

# Search-Based BI

Getting ready for the next wave in Business Intelligence



This paper addresses the next wave of innovation in Business Intelligence. Business Intelligence (BI) has been around for years now, and we would argue that since the introduction of the Data Warehousing concept in the early '90s not much has changed. Innovations in consumer electronics, entertainment, and communications have been revolutionizing the day-to-day lives of millions of people. Through the Internet, consumers have unlimited access to information and tools. Social networks, messaging and new media have changed the way people communicate and exchange information. On the other hand BI has seemed to avoid any disruptive innovation. In this article we will describe what might be the next wave of innovation in BI. We will point out why BI needs to become user friendly and agile to fulfill its potential. Some hints of these new features already exist today. Search-based BI is one of those interesting new innovations capable of changing the landscape of tomorrow's BI.

### Introduction

When Nigel Pendse, author of the famous OLAP (On-Line Analytical Processing) report and industry watcher since the early days, was asked in a recent interview about his thoughts on the level of innovation in the BI industry, he gave the following answer:

*“Actually, I’m disappointed by the level of innovation in the BI tools market, even before the industry consolidation, which will certainly make it even worse. Many of the claimed ‘innovations’ are actually rehashed versions of ideas that have been around for ages. For example, in-memory was the original BI architecture from more than 40 years ago, and column-oriented databases are almost as old. Dashboards, too, date back more than 25 years, to the brief EIS wave. But there’s certainly been a lot of consolidation, which is probably bad news for customers, as they’ll be paying higher prices for products that don’t progress as much as they would otherwise have done (or which may be discontinued altogether). The new owners of the BI products are much more interested in integration between their many products, than innovation in non-core products. The products are also likely to get more technical to install and implement, which may be good for consultants, but not for users. I suppose the best opportunities will come from smaller, creative new vendors who take advantage of the likely mismanagement of the products acquired by large vendors. Many of the BI products now owned by the large non-BI vendors will fail to move forward or will not be promoted aggressively, thus leaving gaps in the market for nimbler vendors.”<sup>[1]</sup>*

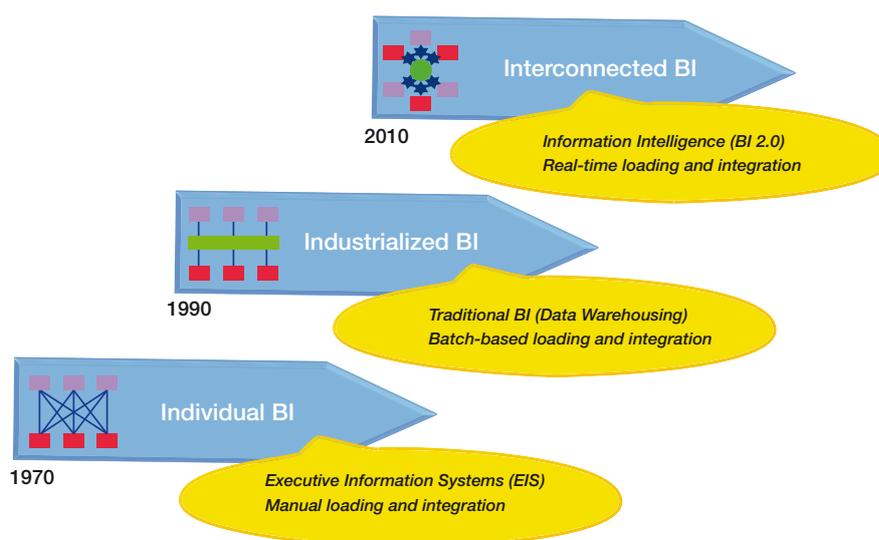
In the following paragraphs we will discuss some new innovations in the BI domain. But before we start discussing some groundbreaking innovations, we will put things in an historical perspective to get a better understanding of both past and future developments in BI.

### The beginning: Executive Information Systems

The evolution of BI in large organizations goes back to the 1970s. In an increasingly competitive and global environment, business managers were looking for tools to support their decision making processes. The early BI tools were focused on extracting data from source systems and on delivering reports displaying performance indicators. The tools used in those early days of BI were most of the time custom-made applications developed by internal IT specialists.

To satisfy the needs of a growing number of business managers, specific queries were integrated in the overnight batch and launched against the production systems. The objective was to get business information out of the production systems in the form of fixed-format standard reports, the so-called “print-outs”. On a regular basis, printed information was manually aggregated and keyed into presentation templates and data sheets. Some years later, the concept of the “Information Support Database” was introduced to offload querying on the transaction systems and to improve the performance of the overall solution [2].

In response to the growing need for management support and reporting tools, software vendors like Pilot Software, Information Resources, and Comshare jumped at the opportunity. The first generation of BI tools is often identified with the term EIS (Executive Information Systems). The early BI tools included ETL (Extract Transformation Load) capabilities, merged data from multiple sources, used relational databases, included what we later called Star-schemas, and built cubes for fast data retrieval.



### The second wave in BI: Data Warehousing

In the early 1990s, the EIS pioneers fell on hard times. The costs of implementing corporate EIS systems were too high. The required technical infrastructure wasn't there, so the EIS tools had to include their own. In addition, EIS didn't target and serve enough end-users because of the “Executive” connotations.

