Journal of Systemics, Cybernetics and Informatics (2022) 20(1), 122-141
https://doi.org/10.54808/JSCI.20.01.122

Strategic Data Pattern Visualisation

Carol E. Cuthbert[1] and Noel J. Pearse

Rhodes Business School, Rhodes University,

Rhodes Business School, Rhodes University,

c.cuthbert@ru.ac.za, n.pearse@ru.ac.za

Abstract[2]

Data visualisation reveals patterns and provides insights that lead to actions from management, thereby playing a mediating role in the relationship between the internal resources of a firm and its financial performance. In this chapter, contingent resource-based theory is applied to the analysis of big data, treating its visualisation as a mode of interdisciplinary communication. In service industries in general and the legal industry in particular, big data analytics (BDA) is emerging as a decision-making tool for management to achieve competitive advantage. Traditionally, data scientists have delved into data armed with a hypothesis, but increasingly they explore data to discern patterns that lead to hypotheses that are then tested. These big data analytics tools in the hands of data scientists have the potential to unlock firm value and increase revenue and profits, through pattern identification, analysis, and strategic action. This exploratory mode of working can increase complexity and thereby diminish efficient management decision-making and action. However, data pattern visualisation reduces complexity, as it enables interdisciplinary communication between data scientists and managers through the translation of statistical patterns into visualisations that enable actionable management decisions. When data scientists visualise data patterns for managers, this translates uncertainty into reliable conclusions, resulting in effective management decision-making and action.

Informed by contingent resource theory and viewing these primary and secondary resources as independent variables and performance outcomes such as revenue and profitability as dependent variables, a conceptual framework is developed. The contingent resource-based theory highlights capabilities emerging from the interrelationship between primary and secondary resources as being central to competitiveness and profitability. Data decision-making systems are viewed as a primary resource, while complementary resources are (1) their completeness of vision (i.e., strategy and innovation) and (2) their ability to execute (i.e., operational capabilities). Data visualisation is therefore crucial as a resource facilitating actionable decisions by management, which in turn enhances firm performance. The balance between expert agents' self-reliance and central control, the

[1] Contact author

[2] Peer-Editor: Professor Karen Bradshaw (Rhodes University)

entity's values, task attributes, and risk appetite all moderate the type of data visualisation produced by data scientists.

Keywords: Knowledge Management, Data Decision-making Systems, Big Data Analysis (BDA), Data Visualisation, and Resource-Based Theory.

1. Introduction

Data analytics seeks to gain insight into strategic and operational management decisions through the creation of heuristic algorithms and models. In an era of the growing availability of big data, data analytics is increasingly viewed as an important competitive capability of firms, including those in the service industry. The service industry has been an early adopter of technology, with a bigger budget percentage-wise than other industries (Alarie et al., 2017). The Professional Service Firm (PSF) sector is one of the most rapidly expanding, profitable and significant sectors of the global economy (Empson, 2021). This includes the accounting, management, legal and architectural sectors. Broadly, service work refers to harnessing one's knowledge for the benefit of others (University of Cambridge et al., 2015). This chapter focuses on the Legal sector. Globally, big data analysis (BDA) expenditure should reach 216 billion dollars in 2021 (Statista, 2021). The size of the business intelligence and analytics software application market is forecast to reach 16.5 billion dollars in 2022 (Statista, 2021) and more specifically the global professional services market, the focus of this study, was $5 trillion in 2020.

The main reason for the growing use of big data is that it is increasingly recognised as a potential source of competitive advantage. Big data analytics (BDA) uncovers hidden knowledge, can improve decision-making, and support strategic planning (Chiang et al., 2018). BDA that can be visually presented and interpreted into actionable strategic decisions, is emerging as a tool for management to achieve a competitive advantage, unlocking firm value, and increasing revenue (Akter et al., 2016). Research on 814 BDA implementations, found a 3% to 7% improvement in revenue after BDA implementation (Müller et al., 2018). Ji-Fan Ren (2017) also linked the application of big data analytics to increased organizational performance, as did Ferraris (2019). Müller Fay and Vom Brocke Müller (2018) found that firms in highly competitive industries gained increased value from BDA assets, relative to other industries. Shabbir and Gardezi (2020) found that after mediating for knowledge management practices, the application of big data analytics accounted for 55.3% of the improved organization performance.

Given its potential and importance, data analytics is increasingly being utilised within the trillion-dollar legal industry. Legal technology is a multidisciplinary field relating to social science, computer science, engineering, statistics, and mathematics. Legal technology can be defined as technologies and IT solutions that provide some legal service (Kerikmäe et al., 2018).

This chapter is focused on predictive analytics and data mining by scientists. A data scientist refers to a professional using scientific methods to liberate and create meaning from raw data. The massive data generated by the Internet of Things (IoT) are considered of high business value, and data mining algorithms can be applied to IoT to extract hidden information from data. In this paper, we give a systematic way to review data mining in knowledge view, technique view, and application view, including classification, clustering, association analysis, time series analysis, and outlier analysis. And the latest application cases are also surveyed. As more and more devices connected to IoT, a large volume of data should be analyzed, the latest algorithms should be modified to apply to big data. We reviewed these algorithms and discussed challenges and open research issues. Prominent tools, techniques, and approaches of data analytics are Artificial Intelligence (AI), Machine Learning, and Data mining, which are all Industry 4.0 technologies (Frank et al., 2019). Data mining involves automated or semi-automated processes relating to, for example, patterns, rules, or knowledge (Chen et al., 2015). These processes are referred to as either supervised or unsupervised learning. Examples of supervised or predictive learning include classification or regression, which is a supervised machine learning technique used to predict continuous values. On the other hand, unsupervised learning is descriptive, meaning it describes a system or entity and its relationship to its environment, by, for example, clustering or association.

