

Elsevier Editorial System(tm) for International Journal of Information Management
Manuscript Draft

Manuscript Number:

Title: Business Intelligence in Magazine Distribution

Article Type: Case Study

Keywords: Business intelligence, information interpretation, small enterprise supply chains.

Corresponding Author: Dr Neil Kenneth McBride, PhD

Corresponding Author's Institution: De Montfort University

First Author: Neil Kenneth McBride, PhD

Order of Authors: Neil Kenneth McBride, PhD

Abstract: This case discusses the use of business intelligence systems in the running and optimisation of magazine distribution by a UK company. The company collects a wide range of data to help it monitor and optimise a supply chain involving subcontractors. The case study raises a number of issues which are discussed. It illustrates the variety of forces which are driving companies to adopt business intelligence systems. It demonstrates how business intelligence systems can help run business processes. It explores the problems and issues with sourcing, collecting and cleaning data. Issues around anonymisation and the concept of a 'single version of the truth' are discussed and ethical issues highlighted. It concludes that an understanding of the role of interpretation in data collection, collation and subsequent decision making is critical to business intelligence and calls for more research in this area.

Cover Letter

Dear Dr Hills,

Here is the final manuscript, submitted as a new submission as you requested in your email

Many Thanks

Yours sincerely,

Neil McBride, PhD.

Business Intelligence in Magazine Distribution

Neil McBride
Centre for Computing and Social Responsibility
De Montfort University
Leicester
United Kingdom

Email: nkm@dmu.ac.uk
Phone: +44 116 207 8500

Business Intelligence in Magazine Distribution

ABSTRACT

This case discusses the use of business intelligence systems in the running and optimisation of magazine distribution by a UK company. The company collects a wide range of data to help it monitor and optimise a supply chain involving subcontractors. The case study raises a number of issues which are discussed. It illustrates the variety of forces which are driving companies to adopt business intelligence systems. It demonstrates how business intelligence systems can help run business processes. It explores the problems and issues with sourcing, collecting and cleaning data. Issues around anonymisation and the concept of a 'single version of the truth' are discussed and ethical issues highlighted. It concludes that an understanding of the role of interpretation in data collection, collation and subsequent decision making is critical to business intelligence and calls for more research in this area.

Keywords:

Business intelligence, information interpretation, small enterprise supply chains.

1.Introduction

Business Intelligence (BI) and analytics concerns the organisation and processing of large databases in order to support decision making in the organisation. Sources of data can be both internal and external. Data can be accessed and processed in a distributed fashion, or gathered into an internal database. BI usually involves statistical software and analysis of trends. The purpose of BI has tended to concern market analysis and decisions on competition, customers and product placement. However, BI can also be used to study and improve business processes, This case study follows an example of such a use of BI.

BI has often been seen as the domain of large companies, accessing large distributed databases and massive sources of customer data, so-called big-data. Some organisations set up departments for BI, centred on a business intelligence competency centre (Laursen and Thorlund, 2010). However, BI can be used by smaller companies to help understand the processes which they manage. In the case study described here, the small company manages a group of suppliers and service providers to ensure the smooth and timely delivery of magazines to newsagent chains and supermarkets.

While BI is not a new idea and has been considered since the beginnings of business computing (Luhn, 1958), it has only recently become a major commercial concern for many companies. A number of factors have contributed to its rise in importance. There is a greater availability of data sources than have been the case in the past. The rapid expansion of large databases offers new opportunities to companies. Improved BI tools which do not require IT expertise are available. This encourages employees at all levels to explore the potential of business intelligence. Additionally, competitive and economic pressures require more efficiency in processes. BI offers a way of analysing operations and identifying improvement.

Finally, through media and trade journals, the profile of BI has increased dramatically in the last few years (Chen et al, 2012).

BI involves three major steps, extraction, transformation and load (ETL). These steps are not necessarily linear and depend on the development of a data design. The metadata model influences what data is extracted, but also is influenced by what data is available, and the nature of the questions being asked by employees and the processes being analysed. Extraction is inevitably a process of selection: selection of sources and data within sources. Data transformation involves resolving inconsistencies, validating data ranges, removing irrelevant data and checking for accuracy. Loading concerns populating the target BI database with the structured information. Each step involves interpretation and the imposing of structure on the data, driven by the goals and purposes of the BI team. The loaded BI target database is then made available to power users and general users through the application of intelligent tools for structuring queries.

