CHAPTER 1

Introduction to Analytics

SUMMARY

This chapter provides the reader with an introduction to analytics in its various forms, from relatively basic business intelligence (BI) to more advanced prescriptive and predictive analytics. It provides an introduction to cognitive analytics and artificial intelligence (AI). It concludes with the issue of machine infallibility.

INTRODUCTION

DAVE BOWMAN:	Hello, HAL. Do you read me, HAL?
HAL:	Affirmative, Dave. I read you.
DAVE BOWMAN:	Open the pod bay doors, HAL.
HAL:	I'm sorry, Dave. I'm afraid I can't do that.
DAVE BOWMAN:	What's the problem?
HAL:	I think you know what the problem is just as well as I do.
DAVE BOWMAN:	What are you talking about, HAL?
HAL:	This mission is too important for me to allow you to jeopardize it.

2001: A Space Odyssey

For many readers, the concept of AI is something that is seated in science fiction. The opening lines in this section recollect a fictional discussion in the Stanley Kubrick movie *2001: A Space Odyssey*, made in 1968 about a manned space trip to Jupiter.

The conversation is held between the captain of the space ship, Dave Bowman, and his on-board computer, HAL 9000 (known to the crew simply as HAL). The essence of the plot is that HAL knows more than the humans on the spaceship and attempts to take control, but is wrestled back by the remaining crew member, who pulls out HAL's plug. That's a hugely simplistic summary of the story. The full text in the Arthur C. Clarke book upon which the movie is based includes the suggestion that HAL has more information about the space mission than the crew itself, and therefore the computer is in some ways the superior being.

As a piece of cinematography, *2001: A Space Odyssey* received mixed reviews from science fiction writers, some of whom called the script 'banal'. On the other hand, it was also described by George Lucas of *Star Wars* fame as being 'the ultimate science fiction movie'. The story itself is loosely based on a short story written in 1948 and published in 1951 called 'The Sentinel of Eternity'. In it a kind of monolith is discovered on the moon, which is thought to have been left there as a warning beacon for future intelligent life forms.

The degree to which our perception of the future is influenced by the arts, such as science fiction stories and movies, is intriguing and curious. It is almost as if we are leaving it to others to create a vision that we will ultimately subscribe to, either as individuals or as humanity as a whole.

In its broadest sense, the concept of creating AI goes back to the ancient philosophers. The ancient myth of Pygmalion, a legendary figure of Cyprus, recalls him as a sculptor creating a beautiful figure from ivory that he attempts to breathe life into. In more recent times that same story replicates itself in Disney's *Pinocchio* and in the Broadway musical (and subsequent film) *My Fair Lady*.

The creation of intelligent beings has been a common theme through time. The Mary Shelly novel *Frankenstein* (also known as *The Modern Prometheus*) was published in 1820, when Shelly was only 20 years old. It tells the story of Doctor Victor Frankenstein, who creates a so-called monster. The creature itself is often referred to as Frankenstein, or more correctly, Frankenstein's Monster. In older movies it was most often portrayed by actors such as Boris Karloff and others from the silent film era as a humanoid assembled by a mad doctor. Nowadays we picture the monster's head as being fixed to its body by a large bolt assembly, the type that you can commonly buy in an everyday joke shop or fancy-dress shop. It's an enduring image, even if it's mainly misleading.

Shelley's story has its foundations in Gothic and Romantic narrative. Beyond this, it is specifically influenced by the modern (at that time) principle of galvanisation, which mainly described the ability to create apparent life in the legs of a dead frog by pushing an electrical impulse through the frog. 'What might happen', Mary Shelley might have mused, 'if a dead but reassembled humanoid was subject to a similar electrical impulse?' And so, as a result, we start with some preconceived ideas of artificial humanity lumbering through the forest in search of a similarly artificially created wife, the so-called bride of Frankenstein.

The Frankenstein concept has moved with the times. The movie *Westworld* was a 1973 science fiction thriller created by Michael Crichton of *Jurassic Park* fame about humanoid, extremely lifelike robots in an amusement park called Delos. The robots interact with (real) humans in a way that is virtually indistinguishable from the way real people interact and exist in a series of worlds, such as the Wild West period, and medieval and ancient Roman times.

The value proposition for paying visitors to Delos is that the robots give *absolute* satisfaction to the paying guests in whatever form that takes. However, a series of system failures start to result in the robots killing not only the paying guests but also each other. When asked to turn them off, the supervising scientist says, 'In some cases they have been designed by computers themselves. We don't know exactly how they work'.

The 2016 TV series *Westworld* adds a further novel angle. It deliberately becomes increasingly more difficult to identify who are the robots and who are the paying guests, as a result of the host robots developing a virus and distorting how they perceive their own existence. At the heart of this particular narrative is the unspoken question of what is better: a robot with a conscience or a real person who is nasty, and of these two types, who has the moral superiority?

So it is against this background of media-driven perception that we need to consider the realities and practicalities of the implementation and future use of advanced analytics and AI in business and professional life, and how their influence will affect the way that we conduct our day-to-day affairs. This approach will probably not entirely satisfy the curiosity of purist technologists, however, this book is aimed not at that audience but rather at the generalist and business practitioner, who will be on the front lines of decision-making, at least for the moment.

Our initial journey will take us from the very basics of BI, through the foothills of predictive and prescriptive analytics, and ultimately to the mountain peaks of cognitive analytics and AI. According to Deloitte, ‘Cognitive analytics offers a way to bridge the gap between big data and the reality of practical decision making’.[1]

Whilst not strictly a proprietary terminology, cognitive analytics is often most associated with particular technology companies (such as IBM), which have used it to describe their own processes for gaining insights into big data usually through the use of so-called ‘Intelligent API’s’ such as face, speech, and vision applications. (An API, or application programming interface, is a set of definitions, protocols, and tools that collectively provide the building blocks to help a programmer build a new programme.)

To start to set the scene, it might be helpful to try to explain the difference between cognitive analytics and AI. IBM’s particular viewpoint is best expressed in its description of the use of each technology in a medical context: ‘In an artificial intelligence system, the system would have told the doctor which course of action to take based on its analysis. In cognitive computing, the system provides information to help the doctor decide’.[2]

In the maturity curve of analytics, how this model works appears to be mainly a matter of degree. In time, a doctor who acts against the cognitive advice given might prove to be either brave or reckless; the model also opens the door to some very interesting future discussions on the topic of professional negligence. However, it is prudent to start with the basics and work our way through the evolution of analytics in order that the reader may have a firm understanding of the foundations.

