National Pedestrian and Bicycle Safety Data Clearinghouse
Phase I: Inventory & Framework

Krista Nordback, Wesley Kumfer, Seth LaJeunesse, Libby Thomas, Katie Heuser

University of North Carolina Highway Safety Research Center

Julia Griswold

University of California, Berkeley
Research Plan

• Literature Review
• Interviews
• Inventory
• Rating
• Identify Data Gaps
• Framework and Website Development
• Conclusions and Next Steps
Literature Review

• Data demands from safety analysis modeling
• Available data and common variables by type:
  – Collision Data
  – Exposure Data
  – Spatial/Socio-Demographic/Socio-Economic Data
  – Infrastructure/Land Use
• 25 databases identified in the literature (international)

<table>
<thead>
<tr>
<th>Mode</th>
<th>Types</th>
<th>Number of databases identified in the literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedestrian</td>
<td>Collision</td>
<td>5</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>Counts</td>
<td>2</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>Infrastructure</td>
<td>3</td>
</tr>
<tr>
<td>Bicycle</td>
<td>Collision</td>
<td>2</td>
</tr>
<tr>
<td>Pedestrian and Bicycle</td>
<td>Collision</td>
<td>7</td>
</tr>
<tr>
<td>Pedestrian and Bicycle</td>
<td>Counts</td>
<td>5</td>
</tr>
<tr>
<td>Pedestrian and Bicycle</td>
<td>Infrastructure</td>
<td>4</td>
</tr>
<tr>
<td>Pedestrian and Bicycle</td>
<td>Other</td>
<td>3</td>
</tr>
</tbody>
</table>
Interviews

Eight Pedestrian/Bicycle Safety Researchers Interviewed

1. What resources do you use?
2. What data do you need?
3. What data are missing that you’d like to have?
4. What % of budget goes to data collection?
5. What % of budget goes to data cleaning/processing?
6. What features do you want?
7. How might a clearinghouse help you?
8. Online datasets we should link to the clearinghouse?
## Interview Summary

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data used</td>
<td>Collisions, context, exposure, infrastructure, behavior, socio-demographics, video, app data</td>
</tr>
<tr>
<td>Data needs</td>
<td>Collision, exposure, infrastructure, socio-demographics, transit data</td>
</tr>
<tr>
<td>Data missing</td>
<td>Hospital, crash typing, risk perception, exposure, sidewalks, infrastructure details, installation dates, AV ped detection, better quality</td>
</tr>
<tr>
<td>% of budget for data collection</td>
<td>10% to 75%, most say less than 30%</td>
</tr>
<tr>
<td>% of budget for data cleaning</td>
<td>5% to 70%, most say less than 30%</td>
</tr>
<tr>
<td>Features desired</td>
<td>Data: Hospital, weather, exposure (counts), infrastructure specifics, speed, perception, socio-demographics, transit usage, time frame of data</td>
</tr>
<tr>
<td></td>
<td>Features: cross tab tool, download data, visualization, glossary, upload and share, mapping, analysis, categories and classification, link to other online tools, standardize quality, ease of navigation, documentation of data cleaned</td>
</tr>
<tr>
<td>Uses of Clearinghouse</td>
<td>Identify long time series data, improve cross-sectional studies, planning and operations, repository for bikes hare and exposure data</td>
</tr>
<tr>
<td>Existing online datasets</td>
<td>FARS, HSIS, NASS, state crash data, NHTS, state hospital records, SHRP2, national/state surveys, Metropolitan Travel Survey Archive, UTC datasets, Bike-Ped Portal, TIMS, Strava, Bikeshare, Census, Google Earth</td>
</tr>
</tbody>
</table>
Inventory of Online Datasets

- Online search of all 50 states, MPOs, cities >100,000, etc
- 4,126 publicly-available ped/bike-safety related datasets found
- Categories by type
  - Collision
  - Counts
  - Infrastructure
- Geographic Scale
  - City
  - County
  - Region
  - State
  - National
Dataset by Type

