Big Data in Business

Easier access to computing power and new analysis methods are enabling the use of big data by businesses, changing the way they operate and communicate with their customers. This note examines the range of applications of big data by UK companies, highlighting the opportunities offered to a number of sectors. It also considers the challenges that businesses and regulators face when big data analytics are implemented.

Overview

- An increase in the quantity and variety of data available to businesses provides opportunities, but managing and analysing this 'big data' can be challenging.
- Big data analytics are being applied in the design, production, distribution and marketing of products and services across a range of sectors, increasing efficiency for many businesses.
- Regulation of personal data use varies between countries and navigating this may require self-certification.
- UK businesses will be affected by the proposed European Union General Data Protection Regulation (GDPR).
- Other challenges faced by businesses and regulators when using big data include finding people with the appropriate skills.

Background

Data can create value for businesses, as insight gained from its analysis can inform decisions. Traditionally, companies have used business intelligence tools for data analysis. These are applied to structured data, which reside in fixed fields, as in a spreadsheet. By comparison, 'big data' analytics typically involve data from a wider variety of sources, which may be rapidly analysed as they are collected (in 'real-time'). These data are often unstructured, such as images or videos. There is no single official definition of big data, but big data specialists often refer to its high volume, velocity and variety. These features make it difficult to readily manage and analyse using desktop computers and traditional data management tools.

Use of big data analytics creates opportunities to increase efficiency and improve predictions, leading to new, more personalised, products and services. It has also led to the creation of new data-driven business models, such as those used by Facebook and Tesco. Estimates suggest that, via efficiencies, innovation and business creation, big data was worth £25 billion to UK businesses in 2011, and may reach an annual value of £41 billion by 2017.[1] The Government has highlighted its potential, investing over £180 million to support big data research.

There are a number of emerging technologies becoming available to business that have led to increased use of big data. A decrease in the cost of data storage and the availability of cloud services (computing delivered as a service using internet technologies, accessed from any device connected to a network) has lowered the cost of adoption for smaller firms wanting to benefit from big data analytics. Storing more data may increase risk, as the potential for damage due to data loss or misuse is greater.

Once data have been collected by a business, sequences of calculations (algorithms) can be run to analyse them. If the data set is large or complex, this analysis may require access to high performance computing (HPC) facilities. There are a number of UK centres run by partnerships between public sector and industry research funders that provide HPC, such as the Hartree Centre (Box 1), HPC Wales and Supercomputing Scotland. BIS estimated that for £1 of public spending on the industrial exploitation of e-infrastructure, £10 of net Gross Value Added (GVA) will be generated within two years, rising to £25 after five years.[2]

Increased uptake of big data analytics by businesses has led to a need for regulators to use similar techniques to provide oversight. For instance, in the financial sector, the requirement for submission of trading data from both sides

of a trade and the fact that there are numerous data repositories can lead to duplication and inconsistencies. The rate of innovation of how businesses are using big data will make it difficult for regulation to stay current.

Applications of big data

Big data has a wide range of uses in businesses in many sectors. Some sectors are only just beginning to recognise the value of big data, while others are more advanced in their use of analytics. For example, the insurance industry is already using data from multiple sources to inform decision making (Box 2). By comparison, the financial sector has implemented big data analytics to deal with data in high volumes and velocities, but is only beginning to apply these techniques to data from a variety of sources (Box 3).

Design

HPC is being used to run computer simulations that model components of a product before the manufacturing process begins. Engineers from Bentley Motors used one such system to create virtual models of vehicles. This enabled faster product development times, decreased the number of prototypes required, reduced costs and eliminated the need for late-stage modification. Design decisions for some new products are being influenced more by data analytics. For example, the ability to monitor social media activity and analyse it in real-time enables companies to gain insight into customer response to a new or proposed product almost immediately. Ford used data from sales and social media to decide the design features on one of its new cars.

Production and distribution

For companies that use perishable goods, faster analysis of sales and supply chain data can be used to adapt production levels to meet demand. Implementation of analytics by the cosmetics company Lush led to a reduction in waste, achieving savings of over £1 million. Applying

the same techniques to its distribution network gave it improved knowledge of regional variation and annual buying patterns, leading to better prediction of future sales and a reduction in the level of stock held within the supply chain.

