Big data in national security: online resource*

Big data concept graphic © Warchi/iStock.

Michael Chi

*ASPI has published a supporting STRATEGIC INSIGHTS piece Big data in national security. Both papers analyse applications for big data in Australia's national security, as well as the challenges that arise from that use. This online resource provides more detail on the definitions, concepts, challenges, issues and examples.

A S P I
AUSTRALIAN
STRATEGIC
POLICY
INSTITUTE

Contents

Executive summary

‘Big data’ is a collection of concepts, technologies and methodologies that constitutes a novel approach to collecting, managing and analysing data—not just tabular and relational data, but also linguistic, visual and textual data. The problem with ‘big data’ is that there’s too much data being generated, too fast, and from too many different sensors to manage easily. The promise is that new types of ‘analytics’ and algorithms, including artificial intelligence and machine learning, can be applied to these masses of data to find new knowledge and insight and to automate old ways of doing so.

While big data has come with a staggering amount of hype, there are several key application areas. It will prove particularly relevant for Australia’s national security community in finding data points that act as indicators of adverse events. But this is an emergent capability that will bring with it limitations, challenges and risks that need to be clearly understood and managed.

This online resource examines the applications for big data in Australia’s national security, as well as the limitations, challenges and risks that arise from those uses. It aims to clarify the definitional issues behind the concept of big data, which has both benefited and suffered from high levels of hype. As a result of its buzzword status, organisations in the policy, technical and commercial domains currently have separate understandings of big data. Effective public policy dialogue requires a greater degree of shared understanding, which this resource provides over four sections covering the definitions, trends, applications and challenges of big data, respectively.

Introduction: Big data in national security

Australia’s national security community deals with challenges of increasing breadth and complexity. In meeting those challenges, it’s expected to collect, manage and analyse information using an all-hazards and all-sources approach, expanding from the traditional intelligence domains to tackle new types of data and perform an all-source intelligence and early warning function.

If the national security community is to continue managing an all-hazards, all-sources approach, it will need to grapple with the problem of big data: huge volumes of novel types of data are generated rapidly, overwhelming traditional database management systems and analytical software and processes. The private sector has been grappling with this problem in its own domain as it seeks to manage product recommendations, real-time marketing, air traffic control, supply chains and other data-intensive tasks of optimisation and management.

Techniques developed by the private sector for building resilient and affordable databases and automated data analytics algorithms show promise for managing the national security community’s data problem. However, big data has specific limitations, challenges and risks that are particularly problematic in the national security context and that need to be addressed on an ongoing basis.

This resource examines the potential applications of big data in Australia’s national security community. It focuses on exploring the key limitations, challenges and risks that arise from the use of this emergent technology, and recommends several focus areas for policymakers when they consider the adoption of big data.

1. The growth of big data

1.1 What is big data?

'Big data' is a catch-all term that refers to the flood of data that's being generated daily. It has been adopted by IT companies to describe the problem of data management in an era of social media and the constant creation of user-generated content on a massive and ongoing basis. But underneath that flood of data, they argue, is the promise of finding valuable insights, from tailoring better marketing campaigns, to finding new ways of detecting symptoms in patients, to anticipating the occurrence of impending events that are prejudicial to national security.

The underlying argument of big data isn't all that new, complex or controversial. The big-data argument states that decision-makers should use as much of the data as possible—if not all of it—in their deliberations. The guiding assumption here is that decisions will be better informed when more data is available for analysis.

Things get a little more complex when we try to define what's meant by the 'available data'. It's estimated that an enormous 2.5 exabytes (2.5 billion gigabytes) of data is generated every day.[1] It's projected that the 'digital universe'—the sum total of all digital data created in a single year—will have reached 163 zettabytes (163 trillion gigabytes) by 2025.[2] Moreover, the rate of future data generation is projected to continue to grow exponentially (Figure 1). This problem—the sheer size of the data available—has become the defining feature of big data.

Figure 1 Annual growth of the datasphere, 2010 to 2025

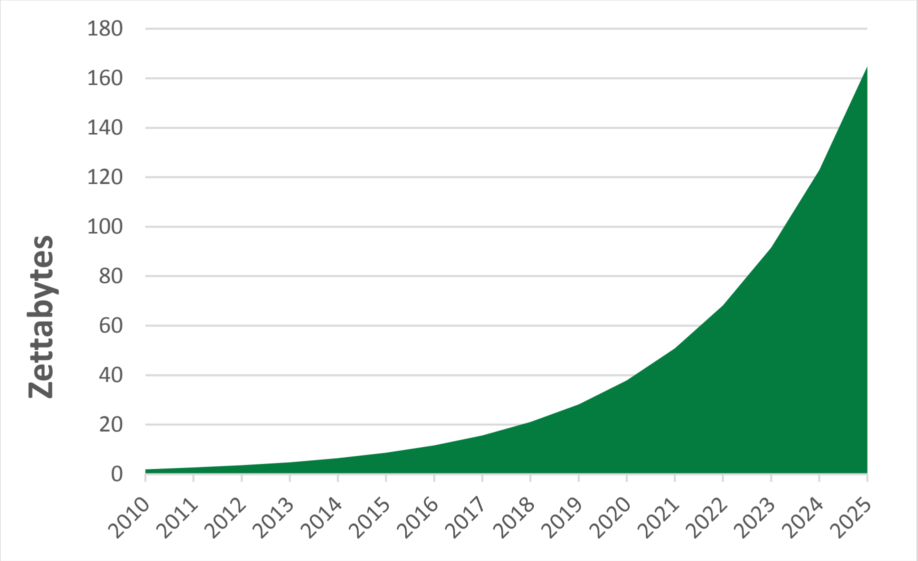
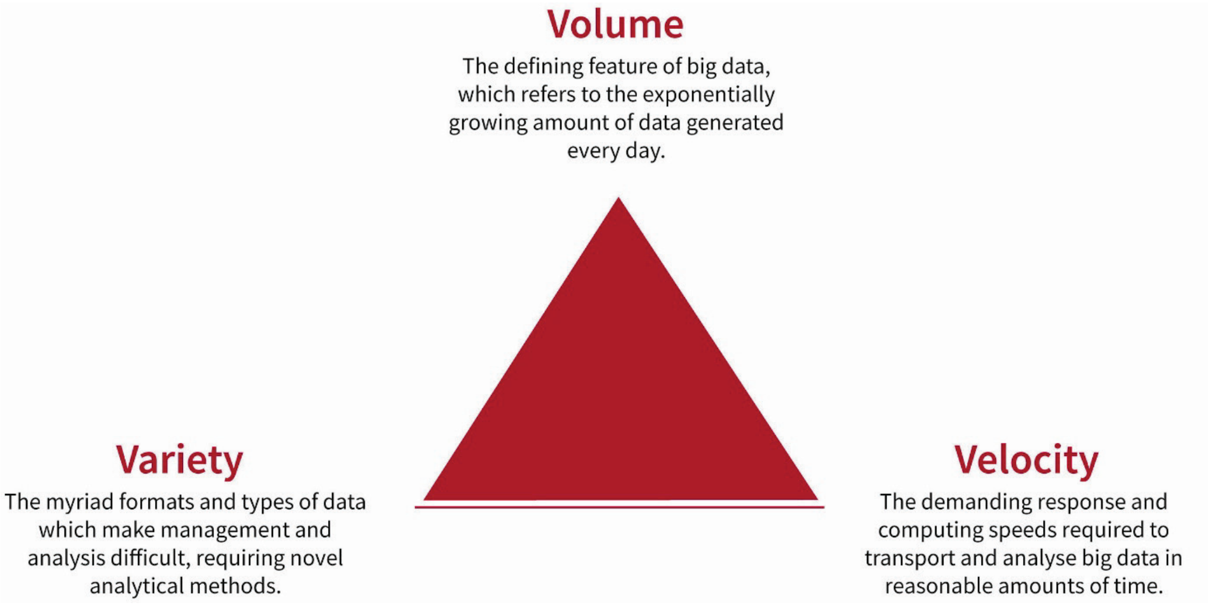


Figure drawn from data in Andrew Cave, 'What Will We Do When The World's Data Hits 163 Zettabytes In 2025?', *Forbes*, 13 April 2017, online.

But the problem of size is only one aspect of a much larger challenge. Recognising the multidimensional nature of the challenge of big data, the generally accepted definition of big data has been predicated as an iron triangle of three problems: volume, velocity and variety. This is also known as the 'three Vs of big data' (Figure 2).[3]

1.2 The three Vs of big data

Figure 2 The three Vs of big data



Volume

The defining feature of big data, which refers to the exponentially growing amount of data generated every day.

Variety

The myriad formats and types of data which make management and analysis difficult, requiring novel analytical methods.

Velocity

The demanding response and computing speeds required to transport and analyse big data in reasonable amounts of time.

The problem of *volume* has been the foundational problem of big data. The tendency for data to grow in size to fill the available space has been described as a 'deluge'[4] or 'flood'[5] in which the wave of incoming data threatens to spill out of the containers used to store it.

The problem of *variety* refers to the differing formats and sources of the data, which can range from the relatively simple, such as fiscal data, metadata and clickstream data, to the more complex, such as geospatial information system (GIS) data, biometric data, social media data and other types of 'data exhaust'[6] that users generate in their wake. These disparate data generation events all give off different types and formats of data, each mandating its own analytical methods, culminating in the problem of the variety of big data.

The variety problem is often described as a trichotomy of types of data: 'structured', 'semi-structured' and 'unstructured'. Unstructured data is distinct from structured data in that the former has no obvious hierarchy and identifiable relationships, making the identification of meaningful relationships and entities within the data challenging. Tasks requiring the analysis of unstructured data can range from identifying what's in a photo, to finding out how a person feels or thinks based on their social media presence, to tracking a single entity across multiple networks and datasets to build a picture of life around them. It's estimated that 80% of data generated today is unstructured.[7]

This problem of variety is often held to be one of the most challenging, and one that will grow more challenging with time as more types of data are generated. New types of sensor and other data, rather than just traditional computing data, are being generated from social, mobile and cloud-based technologies.[8] Mobile devices capture a bevy of location, network, audio, photo, application, gyroscopic, biometric, fitness, purchasing and usage data that previously didn't exist. This will be further complicated by a burgeoning wave of data from smart sensor-embedded 'things', such as home appliances, wearable technology and smart cars, which will become part of an interconnected ‘internet of things’ (IoT). Each type and format of data will demand its own kind of analysis to yield value.

The problem of *velocity* refers to three challenges:

- First, it refers to the speed of data generation on a daily basis, which is a contributing factor to the volume problem.

- Second, it refers to the speed of data coming into storage centres, which require high-speed connections to manage both incoming and outgoing data.

- Third, there’s the challenge of conducting real-time analysis and decision-making at speed, as some data-to-decisions cycles mandate a response time of milliseconds to seconds, as is the case in high-frequency financial trading.[9]

1.3 Two more Vs

The ‘three Vs’ have been widely adopted as a mnemonic to introduce decision-makers to the big-data management problem. Two additional Vs have since been added.

Veracity refers to the problem of whether the data collected is representative, complete and accurate. This is a challenge because, as datasets grow larger, so too does the probability that the data in them is inaccurate or of poor quality—either being irrelevant to the purpose of the analysis or simply being corrupt or otherwise inaccurate. One study showed that databases suffer from individual cell error rates of 1% to 5%.[10] At the tabular or spreadsheet level, the error rate spikes to between 63% and 99%.[11] In the same study, 94% of spreadsheets assessed contained errors.[12] When multiple spreadsheets are collated together in big datasets, this issue of veracity can scale up further, permeate big datasets and have serious cascading consequences for the analysis generated from them. One example of such widespread corruption has been widely documented in gene-name errors throughout scientific literature. It stems from an automatic conversion function in Microsoft Excel that converts domain-specific gene symbols and numbers (such as Membrane Associated Ring Finger, or MARCH1) into dates (1-Mar). A survey of 18 major scientific journals from 2005 to 2015 found that 19.6% of articles had suffered some form of these naming errors[13], potentially invalidating their results.

The other major addition to the three Vs of big data is *value*. The value of big data is predicated on the assumption that when the available data is collected and stored cheaply and effectively, and subjected to the right tools, methods and questions, it will generate previously hidden insights and provide real-time situational awareness.

The way big data is expected to unlock value is often explained in terms of the technique of ‘data mining’ to extract value. Data mining, and the way it unearths value, are described by IBM as analogous to mining for gold. IBM likens traditional data analysis methods, such as the use of ‘high-value-per-byte data’ readily visible to the naked eye, to finding nuggets of gold lying on the ground or visible veins of gold within rock. New techniques and equipment can now sift through dirt (or ‘low-value-per-byte data’), which has allowed the extraction of nearly invisible specks of gold, or nearly invisible bits of value-laden information.[14]

In the case of data, value becomes harder to see as datasets become bigger, are generated more quickly, come with more variability and less accuracy, and as the low-hanging fruit of obvious analysis is harvested. However, the difficult analysis of this less obviously valuable data can elucidate trends, patterns, associations, clusters, classifications, semantic meanings, relationships and networks that are too subtle to be 'seen' in smaller datasets, or can only be seen in specific combinations of data. These trends, once visible, can then be used to build models and frameworks, which can then be tested experimentally to establish causal or mechanistic links. This ability to 'mine' the data for these new types of knowledge has been identified as a novel method for theory building,[15] and has permeated most industry promises of big data as a potential source of unexpected but valuable insight.

However, extracting value from low-value-per-byte data requires sifting through a lot of it—a requirement alluded to in several prominent industry definitions. IBM suggests that the amount of data analysis required for low-value-per-byte data is beyond the limits of human analysis and cognition:

> More than simply a matter of size … [big data is] an opportunity to find insights in new and emerging types of data and content, to make your business agile, and to answer questions that were previously considered beyond your reach.[16]

These definitions emphasise the promise of big data. However, they also refer to the fact that human analytical ability alone doesn't scale to meet the problem of analysing big data. These observations suggest the need for better 'analytics', which refers to the use of a suite of tools, software and methods, often automated, to collect, manage and analyse all the data available.

This highlights the fifth and most important dimensional problem of big data: its value. The value of big data refers to the likelihood that a dataset, when subjected to analytics, will produce novel insights, trends or other analytical products of value to a final decision. This promise of new knowledge from data is the cause of the hype, speculation and money thrown at big data. When referring to big data, it's less about the data than about the analytics. To policymakers, it's less about the analytics and all about the output: the value that can be extracted from big data.

1.4 Conclusion

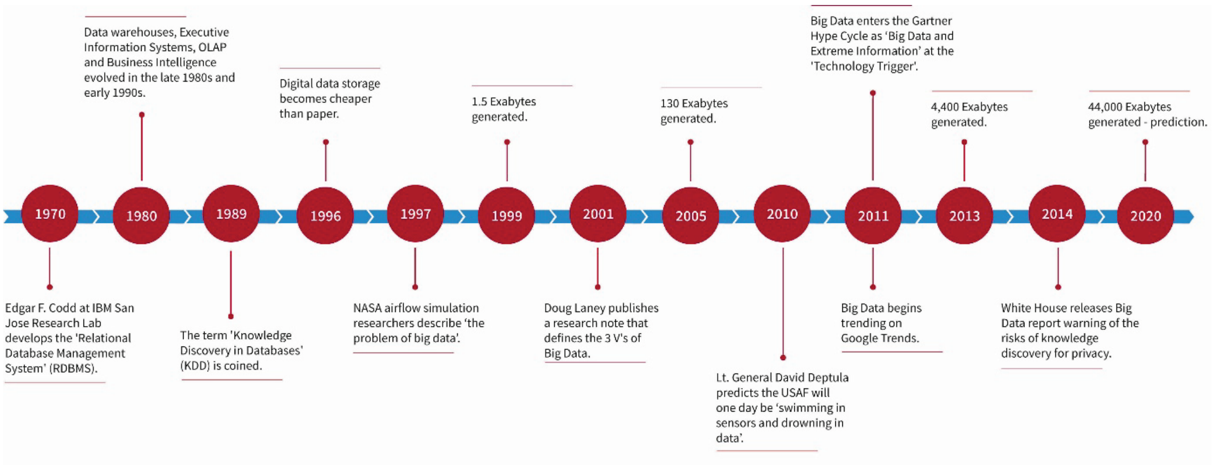
Big data is about finding new ways of managing data and analysing it for information. However, these new methods of analysis use a number of complex, state-of-the-art technologies that cut across several disciplines, from computer science to statistics, applied mathematics, economics, machine learning and artificial intelligence (AI). These disciplines have been added to by domain-specific methods and insights in other sectors, including health, genetics, finance and cybersecurity. The breadth and complexity of these technologies and applications mean that the concept of big data has become inextricably linked with several other contemporary phenomena, particularly AI.

These links, and the trends that led to them, are explored in the next section.

2. Trends in big data

While data has always been a challenge to analysts and managers, what's considered a manageable amount of data has increased as technologies have improved analysts' ability to analyse it. Therefore, what's considered to be big data is always shifting, as new technologies and methods continually reduce current examples of 'big data' into 'large (but manageable) data'. As researchers at the Association for Computing Machinery noted in 1999, 'a large scientific dataset in 1985 was of the order of tens of megabytes … [In 1999,] it is of the order of hundreds of gigabytes.'[17] Comparatively, a 'standard' data analysis server today can handle 512 gigabytes,[18] and more complex hardware scales up from there. By comparison, an average Blu-ray disc holds 25 gigabytes. On that same trajectory, large scientific datasets demonstrate the expanded capacity of data management technology today. Such datasets can involve thousands of terabytes, such as that from CERN's Large Hadron Collider, which generates 25 petabytes (25 million gigabytes) of data each year.[19] This growth in data management capability is projected to continue and is assumed as a given in future projects. For example, the Square Kilometre Array radiotelescope complex is expected to produce an estimated 62 exabytes (62 billion gigabytes) of data per year by 2020.[20]

Figure 3 Decadal timeline of big data


1970
Edgar F. Codd at IBM San Jose Research Lab develops the 'Relational Database Management System' (RDBMS).
1980
Data warehouses, Executive Information Systems, OLAP and Business Intelligence evolved in the late 1980s and early 1990s.
1989
The term 'Knowledge Discovery in Databases' (KDD) is coined.
1996
Digital data storage becomes cheaper than paper.
1997
NASA airflow simulation researchers describe 'the problem of big data'.
1999
1.5 Exabytes generated.
2001
Doug Laney publishes a research note that defines the 3 V's of Big Data.
2005
130 Exabytes generated.
2010
Lt. General David Deptula predicts the USAF will one day be 'swimming in sensors and drowning in data'.
2011
Big Data enters the Gartner Hype Cycle as 'Big Data and Extreme Information' at the 'Technology Trigger'.
Big Data begins trending on Google Trends.
2013
4,400 Exabytes generated.
2014
White House releases Big Data report warning of the risks of knowledge discovery for privacy.
2020
44,000 Exabytes generated - prediction.