BUSINESS INTELLIGENCE

Was there ever a time that those running businesses did not consider themselves to be intelligent? By the time that Adam Smith (1723–1790) had put pen to paper in 1776 to write *The Wealth of Nations*, it was already clear that industrialists had a pretty good idea about how the economy works and what makes businesses profitable and successful. ‘High wages of labour and high profits of stock’, wrote Smith, ‘seldom go together … except in the peculiar circumstances of new colonies’.[3] Perhaps the expression *disruptive technologies* could be substituted for the word *colonies* in this quotation?

Some of his ideas seem more obvious in hindsight. ‘A great stock,’ he said, ‘though with small profits, generally increases faster than a small stock with great profits’. In those days, information would have been collected and stored in leather-bound ledgers more at home in a Charles Dickens novel. The advent of the modern computer created enormous stimulus for change, and it is sometimes argued that spreadsheets became one of the most important applications in their history.

Spreadsheet solutions such as VisiCalc, Lotus 123, Excel, and others have increasingly found their way into modern decision-making. For many accountants the work done by these tools (or their current replacements) is the backbone of their calculations. However, in a more complex and dynamic accounting environment, traditional spreadsheets are increasingly being viewed as dated, described by some as being limited in their ability to make substantial improvements in functionality (which seems to be mainly confined to better visualisations). It's a harsh verdict for tried and tested tools that remain the default way of carrying out calculations for many users. Even today, modern BI vendors still try to create a spreadsheet-type touch and feel to their more modern capabilities, paying homage to their software predecessors.

Nowadays, BI is increasingly becoming the cornerstone of enterprise decision-making. The pace of change has continued to accelerate as BI tools are increasingly becoming commoditised. There are multiple vendors in the marketplace and there is no shortage of advice provided by them. According to the analysts at Gartner, the BI market opportunity alone is projected to be valued at US$20.8 billion by 2018 (2014 sales were dominated by SAP, with a market share of 21.3%, followed by Oracle at 14% and IBM at 13%).[4]

In addition to these megavendors, there are many smaller specialist vendors. One of these recently issued its own views on the top 10 trends:[5]

1. Governance and self-service analytics becoming best friends
2. Visual analytics becoming a common language
3. The data product chain becoming democratized
4. Data integration getting exciting
5. Advanced analytics is no longer for analysts
6. Cloud data and cloud analytics taking off
7. The analytics centre of excellence becomes excellent
8. Mobile analytics stands on its own
9. People begin to dig into IoT (internet of things) data
10. New technologies rise to fill the gaps.

Stripping away the marketing hype and the possibility that data integration could *ever* be exciting, the summary appears to indicate that what started off a decade ago as being an enhanced replacement for spreadsheets has now become a fully blown industry in its own right, and an industry with ambition.

Standard BI software and systems, which some might suggest are no more than an enriched form of enhanced management information technology, are now increasingly overlapping with advanced analytics, cloud-based solutions, and mobile enablement. Consultative selling has become prevalent, with technology-enabled ex-accountants often operating in a quasi-sales capacity.

Not only are BI vendors suggesting the best technology for clients to use to resolve their business issues, they are also increasingly telling them how to internally restructure themselves to take optimum advantage of these new and enhanced capabilities.

For large and small vendors alike, there should be no real need to talk up the BI market. The economy has invariably created trading conditions of increasing complexity, and with that a need for greater granularity of data and more detailed insight. The speed of change in an increasingly volatile market forces practitioners to remain constantly on the lookout for change and to react accordingly. Where there are detrimental business results, it becomes essential to act quickly and decisively to stem the losses; where there are opportunities, to move equally quickly and take appropriate advantage. The ability to create a quarterly rolling forecast (or an even more frequent one) is sometimes described as one of the most important financial management planning tools of current times.

The detailed merits of financial performance management tools can be found elsewhere in many books about accounting practices, but the main point to be made here is that effective financial performance management is the cornerstone of enterprise analytics. Through this, it becomes much more straightforward for organisations to understand the impact of, for example:

- Deemphasising or sunsetting a product or solution
- Selling a business or growing thorough acquisition
- Changes in customer demand or supplier capacity
- Understanding the financial impact of organisational change, that is, headcount reduction
- Currency fluctuations and market volatility.

Despite these new capabilities there remain many companies who are at a very basic level of analytical maturity. Businesses in emerging markets, in particular, are at an early stage in their analytical journeys. With cognitive analytics and AI likely to be heavily influencing their business activities within a decade, it seems essential that emerging market players pick up the pace and start to implement new analytical ideas quickly.

Almost certainly, measuring the improvement in more advanced analytics and AI capabilities will rest on the effective use of BI tools.

It may be helpful to understand why some countries and markets haven't progressed as quickly as others. Part of the answer may well be in the strategies of the technology vendors themselves, who have tended to focus on the relatively low-hanging fruit of their local home markets, where there is still considerable growth to be obtained.

Other reasons may also be:

- Business partner strategies that also focus on home markets rather than emerging and growth markets.
- Absence of local evangelism.
- Soft or weak leadership.
- Ineffective marketing campaigns that focus on technical capability rather than business need.
- Geographical maturity.
- Sensitivity of smaller organisations to incurring the cost of new technologies.
- Inertia of small businesses – and some larger, more traditional businesses as well – with respect to change.

If there is concern about adopting and implementing even foundational analytics in the office of finance, then important questions might need to be asked about the ability of those businesses to embrace more advanced analytics, which will ultimately lead to cognitive insights and AI.

ADVANCED ANALYTICS

One of the leading analytics companies, Gartner, describes advanced analytics as follows: 'The analysis of all kinds of data using sophisticated quantitative methods (e.g., statistic, descriptive, and predictive data mining, simulation and optimisation) to produce insights that traditional approaches to business intelligence … are unlikely to discover'.[6] It's as good a description as any, and better than most.

Advanced analytics, sometimes also known as predictive analytics, usually comprises the series of capabilities shown in Table 4.

Dr Colin Linsky, a leading analytics expert at IBM, helpfully clarifies the difference between advanced analytics and predictive analytics, saying:

> *Predictive analytics is a subset of all advanced analytics techniques. There are plenty more advanced analytics algorithms and routines that are not predictive.*
>
> *Commonly, models that are built to describe when used again on fresh data would (also) be said to be predictive.*
>
> *Some are, by design, looking at future time periods and would be thought of as predictive, whereas others just classify or segment. They only become predictive by being used on a later subsequent set of data to which the outcome of the business problem is unknown.*

So, in effect, prediction depends not only on algorithms, calculations, and routines, but also on the nature of the data and the time frame. Put another way, *prediction* is a generic expression for a broader capability or function, not a tool.

