5-Anon

Data Privacy for Connected Vehicles

Brijesh Mishra
Naveen Janarthanan
Umesh Tanniru
Harrison Pierce
Jose Valadez

Berkeley SCHOOL OF INFORMATION

Data Misuse and Regulation

- Case of Wrongful Arrest
  - LEO Summon Geofenced Records

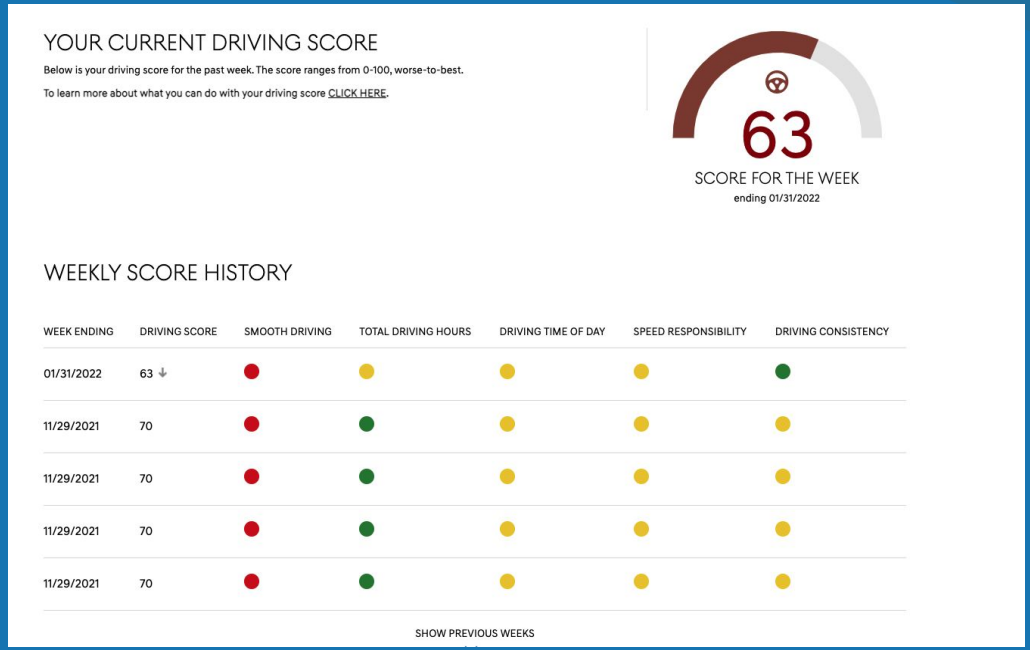
- Presidential Executive Orders
  - Safeguarding Privacy under reproductive health care services

Car Manufacturers NEED to reduce the Privacy Gap between the Data collected and Identity of Car Owners

Berkeley SCHOOL OF INFORMATION

Connected Vehicles & Privacy Concerns

- Connected Vehicles
  - Computer and Software Driven
  - OEM can collect various data from vehicles over-the-air!
- Data Utilization
  - Car Manufacturers:
    - R&D
    - Metrics for Customers
  - Third Parties:
    - Data Brokers
    - Service Providers



YOUR CURRENT DRIVING SCORE

Below is your driving score for the past week. The score ranges from 0-100, worse-to-best.

To learn more about what you can do with your driving score CLICK HERE.

63

SCORE FOR THE WEEK

ending 01/31/2022

WEEKLY SCORE HISTORY

WEEK ENDING	DRIVING SCORE	SMOOTH DRIVING	TOTAL DRIVING HOURS	DRIVING TIME OF DAY	SPEED RESPONSIBILITY	DRIVING CONSISTENCY
01/31/2022	63 ↓					
11/29/2021	70					
11/29/2021	70					
11/29/2021	70					
11/29/2021	70					

SHOW PREVIOUS WEEKS

What can vehicle manufacturers do to raise confidence in vehicle owners about their data privacy?

Berkeley SCHOOL OF INFORMATION

How can Privacy be Addressed?

- Encryption != Privacy Compliance
  - Reversible
  - Backdoor access
  - Insider Threat

- Hashing = Pseudo-Anonymization
  - One way and irreversible
  - Susceptible to patterns
  - Brute force and Replay attack

How can Privacy be Addressed?

