation in federal funding and grants, including review of options based on the changes in federal uniform grant guidance, effective October 2024
- Statewide cost allocation plans
- Existing direct support from federal grant programs, including from Departments of Education and Labor, or matching funds available for state Medicaid agencies
- State information technology capital investment funds

Further improvements can be made by developing a research collaborative for partnerships with academic researchers and evaluators. To engage and understand researcher interest in working with Connecticut administrative data, Yale Tobin Center for Economic Policy conducted interviews, focus groups and surveyed researchers from 15 different institutions of higher education in Connecticut, and presented results to the P20 WIN Data Governing Board in May 2024. The survey identified the following challenges and opportunities working with external researchers:

- 1) **“Sometimes a direct phone number would be helpful.”** Researchers commonly reported uncertainty about pathways to access CT data and that non-participation of some state agencies was reported as limiting the value of the P20 WIN system. Centralizing information on pathways to data access and expanding P20 WIN participation are opportunities here.
- 2) **“We need more opportunities to brainstorm with agencies about salient questions.”** Alignment between researchers and policymakers is critical and researchers reported challenges in understanding and aligning to state priorities. Creating learning agendas and holding focused events are steps to address this concern.
- 3) **“Finalizing the [legal] agreement was a challenge.”** Navigating the complexity of legal agreements remains an obstacle, particularly meeting state data privacy and security requirements. Standardized privacy and security guidance and a secure data enclave – both underway – will help this set of challenges.
- 4) **“The main challenge was timeline.”** The time it takes to access data directly through agencies and through P20 WIN were all expressed as barriers. Regularly repeating data matches through P20 WIN will help along with other measures to reduce manual operations and agency burden.
- 5) **“After a significant investment of time, the data was deemed unusable.”** Data quality is a constant question for researchers looking to work with state data. Analysts want upfront clarity on data integrity before embarking on a lengthy access process. Policies or standards for data quality and transparency on data quality issues is critical here.

The goals of the State Data Plan provide guidelines for data sharing, centralized access to data and clarity on data quality, all of which support research partnerships and a research collaborative.

Artificial intelligence

Goal 6: Develop data governance and data quality policies to support responsible use of AI

Artificial intelligence (AI) was the top development that agency leadership and agency staff anticipated would impact their agencies in the next two years. Critical building blocks for responsible use of AI are policies for data governance and data quality and agency staff identified data governance as a top need. Training and capacity-building for agency staff to make responsible use of AI is also critical. Agency leadership wanted to see collaboration on AI applications and use cases:

Artificial Intelligence (AI) Applications: There is significant interest in using AI for operational improvements, such as call center routing and data analysis, with examples from other jurisdictions illustrating its potential benefits.

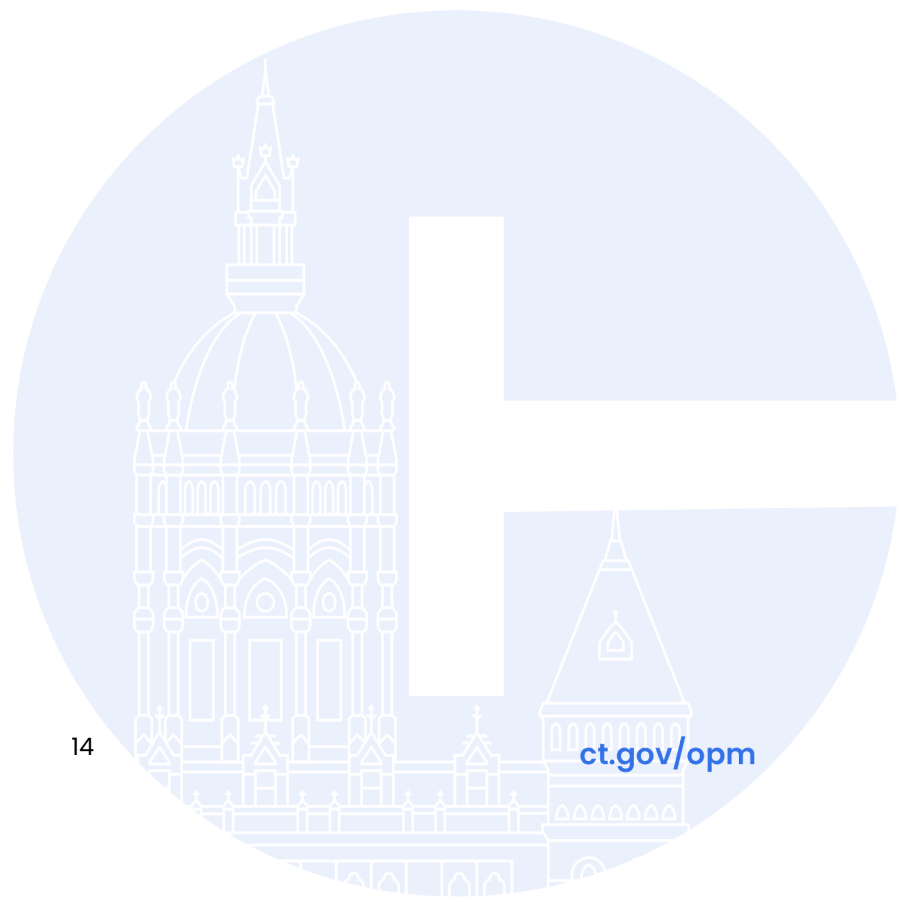
Collaboration on Use Cases: The need for collaborative identification of multi-agency use cases for AI was discussed, highlighting the potential for shared benefits.