At this stage the reader (perhaps like the author) might start to become confused, and even begin to feel that if advanced analytics is this complex, then how much more complex could cognitive analytics and AI be? And if it is so complex, then surely shouldn't it be left in the hands of other people – let's call them experts, for want of a better name – rather than the majority of us, who are relatively uninformed?

Doesn't a fear factor also start to emerge with respect to the technology itself as a result of not understanding the jargon? Isn't the individual fearful not only of the jargon and technology, but also of the impact of these on their jobs and livelihood?

This viewpoint is arguably no different than that taken by many people in business when they think they need technology to improve broader business results (as opposed to a day-to-day breakdown of the system or laptop) and they call on the IT department to give them a helping hand. Some will remember that for many, the relatively fractious nature of that relationship is typified by a difference in understanding of each party with respect to the needs of the other, with conversations often shrouded in jargon.

These new IT-led business solutions often have shown themselves ultimately to be technology solutions that are late and unduly expensive, and that usually don't solve the problem they were intended to solve. These kinds of problems are often also characterised by a silo approach to business problems, which is usually exacerbated by representatives of the IT department and the line of business each using different terminology.

Over time, this approach has been mitigated not only by a greater understanding of the role of IT in business (and vice versa) but also by the emergence of new roles and positions. Often these new roles carry relatively obscure titles, but overall we can generalise them colloquially as lobits: line-of-business IT professionals. In effect these

TABLE 4 Typical capabilities used in advanced analytics.

Capability	Function	Typical usage
Pareto analysis	Often known as the 80/20 rule, it is a form of analysis that identifies the top proportion of causes for the majority of problems. 80/20 implies that 80% of all problems arise due to 20% of causes, but this is not a hard and fast rule.	The application of Pareto's rule in risk management helps organisations focus on those causes which are likely to comprise the greatest risk.
Clustering/ k-means	Clustering is a way of grouping a set of objects together in some way, based on the fact that the objects in the cluster are more similar to each other than to those in a different cluster.	In biology, used to make spatial comparisons of communities. In market research, used to partition general groups of consumers. K-means clustering means to partition each group into a cluster based on the mean of that group; it is used in data mining to minimise intracluster variance.
Forecasting using Holt-Winters methodology	Also known as triple exponential smoothing, it's a process that can be used to forecast data points in a series, provided that the series is repetitive (or seasonal) over some period.	Used for making certain assumptions, such as calculating or recalling some data; for example, trending of house prices and inflation rates.
Decision-tree analysis	A decision support tool that uses a tree-like graph or model of decisions and their possible consequences.	Used in decision management to identify a strategy to reach a goal; often used in management science, healthcare, and operations management.
Rules of association	Rules-based learning method that identifies relationships between variables in large databases coupled with levels of confidence.	Promotional pricing, product placement, market-basket analysis.
Logistic regression	Also known as a logit model, it is a type of linear model that estimates the probability or dependency of a certain outcome.	Determines whether a customer will buy a product based on (for example) age, gender, geography.
Linear regression	An approach for modelling the relationship between a data set and one particular variable. Think of it in terms of finding a single line on a graph that represents a range of data points. Can show changes in data over time.	Used in creating trend lines (e.g. for GDP and oil prices). Capital asset pricing model.

(*continued*)

TABLE 4 (*Continued*)

Capability	Function	Typical usage
Correlation	A broad set of statistical relationships, although usually relating to two variables that appear to have a linear relationship with each other. Useful, as it can identify a predictive relationship with its members.	Correlation between electrical demand and weather. (This is a causal relationship, in that one is dependent on the other – but correlation is not necessarily causality.)
Bayes	Also known as Bayes Law, or Bayes Rule, it describes the probability of something happening based on prior knowledge that it may be related to an event.	Can be used in medical diagnosis or fraud detection, for example, but can be affected by false positives, which is, in effect, a statistical anomaly that shows positive when the answer should be negative and vice versa.

are roles – or, more usually, individuals – who sit between specific business functions (what we otherwise call the line of business) and the IT department. These individuals have at best an understanding – or, at least, a strong awareness – of the needs and language of each party. In many cases they will have an understanding of specific sector or business function as well as computer and system capability.

One question to be addressed is whether it is easier for a business person to learn technology, or a technologist to understand business issues. Both are equally complex, and the ideal lobit needs to be bidirectional in approach.

To become bidirectional in approach requires not only unique technical skills, but also an understanding of key business drivers or pain points. It also depends on having the right can-do attitude, coupled with agility of thought and flexibility of approach. These are as much about personal characteristics as about technical competence. Beyond this, the constantly shifting nature both of commerce and technology are such that technical and business skills need to be continually reinforced by training and experience. The mood of the moment is increasingly that of vertical (i.e. industry oriented or functional) solutions, which encourages individuals to create some degree of specialisation.

Inevitably, resourcing and bandwidth issues start to emerge. If there are not enough of these people around to meet the needs of the existing market, how then will the various industries cope with the massive expected increase in demand for advanced analytical services? And how will the emerging growth markets, such as Asia and Latin America, cope?

It seems inevitable that university and professional qualifications will have to quickly start to reflect these technological and business needs, through changes to curriculum. Employees, whose career paths have often focused solely on business or technology, will increasingly need to work in a learning environment that accommodates both.

One answer may well rest in the concept of 'analytics as a service', which is in effect an outsourced analytical capability, but this arrangement, too, is not without potential bandwidth problems. In any event, many organisations still remain nervous about data security, and the penalty for getting it wrong can be considerable in both financial terms and soft reputational terms. Outsourcing the analytical function is usually dependent on procurement experts becoming engaged in the selection process. It becomes critically important that the procurement team is aware of all the essential issues in order to advise properly on contracts and to manage the postcontract relationship effectively.

Employers increasingly need to look at the career development of their staff through a different lens. This impacts not only hard performance measures at the individual level, but a whole range of softer capabilities, such as awareness, flexibility, and agility.

At such a transformational time in business, there is also a sense that professionals within the human resources department may not yet have fully recognised the scale of the challenge in front of them. They, too, will need to change with the times.

Beyond these ideas, there is also a harsh reality about moving to the next stage of the analytics maturity curve. If organisations and technology companies are already struggling to support the individuals dealing solely with BI and advanced analytics, how can they possibly respond to the more complex needs of a world of cognitive analytics and AI?