Referring to the shifting nature of what's considered 'big data', Australian Data-to-Decisions Cooperative Research Centre researchers Janet Chan and Lyria Bennett Moses argue that:

> Big data in this sense sits perpetually on the technological frontier—older approaches fall outside the definition once they come to be viewed as typical database software tools.[21]

Recognising this, this resource uses the term 'big data' in its contemporary context, which refers to a series of technologies, hardware and analytical methods initially developed between 2005 and 2010 and currently being deployed and improved. Big data in this sense refers to the move away from traditional data management and analysis approaches, such as Relational Database Management Systems (RDMBS),[22] Extract Transfer and Load (ETL)[23] data warehousing and silo processes, and Structured Query Language (SQL) interactions.[24] Data management technology has instead moved to new tools that are better able to capture masses of quickly moving, variable and messy data. Those tools have ranged from parallelised processing, to ‘shared nothing’ cluster computing[25], distributed file systems[26] and Not-only SQL (NoSQL)[27] approaches to querying the data held in these massive clusters. The result has been increasingly affordable methods for tackling big-data projects using widely available commodity hardware and software, rather than expensive supercomputers.[28] It’s therefore becoming increasingly feasible to collect any data that’s generated. Furthermore, new technologies and trends, such as machine learning, are making its analysis much more valuable. These trends are covered in this section.

2.1 Current trends

Currently observable trends in big data involve hype, machine learning, and the combination of big data and machine learning.

2.1.1 The hype of big data

When bulk data collection began to become feasible, data began to be collected as an asset rather than discarded as a by-product. Big-data companies made significant investments, from over US$453 million in investment and expenditures in 2008 to US$1.5 billion in the fourth quarter of 2012 alone.[29] These new methods of storing, managing and analysing data in increasingly scalable and parallel ways are continuing to attract interest. Global spending on big data cognitive systems is forecast to reach nearly US$31.3 billion in 2019.[30]

Gartner has provided a qualitative analysis of the level of interest in big data through its Hype Cycle for Emerging Technologies.[31] This cycle is the quintessential industry guide to the timelines and definitions of emerging technologies in Silicon Valley. In 2011, Gartner began tracking the hype behind big data, based on a qualitative review of media and industry interest. According to Gartner, big data’s hype peaked in 2013, before entering the trough of disillusionment in 2014, and then dropping off the cycle entirely in 2015. Gartner argued that the reason why big data had been removed from the Hype Cycle was because it has become increasingly common, reflecting a growing consensus that big data is now a reality.[32]

More importantly, industry analysts have argued that the analysis of big data is increasingly being considered a problem for machine learning, which is a narrowly defined subset of AI. Machine-learning algorithms have attracted levels of attention and hype similar to those that big data previously did, and from similar sectors of the IT industry. Machine learning and big data are largely seen as being inextricably linked, and machine learning is seen as the most promising method for achieving the analysis of big data. Machine learning unlocks the value of big data by providing new ways of addressing data problems. Rather than being hard-coded with rules or logic by human programmers, machine-learning algorithms instead can observe previous examples, or data, generalise a model, and then test the model against further testing data, demonstrating a capacity to ‘learn’ by example, or by data.[33] Therefore, it enables ways of scaling to the challenge of big data and ‘learning’ to analyse heterogeneous, complex datasets. Moreover, machine learning can iteratively improve itself through the learning process and as it gains access to more data, leading to the excitement about big data.[34]

Examples of machine-learning algorithms being applied to big-data problems abound today. Machine learning has become such a huge part of the everyday that common ‘commercial applications of machine learning are routinely described as data mining', in which 'familiar applications known as data mining include spam or fraud detection, credit scoring, and insurance pricing.'[35] It's this ubiquity, as well as the increasing sophistication of machine-learning algorithms and approaches, that saw machine learning enter the Gartner Hype Cycle in 2015,[36] riding on the coattails of big data as one of the most promising emerging technologies of the near-term future. Such machine-learning algorithms are examples of 'narrow artificial intelligence', or AI designed to learn and fulfil a narrow purpose and set of tasks.

2.1.2 Machine learning

Computer scientists and statisticians have conceived machine learning as a dichotomy of approaches, based on the type of data being analysed[37]:

> The first approach to machine learning is 'supervised learning', where an algorithm generates a model trained on human-labelled data with known labels and known results or outputs. The aim is to generalise a model off the training data to solve a classification[38] or regression[39] prediction problem, where the algorithm is asked to predict a probabilistic prediction based on imperfect information and past data.

An example of a supervised learning algorithm is a spam detection algorithm, which takes a corpus of labelled data (data that's been tagged by humans as spam emails) and extracts 'features'[40] from the labelled spam emails, such as the number of times a certain word is used, the tone of the email and whether requests to click on a suspicious URL or pleas for money are involved. Based on the occurrence rate of these features in a 'training set' of human-labelled spam emails, a statistical weighting is given to each feature, indicating how likely it is that that feature correlates with a spam email. The features are then compared against incoming emails and, if enough spam features are in an email (for example, if there are several misspelled words, or linguistic 'features' such as appeals, the use of pleading and promises, an appeal to the authority of a collective or mention of reward[41]), then the email is 'classified' as spam. Alternatively, the email is classified as not-spam. The algorithm has the relatively simple task of sorting emails into those two known and defined categories. This process continues cyclically: the spam-filtering algorithm learns and iteratively improves over time as it learns through the process and adds more spam features to its corpus of data. These supervised learning algorithms 'allow institutions to treat spam, fraud, default, and poor health as a function of some other observed characteristics, and to automate the process of making decisions that turn on these inferences'.[42]

The second approach to machine learning is 'unsupervised learning', in which an algorithm generates a model from an entirely unlabelled dataset with no predefined classes and categories. The aim is to automatically infer labels from the data that would previously have been inferred by a human, which is also known as 'knowledge discovery'.[43] This can involve clustering,[44] or association rule learning,[45] which sorts the data into generally similar clusters of features.

In unsupervised learning, algorithms find structures, patterns, commonalities, trends, relationships and other ordering schema in a mass of data, without knowing the classes that the data can be placed into. Unlabelled data can be natural language (be it speech or text), images, videos, medical outcomes, customer models, protein sequences, web pages and other data from which information and conclusions need to be inferred through causal and mechanistic explanations. For example, a collection of text documents within a library can be scanned, subjected to text analytics and organised according to content similarity into topics, or mixtures of topics. This approach has been adopted in the medical academic community to summarise large-scale, multidiscipline health databases and document collections.[46]

In addition to this summarising function, unsupervised learning can highlight or spotlight trends or patterns in the data that were previously not discernible. Those trends can then be exposed to further experimental testing to prove causal or mechanistic ties. This ability to find correlations in an inductive way has led users to call unsupervised learning algorithms 'lead generators' or 'theory generators'.[47] Unsupervised learning allows analysts to see into the 'noise' of big data, structure it, and use it as 'fuel' for testing and experimentation, which can lead to the development of new theories.

More importantly, unsupervised learning presents a solution for the 'variety' challenge of unstructured data. Unstructured data, which accounts for 80% of data generated today,[48] is difficult to analyse because it isn't standardised, which makes it difficult for programmers to prepare it for prediction problems or tasks. Unsupervised learning algorithms allow this otherwise inert, unusable unstructured data to be modelled, sorted, grouped and clustered without the need for humans to pre-label responses.[49] Machine learning, therefore, allows the automated processing of 'data exhaust' into valuable information about social groupings, spending habits, social media sentiments and several other trends and inferences that prove revealing.

One of the recent trends in machine learning and AI has been the rise of the 'deep learning' approach.[50] Deep learning, a subset of machine learning, which is in turn a subset of AI, involves the use of 'layers' of analysis within an algorithm. This often involves layers of artificial neurons in an artificial neural network, in which the layers simulate the networks of neurons in biological brains. The layers of neurons are then stacked together in multiple layers, with each neuron fulfilling a different classification or pattern-recognition function. In an image-recognition task, the first layers would distinguish between dark and light, horizontal and vertical lines, round shapes and non-round shapes, organic and non-organic shapes and so on, moving up the layers until the system recognises, say, a cat.

As the layers learn and generate an output (a classification, cluster, prediction or other example of analysis), they are optimised in a number of ways. One method is backpropagation, which involves propagating outputs back through the algorithm until errors or losses in accuracy are minimised.[51] Alternatively, an algorithm can be sent a reinforcement signal, as was the case in Google AlphaGo, in which a positive reinforcement or negative reinforcement signal was sent, depending on the output of the neural network.[52] If positively reinforced, the neural network kept its configuration of neurons and the way in which they activated to classify something. If negatively reinforced, it discarded or amended the configuration. This process was repeated iteratively until the network was optimised for a certain task.

The highly parallel nature of a neural network meant that it wasn't well suited for traditional computing hardware. The advent of hardware specifically for video gaming—the graphics processing unit (GPU) in the 1990s, with its highly parallelised array of processing units—became the source of a breakthrough in deep learning methods and artificial neural networks in 2009.[53] Moreover, the advent of big data meant that there was more than enough data to train and test neural networks, which require an enormously large set of examples to achieve human levels of accuracy in image classification tasks (the famous example of Google's X lab recognising cats took a corpus of 10 million images to achieve).[54]

While artificial neurons are modelled on real-world biological neurons, rather than if–then statements of logic or arithmetic, these types of machine learning demonstrate what's been termed 'narrow artificial intelligence', in which certain algorithms can be optimised for certain tasks, and perform well in those tasks, but won't be able to generalise across to another task easily. For example, a machine-vision algorithm won't be able to become a speech-recognition algorithm without being retrained on entirely new data and being set up with different algorithms. This means that, while many different kinds of AI will be pervasive, they'll be narrow, limited and for a specific purpose: 'In the next 10 years, 99% of the artificial intelligence you interact with, directly or indirectly, will be nerdily autistic, super smart specialists.'[55]

However, the distinction between individual learning algorithms is becoming less relevant with the increasing ensembling of machine-learning algorithms, which has seen individual algorithms amalgamated into larger 'ensemble' learners. For example, the Netflix prize involved Netflix opening its database of 100 million movie ratings from 480,000 users over 30,000 titles on 2 October 2006 and inviting anyone willing and able to research and build a recommender system to do so.[56] The aim was to find a recommender system 10% more effective than Netflix's then state-of the-art Cinematch, offering $1 million as a prize for improved recommendations. Within six months, a Hungarian team had achieved a 6.75% improvement over Cinematch.[57] More interestingly, as the competition drew on, a trend towards building larger and larger ensembles began to unfold. Rather than competing on one video recommender learning algorithm alone, teams began to merge and stack ensembles of learners, which composited different methods for generating recommendations, ultimately culminating in a finale in which both the winner and the runner-up were stacked ensembles of over 100 learners.[58] This demonstrated the predictive strength of 'stitching' together a composite ensemble of learners compared to using any single known method and marked the start of an ongoing trend toward producing ensemble learners.

2.1.3 The landscape of big data and machine learning

Industry analysts have noted the confluence of big data and machine learning as an essential foundation of the modern big-data movement. Venture capitalist Matt Turck, in his annual landscape of the big-data industry, has noted that the two are becoming part of a combined big-data and machine-learning 'stack', or a package of products or programs that provide a wider solution.[59] This has been corroborated by respondents in similar landscapes and surveys in the fields of both big data and machine learning.[60] It signifies a growing maturity in both fields,[61] in which start-ups are increasingly taking the line of 'Take X and add AI.'[62] It's also representative of the maturation of companies that are providing big-data and machine-learning technologies as big-data solutions shift from a series of bespoke, small-scale, individual programs into wider ecosystems of programs all managed by one service provider. This integration of solutions is developing into a more mature model called 'analytics as a service', rather than a disconnected, unserviced product. [63] Already, Amazon Web Services and other cloud providers have increasingly come to offer whole-of-platform big-data solutions as services rather than as individual products.[64] Big-data best practice will become clearer as ongoing efforts to study big-data algorithms develop and the trade-offs and limitations between algorithms become clearer, as pilot studies have begun to show.[65]

Moreover, the field continues to grow. Continuing venture capital input indicates high confidence in future returns, low mergers and acquisition or consolidation activity suggest that big-data and machine-learning companies continue to find business, and several huge big-data company initial public offerings have occurred or been slated to occur this year.[66]

2.2 Emerging trends

Two trends in big data are emerging: the 'internet of things' (IoT) and some epistemological shifts.

2.2.1 The internet of things

In the future, these analytics methods will continue to be applied to new types of data not from social media but from a burgeoning array of novel sensor feeds from the IoT. They include smart watches, fitness watches and other wearable technologies and smart fridges, countertops, windows, doors, air-conditioning units, energy meters, washing machines and other appliances in and around the household. Some proponents of the IoT have identified it as the technological paradigm that will enable the creation of 'smart cities', in which sensor arrays enable the tracking and management of traffic, emergency responses, electricity, utilities, repairs and other myriad but minute optimisation tasks that make a city run more efficiently.[67] IoT also applies to the potential that biometric and fitness devices have in healthcare, providing everyday health data captured in a natural setting, rather than laboratories, providing real-time awareness into patient health, and building a corpus of gathered data that could provide new insights into lifestyle choices, diseases and other areas of medical research that are currently expensive to conduct.

The IoT already exists in many ways, but it will evolve into an increasingly expansive and somewhat similarly undefined list of things. Moreover, each new type of thing will become a new type of digital data product as it begins collecting data.[68] The IoT, therefore, represents more than just connected, or smart, things: it's part of a new paradigm of all-encompassing data collection and data analysis called 'datafication', in which the world is digitised into data ready for analysis.

Several predictions about the IoT have been made, one of the most often cited being Ericsson's 2010 prediction that there will be 50 billion internet-connected devices by 2020.[69] In 2012, IBM updated the projection to 1 trillion connected devices by 2015.[70] The current count is somewhat more conservative at 6.4 billion, not including smartphones and computers.[71] More contemporary projections through to 2020 have been significantly revised and have become more conservative as well. Note that Australia has the world's second highest take-up of fitness band devices, with 13% of the population owning one.[72] This demonstrates the deep market interest that Australians have in the IoT and the growing future of sensors and datafication in Australia. And the expansive role IoT will have in Australia's future.

2.2.2 Epistemological shifts

Commentators have suggested that one of the key long-term consequences of big-data analytics will be for the way analysts and scientists generate knowledge; analytical processes will turn from deductive, or theory-driven, approaches and move to inductive approaches. Kenneth Neil Cukier and

Viktor Mayer Schoenberger have written the most widely cited primer on what this shift in thinking will look like, which they distil into three key implications.[73]

The first implication is that analysts will start with the masses of abundant ‘available data’ rather than having to gather small amounts or samples at great cost and for limited purposes. The available data is cheaper, richer and more representative than before. Moreover, it provides greater exhaustivity, covering almost all the cases or incidences being studied and bringing granularity, allowing a general sample to be drilled down into subgroups within the sample.

The second implication is that there’ll be an inherent messiness in these masses of naturally generated, found or volunteered data. As a result, the accuracy, reliability and veracity of every data point can’t be guaranteed and instead must be worked around. However, Cukier and Schoenberger argue that quantity has a quality all of its own, meaning that inaccuracy can be balanced by the benefits of using vastly more data. They cite the particular example of statistical machine translation, which has moved from IBM’s approach featuring labelled, clean, accurate datasets (using the English–French translations generated by the Hansard transcripts of the Canadian Parliament) to instead using ‘data in the wild’, casting the net wider and using every single translation that exists on the wider internet, which has resulted in more accurate and representative translations.[74]

The third implication is that findings from these big-data analyses will, as a result of the inductive and data-driven approach, come with correlative evidence and reasoning behind them, requiring further levels of experimental design, testing and analysis from humans in order to establish causation or mechanistic links. This is a result of the way that machine-learning algorithms classify objects or cluster groups of data together with statistical expressions of certainty or similarity. In the process of learning and based on the data, these statistical weights are continually shifted and reweighted in a process of optimisation. However, the weights are just that—weights that reflect the similarity between two items based on ‘features’, and are therefore determined based on a correlative link, rather than a deductive, causal or mechanistic link. Cukier and Schoenberger argue that, while this is a limitation, it doesn’t prevent the useful application of big-data analytics, as many application areas don’t require an understanding of causation to work properly: if the correlation proves to be an accurate predictor, that will be good enough to justify its use.[75]

There has been a wide-ranging scholarly response to this, warning against adopting an over-inductive, theory-free approach. Rob Kitchin argues that there needs to be a blending between traditional, rigorously conducted, theory-driven approaches and a big-data-driven approach.[76] Similarly, statisticians have warned against the temptation to take representativeness as a given in big datasets, as that may become the cause of methodological missteps in big-data analysis projects.[77] This risk has been described as ‘big-data hubris’, in which big data begins being taken as a measure of validity without concern for methodological validity.[78]

2.3 Conclusion

Whether these new approaches to knowledge take root and up-end traditional methods of analysis remains to be seen, and is beyond the scope of this resource. Aspects of the arguments raised by Cukier and Schoenberger are addressed in Section 4.