The state's Responsible AI framework proposes creation of an AI advisory group and an AI Engagement Lab for agencies. Expanding use cases and developing the Lab requires agencies to have clear governance for data assets and an understanding of data quality issues that may impact functioning of AI solutions. Establishing data governance and data quality policies will provide foundational support for future focus on data management and analysis with AI.

Metrics

Metrics for the goals listed above will be included in the final report after review and revisions by the DATA Board.

Appendices



Appendix A: Glossary

Data: The final version of statistical or factual information that: (A) is reflected in a list, table, graph, chart or other non-narrative form that can be digitally or nondigitally transmitted or processed; (B) is regularly created or maintained by, or on behalf of, an executive branch agency; and (C) records a measurement, transaction or determination related to the mission of the agency or is provided to the agency by third parties pursuant to law.

Data dignity: The data subject's human dignity, legitimate interests and fundamental rights, with particular regard to the transparency of processing, or the transfer of personal data.

Data governance: The overall management of the availability, usability, integrity, quality, and security of data. The creation of policies, roles and responsibilities ensures clarity and consistency regarding the purpose, use and presentation of data.

Data lifecycle: The different stages from data collection to use. Definitions vary, but can include: planning, data collection, data access, use of algorithms and statistical tools, data analysis and reporting and dissemination,⁵ or acquisition, conception, instrumentation, collection, processing and analysis, dissemination and disposition.⁶

Equity: Access to opportunities and participation in daily living and resultant outcomes are accorded fairly and justly, and not dictated or encumbered by one's race, ethnicity, sex, sexual orientation, gender identity, gender expression, body type, phenotypical ascriptions, ability or religion.

Executive branch agency: Any agency listed in section 4-5 of the general statutes.

High value data: Any data that the department head determines (A) is critical to the operation of an executive branch agency; (B) can increase executive branch agency accountability and responsiveness; (C) can improve public knowledge of the executive branch agency and its operations; (D) can further the core mission of the executive branch agency; (E) can create economic opportunity; (F) is frequently requested by the public; (G) responds to a need and demand as identified by the agency through public consultation; or (H) is used to satisfy any legislative or other reporting requirements.

Open data: Any data that (A) is freely available in convenient and modifiable format and can be retrieved, downloaded, indexed and searched; (B) is formatted in a manner that allows for automated machine processing; (C) does not have restrictions governing use; (D) is published with the finest possible level of detail that is practicable and permitted by law; and (E) is described in enough detail so users of the data have sufficient information to understand (i) the strengths, weaknesses, analytical limitations and security requirements of the data, and (ii) how to process such data.

Public data: Any data collected by an executive branch agency that is permitted to be made available to the public, consistent with any and all applicable laws, rules, regulations, ordinances, resolutions, policies or other restrictions, requirements or rights associated with the data, including, but not limited to, contractual or other legal restrictions, orders or requirements.

Protected data: Any data the public disclosure of which would (A) violate federal or state laws or regulations; (B) endanger the public health, safety or welfare; (C) hinder the operation of the federal, state or municipal government, including criminal and civil investigations; or (D) impose an undue financial, operational or administrative burden on the executive branch agency. "Protected data" includes any records not required to be disclosed pursuant to subsection (b) of section 1-210 of the general statutes.

