

IGNITING ASIA-PACIFIC'S BIG DATA TRANSFORMATION

From Singapore to Australia, the Asia-Pacific region (APAC) is in the midst of a big data transformation. Companies are collecting and storing more data than ever. Emerging technologies are building on older innovations to open up new frontiers in areas including transport, health, safety and urban planning.

Analytics applications that were unheard of just a few years ago have changed how organisations, governments and individuals interact. In Singapore, event organisers use geolocation data to monitor crowd movements. Banks in the Philippines assess loan eligibility based on mobile phone usage patterns. Thai telecommunications companies have analysed customer movement behaviour to optimise store opening hours. Social media companies in Singapore monitored social media sentiment and engagement to gauge the opinions of the silent majority during the General Elections. As a result of recent advancements led by DataSpark, Singapore transport agencies have moved from analysing train performance to using predictive modelling to anticipate queue length.

Gartner analysts predict that over the next few years, "analytics will be pervasive and mission-critical for decisions and actions across the business". At the same time, IDC reports that organisations in APAC have rapidly improved their capabilities for leveraging big data and analytics. More than a third (34 per cent) have even progressed to more mature stages - that's an almost five-fold increase from the seven per cent in 2014.

Combined with the fast pace of technology change, these figures highlight the need for business and political leaders

to embrace data analytics and its potential applications. With Singapore on track to be the world's first smart nation - and neighbouring countries expected to follow suit - the time for action is now.

This white paper comprises key concepts presented at IGNITE, a client engagement forum organised by DataSpark in Singapore. With a focus on the APAC region, it will:

- Examine the role of data analytics in improving public health, mobile interconnectivity, urban transport, safety, security and poverty challenges.
- Identify top privacy issues facing organisations and data scientists, and discuss practical ways to overcome them.
- Outline the most urgent data analytics challenges and list opportunities for change.
- Explore what's next in data analytics, including a forecast of future innovations.
- Share best practices on how data-driven enterprises should navigate the new data analytics frontier.

WHY DATA ANALYTICS MATTERS

Data analytics turns information into actionable wisdom.

From improving a brand's understanding of its customers to monitoring the spread of disease in urban areas, this wisdom is harnessed in many ways.

Speaking at IGNITE, DataSpark's director of data science consulting, Rick Zhen, used Singapore's Formula 1 Grand Prix to demonstrate how data analytics reveals transformative insights. He explained that during major tourism events, data analytics can provide essential information including:

- Visitor's country of origin.
- Visitor behaviour, e.g. length of stay, choice of hotel, dining preferences.
- Nature of the trip, e.g. holiday or business.
- Whether visitors are one-time, first-time or regular Formula 1 attendees.
- Web behaviour, e.g. which social media platforms are most popular with visitors.

"Understanding those questions impacts a billion-, if not tens of million-, dollar industry," said Zhen. "This has implications on the targeting efforts of travel agencies or even the government for next year."

He also stressed the importance of understanding how visitors move around the event and surrounding areas of interest.

"In addition to understanding how people move within the event, how people move outside the event is also needed to identify typical hotspots of interest in the rest of the city, allowing those [with] commercial interests to better target Formula 1 visitors," he said.

On a related note, DataSpark chief technology officer Dr Ying Li pointed out that data analytics matters because it

turns raw data into meaningful, applicable insights. Data analytics goes beyond basic observations, such as the fact that a person is using a mobile device in a particular building, to provide context around activities.

She explained: "In this world you have people at some location at some time and through some device in some context. This is the world (data scientists) study. Mathematically... these are all things you can observe very easily - people with a location in time.

She explained: "Data scientists study where people are at a particular point of time within a certain context. Mathematically speaking, these are things you can observe very easily."

"What we cannot observe, and part of the goal of our analysis, is to make inferences about the activity. This is where the data mining, the technology, the math, the science all combine - to infer about the activities."

Inferences help make sense of data and apply analytics in new ways. In turn, this helps organisations unlock more value from data sets. For example, with data analytics, a retailer may be able to determine the behaviour of a competitor's customers, including when they shop, where they shop and how much they spend. Without analytics, the retailer would have to rely on observational data, which is impractical, potentially invasive and difficult to scale.