Monitoring assets

When a product is in service, its performance can be monitored in real-time by the use of sensors. Rolls Royce collects 40 gigabytes of data a day from its fleet of jet engines, including information on temperature and pressure. These are analysed to determine when an engine requires servicing, or for preventative maintenance before a fault develops, minimising disruption to airline service. The falling cost of sensor technology has enabled smaller companies, such as crane manufacturer Konecranes, to monitor assets that previously would not have been of high enough value to justify investment in this technology. Sensors are also increasingly used in sport. For example, every premier league football club employs up to five analysts. A variety of data, including athletes' physiological data, videos and side-line analysis are combined to monitor performance, identifying areas for improvement and aiding talent spotting.

Marketing

An online transaction usually generates much more data than an in-store purchase. Internet browser type, search terms and shopping behaviour can be recorded to create a customer profile. The amount of data generated in this way is set to increase, as online retail is expected to account for 23% of UK retail in 2016.[5] The volume and variety of data collected in-store could also increase by retailers tracking customer routes using CCTV cameras or mobile applications. Information about individuals allows retailers to group consumers with similar behaviour and implement

---

Box 1. The Hartree Centre
Hartree was set up by a partnership between IBM and the Science and Technology Facilities Council in 2013, following a report commissioned by BIS[2] which argued that the UK was falling behind international competitors in HPC. The centre provides the computing infrastructure and skills needed for a business to extract value from its data, encouraging research and development at a lower cost to the business than if it invested in the technology itself. For example, Unilever worked with Hartree to pre-screen ingredients, making it possible to focus on fewer, better experiments when designing products. Fuel cell manufacturer ACAL Energy used Hartree to simulate and predict the properties of constituent parts, enabling it to make significant research savings and shorten the time to market.

Box 2. Big data in the insurance industry
The increase in data sources available to insurance companies has provided more information about the assets they insure. Combined with a greater understanding of the science of disasters, this has enabled better computer modelling. The ability to use more computationally intensive programs has meant that models take into account their own uncertainty, improving even further their predictions of risk. The re-insurer Willis has increased its use of analytics, and integrated insights gained from this into many areas of its business.

Box 3. Big data in the financial sector
In 2011, financial and insurance services contributed £125.4 billion in GVA to the UK economy, 9.4% of the total.[3] The financial industry has historically analysed large quantities of data to make decisions.
- The ability to process larger volumes of data has led to better computer models and more informed decision making.
- Technological developments have decreased the time it takes to make a trade, introducing high frequency trading (HFT) into the market, which generates over one third of the UK equity trading volume.[4] These trades occur so fast that they are becoming limited by the time it takes light to travel to the exchange. Concerns over the effect of HFT on the financial markets have been raised, but the Government Office for Science found that, in most circumstances, HFT has not increased volatility.
- Data analysis that combines data from a variety of sources is starting to be applied within the financial sector, with unstructured data being increasingly used for informing trading decisions. Analysis of finance-related search terms found patterns that may be interpreted as 'early warning signs' of stock market moves. This research was used by the investment bank Cayman Atlantic, which analyses social media websites using sentiment analysis (POSTnote 460, Box 3) to predict trends in the financial markets. There is evidence that other investors are also using social media data, as a hacked Twitter feed announcing an attack on the White House caused a sudden dip in the markets in April 2013.
- Increasing demand by regulators for banks to provide evidence of cultural reform is beginning to lead to analysis of data from employee meetings and calendars.

behavioural advertising, which includes:
- search discrimination, where consumers' internet search results contain the products they will find most relevant
- targeted discounting, where shopping habits are monitored to provide vouchers and discounts
- dynamic pricing, where prices are changed because of the availability of, and consumer demand for, a product.

Similar techniques can be applied in-store if a retailer has a loyalty card recording an individual's transactions. Catalina Marketing delivers point-of-sale offers on behalf of retailers. Its primary database holds more than 2.5 million gigabytes of information, and is growing by over 300 million retail transactions every week. With tools provided by SAS Analytics, it runs models to determine which offers to deliver to individual customers. Catalina has seen redemption rates of up to 25% compared to an industry average of 6-10%.