Personal data: Any Protected data that contains personally identifiable information or protected health information.

Private data: Any Protected data that is subject to federal or state laws related to individual privacy or confidentiality.

Appendix B: Results of Survey to Agency Leadership

Results from agency leadership (executive branch agency commissioners)

Challenges: To what extent does your agency face any of the following challenges using data?¹

Budget, data availability, data quality, staff skillset, and recruiting qualified staff are reported as the top 5 challenges agencies face using data.

	Agree	Neutral	Disagree
Lack of budget	89%	5%	5%
Data availability	50%	22%	28%
Data quality	43%	32%	25%
Staff lack the appropriate data and analytical skill sets	42%	21%	37%
Inability to recruit qualified staff that know how to interpret and use data effectively	39%	16%	37%
Inability to retain qualified staff that know how to interpret and use data effectively	37%	16%	42%
Lack of executive support	12%	17%	61%
Rate of technological change is too high	16%	42%	43%

Impacts: How likely are the following to have significant impact on your work in the next two years?²

Artificial intelligence, IT optimization, data security, data privacy, and staff turnover were the top 5 predicted impacts reported.

	Likely	Neutral	Unlikely
Artificial intelligence	90%	--	10%
IT optimization	84%	11%	10%
Concerns about data security	79%	11%	11%
Concerns about data privacy	78%	11%	11%
Staff turnover / retirements	74%	21%	5%
Federal changes in race / ethnicity standards	68%	21%	11%
Other changes in state or federal law	58%	32%	11%

¹ Strongly Agree and Agree and Strongly Disagree and Disagree are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

² Very and Somewhat Likely and Somewhat and Very Unlikely are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

Which of the following would you prioritize for training or capacity-building for your agency?³

	Important	Neutral	Unimportant
Data storytelling	100%	--	--
Data visualization	95%	5%	--
Data governance	95%	5%	--
Data privacy	91%	10%	--
Data management	90%	10%	--
Artificial intelligence	81%	10%	10%
Demographic data collection	80%	15%	5%
Accessibility	79%	11%	11%
Performance management	76%	24%	--
Evaluation	76%	24%	--
Geospatial information systems (GIS)	75%	25%	--
Statistical techniques	75%	25%	--
Equity in data use	68%	32%	--
Specific tools (e.g. R, Python, SQL)	65%	30%	5%
Ethics/human subjects training	47%	48%	5%

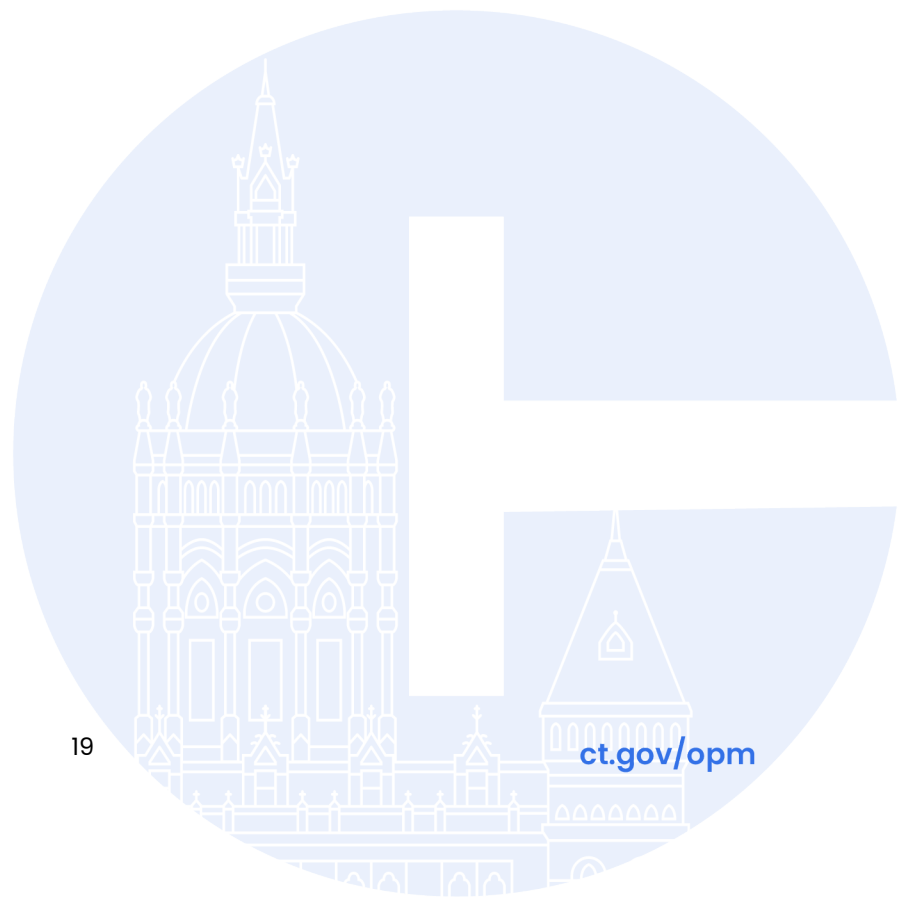
Rankings of the most important data-related activities to achieve agency goals over the next two years⁴