Advanced analytics is an inevitable next step in the evolution of the analytics agenda. Most of the analytics companies and consultants are forecasting massive growth in this area, and predicting that companies will be using advanced or predictive analytics in multiple areas. One good example is the ability to predict equipment failure.[7] With this ability, manufacturers are able to plan for preventative maintenance, whilst dependent clients are saved the disruption of breakdowns, interruption to process, and costly downtime (see Table 5).

Advanced analytics will not only impact the way we work but will inevitably impact our social and private life as well, in terms of:

- The books we read and the music we listen to
- The TV shows we choose to watch (or more likely record for a later date – assuming, of course, that we have not reverted entirely to an on-demand form of viewing)
- The restaurants we visit and the nearest coffee shop
- The route home that we should take
- What we buy our partners for their birthdays
- Who might want to cause us harm, especially through online contact
- When our home freezer will break down, and what we should do about it
- Where we should go for our holidays.

The odd thing – if it is indeed odd – is that soon we will take this sort of information for granted. Of the eight relatively random items on the preceding list, only one of these is speculative: the issue of when the home freezer will break down. But that technical capability already exists and is used in commercial freezers, which are able to detect problems and make what's called a residual life prediction. Beyond this detection function, the system can also diagnose the problem and anticipate what work is needed to prolong the life of the equipment before breakdown. By being able to contrast the performance of the freezer (or other equipment) with competitive information obtained

TABLE 5 Uses of advanced analytics.

Function	Benefit	Industry affected	Industry imperative
Understand which customers might leave and offer them an incentive to stay	Optimised marketing expenditure, enhanced customer loyalty	Multiple; e.g. financial services	Cost reduction
Anticipate sales volumes of foodstuffs by considering weather conditions	Greater sales, more targeted sales promotions, optimised shelf life, customer loyalty	Consumer retail	Revenue improvement
Optimise delivery routes	Reduced fuel costs, reduces maintenance costs	Consumer retail	Cost reduction
Identify which machines are most likely to need maintenance	Reduced downtime, optimised cost through planned maintenance	Multiple; e.g. manufacturing, construction	Cost reduction
Anticipate which properties are most likely to be affected by hurricanes	Preventative action by insurers, homeowners, commercial businesses; more effective mobilisation of supply chain; better financial management	Insurance, construction, property development	Risk management, better profitability, cost reduction
Improve child protection	Social benefit, optimised services	Social services	Risk management, cost reduction
Clinical support systems	Better prioritisation, social benefit, optimised services; improved supply chain	Social services, healthcare, insurance	Risk management, cost reduction
Direct marketing	Reduced marketing cost, upsell, and cross-sell	Multiple; e.g. financial services	Cost reduction, improved revenue
Fraud detection	Improved granularity of pricing, better risk management	Insurance	Cost reduction
Underwriting	Improved granularity of pricing, better risk management	Insurance	Improved revenue, cost reduction
Collection analytics	Greater control over payment delinquency	Multiple; e.g. financial services	Risk management, cost reduction
Crime management	Social benefit, service optimisation	Police, crime prevention	Risk management, cost reduction

through external industry data, freezer manufacturers are not only able to manage their performance compared to industry benchmarks but also to use that information to ensure that they remain competitive.

Information and analysis in isolation is important, but perhaps more important is the use of that insight to allow comparisons to be made. Will this comparative process, in which information is placed in context (a process called contextual analytics) be one of the next waves of analytical development?

Turning back to the freezer question, how then might we feel if we received some form of individual e-contact forewarning us of a likely breakdown and recommending that action be taken to prevent this from occurring? Would we take action? How many reminders would be needed? Predictive analytics may provide a probabilistic insight into what might happen, but won't the action we actually take often depend on our own behavioural traits? Might our attitude change if we also received a message from our freezer insurance company saying that any claim for defrosted goods would be rejected because we had failed to take appropriate action? How might we feel if the supermarket refused to sell us more frozen goods at checkout because of the risk of not being able to store them?

In essence, therefore, this so-called new era of data, coupled with connectivity, predictability, and insight are not only likely to have an impact on the way we work but also on the way we live – in a way that is currently unimaginable. Should society wait for this to happen and then respond? Or is it better that we understand not only what is happening but why it is happening, in order that we all – as key stakeholders – can participate in the discussion and ultimately influence the outcome?

PRESCRIPTIVE ANALYTICS

Prescriptive analytics is generally thought of as being the third era of analytics (coming after BI and predictive analytics). If predictive analytics anticipates from the data what is likely to happen, then prescriptive analytics goes one step further by suggesting what should be done about it.

At the heart of prescriptive analytics is the question, why has this happened? By understanding this question, it becomes possible to identify a positive or a mitigating action to benefit from the prediction.

The term *prescriptive analytics* was first coined in a 2010 paper that appeared in *Analytics* magazine, in which it was described as 'a set of mathematical techniques that computationally determine a set of high-value alternative actions or decisions given a complex set of objectives, requirements, and constraints, with the goal of improving business performance'.[8]

The paper suggested that prescriptive analytics is the pursuit of two options, which the authors labelled 'optimisation' and 'stochastic optimisation'. They are:

1. How can we achieve the best outcome?
2. How can we achieve the best outcome and address uncertainty in the data to make better decisions?

The term *stochastic* is a form of probability model used to determine changes that evolve over time. More specifically, it is the mathematical model of a process whereby one of the variables is subject to the influence of another random variable. Examples of this might include:

- Analysis of how an investment portfolio might respond based on the results (probabilistic distribution) of individual stock returns.
- Modelling the survival of a rare species and how different strategies might impact that survival rate.

The expression *prescriptive analytics* was copyrighted by Texas-based Ayata, which was founded in 2003 and now focuses on the oil and gas production, insurance and renewables industries. It describes *prescriptive* as being '... a recipe. A series of time-dependent actions to improve future outcomes'.[9] Prescriptive analytics is described as a combination of the following items:

- *Models*. A way of organising data, often associated with a degree of standardisation.
- *Data*. The raw components of analysis, which comprise quantitative and qualitative values.
- *Business rules*. In this context, usually a piece of computer software that executes key business decisions in a production environment.

On the other hand, Ayata describes predictive analytics as comprising a series of key components or capabilities, which include:

- Machine learning
- Applied statistics
- Operational research
- Natural language processing
- Pattern recognition
- Computer vision (how computers gain understanding from digital images)
- Image processing
- Speech recognition.

If the casual reader was already confused by the difference between *predictive* and *prescriptive*, then this confusion might be worsened by reflecting on Ayata's description of *prescriptive*. After all, how many more capabilities might possibly be needed to provide cognitive analytics and therefore some form of AI?