Anonymization Techniques

- Data Swapping
- Generalization
- Data perturbation

Technique	Raw Data	Processed Data
Data Swapping	t1={2,5,8,9} t2={11,23,1,7,10}	t1={2,11,1,5} t2={11,8,23,7,9,10}
Generalization	18 mpg	15 - 20 mpg
Data Perturbation	58 mph	2*58mph + 10 = 126 mph

5-Anon

End-to-end data anonymization software development kit (SDK)

- Privacy Analyzer
  - Calculates privacy metrics - probability of disclosures
  - Executes on Server/Storage end
- Parameterization Processor
  - Creates rules for data anonymization
  - Executes on Server/Storage end
- Anonymization Processor
  - Applies the rules on the data making it anonymous
  - Executes at the edge

Demo

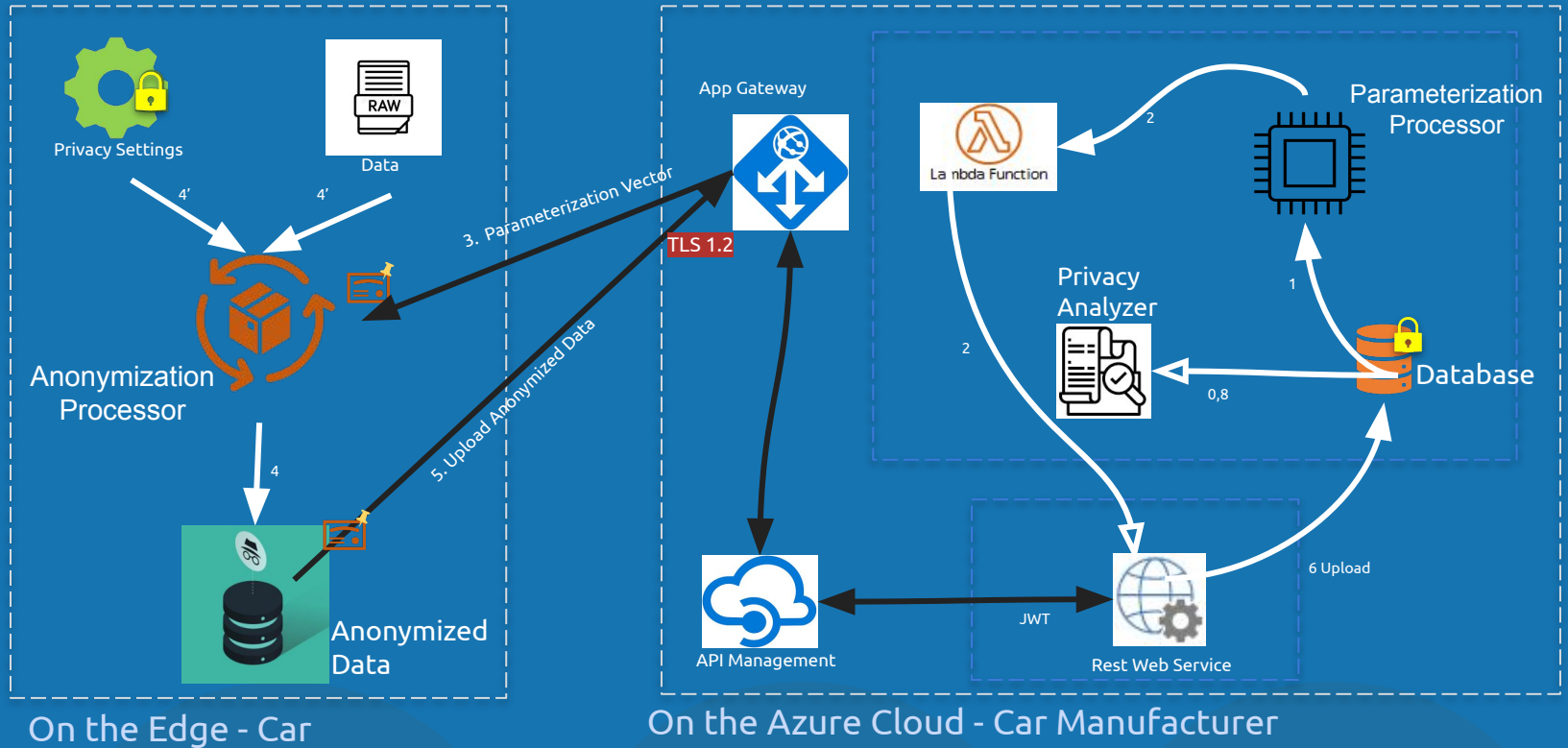
A. Anonymization of BOTH Location Data and Driving Behavior Data

