CANTERBURY BANKSTOWN

# Big Mobile Data

A Data Digest Thought Paper

CANTERBURY BANKSTOWN

## What is a Data Digest thought paper?

This thought paper is a collection of tech trends, future insights, and noteworthy case studies about a given topic.

It's not a definitive statement of *what to do*, it's a collection of what we could do and what we should be *thinking about*.

## Contents



# Part 1: Introduction

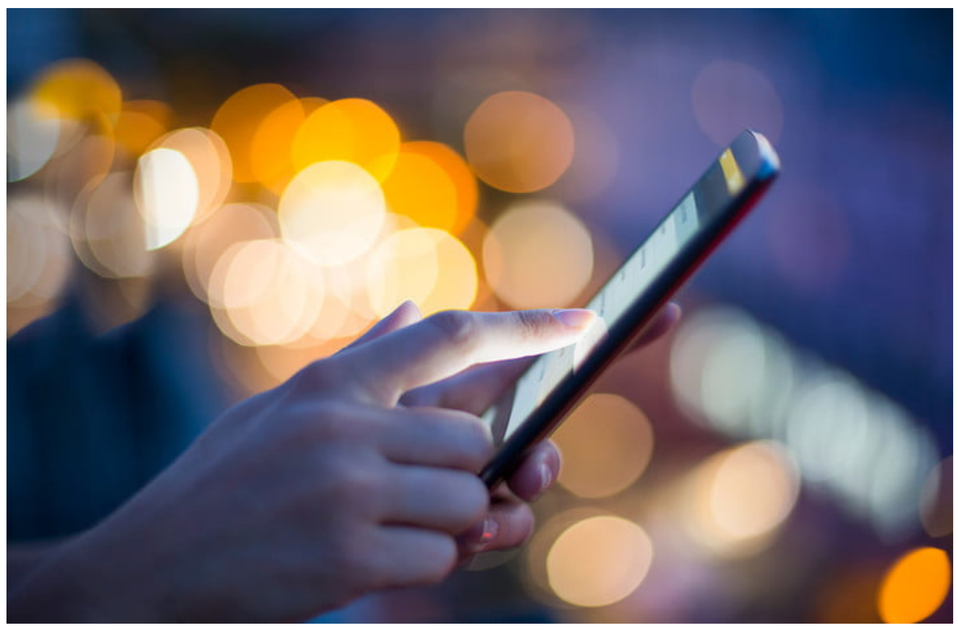
The widespread adoption and daily reliance on mobile phones has enabled a new avenue for understanding human behaviour. Australia's smart phone adoption, with a penetration level of 84% – places it as the fourth highest in the world behind Norway (91%), South Korea (89%), and the Netherlands (87%) (Chanthadavong, 2020).

And not only do more Australians *own* phones than ever, but those that do are *using* their phones more often than ever. In fact, an IDC survey revealed that 79% of phone owners have their smartphones in arm's reach for all but two hours of the waking day, and that 25% have their phones in reach at all times (Isaac, Vandermyde, Loo & Parekh, 2014).

Given the scale of mobile phones and their connectedness with people and their movements, tapping into mobile data enables valuable insights for a significant population.

### *What is 'big mobile data' and mobile location data?*

Big mobile data is data that is generated to support communication, location services, or for advertising purposes. In all cases, information about a device, user attributes, time, and location, is shared with telecommunication providers or ad publishers. The term 'big' refers to volume. A colossal amount of volume is being sent from our devices every second. Consider the continuous communication needed between a phone and a cell tower to stay connected, or the apps that access location data both in-app and continuously in the background (Weather apps, Uber, Facebook and many more we don't realise).



