

DATA SOURCES FOR BUSINESS INTELLIGENCE

Naveen Kunnathuvalappil Hariharan

Sr. Hyperion SME & Department of Information Technology, United States

ABSTRACT

As organizations' desire for data grows, so does their search for data sources that are both usable and reliable. Businesses can obtain and collect big data in a variety of locations, both inside and outside their own walls. This study aims to investigate the various data sources for business intelligence. For business intelligence, there are three types of data: internal data, external data, and personal data. Internal data is mostly kept in databases, which serve as the backbone of an enterprise information system and are known as transactional systems or operational systems. This information, however, is not always sufficient. If the company wants to answer market and industry questions or better understand future customers, the analytics team may need to look beyond the company's own data sources. Organizations must have access to a variety of data sources in order to answer the key questions that guide their initiatives. Internal sources, external public sources, and collaboration with a big data expert could all be beneficial. Companies who are able to extract relevant data from their mountain of data acquire new perspectives on their business, allowing them to become more competitive.

Keywords: business intelligence, data sources, external sources, internal sources

INTRODUCTION

Data is undeniably important in today's business world. Businesses are analyzing a seemingly infinite number of data sources to gain insights into virtually every activity – both internal and external to their organization. Right now, it appears that businesses can't get enough big data for analysis — the potential and benefits on offer are tantalizing, as well as actionable, and can make all the difference for today's businesses.

However, as businesses' appetite for data increases, so does their search for usable data sources. Businesses can obtain and collect big data in a variety of places, both inside and outside their own enterprises. There are currently a number of data brokers who will provide information lists – and while these may appear to be useful, it is up to the organization to examine and make the most use of this data.

Thousands of transactions are tracked and recorded every day by most businesses. Not only customer purchases – which could include information like the customer's name, the products/items sold, the shop from which the purchase has been made, and the date and time of the purchase – but also warehouse activity, inventory transactions, workplace hours and time off, and daily operating costs (Rausch et al. 2013; Grossmann and Rinderle-Ma 2015).

In reality, most businesses are practically drowning in data. If only there was a way to gather all of this information in one place and use a simple report to make sense of it all (or set of reports). Companies who are able to extract relevant data from their mountain of data acquire new perspectives on their business, allowing them to become leaner and more competitive.

Organizations are in the process of gathering, organizing, and analyzing business data in order to turn it into meaningful and actionable information. Organizations can use business intelligence to gain a better understanding of their organization, which can lead to new opportunities, improvements to existing operations or processes, competitive advantages, and more (Chen et al. 2012). For example, they can:

- Identify top-selling products by region, store, or salesperson
- Identify trends, both good and bad, early on
- Create ad-hoc financial reports
- Track competitors in their ad-hoc financial reports

Business Intelligence (BI)

The act of turning data into information and ultimately into knowledge is known as business intelligence (BI). Customer needs, customer decision-making processes, competition, industry conditions, and general economic, technological, and cultural trends are all common sources of data (Sherman 2014). In the early 1990s, business intelligence (BI) was established in the industrial sector in response to managers' requests for

more efficient and effective enterprise data analysis in order to better understand their organization's status and improve decision-making.

Business intelligence enables companies to make well-informed business choices, giving them a competitive advantage. This is especially true when businesses can extrapolate data from external variables and make accurate predictions about future trends or economic situations. Firms can make decisions that benefit them after business intelligence is collected properly and used proactively (Sabherwal and Becerra-Fernandez 2013).

The major goal of business intelligence is to help companies make better decisions. Business intelligence reveals the following:

The company's performance compared to its competitors Customer behavior and buying patterns have changed. Market dynamics, future trends, demographic and economic data; the firm's capabilities the social, regulatory, and political context are all important factors to consider.

Businesses recognize that in today's highly competitive, fast-paced, and ever-changing business world, the speed with which they respond and adapt to change is a critical competitive quantity. Business intelligence enables them to use information obtained to respond to changes rapidly and continuously.

Gathering, preparing, and evaluating data are the most important tasks. The data must be of excellent quality. The data is collected, converted, cleansed, loaded, and stored in a warehouse from numerous sources. The relevant data is pulled from the data warehouse for a given business area. As it moves through multiple phases of informational metamorphosis, a BI organization fully uses data at every phase of the BI architecture. Raw data is created in operational environments, where transactional data is poured in from all corners of the company. As a result, this is the vision of a business intelligence organization: From origin to action, data flows naturally. Furthermore, the data is fully exploited at each step in the flow to ensure that the enterprise's information value grows. Of course, building any organization's vision is a challenge for BI.