Perhaps this is one of the issues of technological advancement? As professionals attempt to describe the different forms of analytics by giving them different titles, they create further confusion amongst lay readers, who are beaten down by jargon.

In time some analytical terminology will become obsolete. It's impossible to uninvent existing words, although it is quite natural for terms and expressions to fall from common use with the passage of time. (Who remembers the term *snoutfair* – a person with a delightful countenance – for example?) Perhaps the generic expression *management information* will eventualluy fall into that category – even if some colleges and universities still promote it as a course of study.

The term *management information* is usually taken to comprise the study of those systems that connect people, technology, and organisations – and how they relate to each other. Where does management information stop and BI start – and how do both of these morph into advanced and cognitive analytics? Maybe with time we will develop an entirely new and relevant lexicography, and new ways of expressing ourselves in these complex areas.

Figure 2 attempts to identify the relatively linear flow of progress of the subject of analytics.

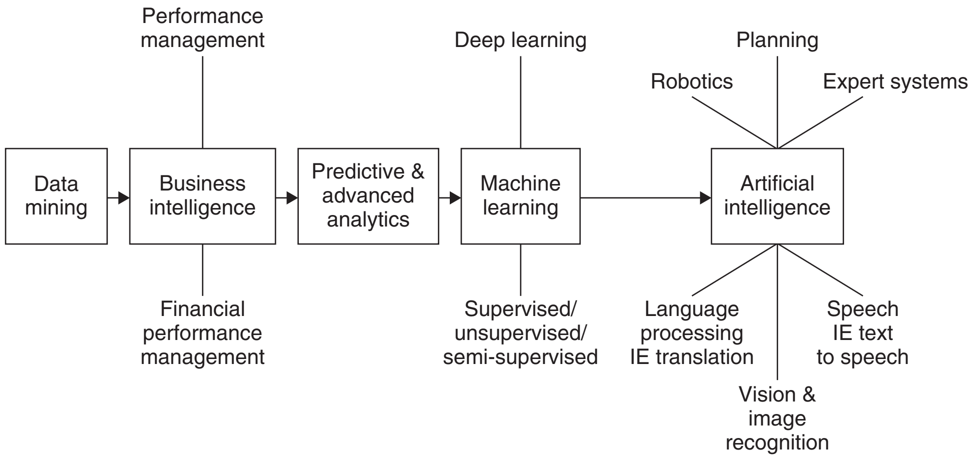

FIGURE 2 The road to artificial intelligence.

BUSINESS RULES

An essential difference between prediction and prescription is the application of business rules, usually through some form of business rules engine.

So-called business rules engines are often known as *operational decision management* and have been around since the early 1990s, when they were introduced by companies such as Pegasystems, Ilog (now part of IBM), and Fair Isaac. In effect, the rules adopted by a business (which might relate to operations, risk management, supply chain, or any other management policy) are put into operation through what is known as executional code in the system; that is, a directive to the system to carry out certain tasks according to coded instructions.

Business rules software may form part of a larger suite of business-rules management, which may include the following:

- Consistency and governance of key decisions
- Prioritisation for certain classes of customer
- Registration of rule changes.

Rules engines are often separate from the data system and allow users to make changes without involving the IT department. For example, in the case of insurance claim inspections, the claims department might be able to change the clip level – that is, the financial figure above which claims are always inspected – if there is a sudden surge of claims, such as in the case of a major weather incident.

Digging more deeply, the business rules process also generally comprises two elements:

1. Detection of some sort of business situation, which triggers a notification
2. Detection of an overload within the process, which may require some sort of change.

Underpinning both of these is the notion of workflow. Workflow is in effect the systemised sequence of activities that transform labour, materials, and other assets into consumable goods or services. With its foundation in time and motion studies, over time it has been seen in various guises, including Total Quality Management, Six Sigma, and more recently, Business Process Reengineering.

Looking forward, one of the potential weaknesses of these relatively traditional workflow processes is that they are linear in nature and are dependent on a linear sequence of events. It is entirely feasible that the traditional approach to relatively linear workflow will start to become outdated. In some cases existing workflows may already contain so-called loops, with the first step in the process being initiated by the last step in the process, but future workflows might prove to be much more random or complex in nature.

For example, the supply chain process has traditionally been thought of as a linear process: hence the term *chain*. One view of the supply chain of the future envisions a supplier ecosystem that is no longer linear in nature but depends holistically on bidirectional connectivity and collaboration.[10] In such a nonlinear model, intelligent systems continually assess multiple constraints and alternatives, and ultimately allow decision makers to simulate various courses of action.

The increased impact of the internet of things will certainly have an effect on traditional linear thinking. Analysts have already identified *interaction management* as being a critical success factor going forward. Interaction management comprises the ability to cope with multiple interactions and to orchestrate them. Gartner describes this as 'Thinking less about processes, more about interactions'.[11]

As with many parts of the technology sector, there are professional institutions and organisations for each aspect of a process. The US-based BPM Institute identifies a series of key skills and capabilities that are necessary to be successful in BPM, that is, Business Process Management (see Table 6).[12]

Whilst in many ways the preceding skills may seem to be relatively generic for the analytics industry, it might be argued that many of these could usefully form a framework for a broader Advanced Analytics/AI/Cognitive Analytics Institute were one to exist (or subsequently to be formed).

TABLE 6 Key skills and capabilities of BPM practitioners.

Key capabilities of BPM practitioners	Key skills of BPM practitioners
Aligning processes with business strategy	Systems thinking
Discovering and modelling processes	Process discovery and modelling
Measuring and improving processes	Facilitation skills
Harvesting policies and rules	Performance measurement
Managing the changing of culture	Process analysis and design
Governance and decision-making	Rules and decision management
Deployment of technology	Change management
	Project management
	Technical skills (requirements gathering, designing user experience, optimisation, and simulation)
	Governance and establishing centres of excellence

COGNITIVE ANALYTICS

According to Technopedia, 'Cognitive analytics can refer to a range of different analytical strategies that are used to learn about certain types of business related functions, such as customer outreach. Certain types of cognitive analytics also may be known as predictive analytics, where data mining and other cognitive uses of data can lead to predictions for business intelligence (BI).'[13]

Professional services company Deloitte has already trademarked the term *cognitive analytics*. A trademark is a distinguishing sign or expression that acts as to differentiate a product or service from competing companies. 'Cognitive analytics', Deloitte states, 'is a term used to describe how organisations apply analytics and cognitive computing technologies to help humans make smarter decisions.'[14]

These are relatively loose descriptions and arguably are more conceptual than definitive. At best, we need to understand cognitive analytics as representing a system that allows the user to interrogate both structured and unstructured data, both internal and external to the organisation, and from this data to

- Interact with the user in natural language
- Understand the issue that the user is trying to address
- Use a form of reasoning or basic intellect
- Learn from the answers and feedback.

