Self-driving Cars & Data Collection

Privacy Perceptions of Networked Autonomous Vehicles

Cara Bloom, Joshua Tan, Javed Ramjohn, Lujo Bauer
Why networked autonomous vehicle (AV) privacy?
Why networked AV privacy?

1. Data collection capabilities
Why networked AV privacy?

1. Data collection capabilities
2. Operated by a private company
Why networked AV privacy?

1. Data collection capabilities
2. Operated by a private company
3. Collection of physical information in public
Research Goal

Discover what is ‘reasonable’ data collection and use for autonomous vehicle (AV) fleets
Research Goal

1. What do people think AV fleets are capable of?
Research Goal

1. What do people think AV fleets are capable of?
2. How comfortable are people with AV fleet capabilities?
Research Goal

1. What do people think AV fleets are capable of?
2. How comfortable are people with AV fleet capabilities?
3. How much effort would people expend to opt out?
Up Next

1. Study Design
2. Findings
3. Policy Applications
Study Design

- Exploratory online survey
Study Design

- Exploratory online survey
- Privacy primed & unprimed groups
Privacy Priming Scenarios

Primary Uses

Secondary Uses
Privacy Priming Scenarios

Primary Uses

- Necessary for autonomous navigation

Secondary Uses
Privacy Priming Scenarios

Primary Uses

- Image Capture
- Aggregation & Storage
- Specific Incident Analysis
- Continuous Analysis

Secondary Uses

Recognition
Identification
Tracking
# Privacy Priming Scenarios

<table>
<thead>
<tr>
<th>Primary Uses</th>
<th>Secondary Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Capture</td>
<td>Non-necessary. Can be achieved with same sensors.</td>
</tr>
<tr>
<td>Aggregation &amp; Storage</td>
<td></td>
</tr>
<tr>
<td>Specific Incident Analysis</td>
<td></td>
</tr>
<tr>
<td>Continuous Analysis</td>
<td></td>
</tr>
</tbody>
</table>
Privacy Priming Scenarios

Primary Uses
- Image Capture
- Aggregation & Storage
- Specific Incident Analysis
- Continuous Analysis

Secondary Uses
- Perceiving People
  - Recognition
  - Identification
  - Tracking
Privacy Priming Scenarios

<table>
<thead>
<tr>
<th>Primary Uses</th>
<th>Image Capture</th>
<th>Aggregation &amp; Storage</th>
<th>Specific Incident Analysis</th>
<th>Continuous Analysis</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Secondary Uses</th>
<th>Perceiving People</th>
<th>Recognition</th>
<th>Identification</th>
<th>Tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perceiving Vehicles</td>
<td>Recognition</td>
<td>Identification</td>
<td>Tracking</td>
</tr>
</tbody>
</table>
Survey Questionnaire

Likelihood of currently occurring

Comfort if currently occurring
Survey Questionnaire

- Likelihood of currently occurring
- Comfort if currently occurring
- General AV questions
- Effort to opt out
- Bias against Uber & demographics
Survey Questionnaire

Likelihood of currently occurring
Comfort if currently occurring
General AV questions
Effort to opt out
Bias against Uber & demographics

Seen by Primed Group
Survey Questionnaire

Likelihood of currently occurring

Comfort if currently occurring

General AV questions

Effort to opt out

Bias against Uber & demographics
Study Design

- Exploratory online survey
- Privacy primed & unprimed groups
- Pittsburgh & four similar cities
Recruitment

- Ads on Craigslist
- Posts on city Subreddits
- Posters *Pittsburgh only*
302 Participants
- 60% male
- 25% in tech fields
- Avg. age 34 [18, 79]
Question: What do people think networked fleets of autonomous vehicles are capable of?
## Privacy Questions

<table>
<thead>
<tr>
<th>Primary Uses</th>
<th>Aggregation &amp; Storage</th>
<th>Specific Incident Analysis</th>
<th>Continuous Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Capture</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Secondary Uses

<table>
<thead>
<tr>
<th>Perceiving People</th>
<th>Recognition</th>
<th>Identification</th>
<th>Tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Perceiving Vehicles</th>
<th>Recognition</th>
<th>Identification</th>
<th>Tracking</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Likelihood of Technical Capability

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

Primary Uses

Secondary Uses
Likelihood of Technical Capability

How likely do you think this scenario is to be happening now?

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles
Likelihood of Technical Capability

How likely do you think this scenario is to be happening now?

Q13. A self-driving car recognizes a vehicle that has been seen by another self-driving car in the fleet
Likelihood of Technical Capability

How likely do you think this scenario is to be happening now?

