Big Data and Analytics

Technology Overview

1.1.1 Big Data and Analytics will provide better insights and decisions for the sectors of interest. Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyse[1]. These extremely large datasets may be analysed computationally to reveal patterns, trends and associations, including relationships with human behaviour and interactions, and is referred to as Big Data and Analytics. The International Data Corporation (IDC)[2] projects that by 2020, the digital universe will grow 10 times – from 4.4 trillion gigabytes to 44 trillion gigabytes – more than doubling every two years. When it comes to data, users demand faster and more seamless connectivity, security and better services. They expect and want hardware and software to be unobtrusive.

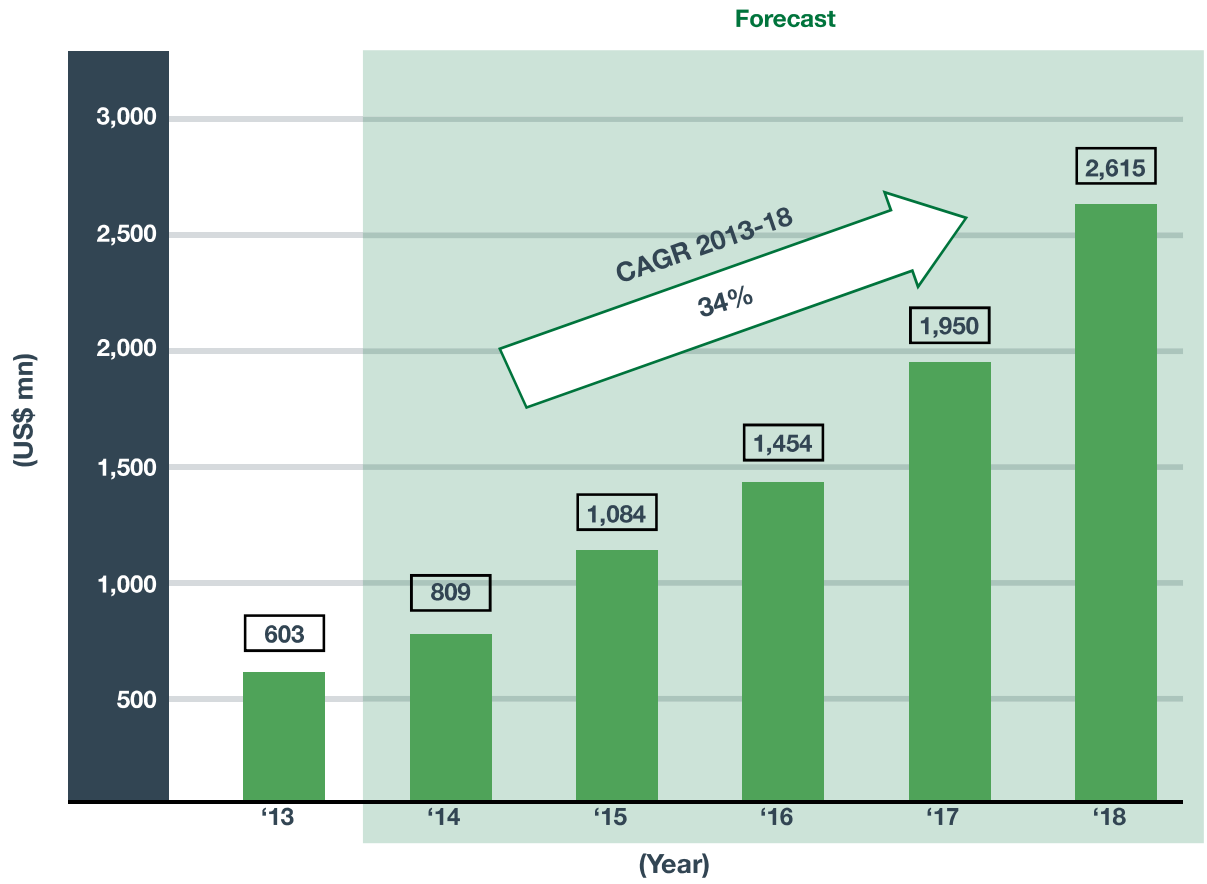
1.1.2 In becoming a Smart Nation built around data, Singapore must draw useful insights from Big Data to improve lives, business and government. It offers opportunities to improve service delivery, predict problems before they come, and anticipate the need for future city services.

Market Size

1.2.1 Businesses are now analysing data to better understand consumer behaviour, increase efficiency, and enhance their competitiveness. The global revenue from Big Data will grow 25 per cent per year from US$14 billion in 2014 to US$34 billion in 2018. In Singapore, our Big Data and Analytics addressable global market is forecast to grow strongly at 34 per cent compound annual growth rate (CAGR) from 2013 to 2018 and reach US$2.6 billion in 2018[3]. We anticipate that data storage and data warehouse infrastructure will be key drivers of the Big Data and Analytics market.