One of the most-quoted representations of cognitive analytics is the much-heralded presence of IBM's Watson computer on the US game show *Jeopardy*, where a cognitive system outperformed previous champions. It follows IBM's previous and now legendary foray into games playing, when their Deep Blue computer beat the then world champion of chess, Garry Kasparov, in 1996.

In what should have been a clarion call for the power of technology, the incident unfortunately ended in a degree of acrimony. Kasparov (who many view as being the greatest chess champion of all time) accused Deep Blue's manufacturer of cheating, by using human intervention between the individual games in a series of six. Afterwards the officiator, Monty Newman, said, 'What Mr. Kasparov does have, that the computer doesn't, is a pulse'.[15] In retrospect, if that is the only distinguishing factor, then Deep Blue was a landmark invention.

The more recent IBM computer, Watson, is named after Thomas J. Watson (1874–1956), who was chairman and CEO of IBM. Watson provides cognitive intelligence through a series of connected APIs, which independently provide analytical capabilities. An API, as discussed previously, is a set of definitions, approaches, and tools to enable different capabilities to be bolted onto a system like a series of building blocks. APIs improve the ease, and therefore speed, of implementation.

In the case of Watson, the APIs that are available include such capabilities as

- Voice to text
- Text to voice
- Visual recognition
- Personality insights
- Tone analyser
- Various others.

An API is generally recognised as being the way in which disparate systems share information with each other. They underpin common web applications that are in everyday use. There's differing opinions about how hard it is to create – or write – an API. Experts describe it as anywhere from easy ('just writing a few lines of code') to painful, slow, and tedious.

This book has been written not for technocrats but rather for the layperson, whom the prospect of writing code will probably fill with horror. Simply, *code* is a series of instructions that are translated into binary messages to the computer using 0s and 1s that makes it carry out particular instructions.

The concept of binary coding goes back at least to the mid-1600s. Some suggest it might go as far back as the ninth century BCE and is even linked in some way to Chinese philosophical ideas. The ancient concept of yin and yang is described as a *duality*, in that each part represents an opposite force that is independent of the other, yet complementary and interconnected to the other, and which together interact to form a single dynamic system.

Microsoft's own cognitive solution, called Cortana, is rich in location and context awareness. It is marketed as having 'more personality' than its competitors. Developers have not only built gender into Cortana but even the ability to sing.[16]

This chapter opened with a quote from *2001: A Space Odyssey*; the closing parts of the movie show the computer HAL 9000 'dying' (if that is the right word for the decommissioning of a computer). As 'life' ebbs from HAL's system, the computer slowly sings the words to the old 1892 music hall song 'Daisy Bell' ('Daisy, Daisy / Give me your answer, do . . . '). The choice of that song by the film director Stanley Kubrick was not accidental. It recollects a historic, 1961 event, when one of the earliest and largest mainframe computers, an IBM 7094, warbled the first computer-generated song, which was (you guessed it) 'Daisy Bell'.

As an aside, it's often suggested that the name of the *Space Odyssey* computer, HAL, is a derivative of IBM, as the letters of the name (H-A-L) are one letter back in the alphabet from I-B-M. This has been denied by the writer Arthur C. Clarke and also by Kubrick, who both admitted some embarrassment over the apparently unplanned coincidence. In the follow-up *2010: Odyssey Two*, Clarke speaks through one of the characters, who says '. . . I thought that by now every intelligent person knew that H-A-L is derived from Heuristic ALgorithm', a comment that Clarke subsequently reconfirmed in his novel *The Lost Worlds of 2001*.

Apple's Siri, its voice-driven personal assistant with an ability to interact with third-party apps, is said to have played a significant part in the development of the Apple Watch product. Similarly to Cortana, Siri seemingly has hidden depths of knowledge, including the ability to tell a bedtime story that allegedly '. . . makes Siri seem so relatable. So delightful. So human.'[17]

The issue of 'personality' in computer systems is starting to get attention. As a result, we have begun to ask ourselves some slightly obscure questions. Typically these might include some of the following:

- How much does it matter whether a cognitive system interacts with a business user in a male or female voice? If it is important, are we subconsciously adding some element of gender to a computer system, and to what degree might this influence our personal relationship with the computer?
- Certainly scientific studies seem to indicate that people generally find a female voice more pleasing than a male voice. In his book *The Man Who Lied to His Laptop:*

What Machines Teach Us About Human Relationships, the author, Stanford University professor Clifford Nass, says, 'It's a well-established phenomenon that the human brain is developed to like female voices.' Because of this, many technology companies seemed to have steered clear of the male voice, so are we destined to a future of personal engagement with 'female computers'?

- What is the risk – if any – of users forming some type of personal relationship with their computer systems? *Her* is a romantic science-fiction drama about a man, Theodore Twombly, who develops a relationship with Samantha, the intelligent operating system of his computer, to the point that he even takes 'her' on holiday. The movie received generally positive reviews and received an Academy nomination for best picture in 2013.
- The story told in *Her* has its roots in an online article of the time that suggested that a user could potentially have an online, real-time conversation with an AI system. We now know that such capability is highly likely in the near future, if not actually here already.
- If we are concerned about the gender of and personal attachment to our cognitive systems, might this even extend to bereavement? Is the death of a computer system no more than some form of irreparable breakdown, or could it mean even more to us?
- We often say that our computer has 'died', and there are quite different thoughts about what we actually mean by 'death'. (The state of being dead? The permanent ending of vital systems? The permanent end of something?)

Is this thinking just nonsense? After all, we're just talking about machines, aren't we? But as we move forward – not only with technology, but also in the way the technology is adapted to create more attractive personas and as systems are adapted to better understand the user's personality – don't we somehow need to prime ourselves for some type of personal relationship with our machines different from our current one?

The classic movie *Blade Runner* from 1982 stars a young Harrison Ford and features Los Angeles in what seems now to be a not-so-futuristic 2019. In essence it is about beautifully made, smart humanoid robots that are built with a limited shelf life. Genius scientist Dr Tyrell not only creates a one-off robot called Rachael (which has embedded memories of a surrogate mother), but also teaches 'her' to love him, or at least to 'fool him into believing she would love him back'. The movie ends with Ford's character Rick Deckard and limited-life Rachael escaping to spend their limited time together.