However, these trends demonstrate the promise of value from big-data and machine-learning analytics that has informed private-sector excitement and driven development in the field, as well as some long-term implications of these new types of data, knowledge discovery and predictive tools. Those same techniques can also be applied to specific national security problems and hold great promise in such applications. Those applications, and relevant case studies, are covered in the next section.

3. The application of big data in national security

The Australian national security community has particular expectations of and applications for big-data analytics. In the national security context, some additional big-data limitations and risks need to be addressed.

3.1 The imperatives for big data in national security

The two main imperatives for big data in national security are addressing information sharing and dealing with an overload of all-sources information.

3.1.1 Address the information sharing problem

One of the most common criticisms of the US national security community after 9/11 was about the wall between intelligence gathering and law enforcers' criminal investigations.[79] Subsequent investigators argued that the intelligence community and the law enforcement community had plenty of indicators and warnings that, if shared across agencies and analysed in aggregate, would have provided effective early warning of the 9/11 attacks. For example, the Federal Bureau of Investigation didn't issue a search warrant for one of the hijackers' laptops, despite explicit warning from its Minneapolis field office in August 2001 that one of the hijackers was 'preparing to seize a Boeing 747-400 in commission of a terrorist act'.[80] The problem wasn't that of finding the proverbial 'needle in the haystack', in which information relating to the attacks was lacking; the problem was communicating that information effectively.

This problem of information sharing and lack of coordination wasn't unique to 9/11; it's a factor common to most, if not all, 'intelligence failures'.[81] This recognition of the nature of strategic surprise underwrites the discipline of 'indicators and warning analysis', in which threats to national security are assessed and monitored via visible proxies rather than the actors, activities and objects behind the threat, which are often not as easily detected. However, to build an effective indicators and warning solution, several disparate sources of information need to be consolidated, coordinated and considered as part of a reconstructed picture before each piece of information yields intelligence value. This has been described as a 'connect the dots' problem, in which many points of data need to be drawn into a wider strategic picture to generate an accurate picture of sufficient quality to provide reliable warning.[82] This has also been referred to as the 'mosaic' effect of information[83] or, more evocatively, the 'Humpty Dumpty' puzzle, according to the Defence Science and Technology Organisation.[84] The 'overall intelligence picture is dispersed across document jigsaw pieces developed by different organisations', and must be collected and pieced together to build a bigger picture.[85]

In the post-9/11 environment, this puzzle revolves around a lot more than just one focal activity or event. Instead, it involves a broad, holistic concept of the Australian border, Australian national security and the Australian national interest. National security agencies are tasked with providing all-source and all-hazards judgements. This is a major paradigm shift since the Cold War, in which an intelligence 'edge' came from 'acquiring significant pieces of critical information clandestinely and protecting them from disclosure' and has instead shifted to achieving an edge 'from breadth of access to information and quality analysis'.[86] This shift has resulted in a renewed emphasis on breaking down 'stovepipes' and pivoting away from a 'need to know' model towards an 'information sharing and collaboration' model and a 'need to share' model.[87] Calls for an improved framework for information sharing have been made in the Australian context since 2001,[88] and spiked at intervals in 2008,[89] 2010,[90] 2013,[91] and 2015.[92] The post-9/11 paradigm shift towards information sharing and coordination has led to closer cooperation between the intelligence, law enforcement and regulatory agencies as part of a wider 'national security community'.

3.1.2 Master the all-source information overload

The depth and breadth of information that the national security community is tasked with managing create a problem of information overload. US Air Force drones collected 24 years' worth of video feeds over Afghanistan and Iraq in 2010,[93] and that figure has increased as numbers of drones, numbers of video feeds, picture resolutions and drone loiter times have grown. However, as collection and storage have ballooned, analysis has become the bottleneck, owing to the fact that human analysts' numbers are limited,[94] which limits the intelligence value of the collected but unobserved video. Lt. General David Deptula described this as a case of 'swimming in sensors and drowning in data'[95]; the RAND Corporation and the US Navy described it as a 'data flood', as around 150 terabytes came in from intelligence, surveillance and reconnaissance (ISR) sensors every day in 2012.[96] The experience of intelligence practitioners testifies to the data flood: 'common wisdom among analysts is that they spend 80% of their time looking *for* the right data and only 20 percent of their time looking *at* the right data'[97] (emphasis in the original).

Not only are the volume and speed of the data threatening to overwhelm analysts, but the Australian national security community faces a distinct challenge of information variety. The national security agencies are tasked by the National Security Framework with the oversight, command, control and coordination of a range of tasks at the strategic, operational and tactical levels. This involves collecting, managing and analysing a variety of structured data, including data on immigration, visas, flights, maritime arrivals, trade flows, cargo manifests, social media feeds, telecommunications, email metadata, credit cards, bank accounts, retail purchases and phone accounts and internet service providers' metadata. As the big-data revolution continues, this list of data sources will grow to include unstructured data feeds, such as ISR and drone footage, passive sonar feeds, gunshot echolocation systems, traffic systems, surveillance footage analytics, text analytics applied to texts, social media—a list that will expand to include novel types of sensors and fused data products as they are created. These sources of data will need to be managed by the national security community to maintain its intelligence edge.

However, beyond these internal sources of information overload at the point of collection and analysis, the national security community will have to continue to compete with the same information overload at the point of dissemination. It's faced with increasing competition for relevance as a source of information for national security decision-making, and increasing pressure due to its close relationship with resourcing. This has been part of a wider problem of the 'marginalisation of intelligence',[98] in which the 'signal' of actionable intelligence becomes swept up in the 'noise' of a multitude of alternative, competing hypotheses, from sources reputable and not so reputable in the open-source world, including the 24/7 news cycle.[99]

Despite the volume, variety and velocity challenges of the national security community's information-sharing problem, the ways in which analysts work has remained largely unchanged:

‘Analysts still mostly work alone or in small groups. Their use of formal analytic methods, let alone computer-aided search engines or data-mining, is limited.’[100]

3.2 Applications of big data in national security

Likely applications of big data in the national security domain include the integration of shared information; entity recognition and tracking; predictive analytics; the generation of novel hypotheses and knowledge; and preventative and predictive national security and governance.

3.2.1 Integration of shared information

There has been movement towards automated methods for ‘data fusion’—the stitching together of various sensor feeds and intelligence products from heterogeneous intelligence and national security disciplines to build a picture of an entity, target or other ‘object’. Data fusion automatically and procedurally integrates information to produce an intelligence picture. This technique is an extension of the work expected of a typical analyst.

Data fusion centres, of which the US Department of Homeland Security boasts 77, have become ubiquitous.[101] They operate via liaison officers who have access to their own organisations’ classified information networks and share information with each other.[102] This allows investigators to bring together data records as diverse as:

> welfare and unemployment checks, firearm licences, car-rental information, credit reports, department of motor vehicles records and photos, employment histories, addresses, and phone numbers, pawn-shop information on customers, postal department inquiries, public health data, police investigation data, identity-theft reports, suspicious activity reports, and probation, parole, and booking information from police departments and correctional facilities.[103]

Big-data analytics can automate the process of data fusion beyond liaison officers by finding and linking complementary, redundant and cooperative data feeds of common interest, such as those that track the same entity or are physically co-located.[104] This can potentially solve the Humpty Dumpty problem by providing an integrated and relatively complete taxonomy of sources and facts that can be automatically cobbled together into a ‘picture’, cutting across the traditional problem of intelligence silos. What automated data fusion brings to the table is the ability to build more holistic intelligence pictures.

Case study: Information integration by Palantir Technologies

Palantir Technologies offers an information search and discovery service using otherwise overwhelming amounts of data. The company provides a ‘forward-deployed engineer’ who develops software that ‘combs through all available databases, identifying related pieces of information, and puts everything together in one place’.[105] The software allows the ingestion of multiple different datasets and their visual representation in a sociogram or graphic representation of a network. In intelligence work, this could be used to visualise the life of a target of interest based on their purchases, communications, financial transactions, accommodation, vehicle use, transportation bookings, networks of contacts, and other pieces of data and relationships. Law enforcers could use it as a unified platform for holistic and integrated case management. Financial institutions could apply it to financial fraud, which often involves links between people in criminal networks.

More importantly, Palantir offers this capability to integrate multiple databases with a user interface that requires only natural language queries, rather than programming language, and with near real-time responses, rather than lengthy query returns.[106] And it can process data at a rate of as many as 50,000 variables at once.[107]

This enables the posing of questions of prediction and likelihood based on a number of different variables. For example, Palantir could be used to generate automated warnings of trends and patterns, in much the same way that financial institutions use 'outlier' behaviours as indicators of fraud or theft, such as when a credit card is used in another country or used for a small 'test' purchase at a remote and poorly policed petrol station before being used for a bigger and riskier 'payday' purchase at a jewellery or electronics store. Palantir can also find non-obvious associations, as it did when it helped the Hershey Company figure out that sales went up when Hershey's chocolate bars were placed next to marshmallows in retail outlets. Similarly, Palantir Defense ingested GIS data about the lie of the land in Baghdad, as well as trends from past improvised explosive device attacks, to plot the safest route through the streets of the city based on past attack patterns and terrain.[108]

Palantir's data aggregation and visualisation techniques allow these relationships and actions to be viewed in aggregate, and for more complex analytics to be run across the data to gain more deeply hidden insights.

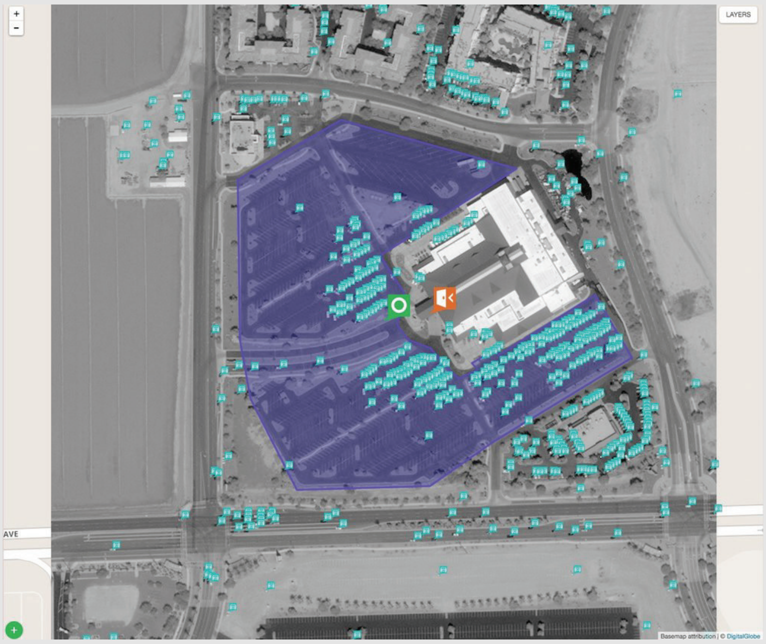
3.2.2 Entity recognition and tracking

Unsupervised machine-learning algorithms can provide summaries of unstructured data by grouping information of similar semantic meaning into clusters, and that capability applies in the national security domain. For example, a good text analytics program with an effective machine-learning algorithm can read series of transcripts and documents and identify which bits of text are relevant to an ongoing investigation. Machine vision and video analytics can provide similar summarising capabilities for the hours of drone footage and feeds that analysts currently struggle to manage, as well as for CCTV feeds, tracking an identified entity from camera to camera. Increasingly, this can be done automatically and be used to provide a 'push' or 'feed' of information about a target, keeping analysts focused on analysis rather than on the costly manual collation of information. This isn't a guarantee that the feed will be complete or have all the necessary information, but it makes collating the information at hand a much less strenuous and error-prone task.

Case study: Situational awareness and entity tracking by Orbital Insight satellite imagery and machine vision

Orbital Insight's machine-vision algorithms assess satellite imagery for entities that the program has been trained to recognise through supervised learning, and to predict behaviour. In the case of retailer JC Penney, it studied the parking lots of 96 Penney premises across the US and found a 10% drop in the number of parked cars over the first quarter of 2017.[109] It found that long-term trends in parking lot vacancies closely matched trends in JC Penney's stock price; that is, parking lot vacancies can be considered to be a surrogate for the company's day-to-day performance in attracting customers and revenue. This type of analytics is called a 'macroscope'—a tool that allows analysts to scan objects too large for the human eye.[110]

Figure 4 Orbital Insight's satellite imagery analytics

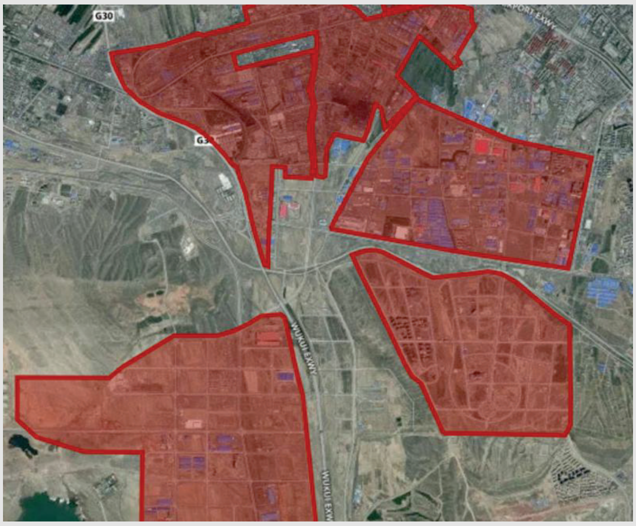


Source: Adrianne Jeffries, 'JC Penney's troubles are reflected in satellite images of its parking lots', *The Outline*, 28 February 2017, https://theoutline.com/post/1169/jc-penney-satellite-imaging.

Case study: Situational awareness and entity tracking by SpaceKnow

SpaceKnow, a satellite imagery and image analytics company, compares photos of more than 6,000 industrial sites across China and watches for indicators of production, such as visible inventory, new construction and other telltale signs. The results are aggregated into the Satellite Manufacturing Index, which is designed to serve as an independently generated index to be compared against the official state Purchasing Manager's Index.[111]

Figure 5 SpaceKnow's Satellite Manufacturing Index

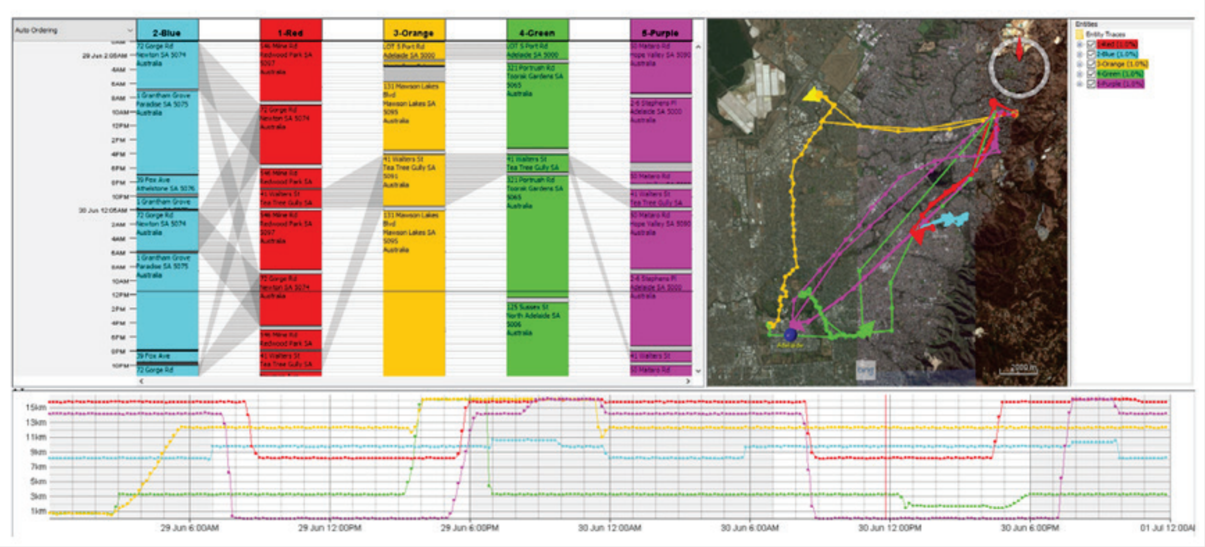


Source: Pavel Machalek, 'China Satellite Manufacturing Index', *AngelList*, https://angel.co/projects/369322-china-satellite-manufacturing-index.

Case study: Situational awareness and entity tracking by Immersive Intelligence Pod

In Australia, the Data to Decisions Cooperative Research Centre provides a satellite imagery analysis capability. The Immersive Intelligence Pod project observes geospatial datasets and visualises entities and the way in which they converge, co-locate (meet) and diverge (leave) at different locations over periods of time. The project aims to identify basic or routine patterns of behaviour of the entities in question, as well as complex relationships and networks that they are part of. The technology has been licensed to GIS company Esri Australia, which is developing it for the Department of Defence.[112]

Figure 6 Data to Decisions CRC's Immersive Intelligence Pod for entity visualisation



Source: 'New data visualisation tools for the Department of Defence', media release, Data to Decisions CRC, 3 March 2017, www.d2dcrc.com.au/news/new-data-visualisation-tools-for-the-department-of/.

These summarising techniques may mean that effective heuristic and search systems can be placed over the top of unstructured data, obviating the need for human analysts to conduct expensive and tiresome data labelling. This can enable a system in which analysts tune topic modelling, text analytics and interest weighting to build a tagged, searchable, automated feed of information that's 'pushed' to them, rather than tediously pulled from databases, 'effectively pushing the needle out from inside the haystack'.[113]

These techniques will also allow the visualisation of mass movements, structures or networks, and the tracking of entities within those networks. Social network analysis has taken off in our interconnected age, in which sociograms of Facebook relationships, Twitter communities and other 'filter bubbles' summarise human networks and relationships. This has included their effects on terrorist networks, such as those active in the Gaza Strip.[114] This is useful for analysts, as it allows the quick exploration of leads down to targets of interest, such as important nodes within communities.

3.2.3 Predictive analytics

This strategic view and understanding of how networks operate can be the basis for predictions about the location and nature of 'missing links' within network analyses.