1.2.2 Technology investment in Big Data and Analytics by public institutions, such as the Defence Advanced Research Projects Agency (DARPA) and the European Commission 7th Framework Programme for Research and Technological Development (EU FP7), and private Mergers and Acquisitions (M&A) mostly focused on predictive/prescriptive analytics technologies, while venture capital firms focused more on marketing and consumer behaviour analysis in the last few years[4].

Singapore's Big Data & Analytics Addressable Global Market, 2013 - 18 (US$ mn)
CAGR = Compound Annual Growth Rate



Trends

1.3.1 Big Data and Analytics technologies accelerate the transition into a data-driven economy. It empowers organisations with the ability to manage and process large volume of complex data to make better decisions.

1.3.2 The improved use of data can generate US$3 trillion[5] in value each year in seven industries - education, transportation, consumer products, electricity, oil and gas, healthcare and consumer finance. More than half of this will directly improve the quality of lives, such as shorter waiting times in traffic and queues or better consumer shopping and media viewing experiences. The rest will help companies innovate and create new content, products and services for competitive advantage.

1.3.3 Today, more than 4.9 billion devices, like sensors, gateways, systems, networks and protocols, connect and communicate with each other, comprising what is known as the Internet of Things (IoT). As these numbers balloon, so will the number of opportunities for Big Data and Analytics[6]. For example, built-in sensors let us gather healthcare information from patients that can be shared anonymously via Wi-Fi and telecom networks from home to remote servers. Analytics conducted at backend servers on these data can provide insightful information on the patient and also trigger alarms when necessary.

1.3.4 As cities become "smarter", we can collect and consolidate significant amounts of data via the infrastructure of a city. For example, we can analyse data obtained from city-wide sensors, traffic cameras together with anonymised data obtained from public travel patterns to draw useful insights and trends that help us better understand how our city works.

1.3.5 The collection and access to data raises concerns about data privacy and confidentiality. There is a risk of data falling into the wrong hands, threatening personal privacy and public safety. To mitigate this risk, data analytics methods must protect data attributes associated with individuals and prevent unauthorised access.

Technology Roadmap

1.4.1 This table shows the industry's view of the likely evolution and mainstream adoption of Big Data and Analytics.

Demand Drivers	1-2 Years	3-5 Years	>5 Years
Larger Volume	**Large Storage Devices/ Systems (in Terabyte/Petabyte)**	**Networked Repository and Services (in Exabytes)**	**Larger Scale & Well Managed Network Repository & Services (in Exabyte/Zettabyte)**
	• Hadoop/HDFS[T] • Enterprise Solid State Drive[T] • Rack-based Flash Storage[T]	• Big Data & Analytics Platform-as-a-Service • Software Defined Storage • Data Federation & Virtualisation • Data Lake	• Data Reservoir
More Variety	**Big Variety of Data Types including Text, Audio, Video, Social Media, Geospatial, etc.**	**More Data Types including Graph, Natural-Language Question Answering, etc.**	**Interdisciplinary Analytics Across Different Domains**
	• Data Governance (e.g. metadata management, domain-specific data models • Geospatial Analytics • Self-service Data Integration • Mobile Business Analytics • Text, Audio, Video and Social Analytics	• Natural-Language Question Answering • Content Analytics Data-as-a-Service • Graph Database Management Services[T] • Augmented & Virtual Reality[T]	• Telematics[T]
Higher Velocity	**Data Processing in Near Real-Time**	**Data Processing in Real-Time**	**Data Processing with Higher Speed and Throughput (e.g. Quantum Speed)**
	• In-Memory Database • MapReduce[T] • Data Deduplication & Master Data Management[T]	• High Performance Message Infrastructure • Complex Event Processing • Real-time data analytics • Key-Value Database Management System[T] • In-Memory Data Grids[T] • Hadoop-Based Data Discovery & SQL interface[T] • Solid State DIMM[T] • Spark[T]	• Quantum Computing[T]
Bigger Value	**Descriptive/Predictive Analytics**	**Prescriptive Analytics**	**Personalised Analytics**
	• Visualisation-based Data Discovery Tool • Predictive Analytics • SaaS-based Business Analytics • Business Intelligence ETL Software • Contextual Analytics • Self-service Analytics	• Supply Chain Big Data Analytics • Graph Analysis • Smart Advisors • Virtual Personal Assistants • Prescriptive Analytics • Deep Learning	• Information Semantics Services • Personal Analytics • Big Data Analytics for Customer Intelligence • Neurobusiness[T]
Better Veracity	**Data Encryption and Copy Management**	**Privacy Preserving Data Analytics**	**Dynamic Methods**
	• Data Encryption Technologies • Continuous Data Protection • Copy Data Management	• Privacy Preserving Data Mining	• Dynamic Data Masking

[T] is classified as Technology, otherwise as Capability.
Industry has differing views on the timeframe for mainstream adoption for some technologies.