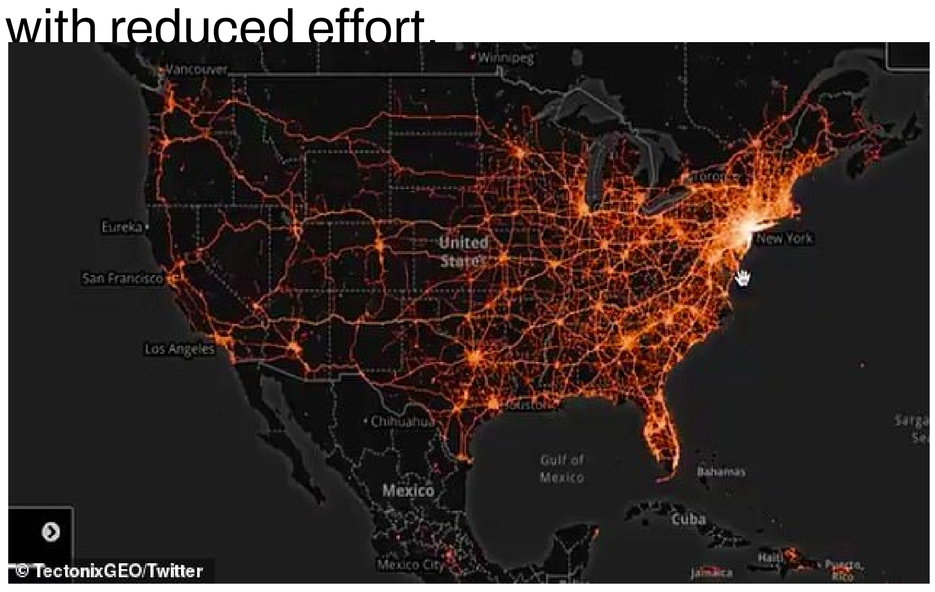
Arguably, the most valuable aspect of this data is locational data. Locational data represents geographical information (coordinates) about a device's position at a specific point in time. This locational data has opened new fields of study in data analytics and service provision, coined *locational intelligence* and *location-based services,* respectively.

### *Why is mobile location data so valuable?*

Using mobile locational data has a range of benefits, with key advantages being:

*Granular insights.* Mobile data has the opportunity of being collected and analysed at a device-level. This can enable device-level messaging or services. When paired with additional datasets, it can also provide a richer understanding of consumer profiles and preferences. This level of detail, however, should be used only with de-identification measures in place to protect privacy.

*Scale.* Depending on the source, big mobile data can capture information relating to a significant proportion of the population under investigation. As opposed to relying on smaller samples or having to wait for full-scale censuses, mobile data can provide large scale data with reduced effort.




*Real-time.* Mobile data can provide real-time, or close to, information to enable efficient decision-making. This real-time data has already equipped multiple governments – Italy, Germany, Austria, UK, and Singapore – in responding quickly to Covid-19 by supporting contract tracing and reviewing social distancing adherence.

*Accurate.* Depending on the source, mobile location data uses *real* data as opposed to modelling. This provides greater accuracy in data, showing real fluctuations to events or environmental changes.

> ***Key takeaway:*** Big mobile data provides a detailed view of the customer. Scaled up, it provides an accurate depiction of aggregated mobility patterns in real-time.

## Part 2: The Tech Stuff

### *How is this data generated?*

As discussed, mobile data is generated to signal a device's location, in addition to other device attributes, to telecommunication providers or app advertising publishers. This communication requires (Tamoco, 2019):

*A location source or signal generator.* It is, in fact, not the device that sends the signal. A location source sends signals out so that devices can detect, or receive, and then 'ping', or communicate, back to.