IGNITING INNOVATION WITH SMALL SPARKS

While data analytics is a decades-old discipline, it is also one of the most innovative and fast evolving. A few years ago, it was near impossible for digital agencies to focus advertising spend in high-traffic areas based on footfall. People in impoverished countries risked financial ruin borrowing from private moneylenders instead of banks, which were reluctant to lend to those without bank accounts or credit histories.

1. Mobile connectivity

DataSpark is working with Singtel-affiliated telcos to expand geospatial analytics use cases across Singapore, Thailand, Indonesia and Australia. The aim is to boost mobile connectivity by overlaying data sources such as subscriber heat maps and channel touchpoints. This enables telcos to better understand their customers, demographics, and demand and deliver a better customer experience.

In Chiang Mai in Northern Thailand, a local telco used geospatial analytics to anticipate and respond to customer needs with greater accuracy. The mobile operator combined multiple sources of location data to assess channel management and coverage, and used dynamic maps to visualise customer movements over time. As a result, it was able to:

- Determine optimal store opening hours.
- Roster enough staff to meet demand during peak times.
- Uncover valuable customer behaviour insights to inform targeted marketing campaign planning.

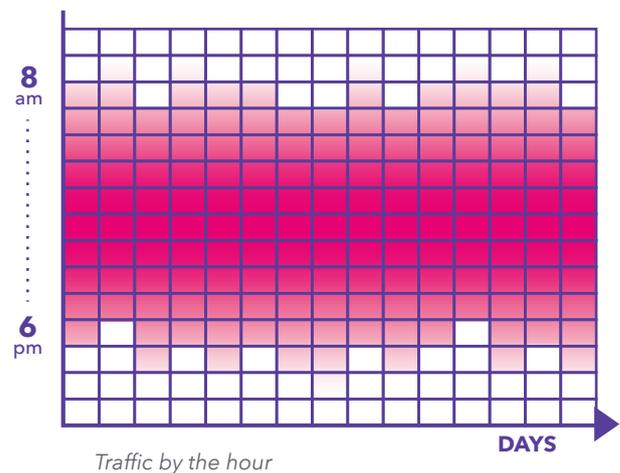
2. Urban transport

When telcos and transport agencies partner to understand how people move around cities, they are able to ask intelligent, dynamic questions based on integrated data sources. The answers lead to an improved ability to leverage transport information for urban planning as well as smarter transport use cases.

DataSpark's flagship transport application, which tracks Singapore's Mass Rapid Transit (MRT) system in real-time, exemplifies these benefits. Using anonymised, aggregated commuter Wi-Fi signals, DataSpark detects train arrivals,

But all that has changed now. As one innovation ignites another, times are changing fast.

The following real-world use cases give a flavour of what's possible as data analytics matures. With a focus on privacy, each use case has tangible benefits – both for businesses and the greater collective good.



departures and headway. This supports comprehensive travel route and travel time analysis.

Measuring MRT traffic during major events can also help redistribute crowds. During the 2015 Singapore National Day Parade, for example, DataSpark found 78 per cent of traffic came from three MRT stations. With insights into crowd movements, organisers could plan for guides and venue directions in prime positions the following year.

3. Public health

In the event of a disease outbreak, geospatial analytics minimises transmission by keeping infected people contained. When the mosquito-borne Zika virus was discovered in Singapore, for instance, telco data and geospatial analytics helped monitor the outbreak among at-risk foreign workers.

DataSpark tracked movement patterns of the affected demographic, particularly in the early morning and late afternoon when disease transmission was most likely. It found that movement patterns correlated with identified Zika hotspots - a positive sign for tackling future disease outbreaks.

Looking to the future, geospatial analytics and wearable technologies such as Fitbits and Apple Watches may also feed into the wider healthcare system.

In the US, wearable device data has already played a role in medical treatment: Emergency room doctors were able to determine the best treatment for a patient suffering a seizure after reviewing his Fitbit data. Companies are also exploring the potential of patient-supplied data to identify problems early and reduce inefficiencies in the healthcare system.