R&D Opportunities

1.5.1 We need to align upstream R&D efforts with the industry and sector development directions in the Infocomm Media 2025 report. This table highlights some examples of technology capabilities in the area of Big Data and Analytics that we need to build.

Targeted Capabilities	Sector	Next Practices (3-5 years)	Transformational Practices (>5 years)
Predictive and Prescriptive Analytics	Health & Wellness Personal Level	**Preventive Health**	**Digitised, Personalised Healthcare**
		Prescriptions' recommendation using Machine Learning algorithms	Prescribing personalised treatments in real-time based on a variety of data sets
	Transport	**Contextual & Data-Driven Intelligence**	**Smart Integrated Transport System**
		Real-time analytics to facilitate transport management and provide personalised intermodal travel advice	High throughput data analytics with complex data models for Machine Learning and simulation
	Financial Services - Trade Finance	**Improved Efficiencies in Trade Financing for Local Companies**	**Greater Efficiencies in Trade Financing in the Region**
		Predictive analytics for trade risk assessment supported by additional human interventions	Automated predictive analytics for trade risk assessment to achieve high level accuracy
	Logistics - Urban Logistics	**Aggregated Distribution**	**Digitised Distribution**
		Supply chain big data analytics to optimise enterprise wide resource planning, utilisation and delivery routing	Prescriptive analytics to determine optimal delivery schedule to allow virtual consolidation of orders and goods delivery
	Digital Advertising	**Levelling Local Capability in Media Analytics**	**Predicting Consumers' Behaviour**
		Predictions using advanced analytical models with statistically reliable data from multiple sources for effective marketing strategies	Highly accurate predictions using robust and sophisticated analytical models with reliable data from multiple sources for effective marketing strategies
	Manufacturing	**Innovation-Led Manufacturing**	**Portfolio Enhancement for Profitable Manufacturing Niches**
		Predictive and prescriptive analytics to analyse sensor data in manufacturing lines to improve production yields	NIL
Distributed Analytics	Logistics - International Trade & Logistics	**Trade Facilitation**	**Trade Multiplication**
		Complex event processing, data federation and visualisation to facilitate the full integration of management processes for international trade	Prescriptive data analytics and information semantics tools/services across geographical distributed locations for multi-modal trade & logistics optimisation
	Manufacturing	**Innovation-Led Manufacturing**	**Portfolio Enhancement for Profitable Manufacturing Niches**
		Data federation, virtualisation, and streaming data processing for data exchanges within a single factory; factory to factory coordination, enterprise data sharing, and multi-layer data analytics	Predictive, prescriptive and big data analytics technologies for domain specific applications such as satellite technologies

Distributed Analytics	e-Commerce	**Development of Strategic e-Commerce Trade Lanes**	**Adoption of Digital Commerce Trade Bloc**
		Data federation & sharing; and, predictive analytics of consumer behaviour	Information semantics services supporting collaborative commerce data-sharing standards and ownership controls of product index information
	ICT - Data Marketplace	**Development of Analytical & Privacy Enhancing Tools for Data Sharing with Better Quality**	**Development of Trusted Third Parties and Infomediaries for Data Insights**
		Data fusion tools to mash datasets from a variety of data sources for further insights	Multi-model data analytics tools (e.g. prescriptive analytics for audio, video, social media analytics) over the network
Video Analytics	Education	**Personalised Learning via Analytics**	**Ubiquitous Connectedness in Learning via Wearables and Sensors**
		Moderate integration to learning accessories (e.g. toy form factors) and/ or other tools (e.g. video analytics to gain insights on learning patterns)	Interactive toys in conjunction with related tools (e.g. video analytics to process video captured from sensors) to interact with learners in real time
3D Modelling and Simulation	Manufacturing	**Innovation-Led manufacturing**	**Portfolio Enhancement for Profitable Manufacturing Niches**
		3D modelling and simulation to offer 3D designs for supporting small-scale manufacturing in additive manufacturing	3D modelling and simulation to advance the designs and supportdomain-specific analytics for advanced additive manufacturing
	ICT - Digital Harbour	**Provision of Sustainable Supply of DC Capacity**	**Improvement of Regional Networks Connectivity**
		Real-time data analytics with 3D models	NIL
Learning Analytics	Education	**Personalised Learning via Analytics**	**Ubiquitous Connectedness in Learning via Wearables and Sensors**
		Basic data-driven analytics of learner and teacher performance	Advanced learning analytics, operating in real-time
Genomics Analytics	Health & Wellness - Personal Level	**Preventive Health**	**Digitised, Personalised Healthcare**
		Analytics tools/packages for basic patient-level genomics, with risk stratification for selected diseases	Treatment plans recommended to the clinician for decision making on personalised medicine
Social Media Analytics	Digital Advertising	**Levelling Local Capability in Media Analytics**	**Prediction of Consumers' Future Behaviour**
		NIL	Real-time analytics covering integrated consumption data across multiple platforms and non-consumption viewership data for targeted advertising, brand building and marketing