DATA Board Meeting

Data Analysis Technology Advisory Board

September 25, 2025

Data & Policy Analytics

CONNECTICUT
Policy and Management

Agenda

1. Welcome & Introductions
2. 2025-2026 State Data Plan – Update on Goals
3. Data Intermediary Report
4. Adjourn

DATA Board powers and duties

CGS Sec. 2-79e: The board shall have the following powers and duties:

(1) To **advise the executive, legislative and judicial branches** of government and municipalities **concerning data policy**, including, but not limited to, best practices in the public, private and academic sectors for data analysis, management, storage, security, privacy and visualization and the use of data to grow the economy;

(2) to **advise the Office of Policy and Management regarding the online repository** established under section 4-67p;

(3) to issue reports and recommendations in accordance with section 11-4a;

(4) upon the request of at least two members of the board, to **request any agency data officer or agency head to appear before the board** to answer questions;

(5) to request from any executive department, board, commission or other agency of the state such assistance and data as necessary and available to carry out the purposes of this section;

(6) to make recommendations to the legislative leaders and the directors of the Offices of Fiscal Analysis and Legislative Research regarding data analysis skills and related expertise that the leaders and said offices may seek to cultivate among their staff through training or as a consideration when hiring staff; and

(7) to establish bylaws to govern its procedures.

State Data Plan Goals

Goal 5

- **Develop an Enterprise Data Sharing Approach:**
  - Establish clear guidelines for secure, standardized data sharing across agencies.
  - Build on successful frameworks like DataLinkCT (formerly P20 WIN) to enhance cross-agency collaboration.

DataLinkCT (P20 WIN) Executive Board Resolution

Approved June 2024 (Summary)

- Authorizes OPM and the DataLinkCT (P20 WIN) Data Governing Board to develop policies and procedures for a hybrid cloud solution that:
  - Combines federated and centralized models;
  - Uses a single cloud platform for matching, analysis and data movement;
  - Ensures agency authority and control over the release and use of data;
  - Allows for auditing and tracking movement, access, and usage of data;
  - Develops improved matching process to accelerate requests and improve data quality;
  - Allows agencies to manage updates to a central cloud-based repository;
  - Reduces movement of data and manual processes; and
  - Complies with all relevant agencies, DLCT, state and federal policies and procedures

Secure Data Enclave Goals

Establish model guidelines for secure, standardized data sharing across agencies to enhance cross-agency collaboration.

- ✓ Reduce manual data movement
- ✓ Securely match PII
- ✓ Host agency data for data requests
- ✓ Provide a modern and secure analytics environment
- ✓ Support evolving agency requirements for data reporting and transparency

Secure Data Enclave Project – Milestone Update

Phase	Timeframe	Description	Status
Planning	January-February	Project plan and timeline are developed and approved. Project priorities, tasks, and deliverables are identified. User requirements are gathered and approved.	**Complete**
Development	March-August	Secure Data Enclave is developed. Data policies are updated, and new policies are developed as needed.	**In Progress**
Testing	August - October	Secure Data Enclave is functional and tested. DOL and OPM users are trained.	**In Progress**
Implementation	October - January	Transition plan begins. Users are trained. DataLadder integration. Data is uploaded into the Enclave.	**Not Started**

Project Status and Key Activities

1. Development of Microsoft Azure Environment

- Azure file storage contain three different folders for routing data request files: red, yellow, green. Each folder represents a different level of data sensitivity and access group.

2. Integration with Axway Secure File Transfer

- Successful proof of concept test between Axway Secure File Transport and MS Azure Environment.
- Collaborating with external Axway consultants to implement enhancements to Axway routing.

3. Partner Agency Engagement

- Execution on initial requirements gathered from agencies based on specific needs for their data.
- Ongoing engagement with participating agencies during DLCT Office Hours and as requested.

4. Policy and Process Updates

- Data Governing Board Subcommittees convened earlier this year; Data Security; Data Privacy. Legal coming soon.
- Updating existing and developing new data policies in alignment with secure enclave.