This brief case study provides a platform for introducing and discussing the range of issues associated with business intelligence and business analytics. The presentation of the case study is followed by a discussion of key areas for reflection, and concludes with challenge to address the dangers of pseudo-objectivity which accompany BI.

2. Case Study

Magroup (pseudonym) is a joint venture between three magazine publishing companies which distributes magazines to retailers and accounts for 28% in value of the magazine market. Magroup manages the process entirely through outsourcing to a major carrier, wholesalers, finishing houses and retailers. Magroup does not own fleets of vehicles or warehouses, but effectively manages relationships and people. It uses the aggregate buying power of three magazine groups to generate economies of scale.

The magazine publishing industry in the United Kingdom is experiencing hard times. *‘For many years there was market growth. It was happy days, with things getting better every year’* (Business Intelligence Manager). A laissez-faire attitude predominated. In the last few years there has been a constant decline in magazine sales.

It was the realisation of the potential of the data daily generated for analysis that led to the development of business analytics at Magroup: *“Some years ago I read Tom Davenport’s Competing on Analytics (2007). I set that out as a vision of what we should be doing. This vision was set out and we have moved forward in the last four years.”* (BI Manager)

Additionally, business analytics offered the possibility of pursuing continuous process and performance improvements in a very competitive market: *‘Now the focus on every aspect of the business is much more intense, all the ins and outs of performance are examined. We have to be leaner and smarter, quicker to respond to market changes. Efficiency and smarter decision-making drive competitive advantage.’*

The process of magazine distribution starts with the printers. Magazines are transported from the printers to two major wholesalers. Some magazines are distributed via a finisher who packages the magazine with inserts and free gifts. The logistics are outsourced to Northern Logistics (pseudonym), which offers supply chain solutions, themselves driven by IT systems and involving the monitoring of Key Performance Indicators (KPIs). The wholesalers manage the logistics between warehouses and their own shops, newsagents, and supermarkets, in all over 50,000 outlets. Once on sale, shops and newsagents have 3-4 weeks to return unsold copies to the wholesalers. The management of this supply chain, split between the printer-to-wholesaler logistics and the wholesaler distribution, requires volumes of information to run and generates volumes of information ripe for analysis.

2.1 Sourcing the BI data

The development of a data warehouse as a basis for business analytics involved Magroup in gathering raw data from a number of sources at different points in the supply chain.

Data is available from the carrier, Northern Logistics, concerning deliveries between printers and wholesalers, including when magazines were picked up, when delivered into the wholesaler.

Data is collected from wholesalers for each magazine, for every store, for every issue: how many copies received, sold and returned. Information is obtained for the rest of the market including supermarkets and newsagents. Information is obtained about other sales which are disguised so that individual sales can't be identified. This information can be used to identify sales of the group's own magazines. However, common industry knowledge of how statistics are collected means that sales for individual magazines can be identified by reverse engineering the data. Some data is available for some retail outlets. Some retailers provide data from EPOS systems which identifies products by retail outlet. This data needs to be cleaned to remove data for non-magazine products by reference to the barcodes. Half a dozen retailers, including a large supermarket chain, provide EPOS data which is paid for through commercial agreements.

2.2 Managing the BI Data

Some data is cleaned using mature enterprise processes in the Extraction, Transformation and Load (ETL) System, which has referential integrity. Other data is less mature, remaining outside the production environment. Magroup are working on bringing all data under central control. Ten data analysts corral data, and manage the ETL processes. All data obtained is stored in an Oracle data warehouse. There is no selectivity at this stage.

Magroup are pursuing one central version of the data which, while open to interpretation, removes argument about what is the right data technically and moves on to consider meaning and consequence. *“If A and B look at different data, from very different sources, then all our time is spent discussing which one is right rather than what it means. If we agree one source, time periods etc., then we can get on with the business analysis. The point about the data is not whether it's right or wrong, but whether it's consistent”* (BI Manager)

The data model evolves over time, growing as needs change and different data sources are used. It is built in the data warehouse, rather than being derived from any corporate data model.