Predictive modeling, which can also be agent-based, uses amongst others, statistical techniques to forecast future outcomes using data mining technology to analyse data to predict future behaviours. In supervised learning, the aim is to predict a target or dependent variable based on a model that includes dependent and independent variables. Classification is used to estimate categorical target variables. Decision trees, Bayesian classification, and nearest neighbour are examples of classification (Han & Kamber, 2012). Regression examples are linear, ridge, lasso, and artificial neural networks (Muayad & Irtefaa, 2016). Predictions from legal judgments can be made using supervised learning, as the trial data, for example, outcome (guiltiness or penalty) can be used as the target variable. This supervised decision-making can affect multiple areas. For example, when trial result prediction is difficult, it could impact pricing analytics and intelligent pricing policy, as well as the acceptance or not, of a case.

In contrast to supervised learning, unsupervised learning examines the structure of data and looks for clustering or association. Examples of unsupervised learning are k-means (where points are partitioned into clusters according to the nearest mean), DBSCAN (where points are divided into clusters if they are close to many points from that cluster), and hierarchical clustering (Han & Kamber, 2012). Association looks at association rules highlighted from cause-and-effect variables in transactional datasets (Agrawal et al., 2018). Apriori (where frequent individual items in the database are identified and extended provided, they appear often) and Frequent Pattern (FP) growth algorithms are examples of association techniques. In Frequent Pattern Growth algorithms, the tree structure maintains the associations between the item sets, and the database is fragmented using one frequent item called the pattern fragment. The number of clusters must be pre-stated by the data scientist and each cluster contains similar records. In the legal arena, a cluster often relates to similar documents. An example of an association in the legal sector relates to the citation relationships between statutes. Furthermore, unsupervised learning techniques have been successfully applied to customer relationship management.

Different disciplines may be driven by visuals to a greater or lesser extent, but all are united in their need for data-driven decision-making. This chapter aims to examine data visualisation as a form of interdisciplinary communication, between data scientists and decision-makers, thereby enabling strategic management decisions. Contingent Resource-Based Theory provides a conceptual framework to discuss harnessing value from data analytics visualisations in legal technology.

2. Data science for strategic decision making

BDA with its predictive analytical capabilities can process and integrate large volumes of data, into familiar and common formats, enabling business decision making (Obitade, 2019). In simple terms, the work of data scientists is threefold. Firstly, they need to assemble a dataset from many sources. Thereafter, this data is analysed and visualised, so that it can ultimately be used by others to make actionable decisions. Coombs et al. (2020) described this work of data scientists as data acquisition and analysis, decision making, and action implementation. This process illustrates how knowledge is built by extracting value from a large variety of data and information sources to make management decisions, or to spot anomalies resulting in competitor differentiation. Several factors contribute to the use of data science for strategic decision-making organisations. These include (1) the availability and functionality of low code systems; (2) data management, or the ability to gather and combine data from a range of sources to create a big dataset; and (3) the availability of various tools to analyse this data.

A Low Code Application Platform (LCAP) can be described as a system supporting rapid application development, deployment, execution, and management. It does so by using programming abstractions such as model-driven and metadata-based programming languages, and easy deployments, providing support in the form of user interfaces such as dashboards (UIs), business processes, and data services (Vincent et al., 2021). The availability and functionality of low code systems, together with access to statistical packages such as R or Statistica, have enabled data scientists to identify patterns in the data. These patterns are then presented in reporting packages such as Tableau and Power PI. Some firms have proprietary systems already configured in cubed data format, and dashboards, such as Qlikview and Tagetik. Given the ease of use of these analytical tools, it has been estimated that in 2021 more than 50% of large enterprises will have adopted an LCAP as a strategic platform (Vincent et al., 2021).

Data acquisition entails combining the disparate data sources and then manipulating or cleaning the data into a format that can be analysed by data scientists. Applications such as Hadoop, NoSQL, MySQL, Hive, and Apache Spark to mention a few, are employed to manage data. Pre-processing techniques include feature selection and construction, missing value imputation, data integration, and transformation.

Moving AI systems to the cloud provides flexible resources and economies of scale in data management. Products run mass analytics through a cloud-based data warehouse, harnessing parallel processing to query petabytes of data, managing large amounts of data in nodes, known as Hadoop clusters (Kyar Nyo Aye, 2013).

Having assembled a dataset, the data scientist can now analyse the data. Traditionally, data scientists have delved into data, armed with a preconceived hypothesis to test. Increasingly, they first explore data to discern patterns, which lead to hypotheses that are then tested. Normalisation of data is required for variables of different scaling. Various techniques can be used to normalise data including linear scaling, clipping, log-scaling, and z-scores. Thereafter, through multi-dimensional mapping of the distance between points, principal component analysis, multivariate analysis of variance, and discriminant function analysis reveal patterns previously undiscerned when investigating one variable at a time in univariate analysis. The patterns identified may also be interpreted as trends. Not only is data analysed for patterns and trends, but also anomalies and outliers. Firms employ BDA to predict and conclude from the data, utilising unusual activity detection as well as pipelines, for the deployment and management of tasks, to support their strategy (Neirotti & Paolucci, 2007; Rivard et al., 2006).

This exploratory mode of data analysis can be complex and inefficient, and so it is not surprising that the data scientist makes use of intelligent automation to assist with this task. Intelligent automation is a combination of artificial intelligence, machine learning, and easy-to-use visual interfaces (DeCanio, 2016) enabling practice management. Coombs (2020) describes intelligent automation as the process whereby artificial intelligence (AI) can learn, adapt, and improve automated tasks found in knowledge and service work. AI uses machine learning to make predictions and reach conclusions from the data, improving results from feedback (Frey & Osborne, 2017), and thereby augmenting human cognitive capabilities. Pipelines (i.e., linear sequence of instructions) are used to train models, deploy, and manage tasks based on the feedback received (Jha et al., 2019). Data is integrated and processed, after which the algorithms are trained and tested by the developers, finally obtaining some output or action in the form of a prediction or an action required. Evaluation metrics such as log loss, absolute error, or precision/recall are used to improve the pipelines (Eraslan et al., 2019). This introduces a cognitive computing element, and this machine learning is used to streamline an organisation's processes (Tarafdar et al., 2017).