5. DataLadder Cloud Migration & Integration

- Data Ladder is the data matching tool used and administered by the Department of Labor to support data matching for DLCT data requests.

DataLinkCT (formerly P20 WIN) Policy Palooza

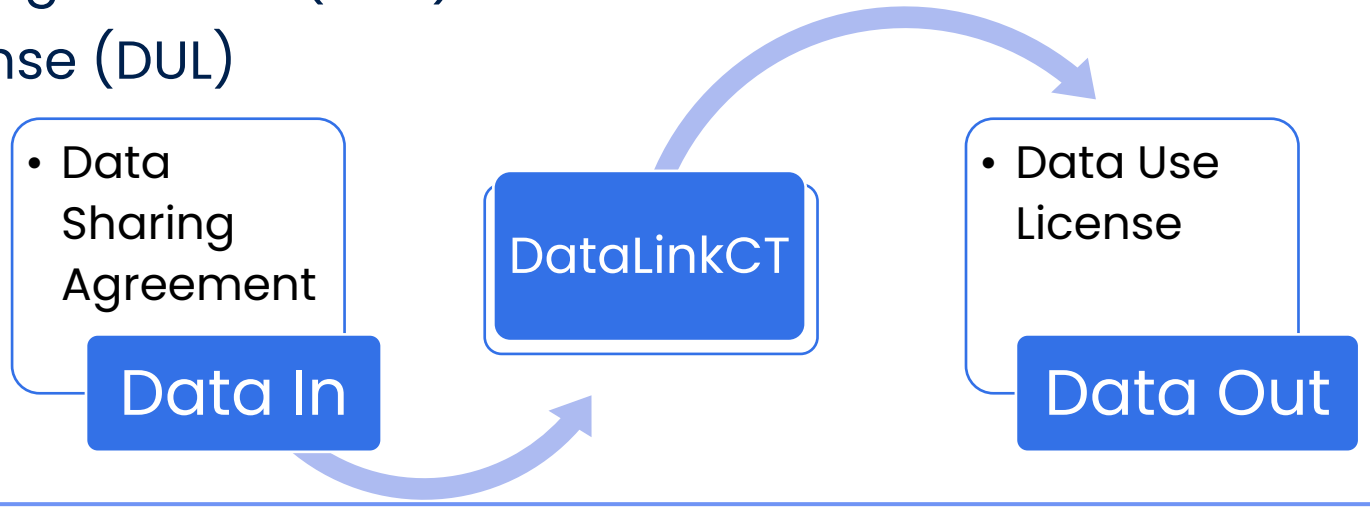
Policy Topic	Status	Purpose
Data Privacy	In Review	Formal document outlining the value and importance of preserving privacy of data used for approved purposes.
Data Security	Updating	Establish standards for data security for DLCT agencies and data requestors.
Data Retention/Destruction	New	Data Destruction Certificate is an exhibit to DSA.
Disclosure Avoidance	New	Ensures PII and other data are not released to unauthorized users.
Data Matching	New	Emphasizes importance and need for documentation matching data
Data Access	New	Document roles and responsibilities necessary to access and use data for approved requests.
Metadata Management	Implemented	Participating agencies maintain a data dictionary.
Pre-publication Review	Implemented	Data requestors will prepare analyses ready for release for review by participating agencies.
Data Quality	Adopted	Ensure agencies have proper documentation and processes to support high-quality data access.
Data Classification	Adopted	Classify data elements based on risk and use in DataLinkCT system.

Data Sharing – Legal Agreements

DataLinkCT uses a tiered structure of legal agreements to share data for approved projects.

Standard but modular agreements can increase the flexibility of legal agreements. Each document has different signatories and serves different purposes.