1. Development of training and capacity building (68%)
2. Providing access and tools and technology for data use (50%)
3. Supporting analysis research and evaluation using state data (32%)
4. Developing guidelines and standards for data use (32%)
5. Communicating the use of data by state agencies (18%)

³ Very and Somewhat Important and Somewhat and Very Unimportant are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

⁴ First and second choice rankings are combined.

Appendix C: State Agency Leadership Meeting Summary Report



State Data Plan Leadership Meeting Summary Report

Prepared by,
Amy Griffin
Nimble Consulting, LLC.

October 2024



CONNECTICUT
Policy and Management

Executive Summary

On September 18, 2024, state agency leaders convened to assist in the development of the 2025-2026 Connecticut State Data Plan (see *Appendix A for a list of attendees*). The meeting aimed to:

- Understand how agencies are currently using their data and identify challenges related to data access and usage.
- Gather feedback, recommendations, and priorities from state agency leaders to inform the new data plan.

Survey Results

Prior to the September 18th meeting, state agency leaders were asked to complete an online survey to share their agency's successes and challenges in using data to inform their work (n=22 responses). The survey also gathered insights on projected data-related training and resource needs for the next two years to guide the development of the plan. Please refer to the Appendix B for a complete list of survey results.

Survey Results Summary: Reported Successes

Successes building staff capacity to use data themes	Successes improving access to data	Successes using data to make decisions
<ul style="list-style-type: none"> • Hiring qualified staff into key roles • Developing skills of junior staff • Acquiring additional resources to grow area • Access to Fellows to increase support • Investing in trainings and technical assistance for staff • Investing in new tools 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Managing call wait times, real-time data to the public via maps, dashboards • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Reviewing call wait times to identify areas for improvement • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication

Survey Results Summary: Challenges and Needs

Top 5 data-related challenges	Top 5 anticipated data-related impacts	Top 5 data-related training or capacity-building needs
<ul style="list-style-type: none"> • Budget, 89% • Data availability, 50% • Data Quality, 43% • Staff skillset, 42% • Recruiting qualified staff, 39% 	<ul style="list-style-type: none"> • Artificial intelligence, 90% • IT optimization, 84% • Data security, 79% • Data privacy, 78% • Staff turnover, 74% 	<ul style="list-style-type: none"> • Data storytelling, 100% • Data visualization, 95% • Data governance, 95% • Data privacy, 91% • Data management, 90%

Recommendations

Based on the discussion, the following are recommendations from agency leadership to include in the next State Data Plan:

1. Develop a centralized hub to provide access to existing online data-related resources, including tools, software, and training materials.
2. Conduct a needs assessment to identify additional data-related training and resource requirements.
3. Offer guidance and support to agencies in hiring qualified data analysts.
4. Establish clear and consistent statewide guidelines for data sharing across agencies.

Existing Resources

The following resource list is provided by the Office of Policy Management, Data and Policy Analytics division in response to requests from State Agency Leaders for specific data-related resources and training. Additional resources and opportunities will be detailed in the 2025-2026 State Data Plan.

