Developing the Hartford Data Collaborative

Amy Hawn Nelson
Actionable Intelligence for Social Policy (AISP)
University of Pennsylvania

Geethika Fernando
Hartford Data Collaborative
CT Data Collaborative
When we talk about Integrated Data Systems, what do we mean?

- We’re talking about people, not tech solutions.
- Efforts that link administrative data across domains/agencies
- Efforts that curate data that is relevant and high quality
- Efforts that serve as a public utility (not research for research’s sake)
- Efforts that have defined governance structures
## AISP’S Role

<table>
<thead>
<tr>
<th>We are:</th>
<th>We are not:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data evangelists</td>
<td>Data holders or intermediaries</td>
</tr>
<tr>
<td>Connectors, community builders, thought partners, cheerleaders, and data sharing therapists</td>
<td>A vendor or vendor recommender</td>
</tr>
<tr>
<td>Focused on ethical data use for policy change</td>
<td>Focused on academic research</td>
</tr>
</tbody>
</table>
What We Do

- Convene a **professional network** for local and state governments working on data integration to share best practices and problem-solve together

- Engage in **advocacy** on behalf of data sharing at the federal, state, and local level

- Provide **resources and sample documents** on data governance, legal considerations, data standards, and linkage technologies

- Offer **training and technical assistance** to help interdisciplinary teams increase state data capacity and use
Usable data

Insights that drive change

People using and owning their agencies data

Analysis!!!

Governance, legal, data cleaning, linking, standardizing
Green Light Data
Data that can be shared openly, either at the aggregate or individual level based on state and federal law. These data often exist in open data portals.

Yellow Light Data
Data that can be shared, but only under specific circumstances with appropriate safeguards in place.

Red Light Data
Data that cannot, or should not be shared, either due to state and/or federal law, or due to context (for example, politically volatile topic, data quality concerns, etc.)

Data Classification Matrix

Can be shared with agreement and approval through governance

Open Data, can be shared with no agreement

Not shareable

Technology and/or data structure limits ability to share data
AISP Network as of 2020

- Network of 22 operational state and local integrated data systems
- Between our three Learning Community cohorts and other sites we support, 14 more sites well on their way
# New Resource: 2019 Integrated Data Matrix

## Table

<table>
<thead>
<tr>
<th>Program</th>
<th>Pre-K</th>
<th>High School</th>
<th>Medical</th>
<th>Health Programs</th>
<th>Law Enforcement</th>
<th>Juvenile Justice</th>
<th>Family Court</th>
<th>Child Support</th>
<th>Early Education</th>
<th>K-12 Education</th>
<th>Employment</th>
<th>Health</th>
<th>UI</th>
<th>Wage Assistance</th>
<th>Federal Data</th>
<th>Administrative Data</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

## Link

[https://www.aisp.upenn.edu/data-integration-matrix/](https://www.aisp.upenn.edu/data-integration-matrix/)
Data is shared at the speed of trust.
Hartford Data Collaborative (HDC) is a shared data infrastructure that facilitates data sharing, integration, and analysis to optimize services and outcomes for Hartford residents, with emphasis on older youth/emerging adults (ages 14-29) at this time. We have a vision that Hartford will become a community where real-time, integrated, client-level data is accessible for regular analysis in support of improved services, operations, policies, and outcomes.
Hartford: Poverty Amid Plenty

Demographics
- Population of 124,770
- Hispanic or Latino (44.5%)
- Black or African American alone (36.9%), White alone (14.8%), Asian alone (3%), Two or more races (5.5%)

Economy
- 30% live in poverty (compared to 10% in CT)
- Median household income is 34K (compared to 76K)
Why share/integrate data?

- Obtain holistic information about population level service access and needs
- 350+ agencies serving children, families and youth in Hartford
- Many agencies working in siloes without a systemic approach to prevalent social problems
- Widespread perception of competition among agencies for funding
- Community is familiar with the concept due to data sharing efforts in the past
How can shared data benefit Hartford

- Saves time by enabling research projects to be conducted in months, rather than years
- Provides holistic information about individuals' trajectories
- Supports systems evaluation and smarter resource planning
- A resource for government agencies and non-profits to better understand collective impact
- Catalyzes collective impact through use of population-level data linked across multiple agencies
- Enables equitable distribution of resources
Hartford Data Collaborative (HDC)

What is HDC? HDC is a network of Hartford area non-profit agencies, Hartford city agencies, state agencies and philanthropic partners that facilitates data sharing and data linking between human service agencies.

Purpose: HDC's purpose is to assemble and link data housed in different agencies ethically and permissibly to create a richer, aggregate view of service utilization and related outcomes over-time.
Connecticut Data Collaborative

We empower an ecosystem of data users by democratizing access to public data and increasing data literacy.

Make data accessible  Liberate data  Increase data literacy  Census State Data Center  Hartford Data Collaborative

ctdata.org
How does HDC work?