2.3 Reporting from the BI Data

Originally, Magroup used Microstrategy, connected to Business Objects This was historically the BI tool. The problem with Microstrategy is in its performance, particularly with large amounts of data. Magroup then moved over to SAS. *“Over the last four years we have developed reporting tools for users. Extracts from the data warehouse are imported to SAS, and data from SAS is provided in Excel worksheets for end users.”*

In 2009, a new architecture was proposed. This was seen as a *“paradigm shift”*. Analysts considered how data could be delivered differently.

Now the reporting tool is directly connected, directly interfaced to the data warehouse. The users will access a web interface. *“We are beginning to use Microsoft BI tools: although we are in an early stage of the learning curve. The surface for the user will be in Microsoft Sharepoint.”*

This will be a thin client solution, an OLAP type solution. For many end-users, they will just receive the results of the query. There are plans to deliver BI results to any device, including mobiles and tablets.

There is still a lot of manual intervention in the BI process. Data must be delivered to remote users. *“We are emailing reports or using FTP processes. It is all very inefficient; we can end up with duplicate data sources and multiple versions of the truth.”*

2.4 Users and Outcomes

There are two audiences, the internal front line and the publishers. The internal team have many different requirements, which rest in different reports being supplied to different parts of the business. Most users access pre-existing reports. Power users can write their own reports. Front-line users are mostly running pre-existing reports, Power users will include the commercial analysts who write bespoke queries

The prime KPI is market share. But each publisher and each function has its own targets, including minimising costs. One KPI concerns how much goes on sale on time. The measuring of the KPI and analysis of logistics helps Magroup to see where the process goes wrong. BI enables Magroup to examine the performance of their suppliers and can lead to negotiation with wholesalers. They can also examine the performance of the carriers and identify underperformance, and isolate problems by drilling down into the data.

SAS is used for forecasting and the allocation of sales up to five years hence. When a magazine comes on sale, the retailer has 3-4 weeks to return unsold copies. In that time EPOS data is used to predict day-by-day what is going to be sold. This is short-term forecasting.

A regression model is used to help publishers. Every magazine issue is different. Isolating factors are different. *“We can provide some insights, strip out underlying trends. Seasonal factors. Impact of a particular gift... is there anything you did differently? Outliers from the norm... things you can predict. It gives us extra awareness, about changing patterns of sales.”*

3. Case Analysis

3.1 Drivers towards BI

It is clear that a number of drivers contribute to the rise of BI. Magroup recognised the potential of the data generated daily by business processes. The availability of this was emphasised by the general management literature and the books written by the management gurus such as Tom Davenport. Newspaper inserts, articles in business journals and advertisements by providers of business analytics software and services contribute to a perception that BI is important and should be pursued

However, the publicity, the resources and the technology will only register with companies if an underlying need is already present, generated by economic and environmental forces. In the case of Magroup it is clear that the interest in BI derives from changing markets and economic forces which direct the company to examine processes and the use of resources in more detail than has previously been the case. The sense of security and happy times in the magazine industry is undermined by a decline in sales, influenced by the Internet and the move to electronic information sources as well as recession and austerity. It is these external forces and threats which put BI on the agenda, rather than the availability of the technology itself. BI provides Magroup with a means of being leaner and smarter in response to market changes. Hence the availability of data and systems does not guarantee its use unless there is a goal or purpose which it can fulfil. An understanding of the business environment and the economic and social forces active within it are required if the spread of technology is to be understood.

3.2 BI and Business Process

Although BI provides the information and tools for statistical analysis of large scale trends, the emphasis may be more frequently on the detail. Here it is the detail of sales of specific issues of individual magazines which is of concern. Also, it is the performance of individual suppliers which is of concern and the ability to drill down to specific occurrences of a process. While datasets may be accumulated in the large, it is the small individual aspects of specific processes occurring in a specific space, at a specific time that are of interest to the decision-makers.

Magroup is effectively a manager of relationships : with magazine suppliers, finishers, logistics companies, wholesalers, newsagents and supermarkets. BI offers a vehicle for managing those relationships and the underlying contracts and demanding improvements in process. BI enables key performance indicators to be examined at every point in the supply chain. There is a significant emphasis on the optimisation and control of processes in the

supply chain which, through economies of scale, drive the profitability and indeed survival of Magroup. This means that both significant decisions and micro decisions concerning the details of deliveries, right down to the background of specific deliveries is dependent on BI.

3.3 Information Sourcing

Data may be collected by as part of a business process, or entered into a system afterwards. In any case the motivation of the collector needs to be examined. If the data is being collected as part of a process, it may be difficult to get the data supplier to extract or format that data for the BI purpose. Data provision for BI may be peripheral to the data provider's interest and hence it may be difficult to get timely and reliable data.