- State data plan overview: <https://portal.ct.gov/datapolicy/state-data-plan>
- Agency data officers (list and duties and responsibilities): <https://portal.ct.gov/datapolicy/knowledge-base/articles/agency-data-officers>
- Resources for agency data users on:
 - Data visualization and accessibility: <https://ctopendata.github.io/data-visualization-guidelines/>
 - Data sharing and legal issues: <https://portal.ct.gov/datapolicy/knowledge-base/articles/data-sharing-resources>
 - Inventorying / cataloging data: <https://portal.ct.gov/datapolicy/-/media/datapolicy/agency-guidance/high-value-data-inventory-guidance.pdf>
 - Metadata / documentations: <https://portal.ct.gov/-/media/ct-data/metadata-guidelines.docx?la=en>
- Data capacity building resources and videos / screencasts: https://portal.ct.gov/datapolicy/knowledge-base/articles/evaluation-and-impact/our-long-term-goals?language=en_US and videos / screencasts: https://portal.ct.gov/datapolicy/agency-guidance?language=en_US
- Responsible AI framework for agencies: <https://portal.ct.gov/opm/fin-general/policies/-/media/d13d6f704fa3408998f20e67ebda8aab.ashx>

Introduction

On September 18, 2024, State Agency Leaders convened to assist in the development of the 2025-2026 Connecticut State Data Plan (see *Appendix A for a list of attendees*). The meeting aimed to:

- Understand how agencies are currently using their data and identify challenges related to data access and usage.
- Gather feedback, recommendations, and priorities from state agency leaders to inform the new data plan.

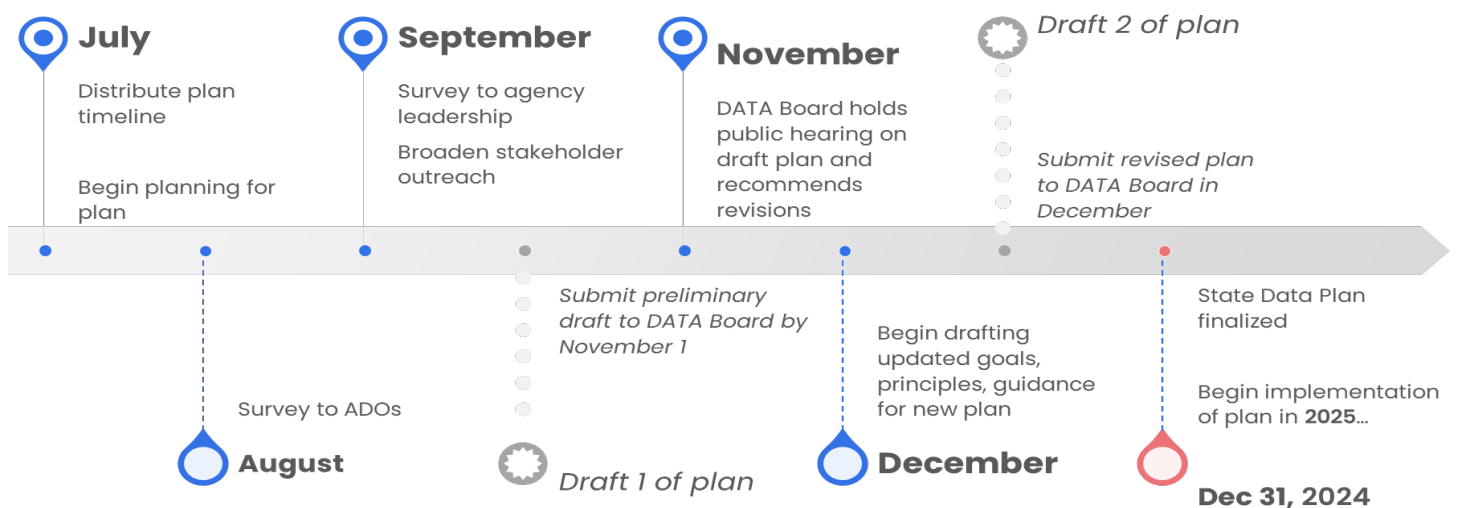
Connecticut State Data Plan

Per [Section 4-67p](#) of the Connecticut General Statutes, the State Data Plan is updated every two years to:

- establish management and data analysis standards across all executive branch agencies,
- include specific, achievable goals within the two years following adoption of such plan, as well as longer term goals,
- make recommendations to enhance standardization and integration of data systems and data management practices across all executive branch agencies,
- provide a timeline for a review of any state or federal legal concerns or other obstacles to the internal sharing of data among agencies, including security and privacy concerns, and
- set goals for improving the online repository [open data portal] established pursuant to subsection (i) of this section.

