



# Data and Privacy: Research Needs and Dockless Mobility Data

Genevieve Giuliano, Professor and Director  
Zakhary Mallett, PhD student  
University of Southern California

Joint Senate Committees on Judiciary and Transportation  
of the State of California

February 25, 2020

# Presentation main points

1. Academic research is in the public interest
2. Transportation research depends on disaggregate data
3. Data collection is not the same as data sharing
4. Privacy concerns can be addressed

# 1. Academic research is in the public interest

- Academic research generates new knowledge
  - Better understanding of how people, organizations behave
  - Better understanding leads to informed public policy, business and consumer decisions
  - New tools, methods, contribute to more efficient public investments, services

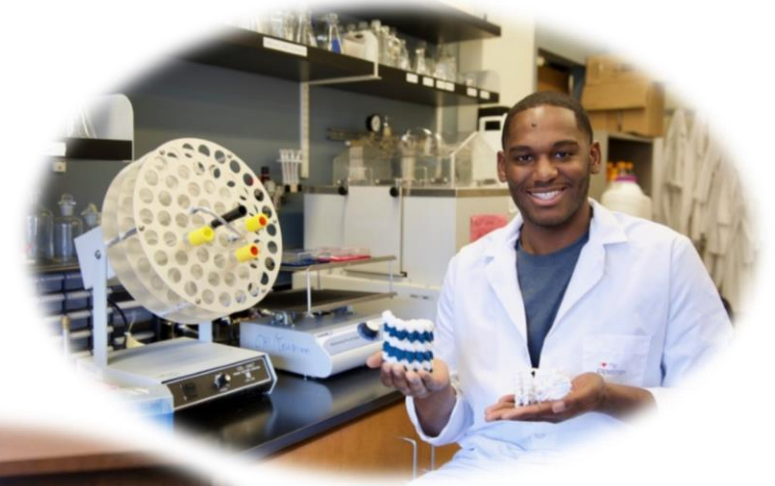


Photo Credits:  
University of North Carolina, Charlotte; The Farmville Herald

# 1. Academic research is in the public interest

## Research Quality Considerations

- Generalizability
- Comprehensiveness
- Independence/Objectivity
- Peer Review

*Academic research process to assure research quality*

# 1. Academic research is in the public interest

- Importance of travel behavior research
  - Changes in travel essential to reach California's GHG reduction targets
- Understandings and data
  - Historic methods/sources no longer adequate
  - Contemporary challenges
    - Emergence of new modes
    - Greater role of private sector



Photo Credits:  
Columbus Neighborhoods, CU Independent, TripSavvy

# 1. Academic research is in the public interest

## Some example policy and planning questions

- Should the public invest in “supportive” infrastructure?
- How should curb space be prioritized?
- Should public transit partner with dockless mobility or ridehail services?
- Should private providers be subject to the same regulations as public providers?

*Answers depend on research....*



Photo Credits:  
Autotrader, Houston Chronicle, LAist

## 2. Transportation research depends on disaggregate data

### Disaggregate data is far superior to aggregate data

- Scalability
  - Provides the base unit; can be aggregated to various levels
- Flexibility
  - Allows for the broadest types of research; answer many different questions
- Discoveries
  - Unanticipated patterns, relationships
- *Disaggregate data has long history in transport research*
  - Travel diaries, fare card traces, stated preference surveys, etc.

# 2. Transportation research depends on disaggregate data

## Some examples of research questions

- Do dockless mobility and ridehail services complement, substitute for, or add to, other travel modes?
- What environmental and sociodemographic factors explain usage patterns?
- What is the price elasticity of the use of ridehail or dockless mobility services?



*These questions best addressed with individual level disaggregate data for analysis*