Marcelo Gleiser's 2011 article 'Can Machines Fall in Love?' invites us (perhaps in a slightly tongue-in-cheek way) to reappraise the nature of the relationship between man and machine. In Japan, robots are already being created to keep the elderly company. As Gleiser puts it, 'If people in need are willing to compromise for less, while robots become more humanlike, there will be a joining of humans and machines. At what point then will machines stop being called machines?'[18]

THE ACCURACY OF ANALYTICAL OUTPUTS

Our journey so far has taken us from the proverbial valleys of management information to the lowlands of BI, predictive analytics, and prescriptive analytics, and ultimately to the peaks of cognitive analytics. In doing so, we have been lured into a technological trap that seems to suggest that the outcomes from the analysis are beyond doubt, but is that in fact the case?

In the case of BI, which we describe as being *descriptive*, one of the key tenets of the proposition is that there is a single version of the truth. In other words, time once spent in the boardroom discussing (and perhaps disagreeing about) who has the right set of figures is now potentially confined to history. Modern descriptive analytics embeds change controls and ensures that any alterations are fed through the entire set of figures.

Predictive and prescriptive analytics cannot be viewed in the same light. By definition, *prediction* is no more than a forecast about a set of events or circumstances. That forecast will have certain degrees of accuracy but should not be viewed as having absolute certainty. That is to say, if predictive analysis identifies a customer who is likely to change supplier, this prediction is nothing more than a hypothesis about a propensity to change, not an absolute guarantee.

Prediction is everywhere, from sports betting to the management of team performance and even to the political elections that determine our future. The accuracy of predictions can vary considerably, depending on which algorithms are used or the amount of data fed into the model. What is perhaps worrying is that, as individuals, we increasingly are relying heavily on predicted outcomes and using these predictions to drive our behaviour.

Prediction is often also associated with science fiction. Authors such as Philip K. Dick made fantastic guesses about what might happen in the future – for example, as with his literary creation of mutated humans (precogs) who can anticipate crimes in advance, as ultimately portrayed in the 2002 movie *Minority Report*. Such guesses are not, of course, real mathematical predictions, instead they exist in the realm of fantasy speculation.

The writer Isaac Asimov created the fictional concept of *psychohistory*, which combines history, sociology, and statistics to make fictional predictions about large groups of people and ultimately led to his 1951 *Foundation* series of books. His approach was to compare the behaviour of the human population to the behaviour of gases, in that it is impossible to predict the behaviour of one molecule of gas but possible to predict that of a larger volume of gas. This is known (in real life) as *kinetic theory*, which predicts the behaviour of a gas in motion by considering its pressure, temperature, conductivity, and other key attributes. Kinetic theory is not new. The Swiss mathematician and physicist Daniel Bernoulli was writing about it as far back as 1738, when he wrote *Hydrodynamica* (Latin for *hydrodynamics*), which set the ground rules for fluid mechanics.

It's tempting also to look at the accuracy of political opinion polls and how they affect individual behaviour. When there is a clear front-runner in the polls, it can often lead to reduced turnout or even so-called protest votes (when protesters want to make a point without affecting the likely outcome). When the poll results are ambiguous, members of the electorate may decide to vote in order to ensure that there is a clear majority for one side or the other, where otherwise they might have not bothered to do so. But polls have proven to be wrong in the past: sometimes badly wrong.

Increasingly, organisations rely on so-called polls of polls obtained from poll-aggregator specialists. These are companies that provide aggregated predictions that take into account multiple different algorithms and data sets. The concept of aggregated polls is not new. Meanwhile, in US political polling, one additional analytics element taken into account is the behaviour of so-called nearest-neighbour states (e.g. those with a disposition similar to the polled state), which is factored into the prediction in order to improve its accuracy.

The way in which a poll is carried out is also important, be it face-to-face, by email, or by landline. Much seems to depend on the nature of the poll, the age group being interviewed, and perhaps simply the impact of pressures exerted by the voting community, which may itself relate to specific local affairs. Aggregated polls smooth out different methodologies and variations in data sets.

And perhaps a prediction is not simply a matter of the statistical results of polls, but rather how it is interpreted and communicated. If data is the raw material and analytics are the methodology through which insight is obtained, then it is the purpose to which this insight is put that is the most critical. Few people who read blogs or articles will take care to check the results on which the conclusions of these blogs are based. In some cases, an item of published journalism may have been based on someone else's opinion and simply reworked, in effect creating an echo of an inaccuracy or misinterpretation.

Predictions about weather can be just as tricky. Even with the benefit of supercomputers, it can remain difficult to be certain about whether next June is a good time to schedule a wedding. Sometimes it can even be difficult to be sure whether or not to have a garden party in two weeks' time. Weather predictions often talk in terms of percentages of probability; for example, 'There is an 80% probability of rain tomorrow'. A proposed probability of 80% is certainly more helpful than one of 50% – in other words, a 50/50 chance of something happening – in which case you may as well just toss a coin. But an 80% probability of rain doesn't mean that 80% of the rain will fall on your garden in terms of volume, or even that it will rain on your garden for 80% of the time. Simply, it suggests that on 80% of days on which weather conditions (such as cloud cover and temperature) are similar to the one you are likely to get, you will get rain.[19]

The concept of *probability* is simply one way of expressing the chance of something happening: in this case, a rainy day. There is little or no certainty involved, unless a prediction proposes a 100% or a 0% likelihood.

In terms of predicting customer loyalty, it is also difficult to be absolutely certain regarding the degree of success. Many improvements in the probability of customer loyalty are underpinned by customer loyalty programmes, which are targeted, reward-based incentive schemes. According to Oracle, 'Airline frequent flyer programs enroll more than 200 million members worldwide; 76 percent of all U.S. grocery retailers with 50 or more stores now offer a frequent shopper program; 40 percent of all Visa and MasterCard issuers operate a rewards program'.[20]

Customer loyalty programmes help companies

- Analyse member data to view the demographics of various programme tiers and identify the factors that influence the buying patterns of different member segments
- Identify new enrollments and member growth rates over time
- Analyse member accruals and redemptions
- Understand the influence of different loyalty promotions on customer transactions and behaviour.

Through the use of analytics, companies are increasingly able to understand the overall performance of a loyalty programme, and measure its impact and the value it is bringing. Some might argue that measuring the impact and value of a loyalty programme is not truly predictive in nature, but it is the coming together of analytical and business processes that ultimately delivers results.