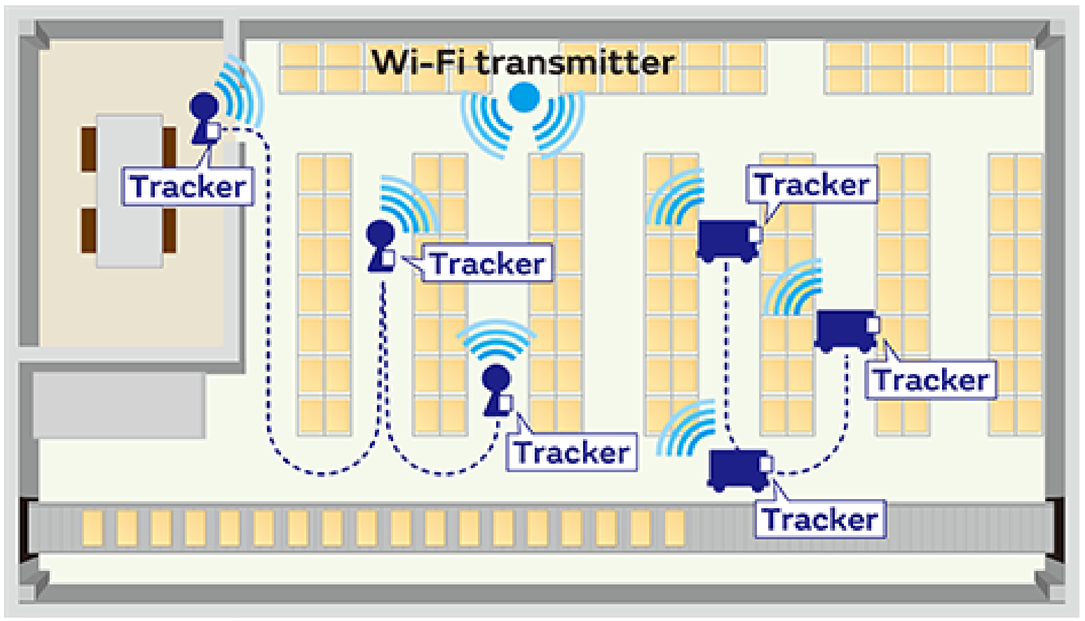



*Figure 1: Wi-fi device detection*

This is similar to how sonar detects objects and their locations under the sea. There are various location sources used by mobile phones (Tamoco, 2019).

- GPS: uses a satellite navigational system to pinpoint locations. It is accurate and precise, especially when used outdoors.
- Wi-Fi: sends signals out to devices in range. Devices don't have to be connected to Wi-fi to have data generated, as being in range enables the router to ascertain distance from source and location. Wi-fi is also accurate and precise but predominantly used indoors when GPS and cell tower access are obstructed.
- Beacons: are small devices, similar to wi-fi routers, placed in areas where information on device locations is sought. Beacons use signal strength to a device's distance away.
- Cell towers: are used for telecommunication purposes. Mobile phones often communicate with multiple cell towers, from which telecommunication companies triangulate and identify a device's location.

*A device (receiver).* A device is needed to receive signals from the location source. Each device is given a unique identifier that allows it to be tracked over time.

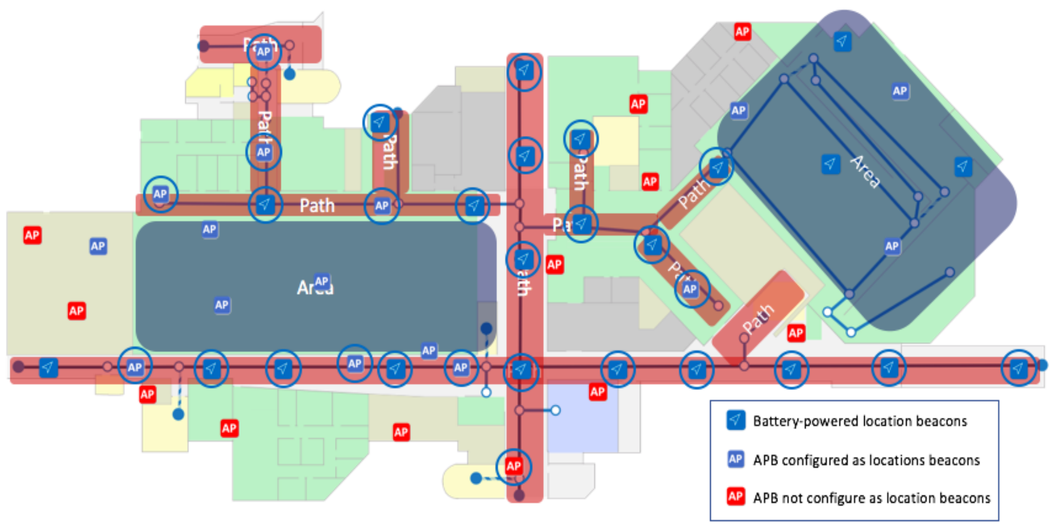



*Figure 2: Outdoor Beacon Deployment*

### *So now I know how this data is being generated, why and by whom is it being collected?*

There are three primary avenues for the collection for mobile location data. Sourcing data from each comes with certain advantages and disadvantages.

*Bidstream:* otherwise known as the 'ad exchange' is a part of the advertising ecosystem. Website and app ad publishers sell space for advertisers to buy real-time. Each purchase (bid request) produces data including device information, time, location, and some user attributes.