Concerns about the use of algorithms to make decisions, such as those applied in behavioural advertising, were raised in a report from the Executive Office of the President in the USA.[6] For example, segmenting people into groups may prevent them encountering information that challenges their biases or assumptions. The report also recommended monitoring for potential discriminatory outcomes that may unfairly target vulnerable people. Concerns that big data may increase the prevalence of differential pricing have also been raised. First-degree price discrimination (offering a different price to each individual) could allow firms to extract more profits, but use has been rare because of a lack of data about individuals. Research has shown a customer with an 'affluent' profile searching for a product on Google may be offered a price up to four times higher than one appearing budget-conscious.

A number of European businesses recently came together with the Internet Advertising Bureau to provide a self-regulatory online behavioural advertising framework to improve transparency for consumers. It requires third party advertisers to notify consumers about the use of behavioural data to deliver relevant advertising, and to provide a way to opt-out of their data being collected for these purposes.

Challenges for businesses using big data
When using big data, a business may face difficulties ensuring data can be both kept secure (POSTnote 468) and accessed reliably. It may need to find new skills and, if using personal data, adhere to the relevant regulations.

Data access
There are a number of factors which can affect the ability of an organisation to maintain access to its data. Businesses using third party cloud computing services can be vulnerable to price increases, reliability problems, and providers going out of business. Because of this, contracts detailing what would happen to data access on acquisition or liquidation are important. Changing business practices may also affect access. For example, charities can collect donations and information about donors through their websites. However, an increased number of donations via third party platforms such as JustGiving has reduced the amount of data they receive. A third way that data access may be affected is an increased control of data by individuals. A potential model is a 'personal data market', where individuals deposit their data with an intermediary, such as MyDex, and determine who has access to it.

Skills
There are a wide range of skills relevant for businesses wanting to use data analytics, including knowledge of statistical techniques, the ability to program and use software, market-specific knowledge and communication.

People
Some companies, such as New Look, create a team of employees with a variety of skill-sets to manage their big data analytics. An alternative approach has been taken by companies such as eBay, which hire individual 'data scientists' with all (or most) of the skills needed. There is a lack of analysts and managers needed to analyse big data and make decisions based on their findings. A UK survey found that over half of businesses recruiting big data staff in 2012 found filling these positions difficult.[7] Data scientists are able to command high salaries that can be prohibitive. There is no formal data scientist accreditation, which can make it hard to identify skilled individuals. Accreditation does exist in similar sectors: the Market Research Society has developed qualifications for certified membership, training people from a range of backgrounds to provide market research insights to businesses with large data sets.

Business-university collaboration
One way to provide the multi-disciplinary skills required for big data analysis is for students to work closely with a company during their studies. Collaboration between a university with analysis expertise and a business with real-world data, can be beneficial for both parties. For example, projects led by the Centre for Doctoral Training in Financial Computing & Analytics at UCL allow companies to gain new insights and students to gain commercial experience.

As the potential value of a data set is often not realised until after analysis, the benefits of collaboration to a business can be difficult to predict. Publication of success stories, such as case studies produced by Hartree (Box 1), may encourage new businesses to take on similar collaborations. The benefits may also be difficult to predict on the academic side, as researchers may face challenges when publishing the results from such projects. This could be due to commercial sensitivity or because the research, such as some Aston University MBA projects, does not use novel methods and so may not be appropriate for a typical academic journal. The Technology Strategy Board's 'Innovation Vouchers' in England, 'Interface' in Scotland, and other schemes provide funding and make introductions between businesses and universities to encourage projects.

Software
Eight-in-ten businesses surveyed by The Economist lacked effective processes for sharing data between departments and among employees. Visualisation software can be used to present company data in an accessible manner. Marks and Spencer uses Spotfire to enable employees to understand data sets relevant to them, reducing the risk of the data being misused because of misinterpretation. Lush cosmetics has implemented similar software (Qlikview), enabling their employees to use data to inform decisions without needing to wait for IT department input. They have found this encourages openness and drives competition.

