How the Hartford Data Collaborative addresses fifteen essential requirements for responsible data stewardship

The Hartford Data Collaborative is a network of Hartford, CT area nonprofit organizations, government agencies, and philanthropic partners that facilitates data sharing and data integration among its partners. Data integrated through HDC provide community and government organizations with a more holistic view of individuals, families, and households in Hartford. Integrated data provides a richer understanding of the services and needs of individuals than the data each organization maintains on its own. HDC was formed in 2019 is managed by staff from the non-profit Connecticut Data Collaborative, with input from community stakeholders.

This document summarizes how the HDC addresses 15 essential requirements for data trusts, data repositories, and other data collaborations.[1]

Legal – Fulfills Legal Requirements

Entities are first asked to sign the Enterprise-wide Memorandum of Understanding (EMOU) which does not bind partners to sharing, but demonstrates the entities commitment to the Hartford Data Collaborative.

- The EMOU documents HDC's vision, mission, and governance process and is signed by all data partners. The EMOU references the data sharing agreement, data use license, and policies and procedures for HDC data access and use.

The Hartford Data Collaborative (HDC) has developed 2 key legal documents that govern the integration and sharing of data among its partners. Developed in consultation with outside legal counsel, these documents help HDC and its partners to comply with applicable federal and state privacy and data use statutes and regulations--Data Sharing Agreement (DSA), and the Data Use License (DUL).

- The DSA is specific to each data project and includes the terms and conditions that govern how data are transferred, stored, and managed within the shared data infrastructure and for end users. The DSA references the MOU and the DUL.

- The DUL outlines the role and responsibilities of the end user--the licensee. This document includes the projects objectives, methodology, data fields, and data security plan including plans to mitigate risk of reidentification, dissemination plan, and timeline of project completion. A completed DUL must be signed by the end user of each approved project prior to data access.

The use and disclosure of data is governed by the Data License Request form, which is completed by the parties that request data and reviewed by the HDC Data Oversight Committee. The parties must indicate whether the data requested is for research or data purposes. Collecting data for research purposes is defined by US Policy for the Protection of Human Subjects and may require Institutional Review Board approval. Legal agreements can be accessed here, https://www.ctdata.org/hdc-governance-legal-agreements.

Governance – Stated Purpose

The Hartford Data Collaborative (HDC) is a shared data infrastructure that facilitates data sharing, integration and analysis to optimize services and outcomes for Hartford residents.

HDC was created to answer big questions about how to improve the lives of Hartford area residents. Traditionally, information about residents—where they live, their enrollment in childcare centers, school, or workforce training programs, and what happens to them after they graduate or leave these programs—is collected by different agencies and stored in different data systems. HDC seeks to securely collect, link, and share data from different agencies to create a richer view of who uses programs, and what the outcomes from these programs are over time. HDC provides clear rules for the storage and use of data to protect the safety and privacy of the individuals served.

Governance – Accountable Governance Bodies

HDC's Executive Board is a community-based committee that provides vision, oversight and leadership for the Hartford Data Collaborative. The Executive Board consists of the executive leadership of community organizations, city administration, and higher education institutions. All Executive Board members will be carefully considered based on their diverse perspectives and representation of agencies that serve the Hartford community.

Governance – Transparency

The HDC data governance process is managed and implemented through a three-tiered model: Executive and Technical Oversight Committees, HDC Legal Framework, and HDC Data License Request Process. A description of this model, including details on the committees, and the data request process is on the HDC website and in the HDC Data Governance Manual, which can be accessed on the HDC website: www.ctdata.org/hdc-resources-and-news HDC holds regular Executive Board and Data Oversight Committee meetings, with meeting dates posted to our website.

Governance – Indigenous Data Sovereignty

HDC and the Connecticut Data Collaborative center equity in our mission and values. We seek to honor the innate value in all people through our behaviors and actions. We are intentional in the way data are collected, analyzed, and shared to ensure data does not exacerbate bias and stereotypes. We strive to provide context for the proper interpretation of results.

Governance – Adaptive and Responsive

HDC developed its committee structure in 2019 but has changed the meeting and format due to evolution of demands during the COVID-10 pandemic. Currently, the Executive Board meets quarterly virtually and developed subcommittees in 2021 to address specific governance concerns, such as funding, membership, and operations. In September 2022, HDC staff conducted a board survey to better understand members' perspectives and improve the effectiveness of the HDC governance structure. As a result, HDC plans to further engage with community priorities in 2023 through data walks and outreach with data partners.