- Enterprise Memorandum of Understanding (eMOU)
- Data Sharing Agreement (DSA)
- Data Use License (DUL)



Data Sharing - Legal Agreements

Enterprise Memorandum of Understanding (E-MOU)

- A foundational agreement among all parties.
- Outlines the rules of the road that all participating agencies agree to.
- The E-MOU establishes the model for DataLinkCT (formerly P20 WIN), including the operational structure, governance structure, and technical processes
- Does not include language specific to any one participating agency

Data Sharing – Legal Agreements

Data Sharing Agreement (DSA)

- A Data Sharing Agreement is a formal document created for an approved data project and the agencies participating in that project. A DSA names the purpose, participating agencies, and details what data are being shared and how the data can be used.
- Data Sharing Agreements should include provisions concerning access and dissemination.
- A DSA should include:
  - Requirements to access to specific data
  - Confidentiality and disclaimers
  - Time limits

Data Sharing - Legal Agreements

Data Use License

- A Data Use License is an agreement between an agency or department and an external entity (in this case, the data requestor) dictating how the requested data can be used.

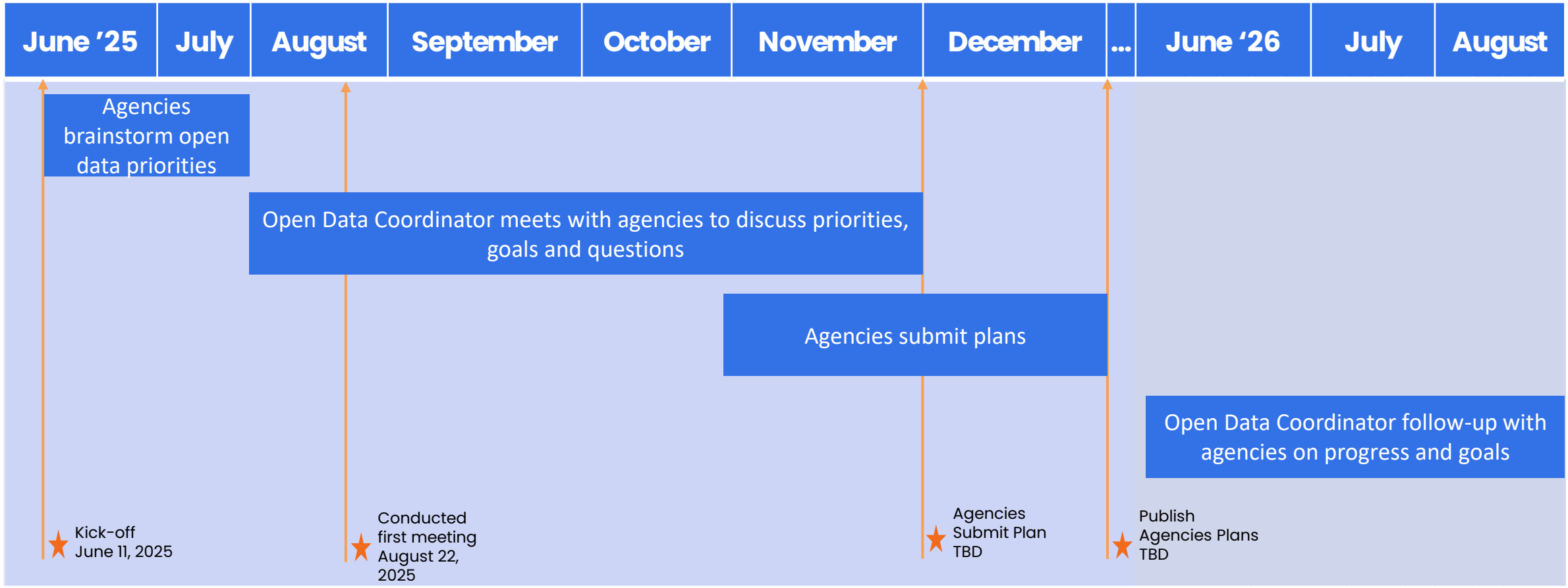
Goal 1

- **Increase Accessibility, Visibility and Transparency of Data Resources:**
  - Expand access to existing tools, software, and training materials.
  - Address barriers that make data difficult to find and use.