Photo Credits:  
San Francisco Chronicle, Star Tribune, Wikipedia



## 2. Transportation research depends on disaggregate data

### Examples of how disaggregate data leads to new knowledge

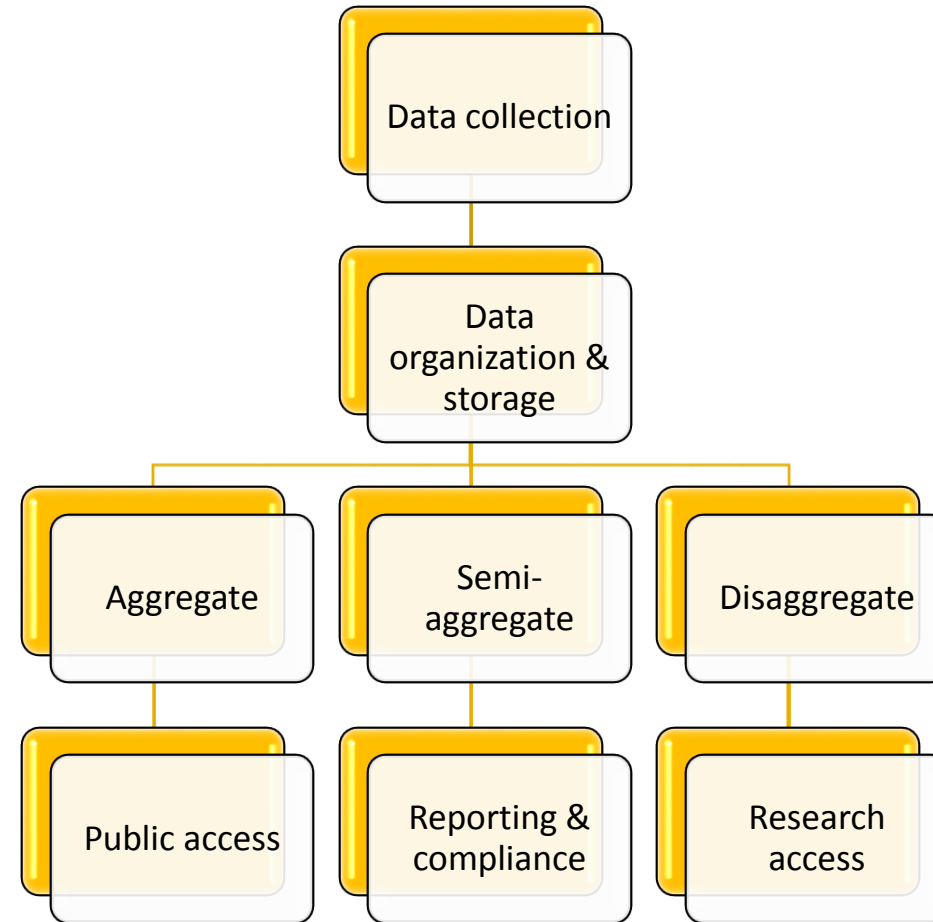
- 10% of Lyft riders, 50% of trips in Los Angeles (Brown, 2017)<sup>1</sup>
- Gentrification in transit-rich communities, when matched with increased density, reduces VMT (Chatman et al., 2019)<sup>2</sup>
- Understanding of *rideshare* (not ridehail) behavior based on disaggregate data

<sup>1</sup>Brown, Anne Elizabeth. *Ridehail revolution: Ridehail travel and equity in Los Angeles*. Diss. UCLA, 2018.

<sup>2</sup>Chatman, Daniel G., et al. "Does transit-oriented gentrification increase driving?." *Journal of Planning Education and Research* 39.4 (2019): 482-495.

# 3. Data collection is not the same as data sharing

- Data can be securely collected and stored
  - Data standards
  - Central data warehouse
- Data sharing can be differentiated
  - Public
  - Local governments
  - Academic researchers



# 4. Privacy concerns can be addressed

- Secure transmission and data storage facilities
- Stripping of identifiers
- Strict data access rules
- Non-disclosure agreements

```
1644014 3 101714 1.0799 6 9 100 6 2 11 3599964
2.1599 2.1599 789 2.1599 0.00022 17639 36738 17724
36805 2.1599 531 116 134 3 2 3 531 116
134 3 2 3 0 60003200
1644014 101714 3 1.0799 6 9 100 6 2 11 3599964
2.1599 2.1599 1563 2.1599 0.00043 17724 36805 17639
36738 2.1599 1155 188 211 3 2 4 1155 188
211 3 2 4 0 60003200
1644015 3 101717 0.7243 6 9 100 6 2 11 3599964
1.4487 1.4487 3 1.4487 0.00000 17639 36738 17614
36806 1.4487 3 0 0 0 0 0 3 0 0 0
0 0 0 60003200
1644015 101717 3 0.7243 6 9 100 6 2 11 3599964
1.4487 1.4487 12 1.4487 0.00000 17614 36806 17639
36738 1.4487 10 2 0 0 0 0 10 2 0
0 0 0 0 60003200 1644009 101724 4 0.2062 4 9
100 6 2 11 3599964 0.4948 0.4948 3857 0.4948
0.00107 17943 35981 17922 35981 0.4948 2886 519
406 16 11 1
```



# 4. Privacy concerns can be addressed

*Universities have extensive infrastructure to protect and secure confidential data*

- Human subject protection
- Non-disclosure agreements
- Secure data infrastructure
- Restricted data access

*Universities have long history of using and storing confidential data*

- California Household Travel Survey (CHTS)
- California individual-level tax data
- US Medicare database

# If we had more time....

- Academics working with private industry
- Personal v service provider data collection
- Examples of public data



Quality Data



Common Knowledge



Informed Decision-Making

Photo Credits:  
Greentech Media; Los Angeles Times; City of Taylor, MI