4. Safety and security

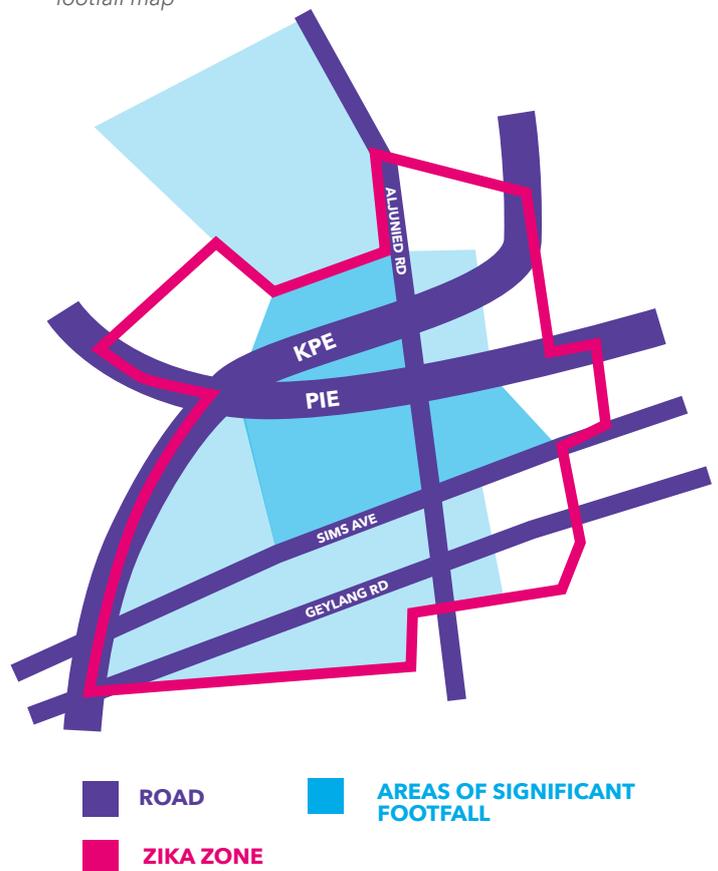
Geospatial analytics has enabled big data applications that make public spaces and events safer. When combined with security camera data and video analytics, footfall data (as illustrated in the Urban Transport section above) can also assist in identifying potential bomb threats. If an unattended

5. Addressing poverty challenges

Financial institutions in the Philippines have started using telco data to help locals climb out of poverty. Just 31 per cent of Filipinos have a bank account, according to the World Bank. With limited information on borrower credit worthiness, banks have traditionally hesitated to approve loans for the remaining 69 per cent. Much of this unbanked segment has turned to private, informal lenders as an alternative.

IDC lists the archipelago as Southeast Asia's fastest growing smartphone market, and this suggested a new way to assess

Parallel study of NEA Zika cluster map and DataSpark's footfall map



backpack is spotted in a sparsely frequented public place, security personnel will receive a notification to investigate it. Such alerts reduce risk, and make cities safer places to live, work and relax in.

an applicant's likelihood of repaying a loan. By leveraging mobile usage patterns to determine credit worthiness, DataSpark was able to improve a bank's loan repayment rates from 58 per cent to 70 per cent. Repayment rates are expected to eventually increase to 85 per cent, as geospatial variables are refined over time during this pilot. With improving confidence in lending, it has now become easier for locals to access credit from the bank.

PRIVACY – THE DEFINING ISSUE OF OUR DIGITAL AGE

The explosion of available information generated by the rise of big data has countless benefits. Big data powers smart city applications, improves customer service and makes urban transport more efficient.

However, as DataSpark chief operating officer Ying Shaowei stated during his IGNITE opening address, big data also comes with risks. "If there is one issue that is the defining issue of our digital age, it is really privacy," he said.

A recent DataSpark survey reinforces Shaowei's sentiments. It found that 78 per cent of Singaporeans are 'extremely concerned about data privacy'. Fortunately, as long as data privacy is built into processes and is not simply the end result, geospatial analytics can bring rich rewards without compromising personal information.

REGION-SPECIFIC CHALLENGES

APAC businesses must maintain high standards of data privacy. But in a region characterised by inconsistent privacy laws and regulations, this is easier said than done.

Legislation differs across developed and emerging economies, and privacy regulations are lagging behind technology developments. Some countries such as Indonesia, Thailand and Vietnam have more regulatory shortcomings than others.

In this scenario, the onus is on businesses and data analytics providers to ensure data privacy. They must be fully aware of

the differences in laws across countries. In addition, they must follow data privacy and security best practices on top of what local regulations demand.

