Practical Advice for Working with 3rd Party Data Providers

Michael Pack, Director of CATT Laboratory
Who we are, and why we know this stuff...

RITIS Data Analytics

• 300+ data feeds from
  • State/Local DOTs
  • First Responders
  • Transit
  • Weather
  • Federal Gov
  • Military

• 8.7+ Billion data points every day

• 3rd Party Data Providers
  • HERE
  • INRIX
  • TomTom
  • WAZE
  • Verizon
  • Airsage
  • Verifone
  • Etc.
The focus of today’s presentation...

- Working with 3rd party data providers
  - First = Crowdsourced event/incident data
  - Second = Probe, O-D, and Trajectory providers
Crowdsourcing

What is crowdsourcing?
• “Outsourcing to the crowd”
• “Voluntary undertaking of a task” by a group of diverse individuals
• Mutual benefit

How is it different from what we did so far?
• Reliance on “unknown” to provide valuable information
• Develop trust in “wisdom of crowds”
• Access to a much wider view

Sources:
Howe, Jeff (June 2, 2006). “Crowdsourcing: A Definition”
Estellés-Arolas, Enrique; González-Ladrón-de-Guevara, Fernando (2012), “Towards an Integrated Crowdsourcing Definition”
Waze Connected Citizens Program (CCP)

The Waze Connected Citizens Program (CCP) is a free, two-way data exchange empowering municipal decisions to achieve concrete community impact. Launched in October 2014 with 10 city partners, the program has expanded to more than 63 partners including city, state and country government agencies, nonprofits and first responders.

**WAZE** provides real-time, anonymous, proprietary incident and slow-down information directly from the source: drivers themselves.

**PARTNERS** provide real-time and advance information on government-reported construction, crash and road closure data.

Source: Waze
Size Matters!

**Note:**
- Waze data excludes jams event type
- 3 Month Period of 3/17 – 5/17 displayed

### Avg Waze Events By Day of the Week

<table>
<thead>
<tr>
<th>Day of Week</th>
<th>CA</th>
<th>DC</th>
<th>FL</th>
<th>IA</th>
<th>MA</th>
<th>PA</th>
<th>VA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>15,000</td>
<td>1,000</td>
<td>3,000</td>
<td>500</td>
<td>2,000</td>
<td>1,500</td>
<td>1,000</td>
</tr>
<tr>
<td>Tuesday</td>
<td>10,000</td>
<td>700</td>
<td>2,000</td>
<td>300</td>
<td>1,500</td>
<td>1,000</td>
<td>800</td>
</tr>
<tr>
<td>Wednesday</td>
<td>5,000</td>
<td>300</td>
<td>1,000</td>
<td>200</td>
<td>750</td>
<td>500</td>
<td>400</td>
</tr>
<tr>
<td>Thursday</td>
<td>1,000</td>
<td>100</td>
<td>200</td>
<td>50</td>
<td>75</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>Friday</td>
<td>50,000</td>
<td>5,000</td>
<td>1,500</td>
<td>250</td>
<td>1,250</td>
<td>1,000</td>
<td>750</td>
</tr>
<tr>
<td>Saturday</td>
<td>25,000</td>
<td>2,500</td>
<td>500</td>
<td>100</td>
<td>1,250</td>
<td>1,000</td>
<td>750</td>
</tr>
<tr>
<td>Sunday</td>
<td>10,000</td>
<td>1,000</td>
<td>250</td>
<td>50</td>
<td>750</td>
<td>600</td>
<td>450</td>
</tr>
</tbody>
</table>