Once data has been analysed, the data scientist must prepare the results of analysis for use by others, who very often do not have specialist skills in the specialist area to which the data is being applied. The data scientist's exploratory mode of working can increase complexity and thereby diminish efficient management decision-making and action. It is therefore critical for the data scientist to translate the results into an accessible form to aid management in decision making and action. This is facilitated by the process of data visualisation, which makes the specialised work of the data scientist accessible to other disciplines. As a result, data pattern visualisation enables interdisciplinary communication between data scientists and managers, through the translation of statistical patterns into actionable management decisions. That is, after analysing results, data scientists can visualise the data patterns for managers, thereby translating uncertainty into reliable and accessible knowledge, resulting in effective management decision-making and action. This is explained further in the following section.

3. Data visualisation as interdisciplinary communication

A visualisation is a presentation of information or concepts as a mental image that is designed to communicate a message (Padilla et al., 2018), aiding clarity (Bresciani, 2019), comprehension time. Visual displays help in the presentation of inferences and conclusions and represent ways of organizing, summarizing, simplifying, or transforming data. Data displays such as matrices and networks are often utilized to enhance data analysis and are more commonly seen in quantitative than qualitative studies. This study reviewed the data displays used by three prestigious qualitative research journals within three years. The findings include the types of displays used in these qualitative journals, the frequency of use, and the purposes for using visual displays as opposed to presenting data in text (Verdinelli & Scagnoli, 2013), and the understanding of complex concepts.

Visualisations are typically the outcomes of statistical analyses that are presented in the form of line, bar, and area graphs. Scatter charts can show the values of two different variables as points and when used in adjacent windows, so that practitioners can compare many variables visually, and instantly. Maps, indicators like gauges, and tickers can show direction, while pivot tables can summarize and highlight critical data by colour, or bullet graphs (which avoid cluttering). Box plots show the distribution of data, while matrixes can show hundreds or thousands of data points at the same time, with colour highlighting the critical points. Power Bi, Tableau, QlikView, Python, and R, as well as other tools, are used to create these data visualisations and present the data in the form of dashboards, graphs, and so forth. Budgeting Planning and Forecasting, Profitability analysis, Cash flow planning, Production Cost planning and control, and other management activities can all be visualised.

Multi-dimensional scenarios can also be modeled using predictive analytics. Potential relationships or correlations can be depicted by leveraging a pair-wise correlation matrix and depicting this as a heatmap. The gradients of the heatmap vary based on the strength of the correlation and in this way the relationships are visualised. Another way is to utilise pair-wise scatter plots to display the relationships between the variables of interest. A third way is to use parallel coordinates where points are represented as connected line segments. Here each vertical line represents a data attribute. One complete set of connected line segments across all the attributes represents one data point. Therefore, points close together tend to cluster visually. One can also easily read the height from the lines depicting which variables are relatively higher or lower on the Y-axis. Two continuous numeric attributes can be visualised using scatter plots and joint plots. Scatter plots visually show where points congregate. Two discrete categorical attributes can be visualised by subplots or facets. Stacked bars and multiple bars can also be used to depict two-dimensional discrete categorical data.

To visualise mixed attributes in two dimensions one could use subplots with histograms or density plots, overlaying the different dimensions in different colours. Box plots are a very visual way of depicting groups of numeric data including their quartile values and outliers. Violin plots are an effective representation of two-dimensional mixed attributes. The ‘fatness’ of the visual depicts the probability density of the data at different values.

For three-dimensional data plot facets, colour, shape, size, and depth can all be used to display the attribute relationships. Time is another dimension that can be depicted. A pair-wise scatter plot can be used with colour to separate the categorical dimension. This means the different plots have different coloured clusters illustrating the relationship. To visualise three continuous numeric attributes, length, breadth, and depth could be used so that the clustered data points would appear to be in a cubed box. Size can be incorporated as a third dimension in a bubble chart, where the higher quantities of data points appear as a bigger bubble, while a kernel density plot can also be used to understand three-dimensional data, where the various shades of colour radiating out or clustering together for the two different variables, create a three-dimensional mixed attributes visualisation.

To visualise data in four dimensions, hue and depth within a 3D box can be applied by mixing two different colours. Similarly, bubble charts and scatter plots can portray the fourth dimension by varying the hue, and size, or number of facets. Five dimensions can be depicted by using bubble charts and hue, depth, and size (with the two main groups in different colours). Six dimensions can be illustrated in scatter charts with hue, depth, shape, and size. Finally, by removing depth (i.e., no cubed box) and using facets (e.g., fat tummy) and hue to represent the numerical axis, six dimensions can be presented. In this way, scatter charts are created with variation in hue, facets, and size.

Uncertainty, which is typically described statistically by calculating the standard deviation, standard error, and confidence intervals can also be visualised and can take the form of error bars and shaded intervals. When differences are small, an effective technique in visualisation is to panel or facet. This involves having multiples of the same visualisation, where the axes and scale are the same, but the data is slightly different. Each panel then represents a change in one variable in, for example, a time-step (Verdinelli & Scagnoli, 2013).

Visualisation represents a form of interdisciplinary communication and assists in the interpretation of statistical results. The goal of the visualisation is to discern and present patterns in an easily recognisable form that provides new insights and informs appropriate action. Once the primary users understand the depiction of the various dimensions these can be distributed throughout the various areas or departments. Data visualisation is also referred to as the art of storytelling and starts with the crucial question of “Why?”. Typically, firms evolve from descriptive, to diagnostic, predictive, and eventually, prescriptive stages of data visualisation.