Management – Policies, Processes, and Procedures Covering the Data Lifecycle

Management of the data sharing process is outlined in the legal documents described above and the HDC Data Governance manual. HDC staff manage the collection, use, retention, disclosure, and destruction of data. The

Data Oversight Committee approves all data requests and the findings from data-linking projects before they are released publicly.

Management – Cybersecurity and Data Protection

All data management is performed by HDC staff, including a data linking agent, depending on the data source and activity. As stipulated in the HDC Data Sharing Agreement, Data Integration Staff will follow guidelines set forth by the National Institute of Standards and Technology (NIST) within NIST Publication 800-53, Security and Privacy Controls for Information Systems and Organizations.

Management – Risk Management

HDC staff work with HDC governance committees to review and evaluate risk management processes.

Management – Data Documentation

HDC's website provides information on the types of data shared through accepted data requests. Internally, HDC maintains information on the data requested, related meeting notes among data partners about data requests, metadata, and the linking process on its internal shared workspace.

Data Users – Privacy and Security Training

HDC provides optional data training through its parent organization, Connecticut Data Collaborative. HDC does not have specific data training required for data access.

Data Users – Consequences for Non-compliance

The Data Use License (DUL) outlines the conditions of the data approved by HDC Data License Request Form. DUL addresses the conditions under which the HDC Data Recipient may use, the Anonymized Data for HDC Projects. Any person or entity that processes or receives the HDC Project Data must be obligated, by contract, to adhere to the terms of the DUL.

By signing the DUL, the HDC Data Recipient acknowledges that as a result of HDC's determination or reasonable belief that unauthorized access, use, or disclosures have taken place, HDC may refuse to release further HDC Project Data to the HDC Data Recipient for a period of time to be determined by HDC, in its sole discretion.

Public & Stakeholder Engagement – Knowledge Users and Data Partners Engagement

HDC does outreach to community partners and is in the process of strengthening its Executive Board through additional representation of community organizations.

HDC is developing plans for public engagement to help build understanding around HDC and its purpose to serve the residents of Hartford.

Public & Stakeholder Engagement – Involving Members of the Public

HDC was launched with support from Hartford community organizations and city and state agencies In 2018, and in 2021 HDC incorporated Hartford community members into this process through the development of youth co-researchers, called the Hartford Youth Researchers. HDC staff trained 6 Hartford youth ages 16-24 in our core workshops in data literacy, analysis, and data storytelling and then analyzed integrated data about youth engagement in city-funded programs post-COVID-19 pandemic. The youth co-led a data walk with members of the community and produced a public report on their findings. In January 2023, HDC launched a more in-depth program called the Hartford Youth Data Fellows. Ten youth from Hartford will spend 5 months with HDC, receiving data training and engaging in community data projects, leading to a public data walk in the spring. For more information, see www.ctdata.org/hartfordyouth

HDC plays an integral role in the research and evaluation of Hartford's involvement in the North Hartford Ascend Pipeline project, collecting and integrating cradle to career data for residents in the North End neighborhoods of Hartford. HDC staff participate in community meetings to discuss the data indicators for this multi-year project.

Public & Stakeholder Engagement – Tailored Approaches for Different Publics

Each HDC project involves different levels of engagement. In addition to direct involvement through the community participatory action research described above, HDC also coordinates with the CTData Equity in Data community of practice to discuss equity related topics including equity and ethics throughout the data life cycle.

---

[1] Paprica PA, Crichlow M, Curtis Maillet D, Kesselring S, Pow C, Scarnecchia T, Schull MJ, Cartagena RG, Cumyn A, Elliston KO, Greiver M, Hill SL, Isaranuwatchai W, Loukipoudis E, McDonald JT, McLaughlin J, Hawn Nelson A., Rabinowitz A, Razak F, Verhulst SG, Verma AA, Victor JC, Young A, Yu J, McGrail KM. (2023). Essential Requirements for the governance and management of data trusts, data repositories, and other data collaborations: a refined set of 15 minimum specification requirements based on the experience of 23 organizations and data sharing initiatives. Manuscript in preparation for submission to *International Journal of Population Data Science*.