Revisions to the data plan are informed by engaging a variety of stakeholder groups including agency data officers, state leaders, and the Data Analysis and Technology Advisory (DATA) Board. This process began July 2024. A draft of the plan will be submitted to the DATA Board in November with a finalized plan slated for December 31, 2024.

2025-2026 State Plan Development Timeline



Results

Prior to the September 18th meeting, State Agency Leaders were asked to complete an online survey to share their agency's successes and challenges in using data to inform their work (n=22 responses). The survey also gathered insights on projected data-related training and resource needs for the next two years to guide the development of the plan. Please refer to the Appendix B for a complete list of survey results.

An external facilitator from Nimble Consulting was hired to facilitate the meeting. The group reviewed a summary of the survey results and shared examples of their experiences using data to provide context for the findings. OPM staff took notes during the meeting which were shared with the facilitator to create this report summary.

Survey Results Summary: Reported Successes

Successes building staff capacity to use data	Successes improving access to data	Successes using data to make decisions
<ul style="list-style-type: none"> • Hiring qualified staff into key roles • Developing skills of junior staff • Acquiring additional resources to grow area • Access to Fellows to increase support • Investing in trainings and technical assistance for staff • Investing in new tools 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Managing call wait times, real-time data to the public via maps, dashboards • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication 	<ul style="list-style-type: none"> • Improving operations and efficiencies • Reviewing call wait times to identify areas for improvement • Improving data sharing, transparency, and access for the public • Developing reporting tools to monitor programs • Enhancing access and internal data sharing to improve staff communication

Survey Results Summary: Challenges and Needs

Top 5 data-related challenges	Top 5 anticipated data-related impacts	Top 5 data-related training or capacity-building needs
<ul style="list-style-type: none"> • Budget, 89% • Data availability, 50% • Data Quality, 43% • Staff skillset, 42% • Recruiting qualified staff, 39% 	<ul style="list-style-type: none"> • Artificial intelligence, 90% • IT optimization, 84% • Data security, 79% • Data privacy, 78% • Staff turnover, 74% 	<ul style="list-style-type: none"> • Data storytelling, 100% • Data visualization, 95% • Data governance, 95% • Data privacy, 91% • Data management, 90%

Group Discussion

General Discussion

Themes from the meeting reflect a collective desire for improved data infrastructure, improved hiring resources, and enhanced collaboration across agencies to effectively use data for decision-making and program evaluation.

- **Data Tracking and Management:** There is a strong emphasis on the need for better data systems to track the success of programs, especially in educational contexts. This includes automated reporting and the ability to query data effectively.
- **Hiring Challenges:** A few leaders highlighted difficulties in hiring qualified data professionals, particularly data scientists. Issues include competition with the private sector, salary concerns, and the need for clearer job classifications.
- **Data Sharing and Collaboration:** The need for clear guidelines on data sharing across agencies was mentioned.
- **Technology and Tools:** There was a focus on the use of tools, like Power BI, for data visualization and reporting. Leaders discussed both successes and challenges related to outdated systems and data quality.
- **Customer Service Improvement:** Analyzing data to enhance customer service, such as measuring call wait times and understanding service preferences, was also a recurring theme.

Data-Related Challenges

Leaders shared the following challenges related to data management, sharing, and service delivery across agencies:

- **Data Availability and Accessibility:** There is confusion and concern over what "data availability" means, with discussions highlighting difficulties in accessing and sharing data across agencies.
- **Collaboration and Support Needs:** A need for a centralized resource or library of tools to aid in data management and sharing was expressed, along with calls for capacity building and training on existing tools.
- **Capacity and Resource Duplication:** The discussion included frustrations over duplicated efforts and the need for better alignment among agencies to optimize resources.
- **Client Experience and Service Delivery:** The burdensome data requests on clients, especially vulnerable populations, were highlighted. Simplifying processes could improve access to services.
- **Data Sharing Challenges:** Statutory and regulatory limitations are believed to hinder effective data sharing. Participants noted the need for agreements and collaboration but faced obstacles due to strict compliance requirements.
- **Legal and Compliance Considerations:** Participants discussed the complexities of adhering to both state and federal laws in data sharing, particularly regarding protected information.
- **Procurement and Data Management Issues:** There were concerns about inefficiencies in data management and the procurement process, with suggestions for better coordination and criteria for vendor selection.

Anticipated data-related impacts

Leaders anticipate a blend of opportunities and challenges associated with artificial intelligence (AI) integration, highlighting the need for ethical considerations, collaboration, and strategic planning.