Q13. A self-driving car recognizes a vehicle that has been seen by another self-driving car in the fleet

For example: Uber knows that different self-driving cars encountered the same vehicle on different days, but does not know who owns the vehicle
Likelihood of Technical Capability

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

Legend:
- Very Unlikely
- Unlikely
- Neither
- Likely
- Very Likely
Likelihood of Technical Capability

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

Legend:
- Very Unlikely
- Unlikely
- Neither
- Likely
- Very Likely
Difference between primary & secondary uses
Likelihood of Technical Capability

- Capture images: 46%
- Aggregate and store info: 38%
- Analyze specific incidents: 22%
- Analyze continuously: 22%
- Recognize individuals: 46%
- Identify individuals: 22%
- Track individuals: 22%
- Recognize vehicles: 46%
- Identify vehicles: 22%
- Track vehicles: 22%

Legend:
- Very Unlikely
- Unlikely
- Neither
- Likely
- Very Likely
Likelihood of Technical Capability

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

Participant misconceptions
Likelihood of Technical Capability

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

Participant misconceptions

Legend:
- Very Unlikely
- Unlikely
- Neither
- Likely
- Very Likely
Question:
How comfortable are people with these potential capabilities?
## Comfort with Capability Scenarios

<table>
<thead>
<tr>
<th>Comfort with Capability Scenarios</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Uses</strong></td>
</tr>
<tr>
<td>Image Capture</td>
</tr>
<tr>
<td>Aggregation &amp; Storage</td>
</tr>
<tr>
<td>Specific Incident Analysis</td>
</tr>
<tr>
<td>Continuous Analysis</td>
</tr>
<tr>
<td><strong>Secondary Uses</strong></td>
</tr>
<tr>
<td><strong>Perceiving People</strong></td>
</tr>
<tr>
<td>Recognition</td>
</tr>
<tr>
<td>Identification</td>
</tr>
<tr>
<td>Tracking</td>
</tr>
<tr>
<td><strong>Perceiving Vehicles</strong></td>
</tr>
<tr>
<td>Recognition</td>
</tr>
<tr>
<td>Identification</td>
</tr>
<tr>
<td>Tracking</td>
</tr>
</tbody>
</table>
Comfort with Capability Scenarios

- Capture images
- Aggregate and store info
- Analyze specific incidents
- Analyze continuously
- Recognize individuals
- Identify individuals
- Track individuals
- Recognize vehicles
- Identify vehicles
- Track vehicles

How comfortable are you with the scenario?
Comfort with Capability Scenarios
Comfort with Capability Scenarios

Differentiation is less clear than for likelihood questions
Comfort with Capability Scenarios

Differentiation is less clear than for likelihood questions
Comfort with Capability Scenarios

Exception: recognition of vehicles
Q25. I would feel _______ if my car was tracked each time it encountered a self-driving car.
Why are people comfortable?

Reasonable benefit
Why are people comfortable?

Reasonable benefit
Ubiquity

P52: “Phones are tracking in the same sense”
Why are people uncomfortable?

P107: “Massive privacy invasion. Not necessary for Uber to navigate their cars.”

Ubiquity
Reasonable benefit

Necessary for AVs
Why are people uncomfortable?

Ubiquity
Reasonable benefit

Necessity for AVs
Consent
Question:
How much effort would people expend to opt out?
Q36. How many minutes would you spend in the system to successfully opt out?
Only Priming had a significant effect on opt-out time (M-W U = 9847.5, p=.02)

Q36. How many minutes would you spend in the system to successfully opt out?
In the Paper:

- More correlation tests
- Comparison of privacy and safety comfort
- Uber-related exposure and bias
Takeaways
Takeaways

1. People differentiate between primary and secondary data uses
Takeaways

1. People differentiate between primary and secondary uses
2. Justifications focused on necessity, consent, and ubiquity
Takeaways 1-2: Policy Application
Takeaways 1-2: Policy Application

Primary uses are reasonable, while secondary uses are not
Takeaways 1-2: Policy Application

Primary uses are reasonable, while secondary uses are not

Possible Exception: Recognition of vehicles
Takeaways

3. Misconceptions about new information
Takeaways

3. Misconceptions about new information
4. Priming had the only significant effect on effort to opt-out
Takeaways 3-4: Policy Application

People will likely react strongly to conversations about autonomous vehicle privacy
Takeaways 3-4: Policy Application

People will likely react strongly to conversations about autonomous vehicle privacy.