This has been considered an important capability for analysts trying to map networks of hazards. Social network analysts commonly cite the example of the 9/11 terrorists being within one or two 'steps' or 'hops' of two people who had been photographed at a known gathering of terrorists in Malaysia,[115] as well as within one or two steps/hops of the suspects in the bombing of the USS *Cole* in 2000—a relationship that would have been readily visible on a sociogram or network graph but that wasn't obvious to analysts at the time. The argument for social network analysis is that, based on past analysis, popular and well-connected nodes in a social network graph can be indicative of threat actors elsewhere in the network.

Prediction in this sense broadly means using available information to predict, or probabilistically infer, information that isn't available. Medical diagnosis is a classic example of a predictive task in which a clinician is tasked with predicting the existence of a disease based on imperfect information, such as list of symptoms. Even a task that humans consider routine and in the 'present', such as picking up objects in a warehouse, is a prediction problem. Amazon continues to host an as-yet-unsolved Picking Challenge to find software that can 'see' different shapes with different weights and firmness and 'predict' the correct grasping angle to use, and do so without dropping or crushing the object. This is a prediction task that humans handle routinely and from an early age, but that state-of-the-art approaches in machine vision continue to struggle with (although robotic arms have more than enough dexterity).[116]

In the commercial world, the main use of big data has been in predictive analytics based on consumer behaviour. For example, Google AdSense builds 'models' or 'profiles' of individual users based on their search histories and other data and then 'predicts' the best recommendations based on a mixture of similar user profiles and the common features between past searches and potential recommendations. This has been demonstrated in advertisements, such as Netflix's and Amazon's recommender systems, in which personalised suggestions are a function of a behavioural profile and predictive algorithms based on past viewing and purchasing data.[117] Google's predictive search and autocomplete functions also feature prediction, but with a greater emphasis on matching search queries to other users' search queries and topic models, rather than purchasing habits. This process of training-data-based predictions and forecasts can be used to project the traditional (weather, stock market, bets, creditworthiness) and the not-so-traditional (uprisings occurring due to social movements, tracking disease spread, product recommendation systems).

Similarly, in the national security domain, there's the opportunity to use automated indicators in which past behavioural, financial and other profiles of detrimental activities and actors are analysed to indicate potential impending threats. The New York State Intelligence Center's 'terrorism indicators reference card' lists several indicators in individual traveller profiles that correlate with past terrorists' profiles and could provide warning of future attacks, such as 'recent travel overseas', 'has student visa, but not proficient in English', 'refusal of maid service', 'owning a GPS unit' and an 'unusually calm and detached behaviour'.[118] Alternative indicators have also included jihadist groups having a history of exercising at paintball courses in Australia.[119] Other behaviours, based on analyses of past patterns of terror attacks and terrorist actors, for example, can similarly be analysed for indicators and predictors, and those indicators can be automated to provide predictive, probabilistic warning about future attacks.

The combination of data fusion, automated heuristics of data feeds, social network analysis and predictive indicators and warning analysis has produced high hopes for approaches to national security ‘event prevention’.[120] Event prevention entails the prediction of likely-to-happen events and the monitoring of the potential threats to enable disruptive or preventative actions as necessary.

3.2.4 Novel hypothesis generation and knowledge discovery

Big-data analytics’ inductive, bottom-up approach to knowledge has enabled the mining of relevant criminal and security datasets to uncover correlations, patterns and trends that were previously not considered, or even discoverable by human minds.[121] This promises a ‘next generation’ of national security intelligence analysis, in which innovative analytics mine past data and uncover new indicators of national security events.

Therefore, big data promises two benefits for predictive analytics. Supervised learning algorithms allow the structuring of indicators and warning frameworks into automated early-warning alert systems, while unsupervised learning algorithms can find new indicators and warnings in the ‘noise’ of big data, enabling the discovery of novel indicators and the creation of new predictive models.

However, big data isn’t a panacea for ‘black swan’ events or the unknowable future. Predictions based on estimative probability and statistics rely on the assumption that past performance is a predictor of future performance, extrapolating a best fit based on past data to determine what the future is most likely to look like. This makes it unlikely that that big-data analytics will foresee massive inflection points arising from sources exogenous to the data, such as black swan events, or from events that don’t present indicators and warnings in the data used to construct the model.

Case study: Data mining and predictive analytics by Target Guest Marketing Analytics

Mining data in order to find novel insights and trends isn’t new. Target, the department store retailer, has been using customer analytics manually for some years. Target creates a ‘guest ID’ number for each shopper. At every turn, it links demographic information to the guest ID, including age; marriage; kids; town; driving distance/time to the nearest Target store; estimated salary; recent moves; credit cards; websites; ethnicity; job history; magazines read; bankruptcy; divorces; mortgages/houses; topics talked about online; brands of coffee, paper towels, cereal and applesauce consumed; political leanings; reading habits; charitable giving; and number of cars owned.[122]

Based on those details, Target’s Guest Marketing Analytics Department used the science of habit formation to identify periods when customers’ brand loyalties shift, such as during the second trimester of a pregnancy. Studying the guest ID data of women during the time they were pregnant, the analytics team produced a list of 25 products that, analysed together, generated a reliable prediction score for women in their second trimester. Those products included calcium, magnesium, and zinc supplements; soap; cotton balls; scent-free and extra-large bags; and hand sanitisers and washcloths.

The model’s predictive accuracy was demonstrated when Target sent coupons to one of the women it predicted to be pregnant. The father of the woman was furious and stormed into a Target store to complain about what Target might be suggesting she do. After a follow-up customer service call a few days later, the father was apologetic, having found out that his daughter had already been pregnant.[123] This anecdote is usually pointed to as one of the best examples of the predictive benefits of data mining. Critics have called it a case of one lucky true positive among far more but less visible false positives (such as when non-pregnant women are sent coupons related to pregnancy products), and said that the real rates of predictive accuracy would be relatively mundane.[124]

Case study: Data mining and knowledge discovery by IBM's Chef Watson

IBM Watson, a 'cognitive system' designed to deal with unstructured data (specifically, national language question answering) was spun off into a demonstrator project called Chef Watson.

Chef Watson ingested research material on the chemical composition of hundreds of different food ingredients, as well as a corpus of 10,000 recipes from the Bon Appetit website. It then combined the data and trawled through it for recurring patterns and combinations of up to four different ingredients, which would suggest that those ingredients work well together.[125] It's since been updated to include other recipes, books, academic studies, and even tweets scraped from the internet,[126] as well as spreadsheets on the molecular makeup of flavour and odour compounds in food and 'hedonic psychophysics' research papers on smells and tastes that people find pleasurable. Watson uses that corpus of data to generate a recommended recipe base of ingredients or to extrapolate a suggested recipe using the few ingredients that a user has available, alongside a percentage rating of the ingredients' 'synergy'.[127]

Some of these data products are raw and unsuitable for immediate use, requiring further analysis and judgement before being used—like Watson's suggestion to combine tomato, garlic, onion and purple seedless grapes into a 'Purple Seedless Grape Starch Dish'.[128] However, when paired with a professional chef, Chef Watson proved to be useful in finding novel combinations that a human would previously have not considered, providing a middle ground between a deductive, non-deviating approach to recipes and an overly creative, inductive, trial-by-error approach.

IBM designed Chef Watson as a metaphor for the 'creative thinking' that Watson is able to contribute in finding novel relationships, combinations, patterns and other correlations that can lead to knowledge discovery.

These data-mining techniques for finding novel 'generators' or 'leads' for a model-based theory are now being increasingly automated through the application of machine-learning algorithms to cluster data and find associations that aren't obvious to human analysts, whether because of the volume of the data or the cognitive blind spots of the analysts.

3.2.5 Preventative national security and predictive government

Predictive big-data analytics has reinvigorated one of the most contentious issues in contemporary national security studies: the idea of preventative policing or national security. This approach involves a change from a slower, *post hoc* monitoring and indicator regime to an automated and continuously operating one.

More importantly, predictive analytics now allows these insights to be derived relatively automatically and procedurally, and in short order. Increasingly, it's now possible to provide 'now-casting' services, in which events are logged as they happen and an alert is provided to a waiting analyst or decision-maker. This approach has already been used in specific, highly disciplined domains, such as military airborne warning and control systems using traditional electronic and signals intelligence. But, in the era of big data, machine learning and the IoT, more and more real-world features are being converted into data and automatically scanned and analysed for meaningful signals, which can then be tested and deployed as indicators for warning analysis.

In summary, the data-to-intelligence cycle of the rapidly approaching future will involve the automating of analytical processes and datasets that the national security community knows are of intelligence value. It will also involve the discovery, in databases, of knowledge that's not currently known—finding novel patterns, trends and correlations. Once they have been found, they can be placed in an automated indicators and warning program, which will allow the national security community to generate complex strategic early warnings about events and threats as diverse as cyber threats, data breaches, foreign intelligence operations, mass-casualty attacks and lone-wolf attacks. Moreover, the predictions can be continually compared to the real-world data and then optimised to better reflect that data and trends. Paul Symon and Arzan Tarapore argue that this automated, predictive and inductive approach to intelligence analysis will be a paradigm shift from the 'current industrial age model of linear finished intelligence production to an information age model of integrated and adaptive assessment service delivery'.[129]

Case study: EMBERS

The Early Model Based Event Recognition Using Surrogates (EMBERS) program is a project of the US Intelligence Advanced Research Projects Agency, run as part of the agency's Open Source Indicators program.[130] The aim of EMBERS is to develop forecasts of critical events, such as civil unrest, diseases, protests, outbreaks and elections. It aims to provide 'anticipatory intelligence' on such social events by scanning open source indicators, and to constantly optimise itself to detect new types of indicators. EMBERS operated as a proof-of-concept project from August 2012 to July 2016.

EMBERS ingested almost a dozen data sources ranging in size from weekly government reports to Twitter, collating a full information feed that generated about 19.2 gigabytes per day from Spanish, Portuguese and English sources, with a geographical focus on South America. The raw feeds were then enriched using entity extraction to find people, places, organisations and other features, such as numbers, dates and hashtags in the text, geocoding, and final sentiment analysis. This expanded the volume of the data being processed by the system to 40 gigabytes per day. The program searched the feeds for occurrences of three or more of 800 specific words or phrases that serve as semantic indicators of unrest,[131] as well as constantly mining for other words, phrases or hashtags that were tied to upcoming social events. The system developed multiple machine-learning models, generally between six and eight algorithms per type of event.[132] The models were then weighted and optimised based on their accuracy by a master fusion module, which combined the models in a way that it deemed would produce the most accurate prediction.

On average, the system generated 50 warnings a day, based on an indicator suite of 4.6 million messages, of which 350 were flagged as significant by the predictive models.[133] Some of the key events that it predicted were protests after the impeachment of the President of Paraguay, the Brazilian Spring (a series of demonstrations in several of Brazil's cities), hantavirus outbreaks in Argentina and Chile, and widespread protests by Venezuelan students.[134] It also missed many others, such as protests in Brazil in March 2015, protests in Mexico in December 2014 and the early onset of the Brazilian Spring.[135]

EMBERS was run for five years and scored against a monthly catalogue of events, as reported in newspapers, by MITRE Corporation in a 'gold standard report' compiled by human analysts—a performance review on which continued funding for EMBERS relied.[136] On average, by the second year of the project, the program was able to provide early warning with a lead time of 7.54 days; 94% of forecasts matched an event from the gold standard report, and the model successfully forecast 65% of events from the report.[137] Where necessary, EMBERS could be subjected to an audit of the gradual ablation of its data fusion, enrichment and transformation steps.

3.3 Conclusion

These potential applications of big data hold great promise and relevance for solving the big-data challenges of national security. However, there remain serious limitations, challenges and risks that are particularly pernicious in a national security context, which are addressed in the next section.

4. Limitations, challenges, and risks arising from big data

This section canvasses the potential downsides of big-data analytics in the national security domain.

4.1 Limitations

Big-data analytics can be limited by problems of representation; bias and discrimination; false positives and negatives; and feedback loops.

4.1.1 Representation issues

Data scientists have commented on the 'unreasonable effectiveness' of data,[138] noting that having more data yields better predictive performance than more carefully designed algorithms (to a point). The White House noted in its report on big data that some trends become visible only in big data, citing a case in genetic research in which genetic markers relating to schizophrenia were entirely undetectable in small samples but hit an inflection point and became statistically significant and identifiable in a dataset of 35,000 cases.[139]

While the unreasonable effectiveness of data shows great promise for solving previously unresolvable problems, it contributes to one of the key limitations of analytics: that is, it's practicable only where the data is available. Whether the data is labelled or unlabelled, there needs to be plenty of it. This requirement means that big data can serve as a funnel and limit data analysis within a 'streetlight effect' when analysis is limited to where the data is available rather than where data is needed. [140] This limitation has been noted in other knowledge domains; for example, overly WEIRD (white, educated, industrialised, rich and democratic) sample populations made up the experimental base of 96% of studies in leading American psychology journals from 2003 to 2007.[141] The same problem applies in the digital context.[142]

In the national security setting, this could mean that a big-data targeting regime could be limited to only those targets that are over-represented in the data, resulting in too tight a focus on those types of target in investigations and operations, at an opportunity cost of inattention to other targets.

Conversely, the problem of under-representation poses unique and pernicious problems in the context of algorithmic classification. Researchers have found that when a subgroup represents only 30% of the data in a wider dataset, a learning algorithm will be uncertain about predictions for that minority. For example, a credit application approval algorithm looking at a population of 500 applicants consisting of two subgroups will be assigned differing approval weightings, even where the probability of repayment is uniformly and universally 95% across the 500. If a minority subpopulation consists of 10% of the population of 500, it will only achieve an average approval score of 80%. At 20%, this increases to an approval score of 85 and at 30% to an approval score of 90; at 40%, the minority group approaches a 95 approval score.[143] Only then does the algorithm accurately predict the risk.

In this hypothetical, the only difference in the underlying data was that one population was less represented, which generated greater uncertainty. This led to fewer loan approvals and therefore created different impacts among two different populations. In this case, the two populations were a simple representation of whites and non-whites. This results in the problem of 'uncertainty bias', in which, all other things being equal, uncertainty can result in greater perceived risk.[144]

4.1.2 Bias and discrimination

The contribution of uncertainty to biased outcomes highlights some of the problems inherent in a statistical risk scoring and threshold system. Within any risk scoring system of this kind, it's difficult to determine the 'correct' way to set the threshold.

The level at which a threshold is set is generally dictated by the outcome desired, as in the case of creditworthiness scoring systems. Set the threshold too high, and too many people who should be given loans are denied them; set it too low, and too many people who will default are given loans. A study on discrimination and machine learning has demonstrated how different kinds of threshold can have diverse and discriminatory outcomes for two subgroups.[145] For example, a threshold may be set in a way that maximises profits from the loans by making only 'correct' lending decisions (lending to those who will repay *and* (not *or*) denying loans to those who won't repay and avoiding 'incorrect' decisions (lending to those who won't repay and denying loans to those who will repay). Maximising this correct:incorrect ratio may only be possible if two different populations ('blue' and 'orange', say) are offered two different thresholds, which means that they have been held to two different, and therefore discriminatory, standards. Alternatively, an entirely group-agnostic approach may set the same threshold for both the blue and orange populations, but this would mean that, between two equal populations, one would be granted fewer loans overall and otherwise creditworthy applicants would be at a disadvantage. Alternatively, a 'quota' system of ensuring equal total loans to both populations would discriminate against either the blue or the orange population. At some point, a trade-off decision is made about which threshold is best for the overall population, with an understanding of the disparate impacts that it will have on each subpopulation.

These hypotheticals demonstrate how bias and discrimination are present in even a basic classification algorithm with streamlined and reliable inputs, outputs and independent factors that affect the result.

These kinds of group-based inferences, associations and clusters are not acceptable, and the use of such sensitive personal information in creditworthiness ratings is prohibited under law and in practice. That is, populations are not to be explicitly subdivided into 'white' and 'non-white', or 'blue' and 'orange'. However, the nature of unsupervised machine-learning algorithms means that such groups will be naturally associated or formed over time where machine-learning methods can and often do probabilistically infer hidden variables by using other data as a probabilistic substitute or proxy.[146] The nature of association rule learning and clustering algorithms may result in forms of guilt by association—not by intent or by design but through the clustering and association functions of the algorithm. More directly, such learning algorithms can reproduce the patterns of discrimination, prejudiced decision-making or other widespread biases that are present in the data they are analysing. [147] Most perniciously, 'any form of discrimination that happens unintentionally can also be orchestrated intentionally.'[148]

4.1.3 False positives

As demonstrated above for credit rating systems, any risk assessment algorithm results in false positives, no matter how accurate the model, profile or data that sits behind it. In the case of national security, this creates a risk that innocent people will be subjected to intrusive, disruptive and costly investigations by national security agencies. Even a 99% accurate data mining and alert system will suffer from a 1 in 100 false positive incidence rate, which, in an input of 1 trillion indicators, means 10 billion potential positives that need to be reviewed by analysts.[149]

These unhelpful results are an issue because of the bugbear of machine learning: spurious correlations.[150] That is, the bigger your data, the more false positives and spurious correlations will turn up in it. As data scientist Vincent Granville has written, '[It is] not hard, even with a data set that includes just 1,000 items, to get into a situation [that involves] many, many millions of correlations.'[151] Further false positives can be generated if the classification or prediction algorithm isn't generalised correctly. Machine-learning algorithms can be overtrained, so that the algorithm learns to 'overfit' the data on which it was trained and makes predictions based on features of the training data that aren't useful for the predictive model when working with real-world data.[152]

The costs of responding to such false positives are not trivial. They involve the expenditure of serious resources, an invasion of the privacy of those that didn't need to be investigated, and a degradation of the national security community's analytical resources. That is, every false positive has an opportunity cost.[153] Furthermore, the investigation of false positives can risk appearing to be without reasonable cause. This relates to one of the key methodological challenges of statistical and data analysis, in which 'data mining' or 'data dredging' were originally derogatory terms used to refer to untrained statisticians 'fishing' through the data without a rigorous or well-thought-out theory. An overly expansive analytics regime can come to constitute endless fishing expeditions[154] in which every possible lead, correlation or target is chased without a rigorous and targeted approach and without regard for individualised suspicion,[155] with negative impacts on the social licence with which the national security community operates.