Percent of Datasets in Inventory

<table>
<thead>
<tr>
<th></th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collisions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>482</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Counts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,355</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,463</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

www.roadsafety.unc.edu
Datasets by Geographic Scale

<table>
<thead>
<tr>
<th>Scale</th>
<th>Percent of Datasets in Inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>City</td>
<td>2,190</td>
</tr>
<tr>
<td>County</td>
<td>521</td>
</tr>
<tr>
<td>Region</td>
<td>348</td>
</tr>
<tr>
<td>State</td>
<td>1,023</td>
</tr>
<tr>
<td>National</td>
<td>89</td>
</tr>
</tbody>
</table>
Rating of Datasets

State-wide, non-PDF datasets were rated on a 5-star scale with 5 being best based on
- Temporal Completeness
- Spatial Completeness
- Linkability

<table>
<thead>
<tr>
<th></th>
<th>1 Star</th>
<th>2 Stars</th>
<th>3 Stars</th>
<th>4 Stars</th>
<th>5 Stars</th>
<th>Total Number of Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collisions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collision Event</td>
<td>0</td>
<td>1</td>
<td>42</td>
<td>40</td>
<td>17</td>
<td>100</td>
</tr>
<tr>
<td>Collision Summary</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>7</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>Counts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ped/Bike Short Duration Counts</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>16</td>
</tr>
<tr>
<td>Ped/Bike Permanent Counts</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>AADT Counts</td>
<td>0</td>
<td>15</td>
<td>78</td>
<td>100</td>
<td>62</td>
<td>255</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>9</td>
<td>230</td>
<td>41</td>
<td>20</td>
<td>7</td>
<td>307</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>247</td>
<td>176</td>
<td>177</td>
<td>92</td>
<td>701</td>
</tr>
</tbody>
</table>
Data gaps

Data needs and gaps by priority:

1. Pedestrian and Bicycle Count Data provide a fundamental metric of exposure. Both high quality Short Duration and Permanent count data are needed, since so little is currently available on exposure.

2. The ability to link datasets needs to be improved.

3. The accuracy of data needs to be better assessed and improved.

4. Funding for data collection and management for pedestrian and bicycle safety related data needs to increase.
Framework and Website: pedbikedata.org
Search Results Expanded

**City: 89 Results**

**County: 5 Results**

**Palm Beach MPO - Regional Greenways**
*Infrastructure (Bicycle Facilities, Pedestrian Facilities)*
*GIS | Palm Beach, CA | Publicly Available | SEE DETAILS*

**Traffic counts and collisions**
*Collisions, Counts (Motor Vehicle), Infrastructure (Street Network/Centerlines, Pedestrian Facilities)*
*GIS | Larimer County, CO | Publicly Available | SEE DETAILS*

**Lee County Bike-Ped Facilities**
*Infrastructure (Bicycle Facilities, Pedestrian Facilities)*
*GIS | Lee County, FL | Publicly Available | SEE DETAILS*

**multi-use pathways and sidewalks in Ada and Canyon counties**
*Infrastructure (Pedestrian Facilities)*
*GIS | Ada and Canyon counties, ID | Publicly Available | SEE DETAILS*

**Transportation Data, Ramsey County, Minnesota**
*Infrastructure (Street Network/Centerlines, Bicycle Facilities, Pedestrian Facilities)*
*GIS | Ramsey, MN | 2015 (every 3 years) | Publicly Available | SEE DETAILS*

**Region: 6 Results**
Palm Beach MPO - Regional Greenways

Agency/Owner: Palm Beach MPO
Data Type: Infrastructure (Bicycle Facilities, Pedestrian Facilities)
Facility/Treatment Types: greenways
Geographic Scale: County
City/County: Palm Beach
State: CA
Date Ranges:
Format: GIS
Availability: Publicly Available
MPO Name: Palm Beach MPO
Static Data: No
URL: https://www.arcgis.com/apps/Viewer/index.html?appid=c031df4e2d44afe98ed42e934579dd
Next Steps

• Phase 2 – User Experience Improvements (2019-2020)
  – Interview and survey users
  – Prioritize and Select Improvements
  – Make Improvements
  – Outreach
    • Report back to interviewees and other stakeholders
    • Spread the word about the Clearinghouse

• Potential improvements
  – Improve geographic search capability (add a map)
  – Improve temporal search capability
  – Expand ratings to include all datasets and add to search
  – Add online submission form for broken links
  – Post a data dictionary
  – Quality check existing Clearinghouse