- Pros: very high volume of requests and therefore data points. Is also the easiest to obtain.
- Cons: quality is mixed; it can often be incomplete, inaccurate or illegitimate.

*Telecommunication providers:* as phones continuously communicate with cell towers, major telecommunication providers hold a vast amount of data for a large proportion

of the population.

- Pros: huge scale of data and also includes some demographic data.
- Cons: contains some issues with accuracy and can be very expensive to obtain.

*Location SDKs:* Software development kits are codes used by app developers to provide third-party functionalities, for example, location services. These codes are embedded during app development and require explicit user permissions via the app.

- Pros: provides accurate and precise information. It also requires obtaining of user consent.
- Cons: is smaller in scale because of the need to integrate with the app and the need to acquire user permissions.

*Beacons:* are hardware transmitters than can be purchased and installed.

- Pros: data is very accurate
- Cons: hard to achieve scale

*Wi-fi:* requires implementation of hardware while leveraging existing Wi-fi infrastructure.

- Pros: can be accurate, dependent on Wi-fi network
- Cons: small-scale.

Based on the advantages and disadvantages of each source of data, it is important to consider certain factors during evaluation: budgeting constraints, accuracy requirements, and the use case for data.

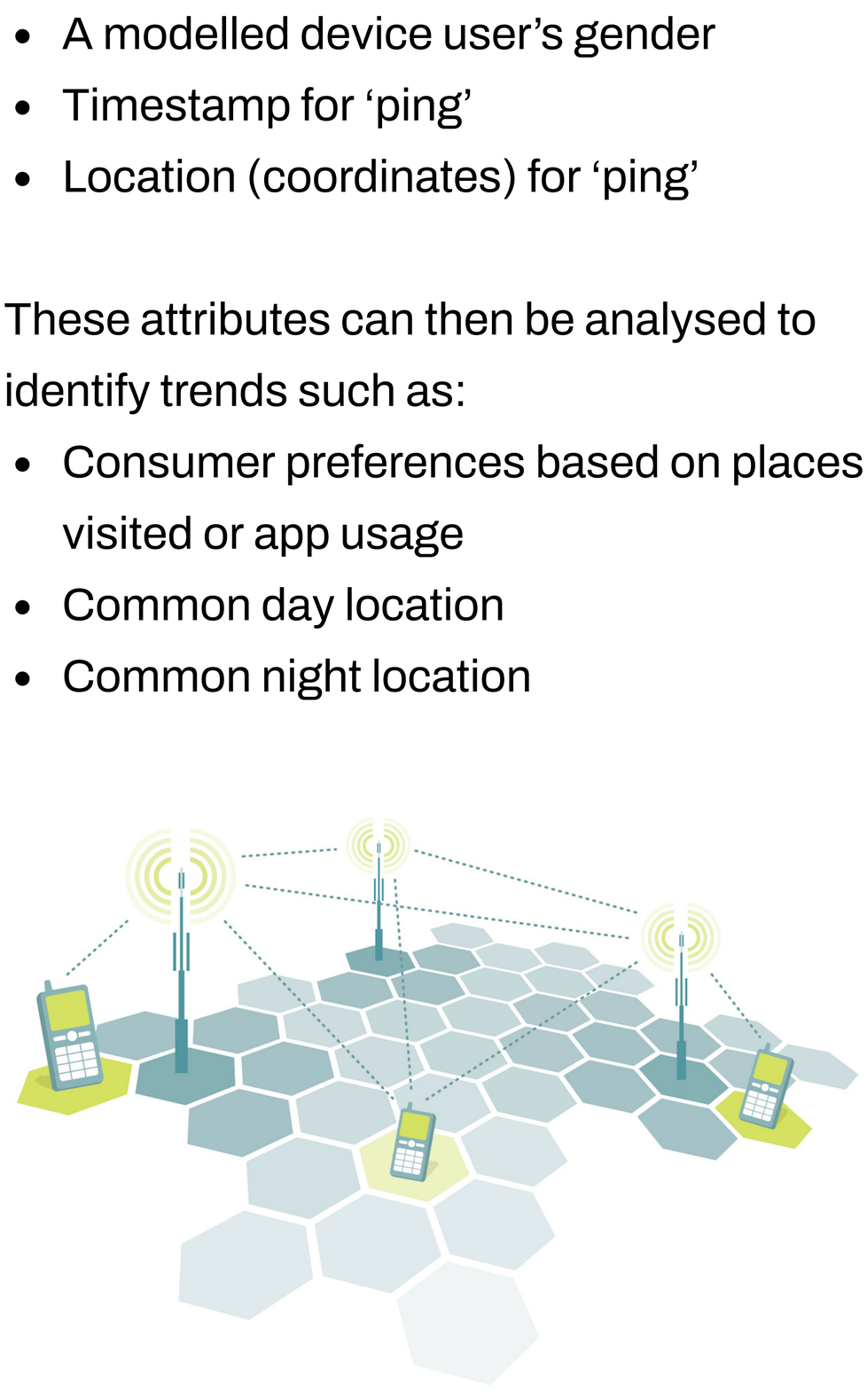
***Key takeaway**:* there are multiple different ways mobile location data is collected. Each represents a source of data for possible analysis, and each has advantages and disadvantages. It's important to carefully evaluate each data source based on use case needs.