B. Anonymization of ONLY Location Data

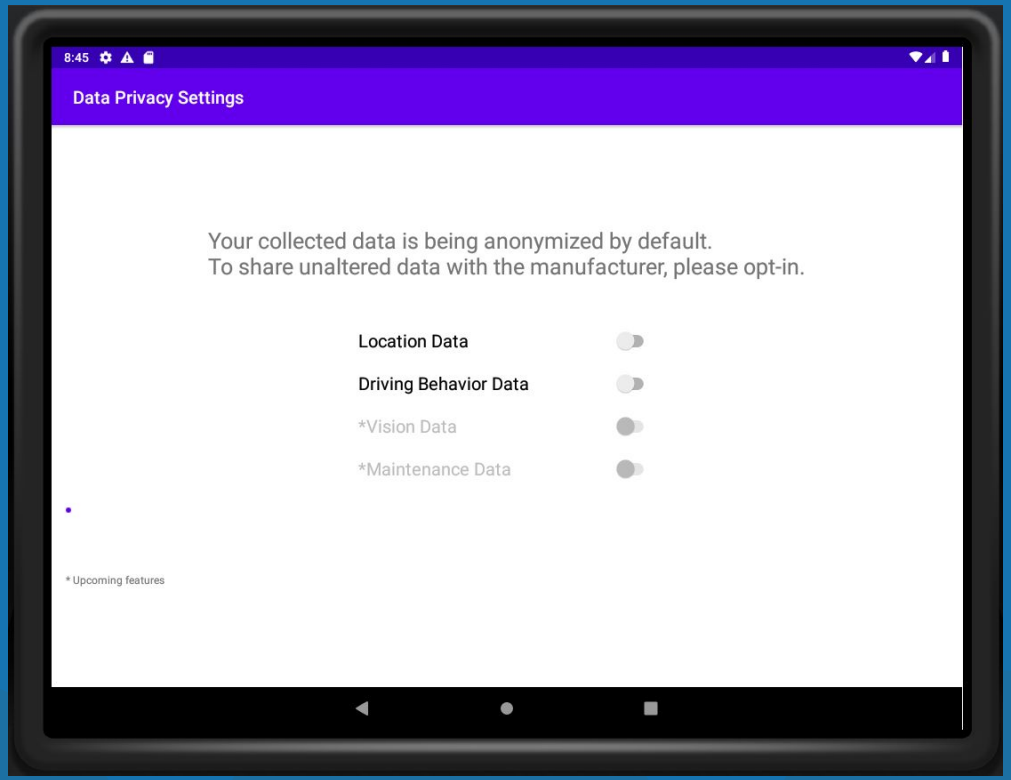
Demo A: Anonymization of BOTH Location Data and Driving Behavior Data

Demo B: Anonymization of ONLY Location Data

5-Anon in Execution



5-Anon in Execution



Advantages

Protection from Data Compromise

- Data Breach
- GM Hack
- Big Tech
- Data Brokerage

Compliance

- Federal Legislation
- Scalable Data Points
- Non-constrained by GDPR

Minimized Cost

- Seamless Integration
- Net Positive Benefit

Looking Forward

- Anonymization of Vision Data and phone synced data [Challenge, not forward benefit]
- Increased testing and validation for continued success with more vehicles
- Enhance anonymization algorithms
- Implement ML techniques for Parameterization processing
- Apply 5-Anon to:
  - Critical Mission Use Cases - Government Fleets (in progress)
  - Integrate it into Ridesharing Services & Rental Companies

Acknowledgements

Dr. Sekhar Sarukkai, Faculty MICS, UC Berkeley

Ryan Liu, Faculty MICS, UC Berkeley

Eric Lybeck, Toyota

Katelyn McCauley, Google

Tom Prevot, Joby Aviation

Families and friends for supporting all the way!

Berkeley SCHOOL OF INFORMATION


Thank you!