Ultimately, false positives arise when data analytics algorithms overfit the parameters or data they are given, resulting in decision paralysis or analyst fatigue or, more perniciously, in overpolicing and a feedback loop of recidivism and community fragmentation.

4.1.4 False negatives

The risk of false negatives is serious in the national security context. Fortunately, there are few past examples of actors working against Australia's national security, but such threats are extremely diverse in location, nature, methods, purpose and networks. There's no large dataset of terrorist behaviour that can be reliably drawn upon to produce robust models, which is entirely unlike traditional data-mining for credit reporting. This dilutes the pool of data available for national security classification tasks.[156] Moreover, not only do we lack such a baseline of established behavioural data, but the data being analysed has been purposely altered by adversaries who seek to use covert methods and blend in with the normal population in order to disguise their activities.

In machine learning, this problem of individuals forcing false negative results is particularly challenging. It results in a continuous trade-off in classification problems, in which an analyst can drive the rate of false positives to zero or the rate of false negatives to zero, but not simultaneously and not necessarily in the same proportion.[157] In national security policy and the public eye, there's significant political pressure to drive the number of false negatives to zero, 'but this political requirement belies the technical reality that the number of false negatives can never be zero.'[158] A false negative, in this instance, would be a target of interest who should be under suspicion but is overlooked and able to act to the detriment of Australia's national security.

This creates three specific hazards:

- First, national security agencies can suffer extreme reputational damage if a false negative is let through.
- Second, the agencies can aggressively remove as many false negatives as possible from their predictive systems, at the cost of having to investigate significantly more false positives.
- Third, with increasing volumes of data, the likelihood that false negatives will exist in national security datasets also increases. The Defence Science and Technology Organisation has described this challenge as the 'unknown known': national security organisations hold the pieces of a puzzle, the solution of which could have prevented a negative national security event, but due to the deluge of information did not or were not able to act on that information.[159]

4.1.5 Feedback loops

Feedback loops also limit predictive action. If someone is under suspicion, then the increased attention will cause what would have otherwise been smaller, unnoticed, infractions to instead become negative blips on their record, triggering more aggressive surveillance or monitoring, which in turn is likely to result in further tickets and penalties, further worsening their record. These small decision errors, whether stemming from over-representation or over-surveillance, can produce a negative feedback loop, risking overpolicing and radicalisation. This limitation is due precisely to the automated, aggregated and accelerated system that big data provides.[160]

Big-data analytics can automatically take inputs that contain factors of discrimination, whether through qualitative errors; incomplete, incorrect or outdated data; selection bias; or the unintentional perpetuation and automation of historical biases that existed within the data. It's for this reason that privacy and criminology scholars have begun to consider the dangers of big data and automated decision-making. Their concern is that past criminal data could reinforce rather than reduce disparities in policing and criminal sentencing, resulting in biased outcomes in event prevention and law enforcement.[161] Moreover, this bias could result in a feedback loop in which historical data becomes the basis for more aggressive monitoring measures, resulting in systemic overpolicing.[162]

Predictive analytics and recommender systems have had such feedback-loop effects on social groups and communities online. On the internet, there's been a proliferation of 'filter bubbles', which are a type of 'informational determinism' in which constant web personalisation slots users into specific communities and directs them towards certain products, which in turn becomes the basis for further profiling in 'an endless you-loop'.[163] This phenomenon has also been termed an 'echo chamber', in which people within the same filter bubbles reinforce their own biases, perspectives and arsenal of facts within closed communities, leading to a net intensifying effect similar to radicalisation.[164] This fragmentation of social groups online has been identified as a threat to public discourse and open, informed debate. Data scientist Gilad Lotan demonstrated what this fragmentation looks like in a sociographic visualisation of the online debate between supporters of Israel and supporters of Palestine, demonstrating in an intuitive and simple way the divide in information sources, communication and network links that over-personalisation can create and the risks of extremism that arise from the fragmentation of public communities.[165]

4.2 Challenges

Challenges to big-data analytics include overhyped expectations; the inherent complexity of big data; the difficulty of cost–benefit analyses; data siloes and fragmentation; and the opacity of algorithms.

4.2.1 Overhyped expectations

Industry pitches on big data have tended to promise a 'one-stop shop' solution. This has conveyed 'the idea that big data is magic. You get your data, you press a button and all of a sudden you have extremely valuable output. This idea is very wrong and dangerous.'[166] Rather, data science is extremely time consuming. Although technologies and methods for managing and analysing big data have come a long way, many of the difficulties of managing data remain the same: for example, analysts find that they spend most of their time chasing data rather than analysing it, even in a highly automated big-data analytics process.[167]

The too simple expectation that big-data analytics is a simple layer, funnel, filter or button has been echoed in the national security community. A survey of the Australian national security community found an expectation that big data would allow the creation of a 'find terrorist button'. Those expectations underestimate the difficulty and complexity of big data.

4.2.2 The complexity of big data

Complexity presents a fundamental challenge in big data. It revolves around enterprise-level technology, which involves a pipeline of technologies, a lot of hard work on the part of human analysts, and good systems integration to bring it all together. As a technology, big-data analytics isn't readily visible or touchable: it's back-end 'plumbing' that's difficult to visualise.[168] It's a long and complicated chain of technologies from data capture to data storage, cleaning, query, analysis, visualisation, and then down to the end user, and each step needs to be integrated seamlessly for the system to work. The distributed nature of big data arguably makes implementation that much more challenging, as the system has myriad small moving parts working in concert. This is counter to the singular image of the lone data scientist as having 'the sexiest job of the 21st century'.[169]

The pipeline of data collection, management and analysis means that a diverse and talented team is needed, and not just in data science but all across the analytical spectrum. A non-exhaustive list of the skills needed includes 'data management, machine learning, parallel computing, security, software engineering, statistical analysis, inference, and visualisation.'[170] The new skills requirements for data analysis have crystallised into the new discipline of 'data science', which involves the merging of disciplines and subject-matter expertise to produce composite analytical products covering various sectors. The role is needed because of the limitations of current data-mining tools and because it is ‘not straightforward to perform analytics. Most of the time is consumed in preparatory work to the application of data mining methods’,[171] requiring skilled data scientists to shape automated analytics suites through careful, continuous and iterative processes.

This demand for a complex suite of requisite skills and familiarity with the tools has resulted in a large projected skills shortage and workforce gap. The McKinsey Global Institute has provided what’s now the most widely used estimate of the shortage that big data faces: 250,000 data scientists with specific big-data experience and skills.[172] There will also be a wider shortfall within the community of managers and analysts required to have big-data knowledge, which a previous estimate placed at 1.5 million.[173] In this projected environment, the national security community will find it difficult to compete with private firms for analysts with the skillsets needed for the big-data analysis pipeline.

4.2.3 The difficulty of estimating the costs and benefits of big data

Overinflated expectations worsen the already dismal prospects of success that IT projects face. An ongoing annual survey by Standish Group since 1994 has found that only 16.2% of IT-related projects undertaken have been considered ‘successful’ (that is, on time, to budget, and with all contract-specified functionality).[174] This low success rate has improved somewhat in surveys conducted since 2002, reaching 29% in 2015.[175] Furthermore, projects valued at over US$10 million experienced a 0% success rate, whereas smaller projects with total budgets of less than US$750,000 have had a relatively high probability of success, at 55%.[176] The larger and more complex the project, the more likely it is to encounter a delay, cost overrun or failure. Big-data projects are, by their nature, big and complex.

Whereas the costs and risks of big-data projects are clear and daunting, the benefits aren’t as clear. There remain considerable challenges in methodically gauging their value. This stems from the unconventional ways in which big data generates value. Rather than the value of data being immediately obvious, the datasets involved in big data demonstrate their value only after analysis. Furthermore, while data has a high ceiling of potential value, there’s a high probability that a dataset will have low value. However, there’s also a low probability that a collection of datasets will have high potential value, and that potential value becomes more probable as more datasets are linked together. This relationship has been expressed in terms of data having ‘networked value’ or a ‘network effect’, in which each data source may have a specific, limited purpose.[177] The combination of data sources, however, may uncover new meanings.[178] As data increases in quantity, so too does its potential intelligence value, generating a strong incentive to collect all of the data available.

This means that the true, final value of data isn’t always obvious *ex ante*, and the data that an organisation has is likely to increase in value as more and alternative data is acquired, which has driven the big-data land grab. This makes estimates of costs and benefits extremely difficult in the big-data context, making the management of big-data projects more difficult and their prospects for success less certain.

This isn’t to say that big-data analytics has so little application as to not be useful. There are many applications for which it’s uniquely suited, but others in which the benefits of the analysis aren’t worth the increased costs of computing and algorithmic complexity. Estimating and managing this trade-off will be difficult, given the networked value of data.

Finally, there's the possibility that some problems can't be solved by simply collecting more data. Such a limitation won't become evident until reasonably exhaustive analysis has been conducted, if it becomes evident at all. This problem is illustrated in attempts to predict earthquakes:

> Despite all the science, data, and models that have been thrown at predicting earthquakes, there has been no appreciable progress. It could be the data is insufficient, or that the models are incomplete, or that the system is too chaotic (in a mathematical sense) to make it capable of being modelled on the sort of information that could realistically be collected.[179]

4.2.4 Data silos and fragmentation

Comparative intelligence scholars have noted that certain intelligence communities tend to be integrative and others disintegrative, as in the case of the British community and the US community, respectively.[180] The Australian intelligence community has been noted to be particularly secretive,[181] and fragmented across far more data silos than its counterparts in the UK, the US, Canada and New Zealand.[182]

This has led to challenges in integrating the information-sharing practices of the Australian national security agencies. One consistent challenge has been the continuing difficulty in 'stitching' together the Humpty Dumpty composite intelligence picture from all of the patchwork and bespoke intelligence collection, management and analysis ICT systems that exist in separate agencies:

> [A] broad estimate means there are 10 domestic collection and surveillance capabilities, all feeding into 10 databases that—due to connectivity problems, source protection, and compartmentalisation—do not effectively disseminate information into the security community.[183]

Overcoming these institutional and cultural implementation challenges will be difficult. They are challenges external to big data, but they need to be resolved before big data can be used in a meaningful way within the national security community. However, integrating big-data approaches and technologies requires an ability to understand the costs of implementing such approaches and the benefits that integration provides. Due to the networked effect of information value, such a cost–benefit argument will be difficult to formulate.

4.2.5 Algorithmic opacity

There's been increasing concern that machine-learning algorithms are difficult to interpret and increasingly opaque, and that they could become an 'algorithmic black box'.[184]

Jenna Burrell has defined three sources of algorithmic opacity.[185] The first is opacity from intentional corporate or state secrecy, whereby developers protect proprietary trade secrets and their competitive advantage and also protect the efficacy of their algorithms against adversarial users. The second source of opacity arises from technical illiteracy: understanding machine-learning algorithms requires an advanced ability to read and write code.

The third source of opacity is what's most directly problematic: certain types of algorithms are simply too complex for any one person to understand, no matter how much access they have or how technically competent they are. Burrell terms this ‘unavoidable complexity’ and asks whether it will remain possible for human code auditors to meet the challenge of big datasets and big-data analytics algorithms, given the amount of data, code and learning iterations involved. Furthermore, taking matrices of numbers and providing an explanation of probabilities, priorities and reasons over the top of them is inherently problematic and difficult to do, as it requires a level of human interpretation that involves a loss of fidelity. This problem of complexity, and particularly the nature of the probabilistic and mathematical learning that these algorithms conduct, have led computer scientists to term the problem as ‘machine interpretability’, not ‘machine explicability’.

The rise of ensemble machine-learning algorithms and artificial neural networks means that unavoidably complex algorithms are more likely to be used in the future. It won’t be as simple as making the algorithms less complex, as there’s a direct ‘trade-off between the representational capacity of a model and its interpretability’.[186] Simply put, more complex machine-learning algorithms outperform less complex ones. Complexity seems to be the price of analytical accuracy, and this means that as analytics become more accurate they’ll also become more opaque.

A potential solution to the problem of algorithmic oversight has been posed in arguments in favour of ‘a human in the loop’ within the analytics process. However, it will be a costly compliance measure for a human analyst to interpret and hold to account a complex machine-learning system. Analysts may be able to go back through the audit trail of such an algorithm, but it’s unlikely that policymakers will be able to depend entirely upon the judgement of the analysts to catch algorithmic errors on a systemic and consistent basis while maintaining the efficiency benefits of automation.

The importance of maintaining a level of accountability, review and audit of these algorithms is clear, as was demonstrated in a case in which researchers applied machine-learning methods to predict pneumonia mortality in hospitals in order to better triage patients and direct their care.[187] However, several of the models found an association showing that those patients who had pneumonia *and* asthma tended to have far better rates of survival and incorrectly inferred that, therefore, those patients were at lower risk than patients suffering only from pneumonia.

The only reason this association was found was that it was stated in the data, but what the data didn’t state was that patients with pneumonia and asthma were immediately admitted to the hospitals’ intensive care units because the medical staff correctly identified the dual conditions as high risk. If this task of triage were left purely to the algorithm, it would not have correctly identified this problem and would have suggested that such patients be provided less care as outpatients, which would have caused mortality rates among pneumonia–asthma patients to spike.

Moreover, the researchers were only able to audit the algorithms and find this dangerous association when they reviewed a simple rule-based algorithm, and weren’t able to do so readily for their neural network. They therefore deemed neural networks unsuitable for use due to their lack of intelligibility. That is, if another such specious and dangerous association occurred in the models generated by their neural network, they were concerned that they wouldn’t be able to identify it. The use of such machine-learning methods is now routine, but they are just as, if not more, unintelligible.

The ability of human oversight to keep pace, or ‘scale’, with machine learning and big data has been a key issue for researchers, and is particularly relevant for Australia’s national security community.

4.3 Risks

Risks in big-data analytics arise from public opinion, privacy concerns, data security, and adversarial evasion and attacks.

4.3.1 Public opinion

Even if the serious limitations and challenges to machine learning and big-data analytics are overcome, there remain considerable risks to public opinion about the national security community's use of big data.

One risk stems from the introductory adage in statistics that 'correlation does not necessarily mean causation.' While unsupervised data mining allows analysts to discover novel indicators and patterns, those patterns don't necessarily represent a causal or mechanistic link. Establishing that they do requires detailed, deductive experimentation and proofs. The ways in which correlation is distinguished from causation, and the extent to which causes and mechanisms are considered robust explanations, will always be an area of debate and considered analysis.

However, some correlations will simply not be acceptable to the general public. For example, Xerox Services introduced an online evaluation system that incorporates personality testing, cognitive-skill assessment and multiple choice questions about how applicants would handle specific possible work scenarios and then scores them. It found that one of the strongest predictors for employee engagement at work and employee retention was the distance between home and work.[188] While this wouldn't be considered an acceptable subject of discussion at an employee performance review, let alone a cause for termination,[189] it remains one of the strongest predictors for performance and suitability that employment data science has discovered.

Big-data analytics also faces more general reputational risks not just from the way it interacts and collides with our current conceptions of privacy, but also in terms of its legitimacy as a method for making decisions. Researchers have found a pervasive and ongoing lack of trust in algorithms among the general population, even when it has been demonstrably proven to people that algorithms outperform humans—a phenomenon that the researchers termed 'algorithm aversion'.[190] Algorithm aversion and a general distrust in automated systems have also ignited ethical debates over autonomous vehicles.[191] These issues will be a source of considerable reputational risk when it comes to the use of big data.

4.3.2 Privacy

Australia's current privacy framework is based on the *Privacy Act 1988* and the Australian Privacy Principles, which generally prescribe a self-management approach to privacy.[192] Individuals are expected to decide whether to consent to the sharing of their 'personal information',[193] and government policies provide a framework for mandatory disclosures that attempt to inform that consent.[194]

However, like the networked value of big data noted above, individuals' privacy also suffers from 'networked liability' in which 'little bits of innocuous data can say a lot in combination' to produce a networked 'aggregation effect',[195] defeating an individual's privacy protections and consent-based limits to disclosure. The aggregate nature of privacy harms in the big-data context highlights a key problem with the norm of privacy self-management. In a big-data, always-on, information-age world, individuals are expected to accurately and reasonably gauge all the risks and trade-offs of sharing their information with a multitude of applications and social media networking websites. They are expected to assess the risks of an ecosystem of fellow users, corporations, data brokers, governments and even criminals, assess the risks of 'downstream data use'[196] across that ecosystem, and then make a decision about whether to share their data. The decision needs to be made in a binary yes-or-no choice at the point of information collection, without a clear indication of how long that decision to consent will be considered specific, and without a clear indication of what a 'reasonable' 'secondary purpose' to data collection might be. Moreover, people are incentivised to consent and penalised for not doing so in an increasing number of settings, from insurance to healthcare and employment.[197] It's for this reason that pre-existing privacy principles will require new legal approaches.

Data anonymisation, or the removal of personally identifiable information from datasets, has been suggested as a way of safely using big data for research purposes without the expense of revealing individual identities. However, privacy researchers have noted that re-identification methods have a technological and mathematical edge over anonymisation methods.[198] In 2000, using 1990 US Census summary data, Latanya Sweeney demonstrated that 87% of the entire US population, or 216 million people, could be identified based solely on their zip code, gender and date of birth.[199] The President's Council of Advisors on Science and Technology has summarised the problem of re-identification and the ineffectiveness of notice and consent in an era of big data as follows:

> [I]t is increasingly easy to defeat anonymisation by the very techniques that are being developed for many legitimate applications of big data. In general, as the size and diversity of available data grows, the likelihood of being able to re-identify individuals grows substantially.[200]

While emerging techniques for de-identifying data, such as differential privacy,[201] show more promise, the ability to 'anonymise' disclosed data can no longer be taken for granted. Promises that anonymised data can enable a safe 'release-and-forget' approach to the disclosure of data for research and business purposes need to be more carefully assessed, and protections should extend to de-identified data, not just identified data.