Regulation on the use of personal data
UK regulation
Any data that can be used to identify an individual is classed as personal data and is regulated by the Data Protection Act 1998 (DPA), which implements the EU Directive 95/46/EC. A new EU General Data Protection Regulation has been proposed (Box 4). Not all big data is personal but, when it is, there are issues that businesses may need to address.
- Personal data held by a company must be accurate, up to date and not kept longer than necessary, requiring data management policies that include regular audits.
- Personal data can only be used for purposes for which consent has been obtained, although concerns have been raised about how terms and conditions are communicated (POSTnote 460). If a company acquires a data set, it must know, and work within, the permissions under which it was collected, or ask for further consent.
- When requested by individuals, a company must provide all the personal data it holds on them at minimal cost to the individual. This can be expensive and time consuming for a business with large data sets collected over different time periods, although big data techniques may make the information easier to find.

The Information Commissioner's Office (ICO) is the UK's independent authority for upholding information rights. Because of limited resources, it is encouraging businesses to act within the law by providing guidance to the business community when violations occur. It has faced challenges in dealing with rogue individuals and is trying to introduce custodial sentences for data misuse. The ICO has run data protection workshops for businesses in Wales that were over-subscribed and has planned a number across England.

Outside the EU
Variation in regulation around the world is challenging for businesses that operate in a number of countries and also has implications for where they store their data. Exact data location determines which regulatory regime applies, but this may not be known by companies using cloud services.

The EU Data Protection Directive prohibits personal data transfer to non-EU countries that do not meet EU privacy standards. Australia and Canada have similar regulations in place. The USA has several privacy laws specific to particular sectors or industries, but has no federal DPA or

---

Box 4: Proposed EU General Data Protection Regulation (GDPR)
Currently, the EU Data Protection Directive is implemented in differing ways by EU member states. The EU Commission has drafted a new regulation that aims to provide greater transparency and consistency, reducing costs for businesses and increasing consumer confidence. It will affect businesses that operate or process data belonging to people in EU member states. The European Parliament gave its backing, suggesting a number of changes to the proposals, and will negotiate with the Commission and the EU Council once the Council has defined its position. If a framework is adopted, businesses in the EU will have two years before enforcement. Adopting the GDPR is estimated to lead to a net cost to UK business of £80-320 million per year. The GDPR differs from current UK regulation in some areas.
- The ICO can impose a maximum penalty of £500,000 on an organisation. The GDPR would raise this to 2% of global annual company turnover, and the EU Parliament proposes a rise to 5%.
- Currently, there is no obligation on UK companies to report breaches resulting in release, corruption or loss of personal data. The GDPR would require all breaches to be reported to the supervisory authority, for example the ICO, within 24 hours where feasible. The EU Parliament proposes raising this to 72 hours.
- Companies with over 250 employees, or whose core activities involve data processing, would be required to appoint a Data Protection Officer; the majority of large businesses already have an equivalent in place.
- Organisations would only have to deal with the authority in the country where they have their main establishment.

The GDPR sets out conditions of 'the right to be forgotten', building on 'the right of erasure' in the current Directive. The European Court of Justice ruled in May 2014 that search engines are responsible for the content they point to, requiring them to follow EU data privacy laws. In the three weeks after this ruling, Google received over 40,000 requests to take down sensitive material, sparking debate about the balance between privacy and censorship.

A study by London Economics consultancy found that there is a lack of understanding about the GDPR across business, and there is a key role for the ICO to increase awareness and understanding of the changes. In 2012, the House of Commons Justice Committee reported concerns that the ICO would be unable to meet future demand from businesses working under the proposed GDPR.

---

privacy act. To enable US companies to comply with the EU Directive, a 'safe harbor' framework was developed. An organisation that self-certifies compliance with this will meet EU standards. The ICO is working with the US Federal Trade Commission to increase cross-border enforcement.

Self-regulation
Large companies with strong global brands can be highly motivated to protect their public image, and have the resources to apply policies that go over-and-above local regulation. Vodafone has implemented a privacy culture, driven by its 'Privacy Commitments'. Smaller businesses with weaker brands may not have this incentive or the resources to administer equivalent policies.

Endnotes
[1] Centre for Economics and Business Research for SAS, "Data Equity", 2012
[2] Tildesly, D., "A Strategic Vision for UK e-infrastructure", 2011
[3] HoC Library, "Financial Services: contribution to the UK economy", 2011
[4] The Government Office for Science, Working Paper 11/1276, 2011
[5] The Boston Consulting Group, "The Internet Economy in the G-20", 2012
[6] Executive Office of the President, "Big Data", May 2014
[7] SAS and e-skills, "Adoption and Employment Trends 2012-17", November 2013