A sensible approach to a continuous improvement loop that includes BI is (Maheshwari 2014): 1. obtaining information 2. Taking decisions and actions based on that information 3. Measuring the success of the project using preset metrics (a fancy phrase for measures). 4. Taking the lessons learned from one decision and applying them to the next.

Any part of an organization can be involved in the process of using data to make better decisions. If operational data can teach us anything, whether it's about consumer behavior, financial data, or something else, BI can help. A team may make better judgments by applying BI practices to transform raw data into relevant insights. The actions made in response to those decisions provide a new set of results, which can be fed back into the system as new empirical evidence for drawing the next set of conclusions.

Companies that use BI reap a slew of benefits. It can help companies respond quickly to changes in financial conditions, customer preferences, and supply chain operations by eliminating a lot of the guesswork within the organization, improving communication between departments while coordinating activities, and allowing organizations to react quickly to changes in financial conditions, customer preferences, and supply chain operations (Grossmann and Rinderle-Ma 2015; Liebowitz 2013). BI boosts a company's overall performance. Information is frequently recognized as a company's second most precious resource (its most precious resources are its people). As a result, when a corporation can make decisions based on current and reliable data, it may improve its performance. BI also speeds up decision-making because acting fast and correctly on information before competitors can frequently lead to competitive advantage. It will also improve the customer satisfaction by enabling prompt and appropriate responses to customer issues and priorities.

In the mid-1990s, academics grew interested in BI, and 10 years of research transformed a collection of basic procedures into a well-founded approach to data extraction and processing.

Sources of business intelligence data

According to the literature, there are three main types of data that enter into a data warehouse: data cane internal data, external data, and personal data.

1) Internal data.

Companies can mine data in a variety of ways from internal sources. These are some of them:

a) Transactional data and POS information: A company's financial and transactional systems contain one of the most powerful sources of data. Companies can use this data to harvest both current and historical data about their own business transactions, as well as information about their consumers' shopping habits. An organization can gain valuable insights from these facts, such as methods to cut costs and stay on budget, as well as key patterns relating to their consumers' purchasing habits and preferences (Sabherwal and Becerra-Fernandez 2013).

b) Customer relationship management system: Businesses can mine data within their CRM systems in addition to their purchasing and shopping data. Client affiliations, localities, and other regional or geographical characteristics might help get a better idea of where the customers are located. These CRM details become much more effective when paired with their transactional data (Todman 2001).

C) Internal records: A company's own internal documents are becoming more valuable than ever, especially in the age of cloud computing. Internal forms that have been digitized can be a valuable source of information, especially when it comes to the company's activities (Shroff et al. 2011), regulations, and procedures. Emails, Word documents, PDF, XML, and a variety of other internal documents may all be mined for big data, according to an infographic from Kapow Software.

D) Archives: When it comes to internal information, companies should not limit themselves to merely the most recent data. Historical data can also be quite telling, which is why Kapow Software suggests searching through the company's archival documents and data streams (White 2005).

E) Other business applications: While CRM is one of the most reliable internal sources of big data, this does not ignore the use of data mining of other internal applications. Other platforms used by employees, including project management, marketing, efficiency, enterprise asset management, human resources, cost management, and automation apps, can also be quite valuable (Khedr et al. 2017). It's in a company's best interest to let the nature of their big data endeavor guide their decisions on which sources to use while mining these sources. For example, if a company wants to learn more about its present budget, tools like spending tracking and resource management will come in handy.

f) Device sensors: The Internet of Things is expanding every day, bringing with it more and more unique data to analyze. Companies that use gadgets with sensors and network connectivity can take advantage of this data as well (Luić and Savić). These include IoT devices that the business employs in its own offices as well as those that it sells to customers. For example, automotive sensors installed in a company's fleet of vehicles can provide a lot of information regarding usage, mileage, gas consumption, and trip costs. Companies that sell fitness or other health monitors can also collect, anonymize, and analyze this data.