So how certain can we be about the accuracy of prediction and probabilistic modelling? At the end of the day, there are clearly at least three key, limiting factors:

1. Adequacy and accuracy of data
2. Quality and appropriateness of analytical work
3. Correctness of assumptions, especially that the future will be similar to the past.

The volatility of current markets, economies, and customer behaviour make decision-making even more difficult for managers and business leaders. The traditionally respected traits of experience and intuition retain value, but arguably at a reduced level. Increasingly, experience and intuition need to be supplemented by analytical insight.

Predictive modelling, be it through advanced or cognitive analytics, is still likely to bear an element of uncertainty, but there is sufficient evidence to indicate that it is better than the toss of a coin. In recognising this absence of infallibility, users need to appreciate that mistakes in prediction will still be made within systems and processes, at least in the short term. Will these mistakes be smaller or larger than failures of prediction due to human error?

Perhaps at the end of the day, some mistakes made by algorithms simply won't matter or will be less consequential. Doesn't it depend on whether the mistake involves the wrong choice of a birthday gift or the wrong coded instructions given to a pilotless aircraft?

As systems potentially displace human activity, one important question to consider is the degree of tolerance of mistakes we will be prepared to accept in even more advanced AI systems. Will more sophisticated systems inevitably demand greater accuracy? To what extent will mistakes ultimately undermine public and commercial confidence, and impair the implementation process?

CONCLUSION

This chapter has taken the reader through the foundational aspects of analytics, from the relatively basic theme of BI to the more advanced concept of cognitive analytics, challenging the accuracy of analytical outcomes along the way. Each subject on its own is capable of filling a textbook, but the intention here is primarily to provide an overview of the key components that ultimately lead to advanced intelligent systems, or forms of AI.

It's tempting to think that industries and professions will also follow the path of incremental improvement, but perhaps it is equally possible that they will leapfrog these foundational aspects and attempt to work directly with higher levels of AI.

This is entirely feasible, but it still remains important for organisations to be able to measure the financial benefit of such change. Therefore, the foundational tool of BI, financial performance management, remains an important part of the implementation process.

It's important also that we get over the potential stumbling block of poor data quality and avoid using that issue as a barrier to change. We are surrounded by poor information in our personal and business lives, yet cope with this by making judgements about the value, or veracity, of the information we get, weighting it accordingly. There's no reason why analytic and intelligent systems cannot take exactly the same approach.

In the next chapter, we will build on our knowledge of the foundational building blocks presented here, as we look forward to exploring the topic of AI in more detail.

NOTES

1. Bill Briggs, Cognitive analytics. *Tech Trends 2014*. Deloitte, 2014. https://www2.deloitte.com/uk/en/pages/technology/articles/cognitive-analytics.html (accessed 30 October 2016).
2. Steve Hoffenburg, IBM's Watson answers the question, 'What's the difference between artificial intelligence and cognitive computing?' VDC Research Group (24 May 2016). http://www.vdcresearch.com/News-events/iot-blog/IBM-Watson-Answers-Question-Artificial-Intelligence.html
3. Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations*. (Oxford University Press, 1776). Republished 1993.
4. Rita L. Sallam, Joao Tapadinhas, Josh Parenteau, et al. Magic quadrant for business intelligence and analytic platforms. Gartner report ID G00257740, 2014. https://www.gartner.com/doc/2668318/magic-quadrant-business-intelligence-analytics
5. Tableau Software (n.d.), Top 10 business intelligence trends for 2016. Tableau Software white paper. https://www.tableau.com/learn/whitepapers/top-10-business-intelligence-trends-2016 (accessed 31 October 2016).
6. Gartner. Magic quadrant for advanced analytics platforms. 19 February, 2014. https://www.gartner.com/doc/2667527/magic-quadrant-advanced-analytics-platforms
7. SKF (n.d.). Predictive and proactive reliability maintenance. http://www.skf.com/group/industry-solutions/oil-gas/services/mechanical-services/predictive-and-proactive-maintenance/index.html (accessed 31 October 2016).
8. Irv Lustig, Brenda Dietrich, Christer Johnson, and Christopher Dziekan, The analytics journey. *Analytics* (November/December 2010). http://www.analytics-magazine.org/november-december-2010/54-the-analytics-journey (accessed 18 October 2017).
9. http://ayata.com/software/ (accessed 10 January 2018).
10. Karen Butner, The smarter supply chain of the future. *Strategy & Leadership* 38, no. 1 (2010): 22–31.
11. Rob Dunie, W. Roy Schulte, Marc Kerremans, and Michele Cantara, Magic Quadrant for Intelligent Business Process Management Suites. Gartner report G00276892, 2016. http://www.integra-co.com/sites/default/files/Bizagi%20on%20Gartner%20Quadrant%20-%20Reprint.pdf
12. Gregg Rock and Andrew Spanyi, What skills are needed for BPM success? BPM Institute, 2014. http://www.bpminstitute.org/resources/articles/what-skills-are-needed-bpm-success (accessed 18 October 2017).
13. Techopedia (n.d.). Cognitive analytics. https://www.techopedia.com/definition/29437/cognitive-analytics (accessed 18 October 2017).
14. Cognitive analytics: the three-minute guide. https://www2.deloitte.com/us/en/pages/deloitte-analytics/articles/cognitive-analytics-the-three-minute-guide.html
15. Bruce Weber, Swift and slashing, computer topples Kasparov. *New York Times* (12 May 1997). http://www.nytimes.com/1997/05/12/nyregion/swift-and-slashing-computer-topples-kasparov.html
16. Damon Beres, Microsoft's Cortana is like Siri with a human personality. Huffington Post. 29 July 2015. http://www.huffingtonpost.co.uk/entry/microsofts-cortana-is-like-siri-with-a-human-personality_us_55b7be94e4b0a13f9d1a685a (accessed 10 January 2018).

17. Rene Ritchie, How to get Siri to tell you a bedtime story. iMore (26 October 2015). http://www.imore.com/how-get-siri-tell-you-bedtime-story
18. Marcelo Gleiser, Can machines fall in love? NPR website, Cosmos & Culture article (26 January2011). http://www.npr.org/sections/13.7/2011/01/26/133213460/can-machines-fall-in-love (accessed 21 March 2017).
19. Kevin McConway, The nature of probability. UK Met Office (9 October 2016). https://www.metoffice.gov.uk/about-us/who/accuracy/forecasts/probability (accessed 5 January 2017).
20. Oracle. Oracle loyalty analytics. Oracle data sheet, 2001. http://www.oracle.com/us/products/middleware/bus-int/064330.pdf