• All data sharing efforts at HDC are managed by governing bodies of data providers and community leadership
  
  ✓ Technical review of all data requests
  
  ✓ Oversight and review of all processes by data providers and executive leadership
  
  ✓ Legal documents that ensure ethical use of data and compliance with applicable laws and regulations

• Full time staff acts as the administrative backbone to all legal and procedural activities related to data sharing and linking
What HDC is and isn’t

- A collaborative that facilitates the linking and sharing of anonymized data for community benefit
- Carefully curated data to support community-based indicators, outcomes, research and evaluation
- A resource for government agencies and non-profits to better understand collective impact

- Not a data management tool, data warehouse or a “quick and easy” data portal
- Not a case management tool or a source you can obtain information on individuals
- Not a resource for evaluating individual programs, agencies or initiatives
HDC Functions

Data Governance
- Legal framework
- Data License process
- Executive and technical oversight

Data Access & Management
- Data security provisions
- Data architecture
- Data matching
HDC Data Governance Structure

**Executive Board**
- Sets research agenda
- Provides fiduciary oversight
- Seeks resources for long-term sustainability
- Builds will and community support for HDC

**Data Oversight Committee**
- Independently and collectively reviews all HDC Data License Requests
- Provides technical support and research expertise to HDC staff
- Reviews all final reports of HDC Data projects

**HDC Staff**
- Facilitates and completes the technical, procedural and relational work of HDC
- Performs secure data integration, storage and transfer

**MHIS staff**
- Doers
HDC Data License Request (DLR) Process

DATA LICENSE REQUEST PROCESS TIMELINE

DLR Submitted
- Project Development (0-12 months)
- Data Requester credentials checked at initial consultation
- IRB submitted (if a research request)
- Research purpose and methodology critically reviewed by experts and data providers

DOC Vote (~1 month)
- Data Use License signed to legally bind the recipient on requested data

Data Pulled
- Data transferred and matched following HDC Data Security Protocols

Consult with Data Partner (0-1 month)
- Data set reviewed by all data providers

DLR Signed
- Requester signs the DLR

Research phase (Months-years)
- Data reports/findings reviewed by data providers and Data Oversight Committee before public dissemination

Data transferred and managed
Data License Request Process

DLR submission and acceptance
- Develop the study/research proposal
- Complete and Submit the DLR to the Data Oversight Committee
- Critical review of the research purpose, methodology, data analysis and statistical modeling by data providers
- Data Oversight Committee vote
- If accepted, sign Data Use License (Data Recipient)
Data Pull, integration and analysis

• Data transfer and match following HDC security protocols
• Matched datasets reviewed by all data providers before transfer to data requestor/recipient
• Data recipient/researcher performs analysis and evaluation
• All Data Providers review final data analysis/reports before dissemination
• Public dissemination of reports
## How is integrated data protected

<table>
<thead>
<tr>
<th>Legal</th>
<th>Technical</th>
<th>Procedural</th>
<th>Physical</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Enterprise Memorandum of Understanding with all participating parties (MOU)</td>
<td>- Encryption while data is at rest and during transfer</td>
<td>- Multiple rounds of reviews by Data Providers and content experts</td>
<td>- Locked cabinets to store all documents</td>
</tr>
<tr>
<td>- Data sharing Agreements (DSA’s)</td>
<td>- Secure servers</td>
<td>- Facilitated Data License Review process</td>
<td>- Secure work stations</td>
</tr>
<tr>
<td>- Data Use License with the Data Recipient (DUL)</td>
<td>- Password protected access</td>
<td>- Limited, monitored access</td>
<td></td>
</tr>
<tr>
<td>- Non-Disclosure Agreements</td>
<td>- Anonymization of data</td>
<td>- Security incident protocol</td>
<td></td>
</tr>
</tbody>
</table>
Challenges

- Overburdened data teams at stakeholder organizations that do not have the time or incentive to assist with extra data requests
- Lack of staffing and resources to handle external data requests
- Lack of internal processes and standards within agencies to request, analyze, and store data extracts
- Lack of resources to handle the legal/governance review
- Some stakeholders have basic systems that are not equipped to process data pulls
- Lack of trust/stigma in the process and individuals involved
- Concern about individual privacy
Strengths

- Strong support from Hartford’s Mayor
- Strong and continued financial commitment and executive-level involvement from philanthropy
- Agency providers that are ready and willing to collaborate to improve the lives of Hartford residents
- Strong data capacity among some community providers and interest in developing data capacity among others
- Desire to establish data informed practices
Contact us

- gfernando@ctdata.org
- www.ctdata.org/HDC
- @CTOpenData
- facebook.com/ctopendata

- ahnelson@upenn.edu
- www.aisp.upenn.edu
- @AISP

Contact information for CTData and AISP.