Subcommittee on Cybersecurity, Privacy and Data: Geolocation Data

Summary
Geolocation data is data on the current or past location of devices such as GPS units and smartphones. AVs will generate geolocation data, and this data will be critical to helping public agencies plan, design and manage the transportation system to accommodate AVs safely and effectively. However, geolocation data can be used to identify people and discern details of where they live, work and travel, potentially enabling stalking and harassment and revealing sensitive destinations such as personal health appointments. A growing number of tools and best practices are emerging, especially with respect to geolocation data from shared bicycles and scooters, that can help inform how Oregon uses geolocation data to maximize the benefits of AVs while protecting people’s privacy. Changes to state privacy and public records law could also help ensure responsible handling of geolocation data by all parties that collect and use this data.

The Subcommittee on Cybersecurity, Privacy and Data makes the following recommendations for how the State of Oregon handles geolocation data generated by AVs:

- Deliver on the public value of AV data while addressing privacy concerns
- Monitor other efforts to handle geolocation data in a way that balances utility and privacy, such as the effort to create common standards and processes for data from shared bicycles and scooters, in order to inform Oregon’s approach to AV data.
- Consider developing new legislation, such as privacy laws that ensure that confidential data is protected both in public- and private-sector uses of geolocation data and/or changes to public records law that define geolocation data as confidential.

What is geolocation data?
Geolocation data is data collected via an electronic communications network or service that indicates the position of equipment used by people who are connected to that network or service. Geolocation data can include information such as the latitude, longitude, and altitude of the equipment. Technologies that are now mainstream – particularly GPS units and smartphones – are being used to collect unprecedented amounts of geolocation data, as well as new types of data such as the time data was collected, direction of travel, and other detailed information. These technologies are expected to be integrated into AVs, so AVs will be capable of collecting and providing information on where they are located and where they have traveled in the past.

Why do public agencies need geolocation data from AVs?
Geolocation data can provide more detailed information on where people traveling to and from than the data sources that transportation agencies traditionally use. Agencies can use
geolocation data from AVs to generally understand travel patterns and to better plan and manage the transportation system as well as to address a number of AV-specific uses that can help make the transportation system safer and more efficient for all travelers, such as:

- Identifying facilities that could benefit from projects focused on AVs, such as connected vehicle infrastructure or dedicated AV-only lanes
- Independently evaluating AV pilot projects
- Investigating AV crashes and safety risks
- Optimizing AV travel patterns in high-demand areas or emergency situations (e.g., designating pick-up and drop-off locations in congested locations, rerouting AVs around incidents or congestion)
- Managing and enforcing regulations for transportation services that operate using AVs

Agencies have other methods of collecting some of the data needed to achieve these use cases, but these methods are all more costly and less effective than collecting geolocation data already being generated by vehicles, and it is unlikely that public agencies would be able to maximize the benefits of AVs without access to geolocation data. AVs will dramatically change how people travel, both over the long term as they inform where people live and what modes they use and also on a day-to-day basis as AVs identify the most efficient travel routes and people choose from a growing variety of transportation services based on current traffic conditions. Public agencies will need rich, up-to-date data to understand and address these rapid changes.

**What privacy risks are associated with geolocation data?**

Geolocation data can be used, either alone or in combination with other data, to identify people and track their movements. A 2013 study found that people can be identified based on their travel patterns alone using as few as four data points.¹ It is also possible to combine geolocation data with the growing amount of other data that is accessible online to identify someone and understand where they go. A 2014 study used New York taxi data, mapping services, and social media data to identify people and their destinations, including people who used taxis to travel to strip clubs.² As a 2019 investigation in the New York Times highlighted, geolocation data that is collected and aggregated through smartphones and apps and that reveals individuals’ detailed travel patterns is available for sale on the private market.³

A growing number of public agencies are collecting geolocation data generated by ride-hailing and shared bike/scooters in order to manage and evaluate these services, in ways that could establish a precedent for how public agencies manage AV data. Some privacy advocates have raised concerns that agencies are collecting more data than they need and that agencies are not implementing adequate safeguards to keep data confidential, including ensuring that people are not able to obtain it through public records requests and that data cannot be used

---

¹ Montjoye et. al., Unique in the Crowd: The privacy bounds of human mobility. [https://www.nature.com/articles/srep01376](https://www.nature.com/articles/srep01376)
by law enforcement to profile people.\textsuperscript{4} Several best practices are emerging to address these concerns, especially with respect to shared bicycles and scooters, that can inform approaches to AV data. Public agencies have updated data policies, systems, and guidelines to better project geolocation data. Open-source data standards and tools like the Mobility Data Specification (MDS)\textsuperscript{5} and SharedStreets Mobility Metrics,\textsuperscript{6} which are governed by bodies that include both public and private-sector members, have emerged to standardize geolocation data formats and aggregate and anonymize data in a way that addresses privacy concerns while maintaining public agency use cases and creating consistency for operators. Several companies are also offering services that collect, clean, and visualize geolocation data for agencies that lack the capacity or necessary policies to manage this data.

**How do existing law and standards treat geolocation data?**
There is an emerging body of laws and guidance that treats geolocation data as confidential, but the laws that currently apply in Oregon do not designate geolocation data as confidential, nor do they limit its collection and disclosure. Public policy is generally designed to keep personally identifiable information, which is information that can be used to identify an individual, confidential, but there are different standards in current law about whether geolocation data is considered personally identifiable information. More recent privacy laws such as the European Union’s General Data Protection Regulation (2019)\textsuperscript{7} and the California Consumer Privacy Act (2019)\textsuperscript{8} both treat location data as personally identifiable information that is subject to protections that limit its collection and disclosure, but none of these apply in Oregon. Oregon public records law requires public agencies to redact certain personal information but does not discuss geolocation data.\textsuperscript{9} Both public agencies and private mobility companies have so far successfully protected geolocation data that Oregon cities are collecting from disclosure. New state legislation that defines geolocation data as confidential could make it easier for public agencies to address some of the privacy concerns discussed above in a consistent way. A more comprehensive update to state privacy law could ensure that privacy is protected both in public- and private-sector uses of geolocation data.

\textsuperscript{4} For an example, see the November 2018 letter by the Center for Democracy and Technology regarding the City of Los Angeles’ [https://cdt.org/insight/comments-to-ladot-on-privacy-security-concerns-for-data-sharing-for-dockless-mobility/](https://cdt.org/insight/comments-to-ladot-on-privacy-security-concerns-for-data-sharing-for-dockless-mobility/)
\textsuperscript{5} The specification is available at [https://github.com/CityOfLosAngeles/mobility-data-specification](https://github.com/CityOfLosAngeles/mobility-data-specification). A non-profit named the Open Mobility Foundation consisting of public and private members was recently formed to govern the MDS and related efforts.
\textsuperscript{6} [https://medium.com/sharedstreets/introducing-sharedstreets-mobility-metrics-37e4ddae8e7b](https://medium.com/sharedstreets/introducing-sharedstreets-mobility-metrics-37e4ddae8e7b)
\textsuperscript{7} Chapter 1, Article 1, [https://gdpr-info.eu/art-4-gdpr/](https://gdpr-info.eu/art-4-gdpr/).
\textsuperscript{9} ORS 192.377, [https://www.oregonlaws.org/ors/192.377](https://www.oregonlaws.org/ors/192.377)