At DataSpark, for example, data is gathered from the use of Singtel's mobile phone and mio TV service. To keep information private, it is encrypted, anonymised and aggregated, according to best practices. Further, care is taken to ensure compliance with Singtel's data governance framework and Singapore's Personal Data Protection Act.

GEOSPATIAL DATA PRIVACY PRACTICES

When it comes to maintaining geospatial data privacy across regions, one thing is certain: Individual records should never be identifiable. Research by DataSpark and UK-based software company Privitar shows these practices decrease privacy loss while maintaining accuracy. The two most common best practices are:

- **Microaggregation:** Occurs when individual records are aggregated into small groups. These are usually called squares or bins. Each square contains a set number of records to prevent disclosure of individual information in sparsely populated areas. Instead of

reporting an individual's exact location, microaggregation reports the location of the center of the square. The larger the square, the smaller the loss of privacy.

- **Noise addition:** Adds random noise (by adding random numbers to confidential data, for example) to a latitude or longitude, minimising the risk of an individual being identified. The larger the noise, the smaller the privacy loss.

A COLLABORATIVE APPROACH TO ANALYTICS

Industry and government partnerships are essential for ensuring that data-driven innovations benefit the greater good. For example, governments can leverage data collected by telcos to inform and guide urban planning and transport development.

Most organisations today collect valuable data – and plenty of it. As analytics becomes more advanced, a cross-disciplinary, collaborative approach can ensure organisations extract as much value as possible from all relevant data sets.

THE FUTURE OF DATA PRIVACY AND SHARING

What does the future of data privacy and sharing look like? According to Microsoft's director of data science, Dr Graham Williams, tomorrow's data gatekeepers will be individuals themselves.

He predicts individuals will soon have access to all personal data that governments and organisations collect about them. This includes everything from tax and genetic information to relationship status and social security records.

Speaking at IGNITE, Dr Williams said that while there are legislative limits on what data can be shared between organisations, change is underfoot.

"In Australia, an individual's federal government health data cannot be shared with the hospitals... Tax data and bank data cannot be readily shared," he pointed out. "There is a movement within the Australian government, and possibly also in Singapore, that looks at allowing individuals to have access to all of the data collected about themselves, from the banks, tax office and healthcare (providers)."

Dr Williams said that for ease of use and accessibility, data would most likely be available to individuals via intelligent APIs. Individuals will have control over their own data, permitting information to be shared from the device as necessary.

"These types of intelligent applications is where I think we're heading... building these collections of models that work individually for us," he added.

OVERCOMING ANALYTICS ROADBLOCKS

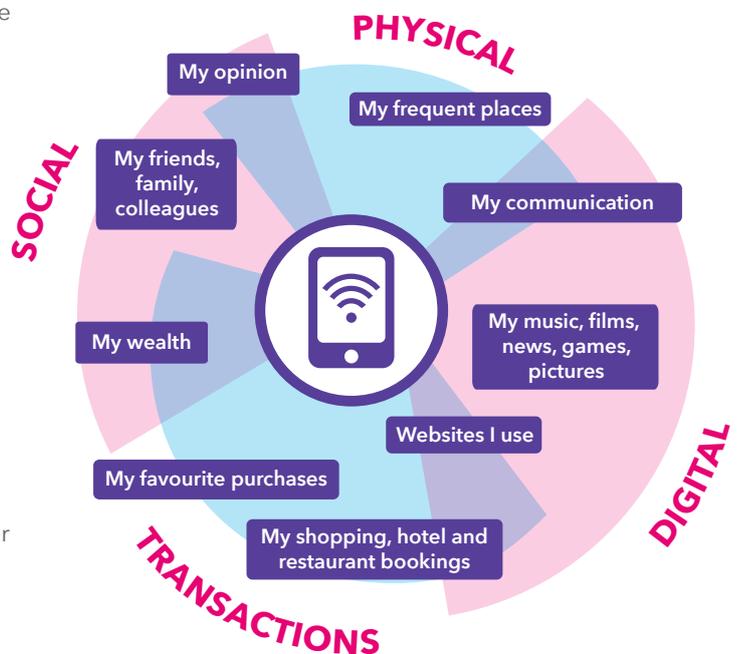
Telcos and their partners have a privileged window into a population's physical and digital lifestyles. However, challenges still remain, particularly when it comes to maximising the use of data assets, harnessing the potential of new technologies and protecting customer privacy.