In the case study, data is collected from the logistics company as part of the supplier relationship, from wholesalers and is bought from various sources. The data is both process data and sales data. The accuracy and motivation for generating the data will vary. Paying for data may not guarantee quality unless the provider understands the customer's goal for needed the data. Reliability is also an issue as well as accuracy.

However it should be noted that none of the data collected by Magroup is specifically generated for BI. Data gathered from process, sales, EPOS may be termed natural data which should provide a more realistic model of the real world than data collected specifically or solely for business analytics. The provision of data by Northern Logistics as part of a supplier contract results in an alignment of goals between the data provider and customer. In the BI application it is being use to optimise the magazine delivery and hence improve efficiency and adherence to the contract.

3.4 Information interpretation

Information interpretation occurs all along the supply line for business intelligence. Hence the effects of interpretation accumulate. Data selection is itself an interpretative step, where sources selected may depend on the questions being asked or the goals being pursued.

The cleaning of data and combining of sources is a further exercise in interpretation. Although Magroup did not select sources, it did remove records which were not of interest through an analysis of barcodes. The ETL process provides a further set of points at which interpretation occurs. The joining together of disparate sources, where keys might not match in order to impose referential integrity is inevitably a selective and interpretive process. Through the technical steps in deriving the BI database, decisions are being made by the BI team, driven by the goals as they understand them. This selection reduces the choices available to users in business analysis and may drive the decision making process in a particular direction.

However, key to the interpretive process are the attempts to meld such data together under a common metamodel may result in a loss of meaning, some simpler common denominator, or a bias towards one particular interpretation of meaning. The concept of a single view of the data is a key theme in business intelligence. In the case of Magroup this is an important consideration, which the BI manager justifies as necessary to curtail debate on content of data

and move the discussion to analysis. This case study suggests that the single version of the truth may emerge from technical considerations, rather than business analysis..

The phrase, 'a single version of the truth' suggests that there are multiple versions and a decision has been made to select one version. But whose version is it? And how is the version influenced by relationships between the BI department and users? The selection of a single version may impose an objectivity on the model and data which is neither real nor justified, but which then drives the subsequent agenda and the direction of the business, perhaps to the disadvantage of the business.

3.5 Information delivery

Business intelligence systems and their precursors such as executive information systems have often involved the extraction of information from the BI system or data warehouse into reports or spreadsheets. Here the centralised, global BI system based on a metamodel which should be used corporate-wide is transformed to localised dataset, contained in personal spreadsheets. These local datasets are then manipulated by the users prior to making decisions. Not only does this involve another layer of interpretation, but it undoes the work of the BI department through a move from consensus and group decision making to individual decision making and the creation of isolated islands of information.

The paradigm shift identified by the BI manager is a shift from a thick client where the user exports data from the data warehouse and then uses it in isolation to a thin client where data is maintained on the server and the reporting tools enable the user to develop reports and explore ideas without creating separate datasets which are held and maintained away from the BI systems. Such personal datasets would often drift away from alignment with the central system as users changed the structure, added their own data and began to maintain the dataset as a separate entity.

In terms of audience, it should be noted from the case study, that most users rely on ready prepared reports. Users may not have the time or inclination to develop their own reports, even when the tools are available. They are leaving the interpretation and analysis to the BI department. Users may have limited understanding of the data and treat it like a black box, where way the data source is constructed is an unknown. Only power users, who will be trying to understand the data and its derivation, are likely to ask crucial question about the construction of the data.

3.6 Ethics of BI

The ethics of business intelligence is an important area which needs development. Business intelligence and analytics result in actions which change organisations and impact customers. These actions have ethical consequences. While issues of privacy and anonymity are important, the ethics of BI will extend well beyond these issues. Critically, BI will impinge on the ethics of the business and social areas. Hence, there will be a need to understand the ethical issues associated with the business in which the BI is being used (McBride, 2012).

Additionally, the ethical use of BI will require transparency such that sources of data, and the way the data has been cleaned and arranged are clear to decision makers. Transparency is a prerequisite for ethical practice in information management (Floridi and Turilli, 2009). As the analysis of BI is frequently driven by key performance indicators (KPIs), the motivation and ethics behind the selection of KPIs also has to be considered.

