Introduction to Data Sharing & Integration

Deja Kemp & Emily Berkowitz
Actionable Intelligence for Social Policy
What we do

- Peer Network
- Guidance & Standards
- Training & Consulting
- Advocacy & Communications
- Actionable Research
## AISPs’ 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</td>
<td>A vendor or vendor recommender</td>
</tr>
<tr>
<td>partners, cheerleaders, and data sharing</td>
<td></td>
</tr>
<tr>
<td>therapists</td>
<td></td>
</tr>
<tr>
<td>Focused on ethical data use for policy</td>
<td>Focused on academic research</td>
</tr>
<tr>
<td>change</td>
<td></td>
</tr>
</tbody>
</table>
Our approach

Data sharing is as relational as it is technical.

We don’t just need to integrate data; we need to integrate people.
What are human service data?

Administrative data:

data collected during the routine process of administering programs

but can also be repurposed to support evaluation, analysis, and research.
Million Dollar Murray

Researcher quoted extensively in this article is Dennis Culhane, the co-founder of AISP.

Read, Million Dollar Murray, The New Yorker, 2/13/2006
Don’t worry, I’m not going to read this to you.

“In the nineteen-eighties, when homelessness first surfaced as a national issue, the assumption was that the problem fit a normal distribution: that the vast majority of the homeless were in the same state of semi-permanent distress. It was an assumption that bred despair: if there were so many homeless, with so many problems, what could be done to help them? Then, fifteen years ago, a young Boston College graduate student named Dennis Culhane lived in a shelter in Philadelphia for seven weeks as part of the research for his dissertation. A few months later he went back, and was surprised to discover that he couldn’t find any of the people he had recently spent so much time with. “It made me realize that most of these people were getting on with their own lives,” he said. Culhane then put together a database— the first of its kind—to track who was coming in and out of the shelter system. What he discovered profoundly changed the way homelessness is understood. Homelessness doesn’t have a normal distribution, it turned out. It has a power-law distribution. “We found that eighty per cent of the homeless were in and out really quickly,” he said. “In Philadelphia, the most common length of time that someone is homeless is one day. And the second most common length is two days. And they never come back. Anyone who ever has to stay in a shelter involuntarily knows that all you think about is how to make sure you never come back.”
When people have a home:

- They spend less time in hospitals, police custody & shelters.

10 people + 10 homes saved $668,000 in 6 months.

Housing First. Think about it.
What is the difference?

Data sharing

Data integration

Person A

Person B

Person C

Person D
When we bring data together we can:

• Understand the complex needs of individuals and families

• Allocate resources where they’re needed most to improve quality and equity of services

• Measure long-term impacts of policies and programs

• Engage in transparent, shared decision-making about how data should (and should not) be used
Our Network

Network of ~36 operational state and local integrated data systems
Our developmental approach
We focus on the kale.
Data Access

**OPEN 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.

**RESTRICTED DATA**
Data that can be shared, but only under specific circumstances with appropriate safeguards in place.

**UNAVAILABLE DATA**
Data that cannot or should not be shared, either because of state or federal law, lack of digital format (paper copies only), or data quality or other concerns.

Data used for “Million Dollar Murray” analysis
Our Framework
The Four Questions

- Is it legal?
- Is it ethical?
- Is it a good idea?
- How do we know? Who decides?

Finding a Way Forward: How to create a strong legal framework for data integration, 2022
1. Is this legal?
Authority and Access
2. Is this ethical?
Social License, Risk v. Benefit

- Mapping indicators to allocate new investments to high-need neighborhoods
- Program evaluation with longitudinal outcomes
- Unduplicated counts of children across early childhood programs
- Open data initiatives that publish aggregate data sets

- Linking individual data on wages & earnings
- Case management algorithms
- Using “risk scores” to target interventions
- Predictive analytics in policing
- Tracking social media on students
- Linking biometric data (e.g., facial recognition)
3. Is this a good idea?  
Data availability, Resources, Action

- Is there available data to answer this question?
- Do we have the resources to respond to the answer?
- Can these data be acted upon?
4. How do we know? Who decides?

Data governance

The people, policies, and procedures that support how data are managed, used, and protected.