Open data access plans

- Section 4-67p of the Connecticut General Statutes requires executive agencies to develop open data access plans.
- These plans should detail the agency's plan to publish open data, including public data, and protected data that can be made public through aggregation or redaction.
- Initial process completed in 2019 but needed follow-up.

Timeline


June '25
July
August
September
October
November
December
...
June '26
July
August
Agencies brainstorm open data priorities
Open Data Coordinator meets with agencies to discuss priorities, goals and questions
Agencies submit plans
Open Data Coordinator follow-up with agencies on progress and goals
Kick-off June 11, 2025
Conducted first meeting August 22, 2025
Agencies Submit Plan TBD
Publish Agencies Plans TBD

Brainstorming questions

1. How does your agency's open data work align with your agency's mission or goals?
2. What data does your agency make available as open data on the CT Open Data Portal, the Geodata Portal, or another open data platform?
3. How has your agency improved access to open data in the last year?
4. What data would be most valuable for your agency to publish as open data that is not already available as open data?
5. What data do you plan to prioritize for publication as open data in the next year?
6. How else do you plan to improve access to your agency's data? (e.g. improved metadata, automated data updates, updating out-of-date data, etc.)
7. What resources do you need to make progress on your open data access plan?

Open Data Access Plans Template

The Ask

- What information would you like to see in the agencies' open data access plans?

Component Feedback Thus Far

- Risks/Challenges/Gaps/SWOT analysis
- Dates
- Lessons Learned

Example #1: U.S. Department of Agriculture's Open Data Plan

Vision

USDA will encourage the collection of data in an open format, using digital collection mechanisms, by default.

Ongoing Activities

- Develop a Data Release Checklist that mandates data owners save public data assets in a machine-readable, open format for release on Data.gov (e.g., JSON, XML, CSV/TXT, KML/KMZ, ESRI Shapefile).
- Coordinate ongoing and planned metadata management efforts across Mission Areas, Departmental Administration and Staff Offices (DASO), and the Office of the Chief Information Officer (OCIO) through Metadata Management Plans to better serve stakeholders that have use for USDA data, fulfill compliance with the Evidence Act, and maintain a Departmental catalog with our most valuable data sets.
- Maintain data in an open format.
- Maintain public data assets as open Government data assets with an open license and make discoverable via Data.gov

Out-Year Plans

- Develop requirements and guidelines for data collection in open formats, under an open license, to increase the accessibility and utility of public data assets.
- Assess active data collection methods for any lack of open format usage.
- Inventory tools currently in use which promote data collection in open formats to support machine-readable public data assets.
- Increase data literacy so the USDA workforce understands the value of open data and is empowered to implement open data policies and processes.

Components

- Goal/Vision
- Related Ongoing Activities
- Related Future Plans

Example #2: U.S. Department of Education's Open Data Plan

KEY ACTION 1: IMPROVE USABILITY OF ALL PUBLIC DATA ASSETS

Task 1.1: Utilize open formats and open APIs for data distribution

Approach

Open formats and application programming interfaces (APIs) makes the data more easily interoperable with other datasets and can be more easily reused. Closed formats or APIs may limit access due to proprietary software requirements or registration fees.

The Department prioritizes open formats for data assets and open APIs. In the CDI, approximately 30 percent of included data assets provide files in open formats. Towards achievement of Task 1.1, the Department will take the following actions.

Short-Term Subtasks

1. Compile and publish a registry of open formats. This will be implemented as an online reference tool and includes plans to provide format conversion tools to facilitate preparing open data assets for public release.
2. Establish a policy for using open formats and open APIs to distribute open data assets.
3. Implement performance measures for the CDI of the percentage of open data assets released in open formats.

Benchmarks for Short-Term Subtasks

1. Compile and publish a registry of open formats:
    - **Count:** Measure the number of open formats listed in the registry.
    - **Updates:** Track the frequency of updates or additions to the registry.