### *What is contained in this data?*

- A pseudonymized ID which is the unique identifier for each device. It is often referred to as a MAID (Marketing Advertising ID). The process of pseudonymization allocates each device a moniker, e.g. 7ddjzl8zamg925n, in place of any personal identification.
- The device manufacturer
- The device's operating system
- A modelled device user's age
- A modelled device user's gender
- Timestamp for 'ping'
- Location (coordinates) for 'ping'

These attributes can then be analysed to identify trends such as:

- Consumer preferences based on places visited or app usage
- Common day location
- Common night location



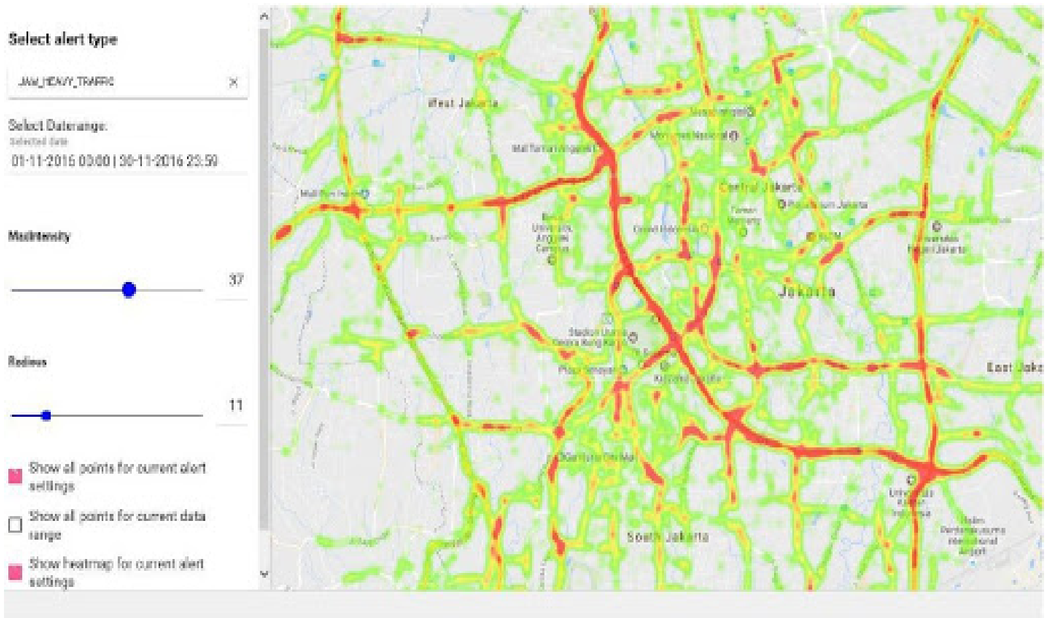



*Figure 3: Traffic Flow Data*

# Part 3: The Non-Tech Stuff

### *What are existing or possible use cases for mobile location data?*

Big mobile location data, and resulting location intelligence and location-based services, can have benefits across all industries, from retail, public safety, healthcare, planning, to transport management. While some use cases are already in practice, others represent great possibilities for utilising mobile location data.