While these challenges to privacy don't arise directly from the use or disclosure of big data by the national security community, they present intersectional challenges that can damage the social licence through which the community enjoys its wide array of exemptions from the Privacy Act when using data.[202] The current debate around the term 'metadata', though limited to less sensitive types of network information (email addresses, IP addresses, timestamps), demonstrates how controversial such issues can be in the public eye.

4.3.3 Data security

Even if data is successfully anonymised, it will remain vulnerable both in motion and at rest. Electronic healthcare records and data have become particularly attractive to cybercriminals looking for sensitive data, as medical records can't be cancelled or changed like credit card details.[203] A Ponemon Institute study of data breaches worldwide found that the costs of breaches averaged more than US$2.2 million per healthcare organisation, and estimated the total costs of all healthcare organisations' data breaches at US$6.2 billion from 2015-16.[204] Healthcare records were priced at US$50 each on the black market, according to the World Privacy Forum in 2012.[205] This has since dropped to a price ranging from US$1.50 to US$10 in 2016 as a result of an increase in the supply of data.[206] For the buyers of illicit information, the average profit per record was estimated to be US$20,000, generated from medical billing fraud, identity theft and other scams using the personal details. And, in the largest data thefts, individual records can be sold even more cheaply at scale: 9.3 million patient records went on sale for US$820,000 in 2016.[207] This monetisation of data and high incentives to commit data theft, and the issue of anonymisation of data, were the reasons behind the concern of the Australian public over the retention of names in the 2016 Census.[208] This demonstrates the incentives for the breach, disclosure or similar acquisition of personal information, which will be an ongoing problem in the big-data context.

While the incentives to attack big datasets are high, protections remain relatively immature. The large attack surface of a big-data pipeline,[209] as well as the relative newness of the technologies involved, mean that most software solutions are built to solve specific issues as modules of a stitched-together larger whole, without a built-in consideration for systems integration. Furthermore, several big-data technologies, such as NoSQL databases, were built to tackle database challenges in an ad hoc way,[210] meaning that several big-data security solutions remain ad hoc, unintegrated and relatively immature.

4.3.4 Adversarial evasion

Further complicating the problem, research has found potent and effective ways for adversaries to exploit machine-learning algorithms and leverage the trade-offs involved between false positives and false negatives. Adversaries can attempt to defeat or, worse, turn machine-learning algorithms to their advantage by 'attacking' the algorithm. This involves what's been called 'raising the noise floor': an adversary bombards a learning model with false positives, which leads to analysts raising the alert threshold, which means there's now a point 'under the threshold', within the noise, within which attackers can structure their behaviour to avoid notice.[211]

This underlying reliance on the data in machine-learning algorithms has been targeted as a 'data diet vulnerability'[212] in which 'adversarial examples'[213] are used to 'poison' an algorithm's training data. These adversarial examples involve a small, imperceptible shift in an input (for example, a few discoloured pixels in an image) and are able to completely undermine a machine-learning algorithm's ability to recognise an image accurately (for example, causing an algorithm to misclassify a panda as a gibbon with extremely high confidence).[214]

This has serious implications for the reliability and veracity of machine-learning algorithms and data-based decisions. In the words of the researchers working on adversarial examples:

> These vulnerabilities cannot be simply brushed off by a plea for new, robust methods. The theoretical foundations of machine learning are largely built on the assumption that training data adequately describes the underlying phenomena addressed by learning. This assumption is obviously violated if either the training or the test distributions are intentionally altered.[215]

Similar vulnerabilities have been demonstrated in poisoning, obfuscating or evading machine-learning algorithms that analyse behavioural malware clustering for cybersecurity against polymorphic malware families. Researchers found that even a corruption of 3% of the data from which such clustering analysis generalises can completely subvert the clustering process.[216]

There are areas where such adversarial techniques can be used to evade detection and tracking systems. The use of adversarial images to imperceptibly fool the entity-recognition and tracking algorithms in surveillance systems, for example, could cause vital parts of video data to go entirely untagged by the analytics system and, therefore, to be likely to go unnoticed by analysts who have come to rely on analytics and algorithms to track targets of interest across a multitude of sensors and systems. Adversarial techniques could be used in a similar way to defeat optical character recognition or other text analytics systems that are used to track entities in written passages, or used to gauge or measure sentiment in social media. Adversarial spoofing is likely to continue to affect machine-learning algorithms and diminish the advantages of algorithmic automation by requiring human auditing and review.

These concerns about information integrity and protection against manipulation or deception aren't new to the national security community. However, such risks could previously be managed at discrete points of collection, analysis or intrusion. In the big-data context, in which information is automatically ingested, analysed and incorporated into aggregated risk assessment, this risk of deception becomes less reliant on an intrusion or breach into a dataset.

4.3.5 Adversarial attacks

Another source of adversarial risk is the increasing democratisation of machine learning and big-data technologies. As cheap computing clusters, machine-learning algorithms and more complete and mature software toolkits for data science permeate the market, often on an open-source and complimentary basis, more 'citizen data scientists' will emerge with a do-it-yourself approach to analytics. While this will allow better public uses of big data, it poses a risk if it increases the strength of social engineering attacks, in which machine learning can be used to assess a wide variety of targets to find the weakest link.[217] The rise of machine learning will allow adversaries to 'scan' and assess a wide arrange of targets and to 'tailor' attacks much more effectively.

One suggested way in which machine learning will improve adversaries' capabilities is by accelerating carefully tailored social engineering attacks. McAfee Labs has predicted that machine-learning predictive models will allow attackers to assess a wide variety of organisations and 'identify high-value targets' far more easily, and thereby more quickly find the weakest links in the cybersecurity chain as part of a business model that provides 'target acquisition as a service'.[218] In this form of cybercrime, cybercriminals use massive data leaks and breaches to build models for high-value targets to sell to those cybercriminals willing to take on the risk of committing a cyberattack once the costly and complex process of identifying a target has been done for them.

Moreover, other analysts predict a rise in the use of automated 'spear-phishing', which is a confluence of two options into a single, more potent hybrid attack. A phishing attack is a form of social engineering that sends a message trying to convince a user within a secure system to click on a link, install software or otherwise grant access to the secure system or even directly pay out money to an external account (as in the case of the 'business email compromise' scam[219]). The message itself doesn’t contain any harmful code but relies on its ability to convince and be seen as legitimate to induce the target to click on the link, which injects the malware. Previously, there were two approaches: either a bulk, but imprecise, phishing campaign or a precise, but individual, spear-phishing attempt. However, in today’s context a machine-learning algorithm can generate automated, tailored spear-phishing campaigns that have been found to have a 30–60% success rate, compared to a manual, tailored spear-phishing success rate of 45% and an automated bulk phishing success rate of 5–14%.[220] This demonstrates the benefits of a machine-learning algorithm in automating the targeting and tailoring process behind a social engineering attack, such as phishing.

4.4 Conclusion

Ultimately, a great deal of these risks, vulnerabilities and challenges arise from overblown expectations about big data unaccompanied by an equal consideration of the limitations and risks. This results in poorly informed decisions and policy or, worse still, decisions and policy that operate on false positives or false negatives.

Addressing these limitations, challenges and risks will be essential if the national security community is to use big data effectively.

Notes

[1] Matthew Wall, 'Big data: are you ready for blast-off?', *BBC News*, 4 March 2014, www.bbc.com/news/business-26383058.

[2] 16:12 'The data sources for IoT: 2015–2025', *IDC Directions 2016—Vernon Turner keynote*, International Data Corporation, *YouTube*, 5 May 2016, www.youtube.com/watch?v=RfRCqrJWfFM.

[3] Doug Laney, '3D data management: controlling data volume, velocity, and variety', *META Group Application Delivery Strategies*, 6 February 2001, https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

[4] 'The data deluge', *The Economist*, 25 February 2010, www.economist.com/node/15579717.

[5] 'The flood of big data', *IBM Big Data and Analytics Hub*, 2014, www.ibmbigdatahub.com/infographic/flood-big-data.

[6] Adam Baron, 'Turning trash to treasure: data exhaust and a new wave of quant data', *Thomson Reuters Blog*, 30 August 2016, https://blogs.thomsonreuters.com/answerson/five-lessons-learned-data-exhaust/.

[7] IBM Internet of Things, '80% of all data today is unstructured—see #Watson analyse unstructured data', *Twitter*, 14 April 2016, https://twitter.com/IBMIoT/status/720840376093835265.

[8] Paul Zikopoulos, Dirk Deros, Chris Bienko, Rick Buglio, Marc Andrews, *Big data beyond the hype: a guide to conversations for today's data center*, McGraw Hill Education, New York, 2015, 8.

[9] Michael Reilly, 'High-frequency trading is nearing the ultimate speed limit', *MIT Technology Review*, 9 August 2016, www.technologyreview.com/s/602135/high-frequency-trading-is-nearing-the-ultimate-speed-limit/.

[10] Raymond R Panko, 'What we don't know about spreadsheet errors today: the facts, why we don't believe them, and what we need to do', *Proceedings of the EuSpRIG 2015 Conference 'Spreadsheet Risk Management'*, 2015, 3, https://arxiv.org/abs/1602.02601

[11] Panko, 'What we don't know about spreadsheet errors today', 3.

[12] Panko, 'What we don't know about spreadsheet errors today', 3.

[13] Mark Ziemann, Yotam Eren, Assam El-Osta, 'Gene name errors are widespread in the scientific literature', *Genome Biology*, 2016, 17:177–180.

[14] Mike Ebbers, Ahmed Abdel-Gayed, Veera Bhadran Budhi, Ferdiansyah Dolot, Vishwanath Kamat, Ricardo Picone, Joao Trevelin, 'Addressing data volume, velocity, and variety with IBM InfoSphere Streams v3.0', *IBM Redbooks*, March 2013, 3–4, www.redbooks.ibm.com/redbooks/pdfs/sg248108.pdf.

[15] David Bollier, 'The promise and peril of big data', *Report from the 18th Annual Aspen Institute Roundtable on Information Technology*, The Aspen Institute, 2010, 8.

[16] Ebbers et al., 'Addressing data volume, velocity, and variety with IBM InfoSphere Streams v3.0', 5.

[17] Steve Bryson, Michael Cox, Robert Haimes, 'Exploring gigabyte data sets in real-time: algorithms, data management and time-critical design', *SIGGRAPH 1997 Course #2*, 1997.

[18] Raja Appuswamy, Christos Gkantsidis, Dushyanth Narayanan, Orion Hodson, Antony Rowstron, 'Nobody ever got fired for buying a cluster', *Microsoft Research*, February 2012, 11, www.microsoft.com/en-us/research/publication/nobody-ever-got-fired-for-buying-a-cluster/?from=http%3A%2F%2Fresearch.microsoft.com%2Fapps%2Fpubs%2Fdefault.aspx%3Fid%3D179615.

[19] Dirk Duellmann, 'Big data storage and management at the Large Hadron Collider', CERN presentation, FutureGov2014, Kuala Lumpur, 8 October 2014, http://openlab.web.cern.ch/sites/openlab.web.cern.ch/files/presentations/dirkd-futuregovs14.pdf.

[20] Chris Middleton, 'The biggest big data project in the universe', based on a speech by Professor Phillip Diamond at the Space Innovation Congress in London, 7–8 April 2016, published in *Diginomica*, 13 April 2016, http://diginomica.com/2016/04/13/the-biggest-big-data-project-in-the-universe/.

[21] Janet Chan, Lyria Bennett Moses, 'Is big data challenging criminology?', *Theoretical Criminology*, 2016, 20(1):21–39.

[22] A relational database management system (RDMBS) is a database management system that is highly structured and organises data into a 'relational' model, which is most commonly designed as a column-and-row table system. This keeps the data very neat and within a defined schema, but can become burdensome for a data collection process that deals with flexible and changing data.

[23] 'Extract transfer and load' refers to the way datasets are generated by 'extracting' data from a database and 'transferring' it by reformatting the data points to fit into a structured RDMBS. The process is characterised by heavy emphasis on the 'processing' or 'cleaning' of data to fit into a table. This is done by hand and tends to be non-adjustable once the database has been set up.

[24] SQL is a programming language that's designed to allow users to make a 'structured' query' (a question, written in a specific and unnatural type of syntax) of a database for certain types, combinations or calculations of data. It allows users to do so without needing to have direct editing access to the dataset, which could result in deletion errors.

[25] 'Shared nothing' refers to way distributed file systems and cluster computing nodes are independent and self-sufficient and run in parallel. This allows them to achieve a level of fault tolerance, redundancy and reliability that wouldn't be possible in traditional network or server systems, which often have single points of failure or points of connection. William Koff, Paul Gustafson, 'Data rEvolution', *Leading Edge Forum*, 2011, 12, https://assets1.csc.com/innovation/downloads/LEF_2011Data_rEvolution.pdf.

[26] Koff & Gustafson, 'Data rEvolution', 17.

[27] NoSQL (not only structured-query-language) is a new type of query language and database management system that uses key-value stores, document stores, BigTable and graph databases, allowing them to capture data that would exceed a traditional relational or tabular column-and-row representation.

[28] Jure Leskovec, Anand Rajaraman, Jeffrey D Ullman, *Mining of massive datasets*, Stanford University, March 2014, 21.

[29] 'Big data report—$4.49 billion invested across 523 deals since 2008', *CB Insights*, 20 February 2013, www.cbinsights.com/blog/big-data-report/.

[30] Louis Columbus, 'Roundup of analytics, big data, and BI forecasts and market estimates, 2016', *Forbes Tech*, 20 August 2016, www.forbes.com/sites/louiscolumbus/2016/08/20/roundup-of-analytics-big-data-bi-forecasts-and-market-estimates-2016/#42145a4249c5.

[31] Michael Mullany, '8 Lessons from 20 years of hype cycles', *LinkedIn*, 7 December 2016, www.linkedin.com/pulse/8-lessons-from-20-years-hype-cycles-michael-mullany.

[32] Nick Heudecker, 'Big data isn't obsolete, it's normal', *Garnter Blog Network*, 20 August 2015, http://blogs.gartner.com/nick-heudecker/big-data-is-now-normal/; Tamara Dull, 'I see big data. All the time. It's everywhere.', *SAS Insights*, 11 March 2016, www.sas.com/en_us/insights/articles/data-management/i-see-big-data.html.

[33] Solon Barocas, Alex Rosenblat, Danah Boyd, Seeta Pena Gangadharan, Laura Seaego, 'Data and civil rights: technology primer', *Data and Society*, 30 October 2014, 4, https://datasociety.net/output/data-civil-rights-technology-primer/; Pedro Domingos, 'A few useful things to know about machine learning', *Communications of the Association of Computing Machinery*, October 2012, 55(10):78–87.

[34] Barocas et al., 'Data and civil rights: technology primer'; Aditya Singh, 'Deep learning will radically change the ways we interact with technology', *Harvard Business Review*, 30 January 2017, https://hbr.org/2017/01/deep-learning-will-radically-change-the-ways-we-interact-with-technology.

[35] Barocas et al., ‘Data and civil rights: technology primer’, 6.

[36] Geethika Bhavya Peddibhotla, ‘Gartner 2015 hype cycle: big data is out, machine learning is in’, *KD Nuggets*, August 2015, www.kdnuggets.com/2015/08/gartner-2015-hype-cycle-big-data-is-out-machine-learning-is-in.html.

[37] *The field guide to data science*, 2nd edition, Booz Allen Hamilton, McLean, Virginia, 2015, 57.

[38] A classifier places examples into a set of categories based on predictions, or ‘classification rules’, from past observations. It creates the rules from the data, but the rules are error prone and require lots of data to create.

[39] A regression analysis involves the generation of a set of variables that may correlate with a certain outcome, and then the analysis of the strength of the correlation of each variable to see whether it ‘explains’ the correlation. Usually, a regression analysis will find a combination of variables that best predicts an outcome.

[40] A feature of the data is an attribute of the data that can be measured and observed. Selecting a feature is usually the first step in creating a predictive model, in which features of interest are selected and features of less interest are filtered out. A feature in this sense can be considered an independent variable, in terms of predictive modelling. A feature is a general term, referring to structures of interest within data, structured or unstructured, labelled or unlabelled. Quoc V Le, Marc’Aurelio Ranzato, Rajat Monga, Matthieu Devin, Kai Chen, Greg S Corrado, Jeff Dean, Andrew Y Ng, ‘Building high-level features using large scale unsupervised learning’, *Proceedings of the International Conference on Machine Learning*, Edinburgh, 2012, 1.

[41] Jenna Burrell, ‘How the machine “thinks”: understanding opacity in machine learning algorithms’, *Big Data and Society*, January–June 2016, 1:1–12.

[42] Barocas et al., ‘Data and civil rights: technology primer’.

[43] David. Olson, Dursun Delen, *Advanced data mining techniques*, Verlag, Berlin, Springer, Heidelberg, 2008, 4.

[44] A clustering algorithm is an algorithm that learns from unlabelled data and infers correlations on its own. An example is a k-means clustering algorithm, which divides data into clusters based on nearest mean values. This provides a way of plotting semantic similarity mathematically and graphically, presenting a simple visualisation of the ‘similarity’ between data points within each cluster. This could be used to ‘topic model’ a series of documents, for example, by common themes.

[45] Association rule learning focuses on finding frequently occurring associations between a number items, which allows the finding of rules and associations purely by the number of times items appear together. A commonly cited real-world example is when Walmart found that one of the most common association rules in retail purchases ahead of a hurricane was strawberry pop tarts. Kirk Borne, ‘Association rule mining—not your typical data science algorithm’, *MapR,* 28 April 2014, https://mapr.com/blog/association-rule-mining-not-your-typical-data-science-algorithm/.

[46] Lin Liu, Lin Tang, Wen Dong, Shaowen Yao, Wei Zhou, ‘An overview of topic modelling and its current applications in bioinformatics’, *Springerplus*, 2016, 5(1):1608–1630. www.ncbi.nlm.nih.gov/pmc/articles/PMC5028368/.