1. Leveraging first-party data more effectively

While telcos collect vast amounts of data about customer locations, communication patterns and application use, they often have limited insight into areas such as income and spending power. To uncover new data sets, telcos must use first-party data more effectively. First-party data might include:

- Social information e.g. opinions, social networks.
- Physical information, e.g. frequently visited places.
- Digital information, e.g. frequently visited websites.
- Transactions, e.g. favourite purchases, hotel and restaurant bookings.

Combining these data sets with existing data can uncover new dimensions, generate powerful use cases and uncover monetisation potential.



2. Mastering data tools to unlock deeper intelligence

As data tools are becoming more advanced, the Apache Hadoop ecosystem (an open-source software platform for storing and processing very large data sets) is growing. Telcos are encouraged to embrace these research technologies and stay at the forefront of predictive analytics capabilities.

Instead of using analytics to describe what is happening, forward-thinking organisations are turning to machine learning and simulations to predict what could happen. As discussed, transport agencies have already moved from

describing how people travel around Singapore to predicting how building new roads or train routes will impact traffic flows. As a key source of geolocation data, telcos play a vital role in helping organisations master these new data tools and unlock deeper intelligence.

3. Creating new business models

Leading telcos have developed new go-to-market business models to capitalise on emerging technologies and access to more extensive data sets. As a result, telecommunications is changing from a service-oriented ecosystem to a data-centric ecosystem.

In China and Europe, for example, telcos are collecting more data. But they are also sharing data sets with developers in their partner ecosystem, with a view to developing new use cases in different industries.

Meanwhile in Australia, a local telco has developed a crowdsourcing network monitoring tool for mobile apps to collect data on network performance and geolocation data. In addition to using the tool for improved customer service,

4. Addressing privacy issues

Telcos may not yet understand the risks associated with big data, analytics and privacy. As many are entering the world of geospatial analytics for the first time, an initial lack of awareness is to be expected. However, both telcos and their partners must get up to speed with potential privacy issues and how to address them - and do this fast.

Would your organisation know how to respond if cyber

the company is also exploring applications in areas including crowd management (to prevent overcrowding scenarios), real-time geo-fencing (for sending promotional messages to consumers as they approach particular stores) and deeper analysis of popular travel routes (allowing for alternative route planning and the development of more roads).

This suggests that innovating within the business is no longer enough. Telcos must also work closely with partners to identify profitable, value-adding analytics applications.

attackers defeated data masking to reveal personal information? What would you do if a customer were denied service and treated poorly because of incorrect analytics? Answering these questions now and becoming familiar with best-practice privacy protocols can save you countless of hours and headaches over the long term.

EMERGING INNOVATION OPPORTUNITIES

The next big data analytics spark could be just around the corner. Here are three ways in which geospatial analytics are igniting futuristic possibilities.