2. EXTERNAL DATA

We looked at internal data sources, such as transactional data, CRM details, business apps, and other company-owned assets. These sources are already under the authority of the company, so they're among the first places data scientists go when obtaining information (Bures et al. 2012). However, this information isn't always sufficient. Whether the company wants to answer wider questions about the industry or better understand future customers, the analytics team may need to explore beyond the company's own data sources. We'll look at the most popular external data sources, including publicly available information that isn't controlled by the company.

a) Social media

Social medias like Facebook and Twitter are among the most reliable external big data sources. These websites have grown extremely popular, not only among individual customers but also among businesses. Businesses may put an ear to the ground, so to speak, through social media profiles and gain a better understanding of their present and potential clients (Muntean et al. 2014).

Overall, social media data may reveal a lot about a company, including good and negative brand feedback, trends, activity patterns, and client preferences. For example, if a business detects that a large number of social media users are looking for a specific type of product, it can move to dominate the market and meet these demands.

b) Data from the government

While social network data is undoubtedly valuable, it isn't the only external data source that businesses should use. The federal government also supplies useful informational resources to assist today's businesses in gaining a better understanding of the public ([Albescu et al. 2008](#)). Some of the greatest areas to look here are ([Marr 2016](#)):

i) Data.gov: This website was just launched by federal officials in order to fulfill the US government's goal to make as much data as possible public. The best part is that all of this information is available for free and on the internet. Companies will find a variety of data here, including consumer data, agricultural data, education data, manufacturing data, public safety data, and much more. Businesses searching for a broader global view might visit Data.gov.uk, where the UK government has gathered an extraordinary quantity of metadata dating back to 1950.

The United States Census Bureau: It has also made a variety of statistics available online, including information on the general population, geographic information, and regional education details.

ii) CIA World Factbook: The Central Intelligence Agency has vast warehouses of data at its disposal, and it has made some of it available online through its Factbook. This database contains information about the world's population, governance, military, infrastructure, economy, and history. The best part is that it not only covers the United States, but also 266 other nations.

Healthdata.gov: Health-care data can be extremely useful for businesses in that area, as well as those in other industries. This site contains more than a century's worth of health-care data from the United States, including Medicare records, population statistics, and epidemiology.

iii) Google

Google has also made a few critical data sources publicly available. Google, being one of the world's most popular search engines, has a plethora of data about search words, trends, and other online activity. One of the better resources here is Google Trends, which provides statistical data on search volumes for practically any phrase — and these databases date back to nearly the start of the internet.

iv) Google Finance. It provides 40 years of stock market data that is constantly updated in real time, is another useful resource supplied by Google. Furthermore, Google Books Ngrams enables businesses to search and analyze the text of Google's millions of books.

There are also other external data sources that can be used to supplement the vast amount of information kept in internal databases. Some authorities, for example, collect and make data on sales data, market share data, and future trend predictions for specific business industries, as well as economic and financial indicators, available to the public. Market surveys and customer opinions obtained through questionnaires are provided by other agencies. Geographic information systems (GIS) ([Azaz 2011](#)), which are a set of applications for gathering, organizing, storing, and presenting territorial data, are another important source of external data. These contain data about entities that have a defined geographic location. Each item is thus linked to latitude and longitude coordinates, as well as a variety of additional attributes, which are often derived from relational databases and vary depending on the application domain. As a result, these data enable subject-specific studies of data connected with geographic components, with the results visually presented.

3. PERSONAL DATA

In most cases, decision makers doing a business intelligence study rely on data and personal evaluations kept in workbooks or local databases on their computers. One of the goals of knowledge management systems is to retrieve such data and integrate it with structured data from both internal and external sources.

CONCLUSION

A data source is the location from which information is gathered. Data can be mined by businesses in a variety of ways, both internal and external. One of the most powerful sources of data is transactional data and POS information. Internal records are more valuable than ever before, particularly in the age of cloud computing. Businesses can use this information to learn more about their own business transactions and purchasing habits. Transactional data, CRM details, business apps, and other company-owned assets can all be collected. Internal data is mostly maintained in databases, which are the backbone of an enterprise information system and are referred to as transactional systems or operational systems. Internal data is collected via transactional applications, which are used to manage a company's activities, such as administration, accounting, production, and logistics. Enterprise resource planning refers to a group of transactional software programs (ERP). Customers, products, sales, staff, and suppliers are among the primary entities involved in a company's processes, and the data contained in operational systems frequently deals with them.

Social media platforms like Facebook, Instagram, and Twitter are among the most trustworthy external data sources. The federal government also offers useful informational resources to help businesses better understand the public. Some of the most useful data sources in the world are Data.gov, Healthdata.gov, and the CIA World Factbook. Businesses can now access a variety of critical data sources thanks to Google. The Census Bureau collects data on the general population, geographic areas, and educational attainment.