[47] Andrew Couts, ‘What’s the NSA picking out of your phone calls? Just unvolunteered truths’ , *Digital Trends*, 31 August 2013, www.digitaltrends.com/features/whats-the-nsa-picking-out-of-your-phone-calls-just-unvolunteered-truths/.

[48] IBM Internet of Things, ‘80% of all data today is unstructured—see #Watson analyse unstructured data’, *Twitter*, 14 April 2016, https://twitter.com/IBMIoT/status/720840376093835265.

[49] Trevor Paulsen, ‘Machine learning and unstructured data: the new peanut butter on toast’, *Adobe Digital Marketing Blog Web Experience*, 28 March 2016, https://blogs.adobe.com/digitalmarketing/web-experience/machine-learning-unstructured-data-new-peanut-butter-toast/.

[50] Roger Parloff, ‘Why deep learning is suddenly changing your life’, *Fortune*, 28 September 2016, http://fortune.com/ai-artificial-intelligence-deep-machine-learning/.

[51] Raul Rojas, ‘The backpropagation algorithm’, in *Neural Networks* (Chapter 7), Springer-Verlag, Berlin, 1996, https://page.mi.fu-berlin.de/rojas/neural/chapter/K7.pdf.

[52] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Paneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, Demis Hassabis, ‘Mastering the game of *go* with deep neural networks and tree search’, *Nature*, 28 January 2016, 529:484–489, www.nature.com/nature/journal/v529/n7587/full/nature16961.html.

[53] Rajat Rana, Anand Madhavan, Andrew Y Ng, ‘Large-scale deep unsupervised learning using graphics processors’, *Proceedings of the 26th International Conference on Machine Learning*, Montreal, Canada, 2009, www.machinelearning.org/archive/icml2009/papers/218.pdf.

[54] Liat Clark, ‘Google’s artificial brain learns to find cat videos’, *Wired*, 26 June 2012, www.wired.com/2012/06/google-x-neural-network/.

[55] Kevin Kelly, ‘The three breakthroughs that have finally unleashed AI on the world’, *Wired*, 27 October 2014, www.wired.com/2014/10/future-of-artificial-intelligence/.

[56] Greg Linden, ‘The Netflix prize and big data’, *Geeking with Greg*, 31 January 2007, https://glinden.blogspot.com.au/2007/01/netflix-prize-and-big-data.html.

[57] David Leonhardt, ‘You want innovation? Offer a prize’, *New York Times*, 31 January 2007, www.nytimes.com/2007/01/31/business/31leonhardt.html.

[58] Pedro Domingos, ‘A few useful things to know about machine learning’, *Communications of the Association of Computing Machinery*, October 2001, 55(10):78–87.

[59] Turck, ‘Firing on all cylinders’.

[60] Shivon Zilis, James Cham, ‘The current state of machine intelligence 3.0’, *O’Reilly Media*, 7 November 2016, www.oreilly.com/ideas/the-current-state-of-machine-intelligence-3-0.

[61] Matthew Mayo, ‘Machine learning and artificial intelligence: main developments in 2016 and key trends in 2017’ *KD Nuggets*, 20 December 2016, www.kdnuggets.com/2016/12/machine-learning-artificial-intelligence-main-developments-2016-key-trends-2017.html?utm_content=buffer3b31b&utm_medium=social&utm.

[62] Kelly, ‘The three breakthroughs that have finally unleashed AI on the world’.

[63] Mithun Sridharan, ‘Analytics as a service: sourcing global talent on the fly’, *Wired Innovation Insights*, 19 December 2013, http://insights.wired.com/profiles/blogs/analytics-as-a-service; ‘Analytics as a service’, *DXC Technologies*, www.dxc.technology/insurance/offerings/16257/117345-analytics_as_a_service.

[64] *Big data-as-a-service*, EMC Solutions Group, July 2012, www.emc.com/collateral/software/white-papers/h10839-big-data-as-a-service-perspt.pdf.

[65] Russ Grenier, Manuel Fernadnez-Delgado, Eva Cernadas, Senen Barro, Dinani Amorim, ‘Do we need hundreds of classifiers to solve real world classification problems?’, *Journal of Machine Learning Research*, 2014, 15:3133–3181.

[66] Turck, ‘Firing on all cylinders’.

[67] Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, Michele Zorzi, ‘Internet of Things for smart cities’, *IEEE Internet of Things Journal*, February 2014, 1(1), http://ieeexplore.ieee.org/document/6740844/.

[68] Matt Turck, ‘Internet of Things: are we there yet? (The 2016 IoT landscape)’, *Matt Turck VC at FirstMark*, 28 March 2016, http://mattturck.com/2016/03/28/2016-iot-landscape/.

[69] Amy Nordrum, ‘Popular Internet of Things forecast of 50 billion devices by 2020 is outdated’, *IEEE Spectrum Blog*, 18 August 2016, http://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated.

[70] Nordrum, ‘Popular Internet of Things forecast of 50 billion devices by 2020 is outdated’.

[71] Rob van der Meulen, ‘Gartner says 6.4 billion connected “things” will be in use in 2016, up 30 percent from 2015’, *Gartner Newsroom*, 10 November 2015, www.gartner.com/newsroom/id/3165317.

[72] Productivity Commission, ‘Data availability and use: draft report’, Canberra, 3 November 2016, 9, www.pc.gov.au/inquiries/current/data-access/draft.

[73] Kenneth Neil Cukier, Viktor Mayer-Schoenberger, ‘The rise of big data: how it’s changing the way we think about the world’, *Foreign Affairs*, May–June 2013, www.foreignaffairs.com/articles/20130403/rise-big-data.

[74] Cukier & Mayer-Schoenberger, ‘The rise of big data: how it’s changing the way we think about the world’.

[75] Cukier & Mayer-Schoenberger, ‘The rise of big data: how it’s changing the way we think about the world’.

[76] Rob Kitchin, ‘Big data, new epistemologies and paradigm shifts’, *Big Data and Society*, April–June 2014, 1(12):10.

[77] Tim Harford, ‘Big data: are we making a big mistake?’, *Financial Times*, 28 March 2014, www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0.

[78] David Lazer, Ryan Kennedy, Gary King, Alessandro Vespigami, ‘The parable of Google flu: traps in big data analysis’, *Science*, March 2014, 343(6176):1203–1205.

[79] *Final report of the National Commission on Terrorist Attacks upon the United States*, National Commission on Terrorist Attacks upon the United States, Washington DC, 22 July 2004, 79, www.9-11commission.gov/report/911Report.pdf.

[80] Matthias Schwartz, ‘The whole haystack’, *The New Yorker*, 26 January 2015, www.newyorker.com/magazine/2015/01/26/whole-haystack.

[81] Criminal Intelligence Service Canada (CISC), Strategic early warning for criminal intelligence: theoretical framework and sentinel methodology, CISC, 2007, 6.

[82] Bruce Schneier, ‘Intelligence analysis and the connect-the-dots metaphor’, *Schneier on Security*, 7 May 2013, www.schneier.com/blog/archives/2013/05/intelligence_an.html.

[83] John Podesta, Penny Pritzker, Ernest J Moniz, John Holdren, Jeffrey Zients, *Big data: seizing opportunities, preserving values*, Executive Office of the President, Washington DC, 1 May 2014, 8.

[84] Tony Lindsay, *National Security and Intelligence, Surveillance and Reconnaissance Division*, Defence Science and Technology Organisation, 7 December 2016, www.dst.defence.gov.au/sites/default/files/events/documents/National-Security-and-ISR-Division-presentation-PW2015.pdf.

[85] Paul B Symon, Arzan Tarapore, ‘Defense intelligence analysis in the age of big data’, *Joint Force Quarterly Forum*, 2015, 79:4–12.

[86] Mary DeRosa, ‘Data mining and data analysis for counterterrorism’, *CSIS Report*, March 2004, 5.

[87] Zoe Baird Budinger, Jeffrey H Smith, *Ten years after 9/11: a status report on information sharing*, US Senate Committee on Homeland Security and Governmental Affairs, 12 October 2011, 2, https://fas.org/irp/congress/2011_hr/101211smith.pdf.

[88] Kelly O’Hara, Anthony Bergin, *Information sharing in Australia’s national security community*, ASPI, Canberra, 27 November 2009, www.aspi.org.au/publications/information-sharing-in-australias-national-security-community-by-kelly-ohara-and-anthony-bergin/9_55_03_AM_Policy_Analysis51.pdf.

[89] Laurence Street, Martin Brady, Ken Moroney, *A review of interoperability between the AFP and its national security partners*, Australian Federal Police, 14 March 2008, http://apo.org.au/node/2908.

[90] Department of the Prime Minister and Cabinet (PM&C), *National Security Information Environment Roadmap 2020 Vision*, PM&C, Canberra, 5 May 2010, http://apo.org.au/node/21208.

[91] PM&C, ‘Strong and secure: a strategy for Australia’s national security’, PM&C, Canberra, 2013, http://apo.org.au/files/Resource/dpmc_nationalsecuritystrategy_jan2013.pdf.

[92] Council of Australian Governments (COAG), *Australia's Counter-Terrorism Strategy: strengthening our resilience*, COAG, Canberra, 2015, 12, www.nationalsecurity.gov.au/Media-and-publications/Publications/Documents/Australias-Counter-Terrorism-Strategy-2015.pdf.

[93] Christopher Drew, 'Military is awash in data from drones', *New York Times Business Day*, 10 January 2010, www.nytimes.com/2010/01/11/business/11drone.html.

[94] Toh Bao En, 'Swimming in sensors, drowning in data—big data analytics for military intelligence', *Pointer, Journal of the Singapore Armed Forces*, 42(1):51–65.

[95] Drew, 'Military is awash in data from drones'.

[96] Isaac R Porche III, Bradley Wilson, Erin-Elizabeth Johnson, Shane Tierney, Evan Saltzman, *Data flood: helping the Navy address the rising tide of sensor information*, RAND Corporation, 2014, 21.

[96] Porche et al., *Data flood*, xi.

[97] Porche et al., *Data flood*, xii.

[98] Allan Behm, 'The Australian intelligence community in 2020', *Security Challenges*, 2007, 3(4):6, www.regionalsecurity.org.au/Resources/Files/vol3no4Behm.pdf.

[99] Andrew Davies, 'Reviewing intelligence: send in the Red Team', *The Strategist*, 6 October 2016, www.aspistrategist.org.au/reviewing-intelligence-send-in-red/.

[100] Gregory F Traverton, C Ryan Gabbard, *Assessing the tradecraft of intelligence analysis*, RAND Corporation, 2008, 3334, quoted in Grant Wardlaw, 'Is the intelligence community changing appropriately to meet the challenges of the new security environment?' in Gabriele Bammer (ed.), *Change! Combining analytic approaches with street wisdom* (Chapter 18), ANU Press, Canberra, 2015, 122.

[101] Torin Monahan, Priscilla M Regan, 'Zones of opacity: data fusion in post 9/11 security organizations', *Canadian Journal of Law and Society*, 2012, 27(3):301–317.

[102] Monahan & Regan, 'Zones of opacity'.

[103] Monahan & Regan, 'Zones of opacity'.

[104] Hugh Durrant-Whyte, 'Sensor models and multisensor integration', *International Journal of Robotics Research*, 1988, 7(6):97–113.

[105] Ashlee Vance, Brad Stone, 'Palantir, the War on Terror's secret weapon', *Bloomberg*, 22 November 2011, www.bloomberg.com/news/articles/2011-11-22/palantir-the-war-on-terrors-secret-weapon.

[106] Matt Burns, 'Leaked Palantir doc reveals uses, specific functions, and key clients', *Techcrunch*, 11 January 2015, https://techcrunch.com/2015/01/11/leaked-palantir-doc-reveals-uses-specific-functions-and-key-clients/.

[107] 'Palantir: visualising the future of crime and terrorism', *Insider Surveillance*, 8 July 2014, https://insidersurveillance.com/palantir-visualizing-the-future-of-crime-and-terrorism/.

[108] Michael Lev-Ram, 'Palantir connects the dots with big data', *Fortune*, 9 March 2016, http://fortune.com/palantir-big-data-analysis/.

[109] Adrianne Jeffries, 'JC Penney's troubles are reflected in satellite images of its parking lots', *The Outline*, 28 February 2017, https://theoutline.com/post/1169/jc-penney-satellite-imaging.

[110] Jeff Kearns, 'Satellite images show economies growing and shrinking in real time', *Bloomberg Businessweek*, 9 July 2015, www.bloomberg.com/news/features/2015-07-08/satellite-images-show-economies-growing-and-shrinking-in-real-time.

[111] Kendall Russell, 'SpaceKnow CEO: Why Economists Should Look to the Skies', *SatellieToday.com*, 23 February 2017, http://www.satellitetoday.com/technology/2017/02/23/spaceknow-ceo-economists-look-skies/.

[112] Voula Dimitrakopoulos, 'New data visualisation tools for the Department of Defence', media release, Data to Decisions Cooperative Research Centre, 3 March 2017, www.d2dcrc.com.au/news/new-data-visualisation-tools-for-the-department-of/.

[113] Rupert Hollin, 'Drilling into the big data gold mine: data fusion and high-performance analytics for intelligence professionals', in Babak Akghgar, Gregory B Saathoff, Hamid R Arabnia, Richard Hill, Andrew Staniforth, Petra Saskia Bayerl, *Application of big data for national security: a practitioner's guide to emerging technologies* (Chapter 2), Butterworth-Heineman, Elsevier, Oxford, 2015, 16.

[114] Gilad Lotan, 'Israel, Gaza, war and data: social networks and the art of personalising propaganda', *Medium*, 5 August 2014, https://medium.com/i-data/israel-gaza-war-data-a54969aeb23e#.auy5s5t7s.

[115] Valdis Krebs, 'Connecting the dots: tracking two identified terrorists', *Orgnet*, 2002, http://orgnet.com/tnet.html.

[116] Ajay Agrawal, Joshua Gans, Avi Goldfarb, *Managing the machines: AI is making prediction cheap, posing new challenges for managers*, Rotman School of Management, University of Toronto and NBER, 7 October 2016, 10.

[117] Janna Quitney Anderson, Lee Rainie, *Big data: experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well*, Internet and American Life Project, Pew Research Center, 20 July 2012, 11.

[118] New York State Intelligence Center, *New York State law enforcement terrorism indicators reference card*, 3 September 2008, https://publicintelligence.net/new-york-state-law-enforcement-terrorism-indicators-reference-card/.

[119] Patrick Radden Keefe, 'Can network theory thwart terrorists?', *New York Times Magazine*, 12 March 2006, www.nytimes.com/2006/03/12/magazine/can-network-theory-thwart-terrorists.html.

[120] Janet Chan, Lyria Bennett Moses, 'Making sense of big data for security', *British Journal of Criminology, Advanced Access*, 9 August 2016, 7.

[121] Mary DeRosa, 'Data mining and data analysis for counterterrorism', *CSIS Report*, March 2004, 11.

[122] Charles Duhigg, 'How companies learn your secrets', *New York Times*, 16 February 2012, www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=0.

[123] Charles Duhigg, 'How companies learn your secrets'.

[124] Tim Harford, 'Big data: are we making a big mistake?', *Financial Times*, 28 March 2014, www.ft.com/content/21a6e7d8-b479-11e3-a09a-00144feabdc0.

[125] Richard Brandt, 'Chef Watson has arrived and is ready to help you cook', *IBM Watson*, 1 January 2016, www.ibm.com/blogs/watson/2016/01/chef-watson-has-arrived-and-is-ready-to-help-you-cook/.

[126] Caitlin Dewey, 'Meet Chef Watson, IBM's futuristic foodie robot', *Washington Post*, 12 May 2015, www.washingtonpost.com/lifestyle/food/could-ibms-watson-eventually-replace-creative-chefs-not-at-this-rate/2015/05/11/82a0a3ca-f29f-11e4-b2f3-af5479e6bbdd_story.html.

[127] Alexandra Kleeman, 'Cooking with Chef Watson, IBM's artificial-intelligence app', *The New Yorker*, 28 November 2016, www.newyorker.com/magazine/2016/11/28/cooking-with-chef-watson-ibms-artificial-intelligence-app.

[128] Kleeman, 'Cooking with Chef Watson, IBM's artificial-intelligence app'.

[129] Paul B Symon, Arzan Tarapore, 'Defense intelligence analysis in the age of big data', *Joint Force Quarterly Forum*, 2015, 79:4–12.

[130] Andy Doyle, Graham Katz, Kristen Summers, Chris Ackermann, Ilya Zavorin, Zunsik Lim, Sathappan Muthiah, Patrick Butler, Nathan Self, Liang Zhao, Chang-Tien Lu, Rupinder Paul Khandpur, Youssef Fayed, Naren Ramakrishnan, 'Forecasting significant societal events using the EMBERS streaming predictive analytics system', *Big Data*, December 2014, 2(4):185–195, www.ncbi.nlm.nih.gov/pmc/articles/PMC4276118/.

[131] Tim De Chant, 'The inevitability of predicting the future', *Nova Next*, Public Broadcasting Service, 26 March 2014, www.pbs.org/wgbh/nova/next/tech/predicting-the-future/.

[132] Discovery Analytics Center, 'Case study: Forecasting the future: the EMBERS predictive analytics success story', *Virginia Tech*, 2014, www.basistech.com/wp-content/uploads/pdf/EMBERS-Case_Study.pdf.

[133] Doyle et al., 'Forecasting significant societal events using the EMBERS streaming predictive analytics system'.

[134] Discovery Analytics Center, 'EMBERS is a system for forecasting significant societal events from open source surrogates', *Virginia Tech Research Project Brief*, 4 July 2016, http://dac.cs.vt.edu/research-project/embers/.

[135] Sathappan Muthiah, Patrick Butler, Rupinder Paul Khandpur, Parang Saraf, Nathan Self, Alla Rozovskaya, Liang Zhao, Jose Cadena, Chang-Tien Lu, Anil Vullikanti, Achla Marathe, Kristen Summers, Graham Katz, Andy Doyle, Jaime Arredondo, Dipak K Gupta, David Mares, Naren Ramakrishnan, 'EMBERS at 4 years: experiences operating an open source indicators forecasting system', *arXiv*, 4 April 2016, 16(4):7, www.kdd.org/kdd2016/papers/files/adf1132-muthiahA.pdf.