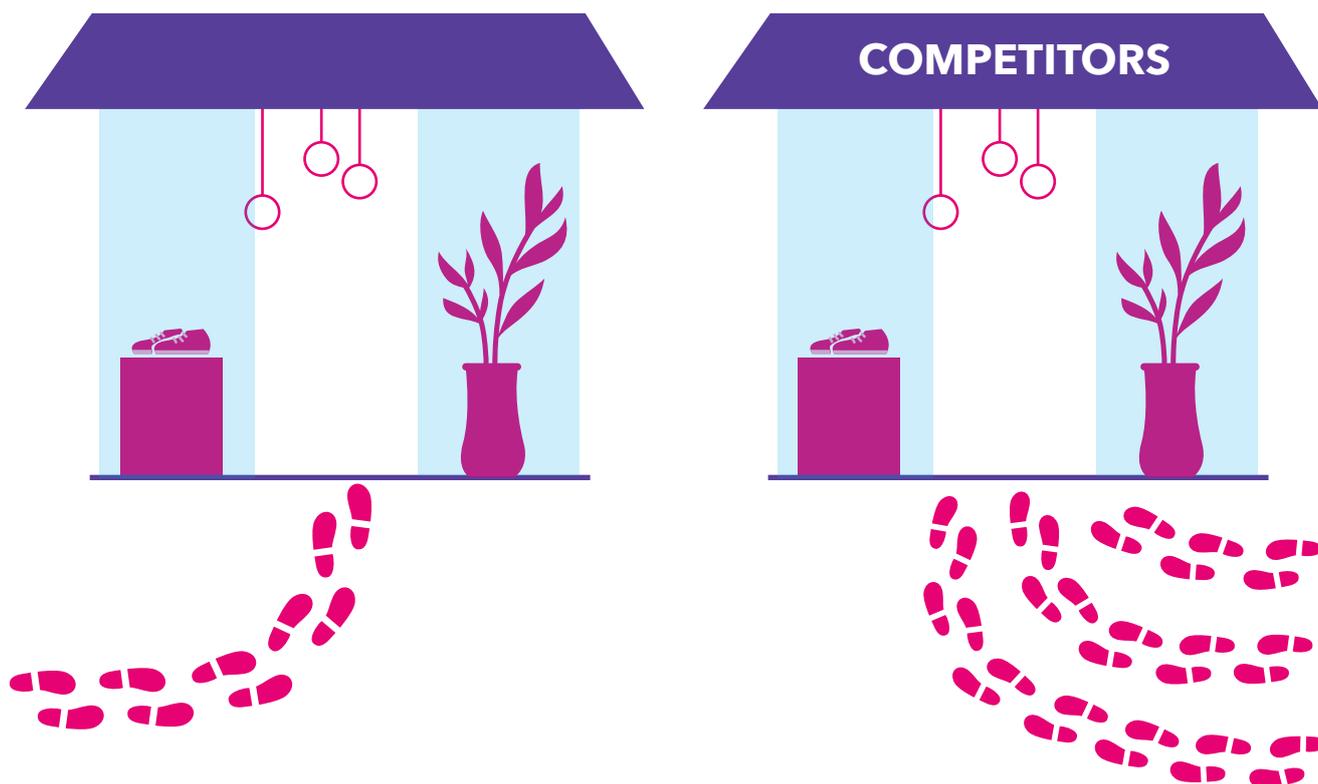
1. Understanding competitors' customers

Geospatial analytics can not only identify and segment your own customers, but also the customers of a competitor. By combining behavioural pattern data collected from mobile phone use (e.g. where customers live, work and play) with market research data (e.g. age, gender, occupation), organisations can gain complete insight into their competitors' customers. You'll be able to answer questions such as:

- How are my competitor's customers different? How are they the same?
- Do my competitor's customers also shop with me?
- What other stores do my competitor's customers visit?

- What is the daily footfall at my competitor's location?
- Which of my customers also visit my competitor's physical and digital stores?
- How do holidays affect my competitor's stores?

These insights can then be used to plan and manage strategies for customer acquisition and retention. They also make it possible to uncover trends early. For example, if a competitor is experiencing higher footfall or more website visitors, an organisation can respond before it impacts their bottom line.



2. Scaling big data analytics to build a smart nation

Data analytics' transformative potential reaches beyond individual organisations. It supports the heart of the smart nation: smart buildings.

Big data analytics presents an opportunity to make homes, offices and public spaces as smart as possible, from intelligent energy-efficient lighting systems to facial recognition for gaining access to buildings. One day in the future, organisations will know exactly who is on their premises at any moment, increasing security and reducing the risk of employees going missing during emergencies.

The challenge will be to ensure that these technologies can scale from small applications – such as in one organisation's office – to powering an entire city.