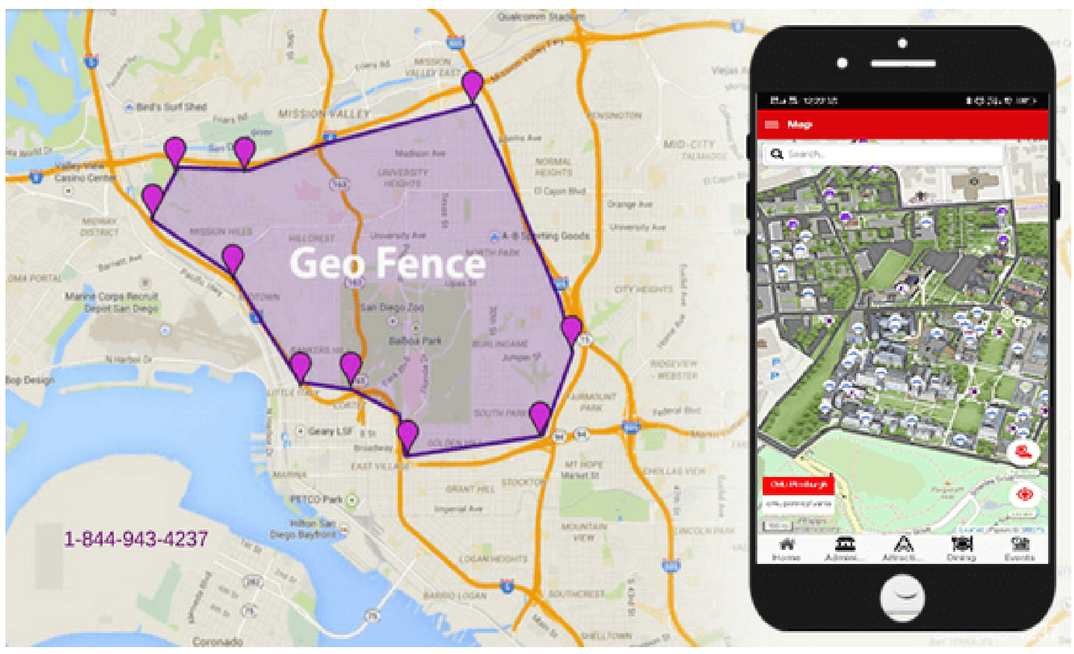
***Transport management.*** Mobile data enables the understanding of how people move throughout the city, in what ways, and at what times. Understanding city travel patterns can facilitate effective services to match demand. Mobile data can be combined with transport network data, such as TransportNSW or Uber to enable greater insight into users of certain modes of transport. Potential transport use cases are limited only by one's imagination. For example, there's opportunity to overlay weather data from IoT sensors to understand how weather patterns and events affect travel.

Data can also inform traffic management decision making, such as where to widen lanes or install new traffic lights. This is invaluable knowledge to provide efficient and relevant services to the public. Using mobile data for analysis of trends and patterns represents location intelligence.

***Emergency relief.*** The analysis of traffic flows overlaid with emergency call locations can assist with identifying optimal locations and routes for emergency services. This is an example of locational intelligence to support better decision-making. Mobile data can also support location-based services in emergency relief. For example, telecommunication providers have the ability to send alerts or instructions to phones in disaster areas, such as those in proximity of a tsunami, to ensure they can move to safer areas. Beacons can also be used with the same affect on a smaller scale.