And, it may be difficult to relay accurate information.
The Big Picture

- Companies should self-regulate
The Big Picture

- Companies should self-regulate
  .. to get ahead of the narrative
  .. to fulfill reasonable expectations
The Big Picture

- Companies should self-regulate
- Policy should restrict secondary uses of AV-collected information
The Big Picture

- Companies should self-regulate
- Policy should restrict secondary uses of AV-collected information

Self-driving Cars & Data Collection

Privacy Perceptions of Networked Autonomous Vehicles

Cara Bloom, Joshua Tan, Javed Ramjohn, Lujo Bauer
10. A self-driving car recognizes a person that has been encountered before by a different self-driving car in the fleet
   ○ Very Unlikely
   ○ Unlikely
   ○ Neither Unlikely nor Likely
   ○ Likely
   ○ Very Likely

11. Individuals are identified by name when they encounter one of the self-driving cars in the fleet
   For example: *Uber knows that the pedestrian next to one of its self-driving cars is Alice*
   ○ Very Unlikely
   ○ Unlikely
   ○ Neither Unlikely nor Likely
   ○ Likely
   ○ Very Likely

12. Individuals are tracked using each time they encounter one of its self-driving cars in the fleet
   ○ Very Unlikely
   ○ Unlikely
   ○ Neither Unlikely nor Likely
   ○ Likely
   ○ Very Likely
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Overall</th>
<th>PGH</th>
<th>Non-PGH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capture images</td>
<td>16% (20)</td>
<td>14% (13)</td>
<td>19% (7)</td>
</tr>
<tr>
<td>Aggregate and store info</td>
<td>42% (54)</td>
<td>43% (40)</td>
<td>38% (14)</td>
</tr>
<tr>
<td>Analyze specific incidents</td>
<td>36% (46)</td>
<td>36% (33)</td>
<td>35% (13)</td>
</tr>
<tr>
<td>Analyze continuously</td>
<td>43% (55)</td>
<td>39% (36)</td>
<td>51% (19)</td>
</tr>
<tr>
<td>Recognize individuals</td>
<td>54% (70)</td>
<td>57% (52)</td>
<td>49% (18)</td>
</tr>
<tr>
<td>Identify individuals</td>
<td>76% (98)</td>
<td>75% (69)</td>
<td>78% (29)</td>
</tr>
<tr>
<td>Track individuals</td>
<td>76% (98)</td>
<td>78% (72)</td>
<td>70% (26)</td>
</tr>
<tr>
<td>Recognize vehicles</td>
<td>43% (56)</td>
<td>46% (42)</td>
<td>38% (14)</td>
</tr>
<tr>
<td>Identify vehicles</td>
<td>71% (92)</td>
<td>68% (63)</td>
<td>78% (29)</td>
</tr>
<tr>
<td>Track vehicles</td>
<td>85% (95)</td>
<td>84% (67)</td>
<td>88% (28)</td>
</tr>
</tbody>
</table>

Table 2: Discomfort with technological capabilities in different scenarios, overall and by whether participants lived in Pittsburgh. The percentage (count) of participants that were uncomfortable or very uncomfortable with a scenario are shown.
Exposure to Uber & AV technology

- 78% Pgh and 42% non-Pgh were exposed to media
- 64% Pgh and 3% non-Pgh had seen one as a pedestrian
Bias against Uber

- 17% would have answered differently if Uber hadn’t been the example
- 18% would trust a different AV company over Uber to have their best interests in mind
Radar

Front facing cameras

LiDAR

Side and rear cameras

GPS and wireless data

Data processing and storage
### Progression of General Scenarios

#### General Proximity

<table>
<thead>
<tr>
<th>Activity</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking Near</td>
<td>24%</td>
</tr>
<tr>
<td>Driving Near</td>
<td>25%</td>
</tr>
<tr>
<td>Cycling Near</td>
<td></td>
</tr>
<tr>
<td>Being Near in Snow</td>
<td>61%</td>
</tr>
<tr>
<td>Riding In</td>
<td></td>
</tr>
</tbody>
</table>

#### General Privacy

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changes in Job Market</td>
<td>85%</td>
</tr>
<tr>
<td>Image Capture</td>
<td></td>
</tr>
<tr>
<td>Aggregation &amp; Analysis</td>
<td>77%</td>
</tr>
<tr>
<td>Accident Liability</td>
<td></td>
</tr>
<tr>
<td>Becoming More Common</td>
<td>30%</td>
</tr>
</tbody>
</table>