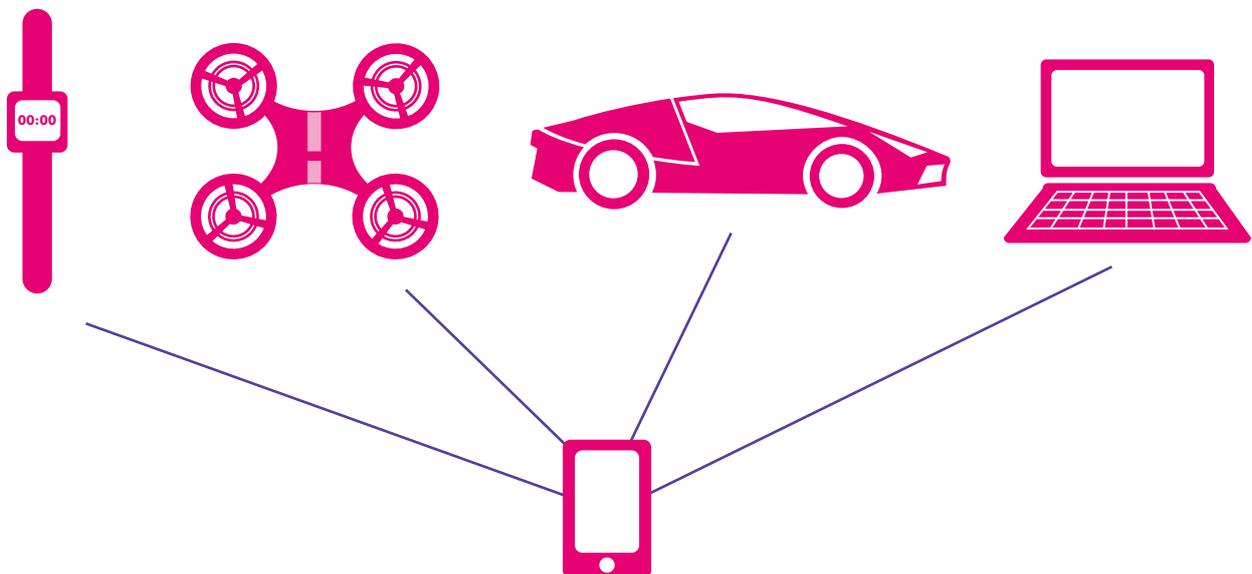


3. Internet of Things

Increased Internet of Things connectivity will create new data sets and ignite future developments, both in Singapore and across APAC.

The uptake of autonomous cars, drones and other technologies that rely on heterogeneous networks (networks that combine Wi-Fi and cellular networks), for example, could open up new frontiers for analytics in areas including traffic management and urban planning.

In Singapore, adoption of these technologies is growing quickly. The world's first self-driving taxis debuted in Singapore earlier this year as part of the nation's plan to become a global leader in autonomous vehicle use. Drones are also becoming more widespread: Singapore's Ministry of Transport recently awarded contracts to three companies to provide unmanned aircraft systems for monitoring and inspection purposes.



WHERE TO NEXT?

In the short to medium term, the future of geospatial analytics lies in more real-time analytics, greater automation and an improved ability to dig deep into data to improve customer experiences.

As DataSpark's Zhen said: "The focus is on actions that can be triggered by analytics in real-time," particularly as organisations move from hardware to the cloud.

"Before the cloud, it was near impossible for SMEs to apply real-time analytics simply because of the hardware cost," he explained. "The pay-per-use system of cloud-based systems reduces this cost significantly, and open-source real-time data processing tools open up real-time analytics to almost everyone."

Zhen also expects to see increased automation in the coming months and years. He said DataSpark is building on its work of developing triggers for manual system actions by

creating triggers for automated system actions. This will make responding to trends and issues faster and more convenient.

Another area in which DataSpark is innovating is collecting data from new systems, such as those that monitor telco network breakdowns. In doing so, service providers will be able to deliver better customer experiences.

"It is now also possible to gather customer locations and web logs, and process customer's real-time usage experience with individual cell towers. This allows operators to identify cell towers with congested traffic, which could impact the customer's experience, and to adjust the experience in real-time," he said.