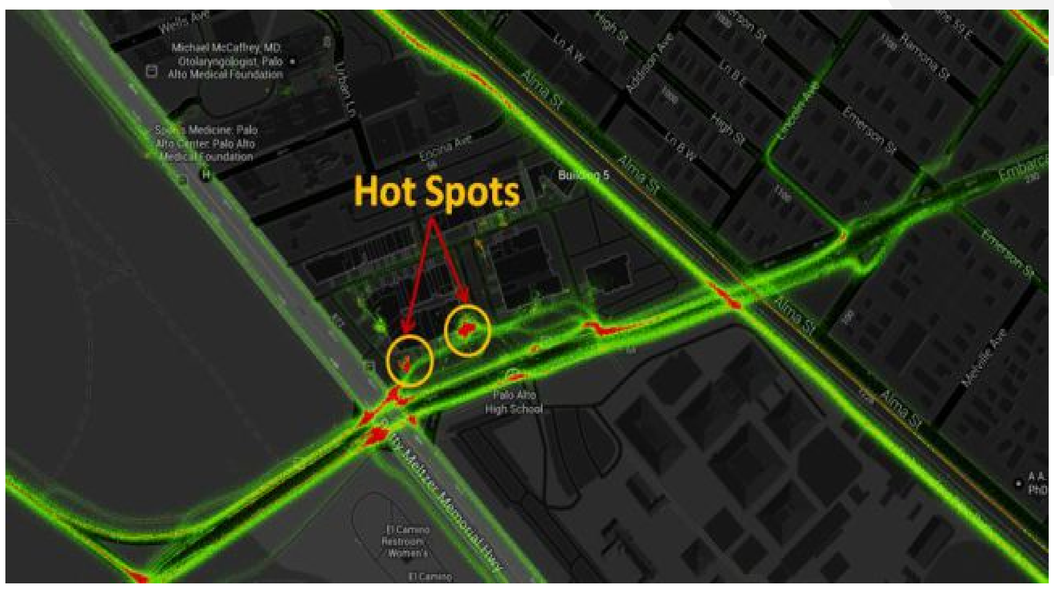

***Location-sensitive billing.*** Licence plate recognition is commonly used to charge vehicles for tolls or parking. These charges are priced based on peak or off-peak rates or length of stay, as was the case with London's 2003 Congestion Charge and in many commercial carparks in Sydney, respectively (Ratti, Pulselli, Williams & Frenchman, 2005).

Mobile data has the ability to provide greater tailoring in location-sensitive billing, by using dynamic road and parking pricing depending on time, location, and other environmental factors (such as nearby events). Telecommunication providers would be key facilitators in the process to unlock this spatial and temporal data, as well as pass on dynamic charges. This would reduce the need for physical meters, boom gates, or on-site technology infrastructure. Mobile data enables geofencing and location-sensitive billing to better manage traffic and parking supply.




***Urban planning.*** "Data based on the location of mobile devices could potentially become one of the most exciting new sources of information for urban analysis" (Ratti, Pulselli, Williams & Frenchman, 2005). Why? Because our cities are becoming increasingly dynamic, where individuals have a greater range of transport modes to reach a greater range of destinations. Urban mobility has become increasingly decentralized. Whereas individuals may have once concentrated movements in their local areas and work locations, the public is now seeking more experiences outside of these two areas. Therefore, real world planning requires real world data (Shukla, 2016).

It requires real time actual patterns of movement, rather than relying on models or estimates (Ratti, Pulselli, Williams & Frenchman, 2005). Mobile location data can help urban planners understand travel patterns, origins and destinations, and preferred routes. It can provide granular insights on pedestrian movements through and usage of public spaces. It can highlight places of interest and their peak days and times and can further ascertain relationships between these places of interest and travel patterns. Mobile location data can further be layered with additional datasets such as weather or events data to enrich the understanding of how these changes affect urban flows. This wealth of data and insights can inform city shaping decision making to provide liveable spaces for the community.




There are also a range of use cases within the private sector that may serve as thought starters.

***Marketing communication.*** Mobile location data when used with beacons or geofencing enable the sending of context-relevant marketing messages via push notifications. Beacons enable mobiles within a certain radius to be detected and based on the user profile or services they're subscribed to, to be communicated to.

Geofencing works in a similar way, however devices within virtual boundaries may be detected and communicated to. Users could be sent notifications of nearby events, points of interest, or promotions and discounts. This mechanism could also be used by Council to send messages of road works or nearby events to those with permissions.

***Retail space design.*** Using Wi-fi signals and mobile location data enables the understanding of how individuals use indoor spaces. Their dwell time, return rate, and routes throughout the space can provide insight into how to optimize these indoor spaces.

***Site planning and selection.*** Using mobile data to understand volume of visitors and therefore exposure can support the site selection of new business locations, or advertising locations.

***Key takeaway:*** there are a range of use cases for mobile location data. Though some aren't directly relevant in local government service provision, many of the same mechanisms, e.g. location-based push notifications, can be used to support Council's services. If nothing else, they highlight the ingenuity in location intelligence and location-based services, and provoke ideas for innovative mobile data use cases.

## Pushing the Boundaries: A C(r)ow(d)sourced Navigation App

An interesting big mobile location data use case comes from Dubai and heavily features cows. Yes, cows.