Descriptive analytics explains what happened in the past and helps practitioners to understand the context. Diagnostic Analytics explains why something happened. Predictive Analytics predicts the most likely scenario, while prescriptive analytics provides recommendations based on the predictions and thereby guides action towards a solution.

4. Examples of data visualisation in law firms

Data visualisation is inter-disciplinary in that it is understood in, for example, management (Sackett et al., 2006) and marketing (Munoz, 2017) disciplines, as well as information science (Eberhard, 2021). The following examples illustrate the use of data visualisation as interdisciplinary communication in several disciplines within law firms, where data visualisation is being used both in the legal work itself and in the management of the legal practice.

In legal work, legal texts and legal outcomes can be analyzed and visualised. Legal texts are also being visualized in norm-graph visualisations. Major aspects of laws are summarised based on semantic legal data, into a legal-concept-ontology. The application Knowlex utilises interactive maps as a form of data visualisation by linking norms and case law outcomes to the relevant legislation or Acts of Parliament. The system then provides a first visual insight in the form of an interactive node graph depicting the properties and relations of the Act. A second visualisation is a tree map depicting the topic, evolution, and the dimension of legal literature.

Quantitative legal prediction (Katz, 2013) where hundreds of thousands of data points, such as cases, holdings, settlements, quality, cost, and client satisfaction are trawled in seconds, aids decision-makers considering pricing, longevity, and risk of a portfolio of matters. The relationships can then be depicted using the techniques for visualising the various numbers of dimensions.

Complementing this use in legal work, in the management of law firms, big data visualisation has been applied to marketing and accounting practises. From a marketing perspective, big data visualisation assists in (1) profiling clients, (2) customer relationship management, and (3) pricing strategies.

To profile or understand their client base, marketing teams can now have client industry experience and client spend displayed geographically. Many visualisations assist with profiling clients, including word clouds, which accentuate specific aspects. Marketing may require the strength of their client relationships or the mix of their clients in terms of industry as well as their practice jurisdiction to be geographically displayed, visually. Future developments include a digital market for those purchasing legal services as clients ramp up their demand for transparency. The relationship between client age and client lifecycle needs to be shown visually, as clients that have been with the firm for longer may spend a different duration of the various phases in the lifecycle. Heat, Flow, Chloropleth, referring to maps that use differences in shading, colouring, or the placing of symbols to indicate averages, and point distribution maps, as well as geo-mapping, assist in discovering patterns resulting in meaningful insights.

For customer relations management, Rosetta, a data-powered solution enabling clients to predict what customers want by turning data into narratives, can also be depicted as word clouds and grouped to provide meaning. This is achieved by displaying data related to past behaviour intentions, payment durations, etc. Histograms with a Gamma Fit depicting the probability distribution can be employed to predict matter portfolio duration. For example, the intensity of interactions with clients, eventually tapering off, can be predicted, and then overlaid over the actuals, to plan various levels of resourcing during the different stages of a legal matter. Knowledge of this would also assist the pricing analysts in pricing the matter and would therefore require marketing and accounting to work together. In some law firms, the pricing analysts are located within the different practice areas, as opposed to being seated in a central analytics team.

Finally, from a marketing perspective, BDA can also address the problem of pricing uncertainty in legal firms by providing transparency of the traditional billable hour delivery model (Dangel et al., 2018). Based on the data it is presented with, pricing analytics, or the analysis of bid data relating to the pricing and quality of legal services, enables management to, for instance, price a matter due to its complexity, chance of getting a good outcome, probable longevity, or risk profile (Agrawal et al., 2018).

Visualisation can also assist in the accounting practises and decision-making in law firms. Firstly, revenue streams and profit margins can be analysed. Understanding the drivers of profitability growth will help with identifying the business units with the highest margins or those requiring cost improvements driven by activity-based costs and billing analysis. Real-time alerts in the form of emails or visualisations using colour or other methods, drive actions required by prompting management to interpret the visualisations and translate their understanding into decisions. Scenarios could also be evaluated using highlighted threshold variances, and once the thresholds are breached can be depicted using colour or trigger alerts.

Secondly, visualisation assists accounting to analyse pricing strategy effectiveness and profitability information. Unified Task-Based Management system (UTBMS) and Legal Electronic Data Exchange Standard (LEDES) together, create codes for types of activities or tasks and types of expenses and provide standardised invoice formats to enable concise billing. This assists data analysis and pattern identification. Outside Counsel Guidelines (OCGs) relating to client compliance and billing requirements heighten e-billing complexity.

By predicting the likelihood of winning a case or its perceived complexity through analysis of the previous case points, pricing analysts can suggest pricing strategies appropriate to the case. Pricing analysts can be seated in either accounting or the various practice areas. In some law firms, they can also work closely with the marketing department.

Thirdly, visualisation can assist accounting in billing-related activities. Vague billing descriptions, listing multiple lawyers performing the same task or having lawyers performing unskilled tasks can result in bill rejections and a diminishing bottom line.

Dashboards displaying bill statuses, client reduction as well as appeal and rejection reasons, can lead to improved leverage, and billing and collection patterns being revealed and addressed. Octagonal or spider diagrams, with various distances from the center reflecting the different dimensions can be used to explain the shift from revenue to profitability to partners, by, for example, highlighting how write-offs, or disbursements (i.e., both hard and soft expenses) that exceed contractual levels are diminishing profits. The visualisations of successful partners, along the various dimensions, can be compared to those less successful, from a bottom-line perspective. Careful consideration needs to be given to the design of the various systems, because without the correct level of coding (e.g., in billing and time entry) the data will not render useful information during the analysis phase. More detailed coding in terms of time categories and project stages can be applied and tested in one practice area, and then rolled out to the wider firm.