### Avg DOT Events By Day of the Week

<table>
<thead>
<tr>
<th>Day of Week</th>
<th>CA</th>
<th>DC</th>
<th>FL</th>
<th>IA</th>
<th>MA</th>
<th>PA</th>
<th>VA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monday</td>
<td>500</td>
<td>50</td>
<td>25</td>
<td>5</td>
<td>25</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Tuesday</td>
<td>300</td>
<td>30</td>
<td>15</td>
<td>3</td>
<td>15</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>Wednesday</td>
<td>150</td>
<td>15</td>
<td>7</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Thursday</td>
<td>50</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Friday</td>
<td>2,500</td>
<td>250</td>
<td>75</td>
<td>15</td>
<td>75</td>
<td>50</td>
<td>35</td>
</tr>
<tr>
<td>Saturday</td>
<td>1,250</td>
<td>125</td>
<td>35</td>
<td>7</td>
<td>35</td>
<td>25</td>
<td>20</td>
</tr>
<tr>
<td>Sunday</td>
<td>600</td>
<td>60</td>
<td>15</td>
<td>3</td>
<td>15</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>

**Note:**
- Waze data excludes jams event type
- 3 Month Period of 3/17 – 5/17 displayed
## Waze Data Background

<table>
<thead>
<tr>
<th>Type of Event</th>
<th>% DOT Matched to Waze</th>
<th>Average Time that a Waze Event was Reported Before a DOT Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freeways/Ramps Crashes</td>
<td>40%</td>
<td>3 Minutes</td>
</tr>
<tr>
<td>Primary/Secondary Crashes</td>
<td>12%</td>
<td>3 Minutes</td>
</tr>
<tr>
<td>Freeways/Ramps Disabled Vehicles</td>
<td>37%</td>
<td>14 Minutes</td>
</tr>
<tr>
<td>Primary/Secondary Disabled Vehicles</td>
<td>4%</td>
<td>16 Minutes</td>
</tr>
</tbody>
</table>
Working with Waze

Working with the Data

- Redundancy
- Feedback loops
- Size
- Credibility and filtering
- Increased Coverage
- Faster Response
- The ability to truly influence route-choice

Working with the Company

- Legal
- Negotiations
- Nothing is *really* free (but it's still worth it)
Probe Data Providers and Services—everyone’s a little different

- Speeds and travel times
- Data Feeds & APIs
- Tiles
- O/Ds
- Trajectory
- Mapping
- Some are working on volumes and turning movements
- Much much more coming soon!!!
- Not all provide the same type of data, the same format, etc.
  - TMC segments/standards differ
  - Capped data at speed limit
  - Additional elements (freeflow or historic average available by some, but not all)
  - Various types of quality indicators provided by each
  - Deviant data only
• But…

YOU the purchaser can ruin it!!!
I mean, really really ruin it.
Procurements can go wrong.
And you can also get played.
Don’t make these mistakes

• DUAs – You have more power than you think!
  • Fight for Great Acceptable use
  • Fight for (and think about) Sharing with partners
  • Don’t just do what your neighbor did (but ask them)
  • Look for model DUAs (I-95 CC for probe data)

• Sharing back with the provider the way YOU want to share it
  • (don’t permanently dumb down your data)

• Treat your provider as part of your team, NOT a whipping boy

• Be open to communication and vendor discussions

• Don’t blend “all” of the features/requirements—EVER!

• Payment terms based on quality and uptime (where applicable)

• Stop focusing on how to pay less. Instead, work to try to get more!
Are You Gonna Go My WAZE?

Practical Advice for Working with 3rd Party Data Providers

By Michael Pack and Nikola Ivanov
Considerations of Current and Emerging Transportation Management Center Data

The purpose of this project is to identify, research and synthesize current, emerging and potential future practice of public and private sectors real-time traffic data collection, sharing and dissemination. This includes an assessment of private sector business models; public-private partnerships and agreements; utilization of private sector data to supplement traffic data collected by public agencies; value considerations of sharing public agency data with private sectors; and consideration of emerging data sources such as crowdsourcing and connected vehicles. This report is intended to serve as a detailed reference that addresses the concepts, business models, methods, processes, techniques, and other related issues for practitioners to consider associated with real time traffic data collection and dissemination.

Project to be completed by Dec. 2018

Consultant = Leidos