In Dubai, traffic jams cost $10.8 billion every year and are often caused by cows (Moreno, 2017). Y&R saw this as both a challenge and an opportunity. While cows are often the cause of these traffic jams, they also present a unique opportunity as they have the ability to find faster routes to bypass traffic through backroads and alleys. To tap into this, Y&R attached smart mobile devices onto cows' collars to relay GPS positioning. This data was transformed into a crowdsourced navigation app that leveraged cows' shortcuts to find faster routes through traffic. This resulted in 15 minutes of travel time saved, lower fuel consumption and pollution levels at peak times (Moreno, 2017).


## *How do we manage privacy and the ethical use of mobile data?*

Unlocking the benefits of mobile location data requires responsible data practices and consumer trust in order to ethically use data and protect privacy. As mobile devices increasingly house more sensitive information and are so intertwined with our personal identity, it comes at no surprise that consumers are increasingly interested in how their data is being collected and used by companies. In fact, 89% of consumers say that they avoid doing business with companies that they think do not protect their privacy online (Power & Guinto, 2014).

It becomes critical to gain consumer trust by clearly outlining the societal benefits associated with enacted use cases. Being transparent and communicating the data being collected and analysed, in addition to the privacy protection measures in place, can foster consumer trust in using data for good.

Some of these privacy protection measures include the deidentification of devices by pseudonymising the device ID - that is providing a non-meaningful string to each device. This strips the ID of any personally identifiable information.

While the ID may not link a device to an individual, other fields, namely location coordinates, have the opportunity to reveal identity if merged with other datasets. To prevent this, several data fields may be removed, or analysis may be reported only in aggregated or approximated form. Other complex safeguards may also need to be implemented to mitigate the possibility of re-identification.

Ensuring these measures are in place and communicating these openly to consumers supports data privacy rights and consumer trust in the ethical use of data.

## *Conclusion*

Big mobile location data, location intelligence and location-based services present an array of use cases to improve Council's services to the community. Through a considered data vendor evaluation process and with privacy measures in place, mobile data has the ability to drastically change the way we provide our services for the better.

**A City Future initiative supporting our mission to:**
***Dream, design and create the most amazing future for our City.***

## References


Chanthadavong, A. (2020). How Australians are using their mobile phones. Retrieved 9 September 2020, from https://www.vodafone.com.au/red-wire/australians-using-mobile-phones

GSMA. (2017). Mobile Privacy and Big Data Analytics. Retrieved 9 September 2020, from https://aiforimpacttoolkit.gsma.com/resources/GSMA-report_Mobile-Privacy-and-Big-Data-Analytics-Paper.pdf

Isaac, C., Vandermyde, A., Loo, S., & Parekh, P. (2014). Mobile data analytics: not just for consumers any more. Retrieved 9 September 2020, from https://www.pwc.com/id/en/publications/assets/ticepublications/data-analytics.pdf

Moreno, H. (2017). 6 Uses Of Location Intelligence: Limited Only By Imagination. Retrieved 9 September 2020, from https://www.forbes.com/sites/forbesinsights/2017/02/09/6-uses-of-location-intelligence-limited-only-by-imagination/#5c071d1787f6

Power, C., & Guinto, G. (2014). Build customer trust through data privacy. Retrieved 9 September 2020, from https://www.pwc.com.au/industry/retail-consumer/assets/data-privacy-jun14.pdf

Ratti, C., Pulselli, R., Williams, S., & Frenchman, D. (2005). *Mobile Landscapes: using location data from cell-phones for urban analysis*. Massachusetts Institute of Technology.

Shukla, S. (2016). Location Data | Predictive Analytics | Data Platforms | Near - Blog. Retrieved 9 September 2020, from https://blog.near.co/technology/how-location-data-can-help-urban-planners-in-their-smart-city-initiatives/

Tamoco. (2019). Best Guide To Location Data 2020 - Everything You Should Know. Retrieved 9 September 2020, from https://www.tamoco.com/blog/location-data-info-faq-guide/


## Author


**Vivien Le**
*Graduate Innovation & Insight Officer at*
*The City of Canterbury Bankstown*

For inquiries or to get involved, contact:
vivien.le@cbcity.nsw.gov.au