Visualisation can also be used to enhance the performance of people and systems. Visualisation enables compensation models to be adjusted to encourage behaviour that is aligned with the firm culture when it is twinned with performance indicators underpinning what is valued. Through visualising trends, predictions are actionable by management. For example, if matters are increasingly complex and difficult to win then they must be priced accordingly. However, simpler matters must be transparently billed with the clients to avoid losing business. Utilisation of what-if scenarios is key as clients are becoming more demanding and will move to those firms that can provide transparency and motivate their actions with facts based on predictions. Finally, by analysing the use of various workflow processes, together with the IT department, management can detect bottlenecks, or dormant activity awaiting authorisations, that could perhaps be processed in parallel to avoid additional time.

5. Contingent resource-based theory

Having explained the concept of big data visualisation and illustrated its application in law firms, it is now possible to present a theoretical explanation for its potential to create a competitive advantage within law firms. This explanation takes the form of developing a conceptual framework that is based on contingent resource theory.

The resource-based theory emphasises the importance of the internal resources (financial, human, information and technology, management) of a firm for the establishing of core competence and thereby providing a competitive advantage to the firm. Resource-Based Theory (RBT) states that resources owned by the organisation differentiate performance levels and provide a competitive advantage (Barney et al., 2001). The ability of an organisation to collect, prepare and analyse its data differentiates them from others (Ferraris et al., 2019). The contingent resource-based theory is a variation of resource-based theory, which emphasises the contingent nature of resources. The contingent resource-based theory addresses the contextual insensitivity of Resource-Based Theory, highlighting capabilities emerging from the interrelationship between primary and secondary resources and focusing on the synergies between these resources, as they rarely act alone (Brandon-Jones et al., 2014).

Kozlenkova et al. (2014) define capabilities as the available resource within the firm, together with the techniques through which the tasks are performed. The contingent resource-based theory, therefore, assumes that situational or contextual variables affect the outcome. Ji-Fan Ren et. Al (2017) and Mikalef (2021) found that process-orientated dynamic capabilities and operational and dynamic capabilities act as mediators. Furthermore, Ferraris et al. (2019) found that BDA capabilities and firm performance are linked and that knowledge management mediates the relationship.

6. Primary and secondary resources enabling decision making and financial performance

The resource-based theory is built on the premise that organisations position themselves based on their resources and capabilities. Resources refer to what an organisation owns while capabilities refer to what an organisation can do. A resource-based model explains how an organisation's unique combination of resources and capabilities designed and utilised. Data Decision Making Systems are the primary resource enabling decision making, supported by two secondary capabilities. These secondary, or complementary resources, are their ability to execute (i.e., capabilities) and their completeness of vision (strategy and innovation). The resources increase competitive advantage through facilitating actionable decisions that ultimately contribute to firm performance, typically measured as revenue or profitability in financial terms or other non-financial terms.

Vincent et al. (2021) used evidence collected through vendor surveys, questionnaires, interviews, demonstrations, as well as sentiment analysis was undertaken by the Gartner Secondary Research team to categorize the LCAP's into four quadrants. The four quadrants were classified based on the two axes of execution and vision (Vincent et al., 2021). Execution relates to both product capabilities and implementation as well as sales execution and pricing; while vision relates to product strategy, innovation, and marketing strategy. The four quadrants are as follows: Leaders (good execution and vision), Challengers (good execution but less vision), Visionaries (good vision but less execution), or Niche Players (neither vision nor execution but good process automation).

The current operational capabilities include planning, investment, coordination, and integration; together with the potential capabilities driven by strategy and vision (i.e., technology management knowledge and business knowledge) (Vincent et al., 2021).

Cuthbert and Pearse (2021) used the Gartner classification and its quadrants to analyse legal firms. Firms featured as Leaders and Niche Players with most classified as Unknowns (i.e., the legal firms that appeared to have not yet implemented LCPs). These three categories were compared according to their total revenue, and it was concluded that legal firms that were implementing niche LCPs were most successful in enhancing their BDA and RPA capabilities. Furthermore, they concluded that Execution and Vision are both required for LCPs to be used successfully in the market.

7. Data visualisation as a mediator

It is argued here that the visualisation of data serves as a mediator between the primary and secondary resources, and actionable decisions, and firm performance. Data visualisation using driver-based modeling harnesses granular insights into actionable strategic decisions. The most obvious aim of visualisations is to communicate information, to shape the understanding and actions of users (Allen et al., 2014). These visualisations provide insights into the current and future, however, care must be taken not to highlight certain viewpoints as being more valuable than others. To illustrate, how many nodes can be reasonably presented in a network graph, and the practice of using colours like red, green, or blue to convey a sentiment immediately is called ‘simple-boxing’, which has sometimes been criticised for hiding complexity. Furthermore, clustering is said to convey the underlying patterns of the data, however, the number of dimensions displayed is subjective.

Highlighting what is interesting has also been called ‘Flatter-boxing’ and highlights differences. Rheindorf (2019) felt that interactive visualisation could act as mediator between human reasoning and artificial intelligence through interactive collaboration. Passive observation does not involve intervention and includes applications for feature maps or dream images. Interactive observation allows users to change the inputs enabling the observation of model reactions or analysis of learned model patterns. Humans can therefore correct or modify internal decisions but should never replace the expert. Visualisation tools can be applied to test hypotheses but lead the researcher to reject a hypothesis if the desired pattern is not displayed. Also, the test distribution could be different from the training distribution the model is trained on and therefore does not reflect the ‘real-world’. Real-time what-if analysis and simulations can be graphically depicted to be decoded by management.