- **Artificial Intelligence (AI) Applications:** There is significant interest in using AI for operational improvements, such as call center routing and data analysis, with examples from other jurisdictions illustrating its potential benefits.
- **Data Integration Challenges:** Leaders highlighted difficulties in integrating and analyzing data, particularly narrative data, and discussed the possibility of using AI to facilitate this process.
- **Ethical Considerations:** Concerns about the ethical implications of AI use were raised, especially regarding its application in sensitive contexts like legal proceedings, emphasizing the need for ethical training and awareness.
- **Policy and Governance:** The existence of a policy on AI use and a dedicated advisory group to evaluate use cases indicates a structured approach to managing AI integration within agencies.
- **Collaboration on Use Cases:** The need for collaborative identification of multi-agency use cases for AI was discussed, highlighting the potential for shared benefits.
- **Training and Workforce Development:** Emphasis was placed on the importance of training staff on AI tools and their applications to enhance capabilities within agencies.
- **Needs Assessment:** A suggestion was made to conduct a needs assessment around AI to identify requirements and opportunities for its implementation across agencies.

Training and Capacity Building Needs

Leaders emphasized the need for training and capacity building that focuses on enhancing data communication, accessibility, and collaboration.

- **Data Storytelling:** Emphasizing the importance of effectively communicating data insights to various audiences, as this is a primary responsibility for agency leadership.
- **Dissemination and Accessibility:** Highlighting challenges in sharing and accessing data, including improved methods to disseminate information.
- **Internship and Recruitment Strategies:** The importance of leveraging internships and partnerships with educational institutions was discussed as a way to build a pipeline of qualified candidates. The benefits of collaborative initiatives, such as the Governor's Fellow program, were also highlighted.
- **Collaboration and Training:** Advocating for joint training efforts among agencies to enhance understanding and utilization of data, especially in community contexts.
- **Centralized Data Resources:** The need for a centralized landing page to facilitate easy access to relevant data across different sectors.
- **Communication and Engagement:** The need to improve transparency and public engagement through social media and improved communication strategies.
- **Unified Data-related Updates:** Proposing a centralized hub for data-related guidance and tools to streamline access and consistency.
- **Tailored Messaging:** The importance of adapting language and communication styles to suit different audiences.

- **Standardization:** Stressing the need for consistent terminology and clear metadata to avoid confusion among users.
- **Training for Consistency:** Training staff to ensure a shared understanding of key data elements and concepts.

Recommendations

Based on the discussion, the following are recommendations from agency leadership to include in the next State Data Plan:

1. Develop a centralized hub to provide access to existing online data-related resources, including tools, software, and training materials.
2. Conduct a needs assessment to identify additional data-related training and resource requirements.
3. Offer guidance and support to agencies in hiring qualified data analysts.
4. Establish clear and consistent statewide guidelines for data sharing across agencies.

Existing Resources

The following resource list is provided by the Office of Policy Management, Data and Polic Analytics office in response to requests from State Agency Leaders for specific data-related resources and training. Additional resources and opportunities will be detailed in the 2025-2026 State Data Plan.

- State data plan overview: <https://portal.ct.gov/datapolicy/state-data-plan>
- Agency data officers (list and duties and responsibilities): <https://portal.ct.gov/datapolicy/knowledge-base/articles/agency-data-officers>
- Resources for agency data users on:
 - Data visualization and accessibility: <https://ctopendata.github.io/data-visualization-guidelines/>
 - Data sharing and legal issues: <https://portal.ct.gov/datapolicy/knowledge-base/articles/data-sharing-resources>
 - Inventorying / cataloging data: <https://portal.ct.gov/datapolicy/-/media/datapolicy/agency-guidance/high-value-data-inventory-guidance.pdf>
 - Metadata / documentations: <https://portal.ct.gov/-/media/ct-data/metadata-guidelines.docx?la=en>
- Data capacity building resources and videos / screencasts: https://portal.ct.gov/datapolicy/knowledge-base/articles/evaluation-and-impact/our-long-term-goals?language=en_US and videos / screencasts: https://portal.ct.gov/datapolicy/agency-guidance?language=en_US
- Responsible AI framework for agencies: <https://portal.ct.gov/opm/fin-general/policies/-/media/d13d6f704fa3408998f20e67ebda8aab.ashx>