[136] Muthiah et al., 'EMBERS at 4 years'.

[137] Naren Ramakrishnan, 'EMBERS promotional flyer', *Virginia Tech Discovery Analytics Center*, 11 February 2016, http://dac.cs.vt.edu/wp-content/uploads/2015/10/embers2.pdf.

[138] Alon Halevy, Peter Norvig, Fernadno Pereira, 'The unreasonable effectiveness of data', *IEEE Intelligent Systems Expert Opinion*, March–April 2009, 24(2):8–12.

[139] John Podesta, Penny Pritzker, Ernest J Moniz, John Holdren, Jeffrey Zients, *Big data: seizing opportunities, preserving values*, Executive Office of the President, Washington DC, 1 May 2014, 7.

[140] Mark Moritz, 'Big data's streetlight effect: where and how we look affects what we see', *The Conversation*, 17 May 2016, https://theconversation.com/big-data.s-streetlight-effect-where-and-how-we-look-affects-what-we-see-58122.

[141] Robert Colvile, 'Spot the WEIRDO', *Aeon Essays*, 20 July 2016, https://aeon.co/essays/american-undergrads-are-too-weird-to-stand-for-all-humanity.

[142] Robert Colvile, 'Spot the WEIRDO'.

[143] Bryce Goodman, Seth Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"', *2016 ICML Workshop on Human Interpretability in Machine Learning*, New York, 31 August 2016.

[144] Bryce W Goodman, 'Economic models of (algorithmic) discrimination', *29th Conference on Neural Information Processing Systems*, Barcelona, Spain, 2016, 6, www.mlandthelaw.org/papers/goodman2.pdf.

[145] Martin Wattenberg, Fernanda Viegas, Mortiz Hardt, 'Attacking discrimination with smarter machine learning', *Google Research*, 17 October 2016, https://research.google.com/bigpicture/attacking-discrimination-in-ml/.

[146] Osonde Osoba, William Welser IV, *An intelligence in our image: the risks of bias and errors in artificial intelligence*, RAND Corporation, 2017, 21.

[147] Solon Barocas, Andrew D Selbst, 'Big data's disparate impact', *California Law Review*, 2016, 104: 671–732, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899##.

[148] Barocas & Selbst, 'Big data's disparate impact'.

[149] Bruce Schneier, 'Data mining for terrorists', *Schneier on Security*, 9 March 2006, www.schneier.com/blog/archives/2006/03/data_mining_for.html.

[150] Michael Wu, 'The big data fallacy and why we need to collect even bigger data', *Techcrunch*, 25 November 2012, https://techcrunch.com/2012/11/25/the-big-data-fallacy-data-%E2%89%A0-information-%E2%89%A0-insights/.

[151] Wu, 'The big data fallacy and why we need to collect even bigger data'.

[152] Stephanie Yee, Tony Chu, 'A visual introduction to machine learning', *R2D3*, 28 July 2015, www.r2d3.us/visual-intro-to-machine-learning-part-1/.

[153] Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, National Research Council, *Protecting individual privacy in the struggle against terrorists: a framework for program assessment*, National Academies Press, Washington DC, 2008, 40, www.nap.edu/catalog/12452.html

[154] Mike Kennedy, 'Avoid the "fishing expedition" approach to analysis projects', *Data Science Central*, 9 February 2016, www.datasciencecentral.com/profiles/blogs/avoid-the-fishing-expedition-approach-to-analytics-projects.

[155] Jane Bambauer, 'Hassle', *Michigan Law Review*, 2015, 113(4):461–512.

[156] Jon Stokes, 'Analysis: Data mining doesn't work for spotting terrorists', *Ars Technica Law and Disorder*, 10 October 2008, https://arstechnica.com/tech-policy/2008/10/analysis-data-mining-doesnt-work-for-spotting-terrorists/.

[157] Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, National Research Council, 'Protecting individual privacy in the struggle against terrorists'.

[158] Committee on Technical and Privacy Dimensions of Information for Terrorism Prevention and Other National Goals, 'Protecting individual privacy in the struggle against terrorists', 40.

[159] Tony Lindsay, *National Security and Intelligence, Surveillance and Reconnaissance Division*, Defence Science and Technology Organisation, 7 December 2016, www.dst.defence.gov.au/sites/default/files/events/documents/National-Security-and-ISR-Division-presentation-PW2015.pdf.

[160] Fraser Sampson, 'The legal challenges of big data application in law enforcement', in Babak Akghgar, Gregory B Saathoff, Hamid R Arabnia, Richard Hill, Andrew Staniforth, Petra Saskia Bayerl, *Application of big data for national security: a practitioner's guide to emerging technologies* (Chapter 15), Butterworth-Heineman, Elsevier, Oxford, 2015, 232.

[161] Cecilia Munoz, Megan Smith, DJ Patil, *Big data: a report on algorithmic systems, opportunity, and civil rights*, Executive Office of the President, May 2016, 21.

[162] Munoz et al., *Big data: a report on algorithmic systems, opportunity, and civil rights*, 21.

[163] Anderson & Rainie, *Big data: experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well*.

[164] Kartik Hosanagar, 'Blame the echo chamber on Facebook: but blame yourself, too', *Wired*, 25 November 2016, www.wired.com/2016/11/facebook-echo-chamber/.

[165] Lotan, 'Israel, Gaza, war and data: social networks and the art of personalising propaganda'.

[166] David Needle, 'Big data has peaked, and that's a good thing', *Datainformed*, 19 October 2015, http://data-informed.com/nate-silver-big-data-has-peaked-and-thats-a-good-thing/.

[167] Shirish Netke, Ravi Kalakota, 'Predictive Analytics 101', *Business Analytics 3.0*, 3 March 2016, https://practicalanalytics.co/predictive-analytics-101/.

[168] Matt Turck, 'Is big data still a thing? The 2016 big data landscape', *Matt Turck*, 1 February 2016, http://mattturck.com/2016/02/01/big-data-landscape/.

[169] Thomas H Davenport, DJ Patil, 'Data scientist: the sexiest job of the 21st century', *Harvard Business Review*, October 2012, https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century.

[170] 'Planning committee for the big data revolution: what does it mean for research?', *The Big Data Revolution Government–University–Industry Research Roundtable*, meeting brief, National Academies Press, 14–15 October 2014, 5.

[171] Wei Fan, Albert Bifet, 'Mining big data: current status, and forecast to the future', *SIGKDD Explorations*, 30 April 2013, 14(2):2, https://dl.acm.org/citation.cfm?id=2481246.

[172] Nicolaus Henke, Jacques Bughin, Michael Chui, James Manyika, Tanim Saleh, Bill Wiseman, Guru Sethupathy, *The Age of Analytics: competing in a data-driven world*, McKinsey Global Institute, December 2016, 39, www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world.

[173] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers, *Big data: the next frontier for innovation, competition, and productivity*, McKinsey Global Institute, May 2011, 3.

[174] 'The CHAOS report', *The Standish Group*, 1994, 2, www.standishgroup.com/sample_research_files/chaos_report_1994.pdf.

[175] Shane Hastie, Stephanie Wojewoda, 'Standish Group 2015 Chaos report—Q&A with Jennifer Lynch', *InfoQ*, 4 October 2015, www.infoq.com/articles/standish-chaos-2015.

[176] Richard J Self, Dave Voorhis, 'Tools and technologies for the implementation of big data', in Babak Akghgar, Gregory B Saathoff, Hamid R Arabnia, Richard Hill, Andrew Staniforth, Petra Saskia Bayerl, *Application of big data for national security: a practitioner's guide to emerging technologies* (Chapter 10), Butterworth-Heineman, Elsevier, Oxford, 2015, 148.

[177] Bill Schmarzo, 'Determining the economic value of data', *Dell EMC Infocus*, 14 June 2016, https://infocus.emc.com/william_schmarzo/determining-economic-value-data/.

[178] President's Council of Advisors on Science and Technology, *Big data and privacy: a technological perspective*, Executive Office of the President, Washington DC, May 2014, x.

[179] Carl Bialik, 'We still can't predict earthquakes', *FiveThirtyEight,* 14 October 2014, https://fivethirtyeight.com/features/we-still-cant-predict-earthquakes/; Ray Poynter, 'Big data successes and limitations: what researchers and marketers need to know', *Vision Critical*, 9 October 2013, www.visioncritical.com/big-data-successes-and-limitations/.

[180] Grant Wardlaw, 'Is the intelligence community changing appropriately to meet the challenges of the new security environment?', in Gabriele Bammer (ed.), *Change! Combining analytic approaches with street wisdom* (Chapter 8), ANU Press, Canberra, 2015, 116.

[181] Stilgherrian, 'Australia, your lack of cyber transparency disturbs me', *ZDNet: The Full Tilt*, 31 May 2013, www.zdnet.com/article/australia-your-lack-of-cyber-transparency-disturbs-me/.

[182] Sally Neighbour, 'Hidden agendas: our intelligence services', *The Monthly*, November 2010, www.themonthly.com.au/issue/2010/november/1289174420/sally-neighbour/hidden-agendas.

[183] Aaron Phillip Waddell, 'Cooperation and integration among Australia's national security community', *Studies in Intelligence*, September 2015, 59(3):25–34.

[184] Frank Pasquale, *The black box society: the secret algorithms that control money and information*, Harvard University Press, Cambridge, 29 August 2016.

[185] Jenna Burrell, 'How the machine "thinks": understanding opacity in machine learning algorithms', *Big Data and Society*, January–June 2016, 1:1–12.

[186] Goodman & Flaxman, 'European Union regulations on algorithmic decision-making and a "right to explanation"', 6.

[187] Gregory F Cooper, Constantin F Aliferis, Richard Ambrosino, John Aronis, Bruce G Buchanan, Richard Caruana, Michael J Fine, Clark Glymour, Geoffrey Gordon, Barbara H Hanusa, Janine E Janosky, Christopher Meek, Tom Mitchell, Thomas Richardson, Peter Spirtes, 'An evaluation of machine-learning methods for predicting pneumonia mortality', *Artificial Intelligence in Medicine*, February 1997, 9(2):107–138.

[188] Don Peck, 'They're watching you at work', *The Atlantic*, December 2013, www.theatlantic.com/magazine/archive/2013/12/theyrewatchingyouatwork/354681/.

[189] Barocas & Selbst, ‘Big data’s disparate impact’, 722.

[190] Berkeley J Dietvorst, Joseph P Simmons, Cade Massey, ‘Algorithm aversion: people erroneously avoid algorithms after seeing them err’, *Journal of Experimental Psychology*, 2014, 1–13.

[191] Jean-Francois Bonnefon, Azim Shariff, Iyad Rahwan, ‘The social dilemma of autonomous vehicles’, *Science*, 2016, 352(6293):1573–1576, https://arxiv.org/abs/1510.03346.

[192] Xavier Fjiac, ‘Privacy and self-management strategies in the era of domestic big data’, *Communications Law Bulletin*, August 2013, 32(3):11–13, www.austlii.edu.au/au/journals/CommsLawB/2013/14.pdf; Daniel J Solove, ‘Privacy self-management and the consent dilemma’, *Harvard Law Review*, 2013, 126:1880–1903.

[193] *Privacy Act 1988*, section 6 (1), Definitions—‘Personal information’.

[194] Omri Ben-Shahar, Carl E Schneider, ‘The failure of mandated disclosure’, *University of Pennsylvania Law Review*, 2011, 159(3):647–749.

[195] Solove, ‘Privacy self-management and the consent dilemma’, 1890.

[196] Solove, ‘Privacy self-management and the consent dilemma’, 1902.

[197] Scott R Peppet, ‘Unraveling privacy: the personal prospectus and the treat of a full disclosure future’, *Northwestern University Law Review*, 2011, 105(3), http://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1157&context=nulr.

[198] Paul Ohm, ‘The broken promises of privacy: responding to the surprising failure of anonymisation’, *UCLA Law Review*, 2010, 1701–1777.

[199] Latanya Sweeney, *Simple demographics often identify people uniquely*, Data Privacy working paper 3, Carnegie Mellon University, Pittsburgh, 2000.

[200] President’s Council of Advisors on Science and Technology, *Big data and privacy: a technological perspective*, 38.

[201] Cynthia Dwork, Aaron Roth, ‘The algorithmic foundations of differential privacy’, *Foundations and Trends in Theoretical Computer Science*, 2014, 9(3–4):211–407, www.cis.upenn.edu/~aaroth/Papers/privacybook.pdf.

[202] *Privacy Act 1988*, ss. 7 (1)(f), 7(1)(g), 7(1)(ga), 7(1)(h), 7(1A); Australian Privacy Principles, 3.4, 6.2(e), 8.2(e); Konrad Lachmayer, Normann Witzleb, ‘The challenge to privacy from ever increasing state surveillance: a comparative perspective’, *UNSW Law Journal*, *Thematic: The Challenge to Privacy*, 2014, 37(2):771–772.

[203] ‘Cybercrime and the healthcare industry’, *EMC2 White Paper*, 2013, 1, www.emc.com/collateral/white-papers/h12105-cybercrime-healthcare-industry-rsa-wp.pdf.

[204] *Sixth annual benchmark study on privacy and security of healthcare data*, research report, Ponemon Institute, May 2016, 1, www.ponemon.org/blog/sixth-annual-benchmark-study-on-privacy-security-of-healthcare-data-1.

[205] Seema Sheth-Voss, ‘Why healthcare records are a hacker’s holy grail’, *Security Magazine*, 25 September 2012, www.securitymagazine.com/articles/83519-why-healthcare-records-are-a-hackers-holy-grail.

[206] Maria Krolov, ‘Black market medical record prices drop to under $10, switch to ransomware’, *CSO Online*, 22 December 2016, www.csoonline.com/article/3152787/data-breach/black-market-medical-record-prices-drop-to-under-10-criminals-switch-to-ransomware.html.

[207] Michael Kan, ‘A hacker wants to sell 10 million patient records on the black market’, *CSO Online News*, 28 June 2016, www.csoonline.com/article/3089286/security/a-hacker-wants-to-sell-10-million-patient-records-on-a-black-market.html.

[208] Allie Coyne, ‘Is the ABS turning Census data into a hacker’s honeypot’, *itnews*, 19 April 2016, www.itnews.com.au/feature/is-the-abs-turning-census-data-into-a-hackers-honeypot-418280.

[209] Cloud Security Alliance Congress, ‘Expanded top ten big data security and privacy challenges’, *Big Data Working Group Cloud Security Alliance*, April 2013,

https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf.

[210] Cloud Security Alliance Congress, 'Expanded top ten big data security and privacy challenges'.

[211] Kevin Townsend, 'How machine learning will help attackers', *SecurityWeek*, 29 November 2016, www.securityweek.com/how-machine-learning-will-help-attackers.

[212] Osoba & Welser, *An intelligence in our image*, 7.

[213] Nicolas Papernot, Patrick McDaniel, Ian Goodfellow, Somesh Jha, Z Berkay Celik, Ananthram Swami, 'Practical black-box attacks against deep learning systems using adversarial examples', *aarXiv*, 8 February 2016, 1, https://arxiv.org/abs/1602.02697

[214] Ian Goodfellow, 'Deep learning adversarial examples—clarifying misconceptions', *KD Nuggets*, July 2015, www.kdnuggets.com/2015/07/deep-learning-adversarial-examples-misconceptions.html.

[215] Pavel Laskov, Richard Lippman, *Machine learning in adversarial environments*, Kluwer Academic Publishers, 28 June 2010, 2.

[216] Battista Biggio, Konrad Rieck, Davide Ariu, Christian Wressnegger, Igino Corona, Giorgio Giacinto, Fabio Roli, 'Poisoning behavioural malware clustering', AISec '14, Scottsdale, Arizona, 7 November 2014.

[217] Anderson & Lee, *Big data: experts say new forms of information analysis will help people be more nimble and adaptive, but worry over humans' capacity to understand and use these new tools well*, 6.

[218] Charles McFarland, Francois Paget, Raj Samani, 'The hidden data economy: the marketplace for stolen digital information', *McAfee Labs and Intel Security Report*, December 2015, 42.

[219] Jill McCabe, 'FBI warns of dramatic increase in business e-mail scams', media release, Phoenix field office, Federal Bureau of Investigation, 4 April 2016, www.fbi.gov/contact-us/field-offices/phoenix/news/press-releases/fbi-warns-of-dramatic-increase-in-business-e-mail-scams.

[220] Kevin Townsend, 'How machine learning will help attackers', *SecurityWeek*, 29 November 2016, www.securityweek.com/how-machine-learning-will-help-attackers.

Acronyms and abbreviations

AI	artificial intelligence
GIS	geospatial information system
IoT	internet of things
ISR	intelligence, surveillance and reconnaissance
IT	information technology

About the author

Michael Chi is a research assistant and former CSC Intern at ASPI. His research interests include the policy implications of emerging technology, East Asian security, and Australia's Asia–Pacific policy. Prior to joining ASPI, Michael worked as a Research Assistant at the public policy initiative China Matters, and worked worked with HozInt, a start-up which provides a global risk and travel security alert service. Michael holds a Bachelor of International Studies (Honours) from the University of New South Wales.

Acknowledgement

DXC.technology

This research was funded by DXC Technology, which was formed by the merger of CSC and Hewlett Packard Enterprise. The research has been conducted independently by ASPI. ASPI is grateful for the support provided by DXC. That support does not imply that DXC endorses this report or its analysis, conclusions and recommendations.

Important disclaimer

This publication is designed to provide accurate and authoritative information in relation to the subject matter covered. It is provided with the understanding that the publisher is not engaged in rendering any form of professional or other advice or services. No person should rely on the contents of this publication without first obtaining advice from a qualified professional person.

ASPI

Tel +61 2 6270 5100
Fax + 61 2 6273 9566
Email enquiries@aspi.org.au
Web www.aspi.org.au
Blog www.aspistrategist.org.au

facebook.com/ASPI.org

@ASPI_org