Effective visualisation of data can enable the taking of actionable decisions. In this instance, visualisation is acting as an interdisciplinary communication and enabling the conversion from insight into action. This is supported by others such as Eberhard (Eberhard, 2021) who found that information visualisation improves decision accuracy and quality. Furthermore, Yildiz and Boehme (2017) found that visualisation directly impacts managerial decision-making.

8. Moderators of data visualisation

The use of primary and secondary resources by data scientists to produce data visuals is moderated by variables such as (1) the control versus self-reliance of the team, (2) the organization's values, (3) the task attributes, and (4) risk appetite.

Firstly, data scientists should be self-managing, self-organising, or autonomous to enhance output. This, however, implies a team goal or "why?" factor to ensure that the team is working toward a common goal, in a trusting culture that values the "new". To direct this research data scientists could be guided by an information plan outlining the collection, curation, and description and dissemination of their research.

The approach based on maximising effort, is one of the most pressing underlying problems, as the research problem, is also known in data circles as the "big idea". An autonomous environment, against the backdrop of a unified research question, ensures optimum output (Lopez et al., 2021).

Secondly, the organisation's values drive what people are rewarded for, and therefore change is not possible without rewarding both tangible and intangible behaviour (Harder, 2008). For example, remuneration can be linked to measured key performance indicators displayed on dashboards. Communicating what is desirable, is, however, considered to be the first step and indicates a transparent environment. An organisation needs to value the new, creative, or 'different', and risk-taking within safe parameters needs to be encouraged. Creativity and freedom were found to augment output within a data governance framework, communicated throughout the organisation. A recent report from Walker Sands found that 56% of marketers found that creativity and technology were equally vital to developing strategy (McKinsey & Company, 2020).

Thirdly, task attributes serve as moderators. For example, task complexity moderates (Eberhard, 2021) the visualisation of big data analysis. Task characteristics such as visual working memory (Tintarev & Masthoff, 2016) and numeracy (Honda et al., 2015) act as moderators enhancing or amplifying the value derived from the decision-making system

Data governance defines roles, responsibilities, and processes for ensuring accountability for and ownership of data assets across the firm. This includes a component of risk appetite (i.e., how much risk an organisation is willing to assume), together with a framework for dealing with risk management.

9. A conceptual framework for data visualisation

The potential use of data analysis and visualisation in legal firms has been illustrated. An argument has also been developed emphasising the centrality of data visualisation to the work of the data scientist as a mechanism to mobilise resources for effective decision making and action taking that contributes to firm performance. Several moderating variables have also been identified. This discussion is summarised in the following conceptual framework.

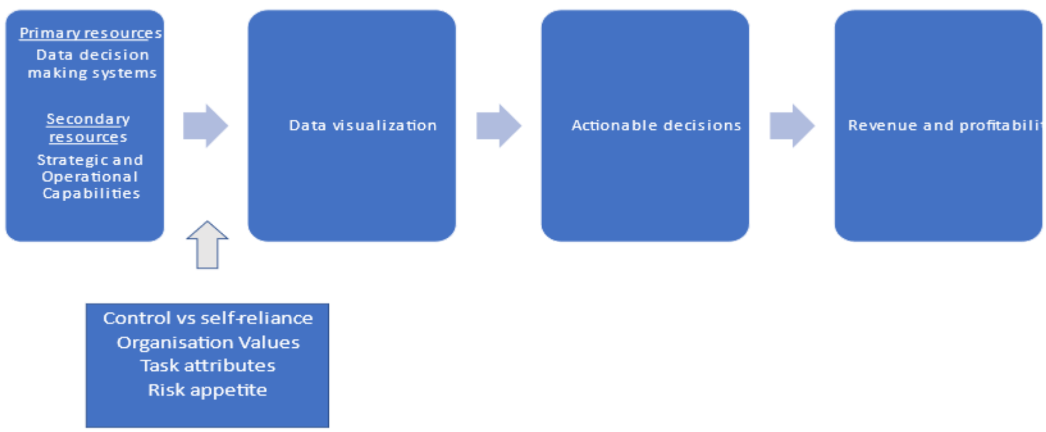

Figure 1. Data Visualisation Conceptual Framework

10. Conclusion

This chapter has explored the role of combining data decision-making systems with strategic and operational capabilities to facilitate data visualisation as a form of interdisciplinary communication between data scientists and decision-makers. Data visualisation enabled actionable decisions which ultimately contributed to firm performance. Furthermore, several moderators of the relationship between the primary and secondary resources and the data visualisation produced were identified.

The novelty of this chapter can be demonstrated in two major ways. Firstly, the chapter assesses the mediating role of data visualisation as a form of interdisciplinary communication to facilitate strategic management decision-making. Secondly, the chapter assesses the theory around data-driven decision-making systems and strategy to enhance competitive firm advantage and clarifies the understanding of the operational and strategic capabilities required for managing a firm and ultimately increasing revenue.

The conceptual framework that has been constructed through the review of the literature has several implications for interdisciplinary communication. Visualisation unites interdisciplinary research agendas providing clarity and integration throughout the firm, with the data scientist at the heart of the system. It has been noted that the various departments or sections within the firm have slightly different data requirements and visualisation preferences.

Data visualisation plays a role mediating the relationship between firm resources and profitability by revealing data patterns that provide insights leading to actions from management. There are several applications of this conceptual framework for the utilisation of data visualisation by decision-makers. Management needs to be aware of the moderators that amplify actionable decisions from data to make sure they create an enabling environment for data scientists and expert decision-makers, with enough freedom to pursue research agendas within data research planning outlines. The organisational values need to drive what is tangibly and intangibly rewarded through the dashboard and the use of KPI's. Cognisance of the task attributes needs to feed into the research agenda, as not all tasks are equally suited to visualisation. Finally, risk appetite for new research agendas needs to prevail within a risk management framework.