It's vital to guarantee that all of a company's internal data sources are mined, examined, and utilized for the company's benefit before decision-makers and data scientists hunt for external sources. While external data has its advantages, internal data sources are usually easier to obtain and more useful for the company's own aims and insights.

REFERENCES

- 1) Albescu, Felicia, Irina Pugna, and Dorel Paraschiv. 2008. "Business Intelligence & Knowledge Management--Technological Support for Strategic Management in the Knowledge Based Economy." *Revista Informatica Economica* 4 (48): 5–12.
- 2) Azaz, Lotfy. 2011. "The Use of Geographic Information Systems (GIS) in Business." In *Int. Conf. Humanit*, 299–303. mlsvc01-prod.s3.amazonaws.com.
- 3) Bagambiki, Erika. 2018. "Enterprise Data Warehouse and Business Intelligence Solution." In *Proceedings of the 11th International Conference on Theory and Practice of Electronic Governance*, 665–66. ICEGOV '18. New York, NY, USA: Association for Computing Machinery.
- 4) Bures, Vladimir, Tereza Otcenaskova, and Veronika Jasikova. 2012. "The Evaluation of External Data Resources for Business Intelligence Applications: The Example of the Czech Republic." *Journal of Systems Integration* 3 (1): 32–44.
- 5) Chen, Hsinchun, Roger H. L. Chiang, and Veda C. Storey. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact." *The Mississippi Quarterly* 36 (4): 1165–88.
- 6) "Competitive Business Intelligence Gathering and Analysis." 2000. *Nature Biotechnology* 18 Suppl (October): IT5–6.
- 7) Grossmann, Wilfried, and Stefanie Rinderle-Ma. 2015. *Fundamentals of Business Intelligence*. Springer, Berlin, Heidelberg.
- 8) Khedr, Ayman, Faculty of computers and information systems, Information systems dept., Helwan University, Cairo, Egypt, Sherif Kholeif, Fifi Saad, Faculty of computers and information systems, Information systems dept., Helwan University, Cairo, Egypt, and Faculty of computers and information systems, Information systems dept., Helwan University, Cairo, Egypt. 2017. "An Integrated Business Intelligence Framework for Healthcare Analytics." *International Journal of Advanced Research in Computer Science and Software Engineering* 7 (5): 263–70.
- 9) Liebowitz, J. 2013. "Big Data and Business Analytics."
- 10) Luić, Ljerka, and Goran Savić. n.d. "BUSINESS INTELLIGENCE IN MANAGING OF TECHNICAL-INFORMATION SYSTEM." Bib.irb.hr. https://www.bib.irb.hr/949165/download/949165.ICIL2016_POSTER_-_Lui_Savi.pdf.
- 11) Maheshwari, Anil. 2014. *Business Intelligence and Data Mining*. Business Expert Press.

- 12) Marr, Bernard. 2016. "Big Data: 33 Brilliant And Free Data Sources Anyone Can Use." Forbes Magazine, February 12, 2016. <https://www.forbes.com/sites/bernardmarr/2016/02/12/big-data-35-brilliant-and-free-data-sources-for-2016/>.
- 13) Muntean, Mihaela, Liviu Gabriel Cabău, and Vlad Rînciog. 2014. "Social Business Intelligence: A New Perspective for Decision Makers." *Procedia - Social and Behavioral Sciences* 124 (March): 562–67.
- 14) Rausch, Peter, Alaa F. Sheta, and Aladdin Ayesh. 2013. *Business Intelligence and Performance Management: Theory, Systems and Industrial Applications*. Springer Science & Business Media.
- 15) Sabherwal, R., and I. Becerra-Fernandez. 2013. "Business Intelligence: Practices, Technologies, and Management."
- 16) Sherman, Rick. 2014. *Business Intelligence Guidebook: From Data Integration to Analytics*. Elsevier Science.
- 17) Shroff, Gautam, Puneet Agarwal, and Lipika Dey. 2011. "Enterprise Information Fusion for Real-Time Business Intelligence." In *14th International Conference on Information Fusion*, 1–8. ieeexplore.ieee.org.
- 18) Todman, Chris. 2001. *Designing a Data Warehouse: Supporting Customer Relationship Management*. Prentice Hall PTR.
- 19) White, Colin. 2005. "The Next Generation of Business Intelligence: Operational BI." *New York* 15 (5): 34.