The ethical risks of BI are not only concerned with the selection of a single version of the truth, but also with the use of BI for forward prediction. The use of business intelligence in prediction, by attempting to extrapolate from current information, produces a supply chain whose end product is not just reports about the present but predictions about the future. The problem with future predictions is that they are inherently unreliable in an uncertain world.

4. Lessons Learnt

This case study indicates some important issues that need to be addressed by both researchers and practitioners. Questions concerning how accurate BI data is and how well it models the real-life events need to be asked. If critical decisions are being made on the basis of reliance in the truth of the meta-model and the data it contains, then an unrealistic BI representation will result in wrong interventions. Information at the end of the supply chain, from which decisions are made, has been subjected to an number of processing steps and layers of interpretation. There may be some distance between the original data and the processed data on which decisions are being based. The BI supply chain requires cooperation from data suppliers.

The 'single version of the truth' concept particularly needs to be questioned. Whose version of the truth is acceptable? How do we decide to reject other versions? What is lost in terms of data and metadata in pursuing a 'single version of the truth'? In this case study the danger is that the selection of data is being influenced by the requirements of the technical architecture, which also influences how data is distributed and used and may affect decision making.

The case study underlines the importance of trust between the data provider and customer. Data received must be reliable. This requires that we trust the supplier not to falsify data in order to make delivery times look better. Such trust must then extend to the supplier's drivers. They must be trusted not to falsify tachometers or alter departure or arrival times. Trust in the competency, motivation and ethics of the data supplier will be key to BI. If the data is falsified, the analysis will be flawed and the resulting decisions wrong.

The case study highlights the additional important issue of anonymisation. Data is collected from some supermarkets where the identity of the publication is supposed to be hidden, but can easily be identified by reverse engineering. Here the anonymisation is ineffective and known to be ineffective. Adequate anonymisation will require aggregation to a point at which there is insufficient detail to infer specifics. However, the value of business intelligence lies in the ability to extract detailed and unexpected connections and to 'drill down' into the details of magazine distribution and identify concerns with specific magazines and deliveries. Anonymisation will work against this; so it is inevitable that anonymisation will be resisted or skirted round.

5. Conclusions

The purpose of this case study has been to identify some themes and concerns associated with the practice of business intelligence and analytics. The case study illustrates the value of BI in managing a supply chain and its value outside the confines of large companies and massive databases. The analysis of data affects relationships between businesses and customers. But BI carries the risk of mediating a relationship solely by quantitative values and disregarding qualitative and contextual factors.

Creating a single version of the truth introduces an unjustified certainty as to the nature and accuracy of the data and may suppress other views. In a management situation, consensus about what is the right and accurate data may be difficult to achieve. Multiple worldviews cannot easily be reconciled. The danger of BI is that data which is inevitably a flawed and dynamic model of the world is treated as facts which are considered more real than the actual world they are trying to describe. Data is then irrefutable, unchallengeable and 'scientifically accurate'. Decisions are then considered scientific and objective back up by business intelligence 'facts'.

Business intelligence and analytics involves a data supply chain. Data drawn from raw sources, is selected, cleaned, treated and integrated. The structuring of the data then enables reports to be provided to consumers. The supply chain involves a number of organisations, each contributing not only value to the product, but meaning. Hence layers of meaning accrete which are embedded in the data. The influence of the meanings attached by stakeholders in the supply chain may not be apparent to the consumer reading the reports and trying to make decisions based on them. Information provenance and transparency are really an essential requirement in business intelligence. In the same way food should be traceable to the farm that supplied it, so data should be traceable, and the additives clearly stated.

References

Chen, H., Chaing, R.H.L. & Storey, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, **36**, 1-22.

Davenport, T.H. & Harris, J.G. (2007). *Competing on Analytics: The New Science of Winning*. New York : Harvard Business Press.

Laursen, G.H.N. & Thorland, J. (2010). *Business Analytics for Managers*. Chichester : John Wiley.

Luhn, H.P. (1958). A Business Intelligence System. *IBM Journal of Research and Development* **2**, 14-31.

McBride, N. (2012). The ethics of software engineering should be an ethics of the client. *Communications of the ACM*, **55**, 39-41.

Turulli, M & Floridi, L. (2009). The ethics of information transparency. *Ethics and Information Technology* **11**, 105-122.