Components

- Goal
- Task(s)
- Approach
    - Subtasks
- Benchmarks

Goal 2

- **Improve Metadata and Documentation:**
  - Enhance metadata to improve data discoverability and linkage.
  - Provide better documentation on data collection methods and limitations.

Goal 3

- **Identify Training and Resource Needs for Agency Staff:**
  - Conduct a needs assessment to strengthen data literacy, visualization, automation, and analytical skills.
  - Address gaps in onboarding, resources, and professional development.

Open Data Cohort Training

- Workshops for CT open data users from June 10-25
- Led by Tyler Technologies Education team
- Open to all users of the Open Data Portal in CT state agencies
- Most courses have about 15 participants registered, with some courses with 30+ people in attendance

Courses included in training

- Clean and Tidy Data
- Explore Data with Charts
- Map Your Data
- Tell a Story with Perspectives
- Collaborate with Enterprise Data
- Create and Manage Your Dataset
- Data Automation and Management with Gateway
- Transform Data with SoQL
- Shape and Join Data in Exploration Canvas

Goal 4

- **Review Job Classifications for Data and Analytics Roles:**
  - Update job specifications to align with current trends and needs.
  - Attract and retain skilled data professionals through modernized roles and competitive compensation.

Attracting and Retaining Talent

"The solution? Build strong partnerships with HR. Work together to modernize job classifications, craft clear role definitions, and create flexible, targeted hiring strategies. Beyond these strategies, many DSN leaders have found success using internships, fellowships, and direct hiring authorities to move faster and reach nontraditional candidates. Partnering with local work training organizations can also create a new pathway for hiring."

- "Hiring, Retaining, and Upskilling Digital Service Talent in Government: Stories and Strategies from State and Local Leaders," Beeck Center, Digital Service Network

First Step – Initial Draft of Data Positions

1. Chief Data Officer
2. Data Services Manager
3. Enterprise Data Architect (Big Data, Data Warehouse, Data Lakehouse)
4. Data/Big Data Engineer
5. ELT Engineer
6. Data Analyst
7. Data Modeler
8. Data Scientist
9. Data Steward/Chief Data Steward
10. Digital Data Assets Specialist
11. BI Analyst
12. AI/ML Engineer
13. GIS Manager
14. Data Integration Specialist
15. Information Flow Architect
16. Data Librarian
17. Open Data Coordinator
18. Analytics Developer
19. Data Governance Coordinator
20. DataOps Manager
21. Data Literacy Coordinator and/or Data Literacy Analyst/Developer
22. Data Security/Privacy

Sample from Oregon!

Potential/Proposed Data Job Classes

sample from Oregon!

Data Managers/Leaders

Chief Data Officers, Deputy CDOs, DataOps & Analytics managers, Data Governance & Architecture managers, GIS managers, etc.

Data Analytics & Business Insights

Data Analysts, Data Scientists, BI Analysts, AI/ML Engineers, etc.

Data Strategy & Governance

Data Stewards, Data Librarians, Data Literacy Coordinators, Data Governance Coordinators, etc.

DataOps & Architecture

Data Architects, Data Engineers, Data Modelers, Data Security/Privacy roles, etc.

Agency Data Officer Focus

- The goals of this review are to:
  - Attract talent to Connecticut analytical position
  - Keep/make Connecticut competitive in the analytical job market
  - Retain current and future talent in analytical positions
  - Review Connecticut's current job classifications and modify as needed

Current (2024) Employee Counts

- There are 1,598 unique employees with job descriptions containing
  - Research
  - Data
  - Info
  - IT
- Of the 1,598 there are 93 unique job descriptions (taken from Comptroller's Office) and 47 agencies (also sourced from CO)

Goal 6

- **Support Responsible Use of Artificial Intelligence (AI):**
  - Develop policies for responsible and ethical AI use, focusing on transparency and accountability.
  - Launch training and resources for agencies to explore AI-driven solutions and utilize these tools effectively

Bridging the Data Gap:

Creating State-Wide Data Pathways to Support Connecticut's Disconnected Youth

Adjourn