Appendix A: State Leadership Attendance

- Thea Montanez, Office of The Governor
- Scott Gaul, Chief Data Officer, Office of Policy and Management
- Ellen Solek – Connecticut Technical Education and Career Services Executive Director
- Mark Boughton – Department of Revenue Services Commissioner
- Bryan Cafferelli – Department of Consumer Protection Commissioner
- Shantelle Barnes – Department of Social Services Dept Commissioner
- Amy Porter – Aging and Disability Services Commissioner
- Andrew Hoskins – Department of Energy and Environmental Protection Chief of Staff
- Kelli-Marie Vallieres – Chief Workforce Officer, Office of Workforce Strategy
- Charlene Russell-Tucker – State Department of Education Commissioner
- Angel Quiros – Department Of Correction Commissioner
- Paul Potamianos – Office of Policy and Management Deputy Secretary
- Nancy Navarretta – Department of Mental Health and Addiction Services Commissioner
- Deidre Gifford – Office of Health Strategy Commissioner
- Tim Larson – Office of Higher Education Commissioner
- Joyce Lee Taylor Department of Children and Families Deputy Commissioner
- Manisha Juthani – Department of Public Health Commissioner

Appendix B: Leadership Survey Results

Which agency do you represent?

- ADS
- CETCS
- DAS
- DCF
- DCP
- DECD
- DEEP
- DMHAS
- DOC
- DOH
- DOL
- DOT
- DPH
- DRS
- DSS
- OEC
- OHE
- OHS
- OPM
- OWS
- SDE
- Other (not defined)

What is your role within your agency?

- Agency Data Officer, 1
- Agency Leadership, 21

Challenges: To what extent does your agency face any of the following challenges using data?¹

Budget, data availability, data quality, staff skillset, and recruiting qualified staff are reported as the top 5 challenges agencies face using data.

	Agree	Neutral	Disagree
Lack of budget	89%	5%	5%
Data availability	50%	22%	28%
Data quality	43%	32%	25%
Staff lack the appropriate data and analytical skill sets	42%	21%	37%
Inability to recruit qualified staff that know how to interpret and use data effectively	39%	16%	37%
Inability to retain qualified staff that know how to interpret and use data effectively	37%	16%	42%
Lack of executive support	12%	17%	61%
Rate of technological change is too high	16%	42%	43%

¹ Strongly Agree and Agree and Strongly Disagree and Disagree are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

My agency effectively utilizes data to achieve our agency goals.

- Strongly agree, 4
- Agree, 18
- Neither agree or disagree, 0
- Disagree, 0
- Strongly disagree, 0

Impacts: How likely are the following to have significant impact on your work in the next two years?²

Artificial intelligence, IT optimization, data security, data privacy, and staff turnover were the top 5 predicted impacts reported.

	Likely	Neutral	Unlikely
Artificial intelligence	90%	--	10%
IT optimization	84%	11%	10%
Concerns about data security	79%	11%	11%
Concerns about data privacy	78%	11%	11%
Staff turnover / retirements	74%	21%	5%
Federal changes in race / ethnicity standards	68%	21%	11%
Other changes in state or federal law	58%	32%	11%

Which of the following would you prioritize for training or capacity-building for your agency?³

	Important	Neutral	Unimportant
Data storytelling	100%	--	--
Data visualization	95%	5%	--
Data governance	95%	5%	--
Data privacy	91%	10%	--
Data management	90%	10%	--
Artificial intelligence	81%	10%	10%
Demographic data collection	80%	15%	5%
Accessibility	79%	11%	11%
Performance management	76%	24%	--
Evaluation	76%	24%	--
Geospatial information systems (GIS)	75%	25%	--
Statistical techniques	75%	25%	--
Equity in data use	68%	32%	--
Specific tools (e.g. R, Python, SQL)	65%	30%	5%
Ethics/human subjects training	47%	48%	5%

² Very and Somewhat Likely and Somewhat and Very Unlikely are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

³ Very and Somewhat Important and Somewhat and Very Unimportant are combined. Not applicable is not reported in this table and therefore, totals may not equal 100%.

Rankings of the most important data-related activities to achieve agency goals over the next two years⁴

1. Development of training and capacity building (68%)
2. Providing access and tools and technology for data use (50%)
3. Supporting analysis research and evaluation using state data (32%)
4. Developing guidelines and standards for data use (32%)
5. Communicating the use of data by state agencies (18%)

⁴ First and second choice rankings are combined.