



CHANGING WHAT'S POSSIBLE

Breakout Group: Traveler & Traffic Data Discussion Summary

March 10, 2014

Key Questions for Breakout Session

1. Where are the gaps in the available travel data? (including data from GPS-enabled devices, roadway sensors, transponders, etc.)
2. What additional data do we expect to be widely available in the next three years?
 - Lots of data available or potentially available. (Little actually used today.)
 - Transportation energy optimization questions have rarely been asked.
 - ▶ Transportation research community usually safety or congestion focused.
 - Data Categories:
 - ▶ Physical Data Elements
 - Energy consumption and efficiency data of vehicles (real-time)
 - Acceleration / Deceleration
 - Pedestrian counts / Bicycle counts (substantially behind vehicle counts)
 - Vehicle occupancy
 - Detailed road geometry / real-time conditions (temp. construction)

Key Questions for Breakout Session

1. Where are the gaps in the available travel data? (including data from GPS-enabled devices, roadway sensors, transponders, etc.)
 2. What additional data do we expect to be widely available in the next three years?
- **Data Categories:**
 - ▶ **Data sets**
 - OEM data sets (differences between OEMs, new vs. old vehicles)
 - Real-time O/D pairs increasingly available but difficult to access
 - Very difficult to anonymize
 - Law enforcement data
 - ▶ **Data-derived information**
 - **Trip purpose**
 - Trip chaining based on behavior patterns
 - Correlation among individuals (to coordinate delays)

Key Questions for Breakout Session

3. How is the usage of this data (current and future) restricted, due to security considerations, privacy rules or other regulations?

4. How should data from disparate sources be integrated for use in network models?

- **Individual privacy is dead**
 - ▶ Individuals consent to sharing data commonly. (Value propositions)
 - ▶ BUT, privacy in aggregated data sets is complicated.
 - ▶ AND, privacy of data is really complex anytime a government, public agency is involved.
- Synthetic, anonymous data sets are very difficult to create and validate
- Disparate source integration:
 - ▶ **What problem are you trying to solve? What is the use case?**
 - ▶ Probably clever ideas that can create “magical” solutions
- **FINAL THOUGHT: People do not trust data. How do you make them trust data?**