Strong and inclusive data governance for cross-sector data sharing and integration should be:

• Purpose-, value-, and principle-driven
• Strategically located
• Collaborative
• Iterative
• Transparent
Benefit/Risk Matrix

GOVERNANCE is how you know that you’re operating in the green
Data & Violence Prevention
Violence Against Women Act

• Federal statute is the main mechanism that provides funding for housing, legal assistance and other supportive measures (ie. this act is the reason that it is usually free to file a restraining/protective order)

• Protects the confidentiality of “personally identifiable information” (PII) collected in connection with services

• Incredibly restrictive
## Data Access

<table>
<thead>
<tr>
<th>OPEN DATA</th>
<th>RESTRICTED DATA</th>
<th>UNAVAILABLE DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>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.</td>
<td>Data that can be shared, but only under specific circumstances with appropriate safeguards in place.</td>
<td>Data that cannot or should not be shared, either because of state or federal law, lack of digital format (paper copies only), or data quality or other concerns.</td>
</tr>
</tbody>
</table>

VAWA
**Violence Against Women Act**

Can be shared under very limited circumstances:

- Informed written consent
- Aggregate de-identified
- Judicial or legal process (court order or mandate, subpoena, warrant for limited purposes)
Survey + admin data paint a picture

Homeless Services. In the Outreach Encampment Survey, 57 percent of the respondents indicated having spent time in a Philadelphia homeless shelter. In contrast, Table 5d shows that 38 percent of the people on the BNL had a record of a shelter stay prior to the ERP implementation. That proportion increased substantially after the ERP started, as over half of those on the BNL (98 people or 51.9 percent) used some shelter or temporary housing during the ERP implementation and sustainment periods. Much of this latter proportion reflects use of the navigation and respite centers, although the finding reflects use of any shelter in the City.

Table 5d. Prevalence of Involvement in Services Provided OHS by People on the BNL (n=189)

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Shelter or Temporary Housing</th>
<th>Safe Haven</th>
<th>Journey of Hope</th>
<th>Outreach Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-ERP (before May 2018)</td>
<td>72 (38.1%)</td>
<td>9 (4.8%)</td>
<td>3 (1.6%)</td>
<td>108 (57.1%)</td>
</tr>
<tr>
<td>ERP period (after May 2018)</td>
<td>98 (51.9%)</td>
<td>16 (8.5%)</td>
<td>13 (6.9%)</td>
<td>103 (54.5%)</td>
</tr>
<tr>
<td>Lifetime</td>
<td>119 (63.0%)</td>
<td>22 (11.6%)</td>
<td>15 (7.9%)</td>
<td>131 (69.3%)</td>
</tr>
</tbody>
</table>
Lessons learned

| Table 5k. Engagement in Services for People on the BNL: June 1 and June 26, 2018 (n=192) |
|---------------------------------|-------------|-------------|
|                                | June 1      | June 26     |
| Total on BNL                   | 110 (100%)  | 192 (100%)  |
| Placements:                    |             |             |
| PPP Navigation Center          | 35 (31.8%)  | 39 (20.3%)  |
| PPP Respite Center             | 8 (7.3%)    | 26 (13.5%)  |
| ODAAT Respite Center           | 2 (1.8%)    | 7 (3.6%)    |
| Drug Detox or Treatment Center | 7 (6.4%)    | 19 (9.9%)   |
| Safe Haven Facility            | 6 (5.5%)    | 9 (4.7%)    |
| Recovery Housing               | 1 (0.9%)    | 1 (0.5%)    |
| Return Home                    | 1 (0.9%)    | 1 (0.5%)    |
| VA Housing                     | 1 (0.9%)    | 1 (0.5%)    |
| Salvation Army                 | 0           | 1 (0.5%)    |
| No Placements                  | 49 (44.5%)  | 86 (44.8%)  |

Note. The 192 names listed in the June 26 report was subsequently unduplicated to 189 names.

| Table 5t. Summary of Placements for People on the BNL: October 15, 2018 (n=189) |
|---------------------------------|-------------|-------------|
|                                | October 15  |
| Long-term placement or permanent housing | 36 (19.0%)  |
| In time-limited placements      | 19 (10.1%)  |
| Unable to access placement      | 12 (6.3%)   |
| Contact but no ongoing placement| 77 (40.7%)  |
| Unengaged                       | 45 (23.8%)  |
Questions?

Deja Kemp, dejak@upenn.edu
Emily Berkowitz, eberko@upenn.edu

Check out our publications & quality framework: www.aisp.upenn.edu

Subscribe to our newsletter: https://bit.ly/signupAISPnewsletter