What does your smart car know about you?

Charles S. Givre
Senior Lead Data Scientist
Dark Labs, Booz | Allen | Hamilton
givre_charles@bah.com
@cgivre, thedataist.com
Stuff my company wants me to say

- The techniques I demonstrated here are the results of my own research. I have no knowledge of anyone using or not using the techniques demonstrated here.

- The data I gathered all belongs to me and was gathered from devices that I own. Please remember that unauthorized access to someone else’s computer or network IS A CRIME.

- The views presented here represent only my own and not those of my company or anyone else.

- I have no financial interest in any of the products you are seeing here, nor do I have any connection with their parent companies, aside from having purchased their products.

- Please use caution when gathering data from moving vehicles. :-)

darklabs.bah.com
A bit of background
Why cars?

Data gathered by connected devices reveals a lot about its owner.
Data gathered by connected devices reveals a lot about its owner, and most people don’t even know it.
When this data is aggregated and combined with other data, it can be extremely revealing:
When this data is aggregated and combined with other data, it can be extremely revealing:
When are you home?
When are you home?
When are you home?
IoT is just for geeks
A basic car has ??? sensors
A basic car has between 60–100 sensors
A basic car has between 60-100 sensors

- O₂ Sensors
- Throttle Position
- Engine Coolant Temperature
- Catalyst Temperature
- EVAP System Vapor Pressure
- Fuel Pressure
- Accelerator Pedal Position
- Camshaft Position
- Mass Airflow (Measures engine load)
These sensors record ??? fields of data every second
These sensors record nearly 600 fields of data every second
In the US, on board diagnostics (OBD) are required by the Clean Air Act Amendments of 1990.

https://www3.epa.gov/obd/
In Europe, all cars are required to have on-board diagnostics to monitor engine emissions.
But... the data is stuck in your car.
But... the data is stuck in your car, or is it?
The test bed
The Automatic Car Dongle
The Automatic Car Dongle

WTH does it do?
The Automatic Car Dongle

WTH does it do?

- Puts your car’s data to work
- Decodes check engine light diagnostics
- Improves driving with real-time feedback
- 24/7 Crash Response
- See your driving in their dashboard
The Automatic Car Dongle
The Automatic Car Dongle
The Automatic Car Dongle
https://developer.automatic.com/api-reference
https://github.com/adamwulf/automatic-php-api
The Automatic Car Dongle

{
  "url": "https://api.automatic.com/vehicle/XXXXXXXXXXX/",
  "id": "XXXXXXXXXXX",
  "vin": "JM1BL1SG9AXXXXX",
  "make": "Mazda",
  "model": "3",
  "year": 2010,
  "submodel": "Base",
  "color": "#000000",
  "display_name": "Mazda 3",
  "created_at": "2015-03-15T14:15:32.588000Z",
  "updated_at": "2015-03-15T14:15:32.588000Z",
  "fuel_level_percent": 60.54
}
The Automatic Car Dongle

{  
  "url": "https://api.automatic.com/vehicle/XXXXXXXXXX/",
  "id": "XXXXXXXXXXX",
  "vin": "JM1BL1SG9AXXXXX",
  "make": "Mazda",
  "model": "3",
  "year": 2010,
  "submodel": "Base",
  "color": "#000000",
  "display_name": "Mazda 3",
  "created_at": "2015-03-15T14:15:32.588000Z",
  "updated_at": "2015-03-15T14:15:32.588000Z",
  "fuel_level_percent": 60.54
}
What we learned:

- The subject owns a 2010 Mazda 3 and a 2005 Honda Odyssey
- Complete vehicle history (If you want to pay...)
- Subject removed a Hyundai Santa Fe in May 2015 and replaced it with a Honda minivan.
Allows you to build a spreadsheet of all your trips...automatically.
Ultimately, IFTTT is only protected by your username/password.
- Trip Start/Stop
- Coordinates
- Paths
- Duration
- Vehicle Used
Calculate the target’s pattern of life
Trips per day

- Monday: 98
- Tuesday: 110
- Wednesday: 111
- Thursday: 126
- Friday: 102
- Saturday: 3
- Sunday: 84
Where are you going?
Where are you going?

Most common destinations on Thursdays, after 19:00h
• 201 Reisterstown Road, Pikesville MD: 12 Trips
• 2801 Smith Avenue, Baltimore MD: 7 trips
What we learned:

- The subject's work and home addresses
- The subject goes grocery shopping on Thursday nights
- The subject doesn't drive on Friday nights and Saturdays
Trips
pip install polyline

import polyline
polyline.decode('u{~vFvyys@fS}')
<table>
<thead>
<tr>
<th>Distances</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000000</td>
<td>39.37823</td>
<td>-76.67073</td>
<td>0.000000</td>
</tr>
<tr>
<td>0.008891</td>
<td>39.37835</td>
<td>-76.67067</td>
<td>0.016466</td>
</tr>
<tr>
<td>0.011058</td>
<td>39.37851</td>
<td>-76.67067</td>
<td>0.020478</td>
</tr>
<tr>
<td>0.017969</td>
<td>39.37877</td>
<td>-76.67067</td>
<td>0.033277</td>
</tr>
<tr>
<td>0.050259</td>
<td>39.37880</td>
<td>-76.66973</td>
<td>0.093073</td>
</tr>
<tr>
<td>0.027666</td>
<td>39.37840</td>
<td>-76.66971</td>
<td>0.051233</td>
</tr>
<tr>
<td>0.015289</td>
<td>39.37818</td>
<td>-76.66974</td>
<td>0.028313</td>
</tr>
<tr>
<td>0.019731</td>
<td>39.37791</td>
<td>-76.66986</td>
<td>0.036539</td>
</tr>
<tr>
<td>0.028725</td>
<td>39.37756</td>
<td>-76.67015</td>
<td>0.053195</td>
</tr>
<tr>
<td>0.036604</td>
<td>39.37713</td>
<td>-76.67055</td>
<td>0.067785</td>
</tr>
</tbody>
</table>

https://developers.google.com/maps/documentation/utilities/polylineutility
The 60-100 sensors generate nearly 600 columns of data

One row PER SECOND
The **60-100** sensors generate nearly **600** columns of data **per second**

- Accelerator position
- What gear the car is in
- Engine speed
- Vehicle speed
- Clutch/Brake position (some vehicles)
- Accelerometer Information
This data can be used to identify a driver
This data can be used to identify a driver

Driver 1: 2nd Gear Usage & Acceleration
This data can be used to identify a driver.

Driver 2: 2nd Gear Usage & Acceleration

RPM (Thousands)

0.25  1.5  2.75  4
This data can be used to identify a driver

Driver 2: 2nd Gear Usage & Acceleration (Trip 2)
This data can be used to identify a driver.

Driver 2: 2nd Gear Usage & Acceleration (Trip 3)
This data can be used to identify a driver

Driver 1: 2nd Gear Usage & Acceleration

Driver 2: 2nd Gear Usage & Acceleration
Still a work in progress…
Other uses
Other uses: Fuel Optimization
Other uses: Risk Management
Other uses: Fraud Detection

http://www.nytimes.com/2013/02/10/automobiles/stalled-on-the-ev-highway.html

http://www.teslamotors.com/blog/most-peculiar-test-drive
What does this mean?
They’re watching…
Consumers do not understand the implications of ubiquitous data collection
Obvious privacy implications
Data collection is based on a **social contract** between consumers and data collectors.
If car companies are going to collect and retain all this data, they only do so if there is a mutual benefit to the consumer.
Questions?
Thank you!

Charles Givre
@cgivre
givre_charles@bah.com
thedataist.com