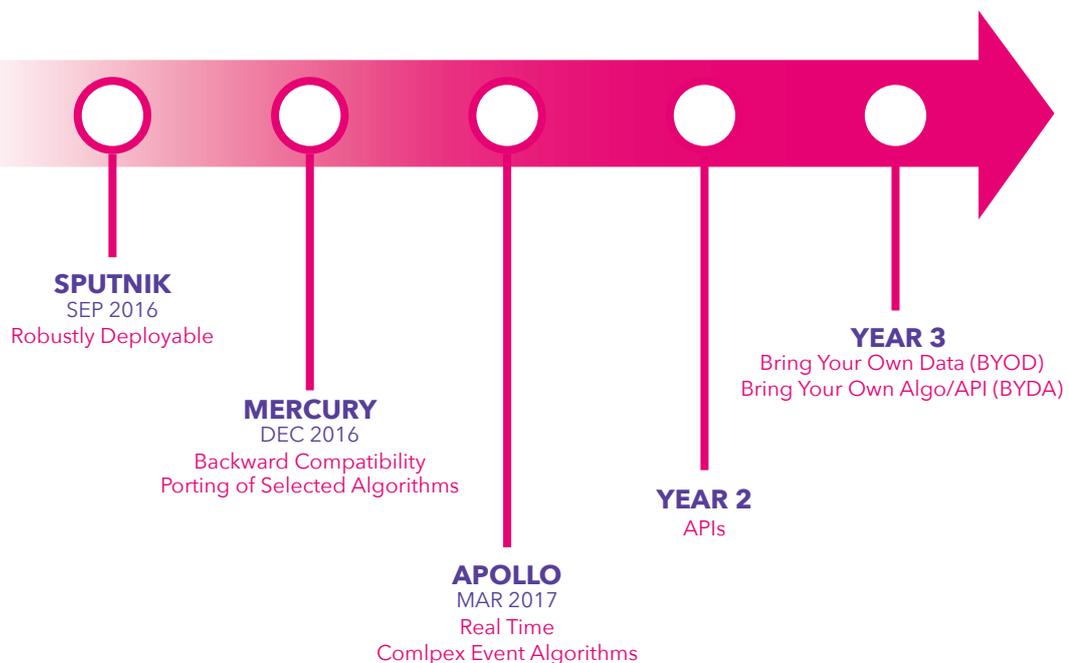
DATASPARK'S ROADMAP FOR THE FUTURE

To meet the evolving needs of current and future clients, DataSpark's roadmap prioritises flexibility, performance and security. In the face of increased data complexity, a standardised platform is necessary to ensure high performance, stability and predictable outcomes. For this reason, DataSpark has moved from project development to product development, with quarterly product releases.

Beyond March 2017, product development will focus on APIs, bring your own data, and the Internet of Things. IT teams will be empowered to build algorithms from scratch, query data without dashboards and rigorously analyse their own data.

Future releases will focus on:

- An improved ability to understand mindsets and behaviours.
- Porting of algorithms from earlier analytics projects.
- Complex event algorithms.
- Real-time analytics.



DATASPARK CURRENT AND FUTURE RESEARCH PRIORITIES

To stay at the leading edge of the data analytics innovation, DataSpark researchers have concentrated efforts in four key areas:



People

Demographics, interests and hobbies, social networks, inbound visitors, outbound travellers.



Mobility

Footfall, speed, distance, mode of transport, route, train detection.



Content

Web topic classification, web entity and location identification, web consumption patterns, TV topic/genre classification, TV consumption patterns, association analysis.



Network

Cell usage statistics, mobile network experience, fixed broadband experience.

DATA ANALYTICS IN ACTION – A GLIMPSE INTO THE FUTURE

As Singapore strives to become the world's first smart nation, everything will become smarter and more integrated: public health, urban transport, security and even housing.



6.05am: Wake at optimal point in sleep cycle with vibrating alarm on a smart watch or fitness tracker.



6.15am: Go to bathroom. Mirror displays information including sleep quality and calendar, as well as a recommendation to leave home in 45 minutes to avoid delays.



7.00am: Drive to work. Car suggests alternate route based on traffic bottlenecks.



8.00am: Arrive at work. Receive smartphone alert that teenage children have moved from one geo-fencing location (home) to another (school).



8.45am: Personal virtual assistant suggests eating breakfast before 9am meeting.



10.30am: Receive smartphone alert to check on elderly family member. Infrared sensors have detected minimal movement in the home, indicating a possible fall. The fastest way to their home is by train, leaving in five minutes.



12.00pm: Return to work. Personal virtual assistant sends a lunch meeting reminder and recommends taking an umbrella in chance of rain.

CONCLUSION

APAC's geospatial data analytics revolution presents enormous opportunities for businesses, governments and individuals. While big data analytics has been around for some time, organisations such as DataSpark are opening up new frontiers by building on small sparks.

From mobile connectivity to insights revealed by self-driving cars, each innovation opens up new possibilities. When harnessed in the right way with best-in-class security and privacy practices implemented, organisations that proactively adopt the latest analytics capabilities will see positive impact on their bottom lines as well as on society.

Those that understand where innovations are likely to emerge and are open to embracing emerging tools will also be well placed to unlock valuable insights that improve business and community outcomes.