Given the conceptual framework that has been developed here, the following research recommendations are presented. The conceptual model needs to be tested in the field through a structured equation model. This could be supported by a case study. The research also needs to be widened to other decision-making and visualisation systems to enhance interdisciplinary communication. Future research could also look at the ethical angle of automated decisions including author bias and discrimination.

11. Acknowledgements

I would like to thank Prof D. Sewry (Rhodes University) and Prof K. Bradshaw (Rhodes University) for their peer-reviewing. Your time and insight were much appreciated. I would also like to thank T. Hayes who dedicated her time to beta-reading this chapter. Your fresh pair of eyes from a different perspective was much appreciated.

12. References

Agrawal, A., Gans, J., & Goldfard, A. (2018). Prediction machines: The simple economics of artificial intelligence. Harvard Business Review Press.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? International Journal of Production Economics, 182, 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Alarie, B., Niblett, A., & Yoon, A. (2017). How Artificial Intelligence will affect the Practice of Law. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3066816

Allen, P. M., Edwards, J. A., Snyder, F. J., Makinson, K. A., & Hamby, D. M. (2014). The effect of cognitive load on decision making with graphically displayed uncertainty information: Effect of cognitive load on decision making. Risk Analysis, 34(8), 1495–1505. https://doi.org/10.1111/risa.12161

Balfe, N., Sharples, S., & Wilson, J. R. (2015). Impact of automation: Measurement of performance, workload, and behaviour in a complex control environment. Applied Ergonomics, 47, 52–64. https://doi.org/10.1016/j.apergo.2014.08.002

Barney, J., Wright, M., & Ketchen, D. J. (2001). The resource-based view of the firm: Ten years after 1991. Journal of Management, 27(6), 625–641. https://doi.org/10.1177/014920630102700601

Becerra-Fernandez, I., & Sabherwal, R. (2010). Knowledge Management: Systems and processes. Sharpe.

Brandon-Jones, E., Squire, B., Autry, C. W., & Petersen, K. J. (2014). A Contingent Resource-Based perspective of Supply Chain resilience and robustness. Journal of Supply Chain Management, 50(3), 55–73. https://doi.org/10.1111/jscm.12050

Bresciani, S. (2019). Visual design thinking: A collaborative dimensions framework to profile visualisations. Design Studies, 63, 92–124. https://doi.org/10.1016/j.destud.2019.04.001

Cariceo, O., Nair, M., & Lytton, J. (2018). Data science for social work practice. Methodological Innovations, 11(3), 205979911881439. https://doi.org/10.1177/2059799118814392

Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A. V., & Rong, X. (2015). Data Mining for the Internet of Things: Literature Review and Challenges. International Journal of Distributed Sensor Networks, 11(8), 431047. https://doi.org/10.1155/2015/431047

Chiang, R. H. L., Grover, V., Liang, T.-P., & Zhang, D. (2018). Special Issue: Strategic Value of Big Data and Business Analytics. Journal of Management Information Systems, 35(2), 383–387. https://doi.org/10.1080/07421222.2018.1451950

Coombs, C., Hislop, D., Taneva, S. K., & Barnard, S. (2020). The strategic impacts of Intelligent Automation for knowledge and service work: An interdisciplinary review. The Journal of Strategic Information Systems, 29(4), 101600. https://doi.org/10.1016/j.jsis.2020.101600

Cuthbert, C. E., & Pearse, N. J. (2021). Strategic use of Low Code Platforms. International Conference on Education and Information Systems, Technologies and Applications (EISTA 2021), ISBN-Volume I: 978-1-950492-58-9

Dangel, S., Hagan, M., & Bryan, J. (2018). Designing today's Legal Education for tomorrow's lawyers: The role of legal design, technology, and innovation. https://doi.org/10.13140/RG.2.2.13776.56326

DeCanio, S. J. (2016). Robots and humans – complements or substitutes? Journal of Macroeconomics, 49, 280–291. https://doi.org/10.1016/j.jmacro.2016.08.003

Eberhard, K. (2021). The effects of visualization on judgment and decision-making: A systematic literature review. Management Review Quarterly. https://doi.org/10.1007/s11301-021-00235-8 Empson, L. (2021). Researching the Post-Pandemic Professional Service Firm: Challenging our Assumptions. Journal of Management Studies, 58(5), 1383–1388. https://doi.org/10.1111/joms.12697

Eraslan, G., Avsec, Ž., Gagneur, J., & Theis, F. J. (2019). Deep learning: New computational modeling techniques for genomics. Nature Reviews Genetics, 20(7), 389–403. https://doi.org/10.1038/s41576019-0122-6

Ferraris, A., Mazzoleni, A., Devalle, A., & Couturier, J. (2019). Big data analytics capabilities and knowledge management: Impact on firm performance. Management Decision, 57(8), 1923–1936. https://doi.org/10.1108/MD-07-2018-0825

Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics, 210, 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Han, J., & Kamber, M. (2012). Data mining: Concepts and techniques (3rd ed). Elsevier.

Harder, J. (2008). Organizational Reward Systems. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1281273

Honda, H., Ogawa, M., Murakoshi, T., Masuda, T., Utsumi, K., Park, S., Kimura, A., Nei, D., & Wada, Y. (2015). Effect of visual aids and individual differences of cognitive traits in judgments on food safety. Food Policy, 55, 33–40. https://doi.org/10.1016/j.foodpol.2015.05.010

Jha, K., Doshi, A., Patel, P., & Shah, M. (2019). A comprehensive review on automation in agriculture using artificial intelligence. Artificial Intelligence in Agriculture, 2, 1–12. https://doi.org/10.1016/j.aiia.2019.05.004

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. International Journal of Production Research, 55(17), 5011–5026. https://doi.org/10.1080/00207543.2016.1154209

Katz, D. M. (2013). Quantitative Legal Prediction-or-how I learned to stop worrying and start preparing for the Data-Driven future of the Legal Services Industry. Emory Law Journal, 62(4), 909-66.

Kerikmäe, T., Hoffmann, T., & Chochia, A. (2018). Legal technology for Law Firms: Determining roadmaps for innovation. Croatian International Relations Review, 24(81), 91–112. https://doi.org/10.2478/cirr-2018-0005

Kozlenkova, I. V., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. Journal of the Academy of Marketing Science, 42(1), 1–21. https://doi.org/10.1007/s11747-013-0336-7

Kyar Nyo Aye. (2013). A platform for Big Data Analytics on distributed scale-out storage system. https://doi.org/10.13140/RG.2.1.4760.5203

Legal Insider UK Top 200 2021. (2021). Top 200. https://legaltechnology.com/top-200/.

Lopez, C.-P., Aguilar, J., & Santorum, M. (2021). Autonomous VOs management based on industry 4.0: A systematic literature review. Journal of Intelligent Manufacturing. https://doi.org/10.1007/s10845-02101850-8

McKinsey & Company. (2020). How to unlock marketing-led growth: Data, creativity, and credibility. https://fredfarid.medium.com/data-creativity-markhttps://fredfarid.medium.com/data-creativitymarketings-brave-new-world-2f1dedacf003

Mikalef, P., van de Wetering, R., & Krogstie, J. (2021). Building dynamic capabilities by leveraging big data analytics: The role of organizational inertia. Information & Management, 58(6), 103412. https://doi.org/10.1016/j.im.2020.103412

Muayad, A., & Irtefaa, A. N. (2016). Ridge regression using Artificial Neural Network. Indian Journal of Science and Technology, 9(31). https://doi.org/10.17485/ijst/2016/v9i31/84278

Müller, O., Fay, M., & vom Brocke, J. (2018). The effect of Big Data and Analytics on firm performance: An Econometric analysis considering industry characteristics. Journal of Management Information Systems, 35(2), 488–509. https://doi.org/10.1080/07421222.2018.1451955

Munoz, J. M. (Ed.). (2017). Global Business Intelligence (1st ed.). Routledge. https://doi.org/10.4324/9781315471136

Neirotti, P., & Paolucci, E. (2007). Assessing the strategic value of Information Technology: An analysis on the insurance sector. Information & Management, 44(6), 568–582. https://doi.org/10.1016/j.im.2007.05.005

Obitade, P. O. (2019). Big data analytics: A link between knowledge management capabilities and superior cyber protection. Journal of Big Data, 6(1), 71. https://doi.org/10.1186/s40537-019-0229-9

Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: A cognitive framework across disciplines. Cognitive Research: Principles and Implications, 3(1), 29. https://doi.org/10.1186/s41235-018-0120-9

Rheindorf, M. (2019). Visualization, interactive visualization, and Open Science. In M. Rheindorf, Revisiting the Toolbox of Discourse Studies (pp. 223–253). Springer International Publishing. https://doi.org/10.1007/978-3-030-19369-0_6

Rivard, S., Raymond, L., & Verreault, D. (2006). Resource-based view and competitive strategy: An integrated model of the contribution of information technology to firm performance. The Journal of Strategic Information Systems, 15(1), 29–50. https://doi.org/10.1016/j.jsis.2005.06.003

Sackett, P. J., Al-Gaylani, M. F., Tiwari, A., & Williams, D. (2006). A review of data visualization: Opportunities in manufacturing sequence management. International Journal of Computer Integrated Manufacturing, 19(7), 689–704. https://doi.org/10.1080/09511920500504578

Shabbir, M. Q., & Gardezi, S. B. W. (2020). Application of big data analytics and organizational performance: The mediating role of knowledge management practices. Journal of Big Data, 7(1), 47. https://doi.org/10.1186/s40537-020-00317-6

Statista. (2021). Revenue from big data and business analytics worldwide from 2015 to 2022. Statista. https://www.statista.com/statistics/551501/worldwide-big-data-business.

Susskind, R., & Susskind, D. (2017). The future of the professions: How Technology will transform the work of human experts. Journal of Nursing Regulation, 8(2), 52. https://doi.org/10.1016/S21558256(17)30099-6

Tarafdar, M., Beath, C. M., & Ross, J. W. (2017). Enterprise Cognitive Computing Applications: Opportunities and Challenges. IT Professional, 19(4), 21–27. https://doi.org/10.1109/MITP.2017.3051321

Tintarev, N., & Masthoff, J. (2016). Effects of individual differences in working memory on plan presentational choices. Frontiers in Psychology, 7. https://doi.org/10.3389/fpsyg.2016.01793

University of Cambridge, Barrett, M., Davidson, E., University of Hawai'i at Mãnoa, Prabhu, J., University of Cambridge, Vargo, S. L., & University of Hawai'i at Mãnoa. (2015). Service innovation in the digital age: Key Contributions and Future Directions. MIS Quarterly, 39(1), 135–154. https://doi.org/10.25300/MISQ/2015/39:1.03

Verdinelli, S., & Scagnoli, N. I. (2013). Data display in qualitative research. International Journal of Qualitative Methods, 12(1), 359–381. https://doi.org/10.1177/160940691301200117

Vincent, P., Yefim, N., Kimihiko, I., Jason, W., Saikat, R., Akash, J., & Adrian, L. (2021). Magic quadrant for enterprise Low-Code Application Platforms. https://www.gartner.com/doc/reprints?id=12

Yildiz, E., & Bšhme, R. (2017). Effects of information security risk visualization on managerial decision making. Proceedings 2nd European Workshop on Usable Security. European Workshop on Usable Security, Paris, France. https://doi.org/10.14722/eurousec.2017.23010