At the same time, new innovations like Data Warehousing and OLAP began broadening the realm of decision support and initiated a larger category of BI tools. The so-called “Data Warehousing” model was further popularized as a means to describe a new set of concepts and methods to improve decision making by using fact-based decision support systems. New industry leaders like Bill Inmon (“The father of the Data Warehouse”) [3] and Ralph Kimball (“The father of the Star-schema”) [4] actively promoted the Data Warehouse by using relational and multi-dimensional database technologies.

During the second wave of innovation in BI, the production of management information was being industrialized by means of sequentially scheduled batch-processes (Information Logistics). The entire production process, from the extraction of source data to the generation of reports, was being automated by means of specialized BI tools. The Data Warehousing model, as introduced in the early 1990s, has shaped the BI landscape ever since. Today, because of its proven concepts and technologies, the traditional BI model (ETL, Data Warehouse, Reporting, etc.) is still the guiding principle for designing new BI architecture in large organizations.

Despite its success during the past two decades the Data Warehousing model is nowadays being challenged by new concepts and technologies. Modern business managers are pointing to the shortcomings and drawbacks of the current model both from an organizational and structural point of view. In other words: the Data Warehousing model as we know it has become too complex and expensive to maintain and too rigid to provide the required speed of decision needed in today's 24/7 economy.

Developing a traditional multi-layered BI system is an expensive and labor-intensive exercise. To design and build interfaces, ETL jobs, Star-schemas, Data marts, reports, etc. takes a lot of time. In addition, highly qualified experts from various disciplines are required to deliver and build a new version on time. Delivery cycles ranging from 6 to 12 months are not an exception because of the various teams and tools involved.

### **The next wave in BI: Information Intelligence or BI 2.0**

Since the introduction of the Data Warehouse in the early 1990s, the developments in BI seem to lack any disruptive innovation. To fulfill its promise and to respond to future requirements, BI needs to become more intelligent, user friendly and flexible. Clearly, BI has not fulfilled its promises yet. Just ask yourself the following questions:

- How good is your BI solution in predicting the future?
- How intelligent is your BI in anticipating events?
- How much does it cost to maintain your BI?
- How good is the return on investment on your BI?
- How agile and flexible is your BI system?
- How good is your BI system in supporting end-users?
- How easily does your BI provide access to relevant business information?
- How rapidly and cost efficiently can you add new data to your BI system to adapt to changing market conditions?

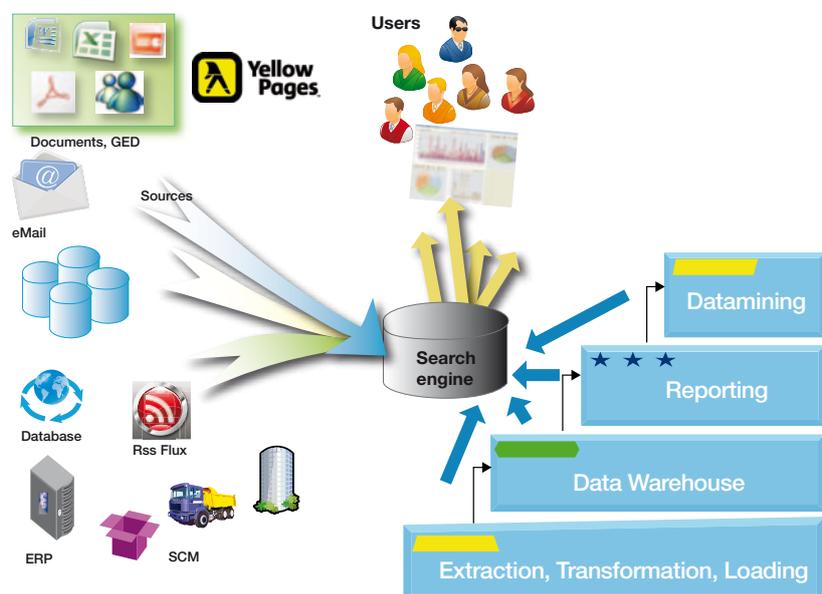
Today's BI, based on the Data Warehousing model, is lacking some very basic features and functionality. Adding another BI tool will only increase complexity and costs and is therefore not an appropriate solution. We need to reconsider the basics of the current model and identify areas and technologies with the potential to improve things structurally. While there are precious few references that paint a vision for the future of BI, there have been some authors that describe the next wave. [5, 6]. The following list describes some improvement areas:

- Artificial Intelligence (AI)/Predictive Analytics
- Pro-active alerts and notifications
- Event-driven/Real-time access to information
- Integration of non-structured and external information
- Enterprise Integration/Closed loop BI
- Portal integration/Mobile/Ubiquitous access
- Improved visualization/Rich Interfaces (RIA)
- Management Automation/Decision Engines
- Collaborative tools to leverage collective intelligence.

The BI of the future is becoming the brain and the central nervous system of organizations. Management information doesn't find itself locked in a Data mart or in a management report any more; instead it is automatically being re-injected in operational source systems to adapt to ever changing market conditions. In this sense BI is becoming a critical business function interconnected and at the core of the enterprise information architecture. The next wave in BI, Information Intelligence, will be the lifeblood of organizations.

Information Intelligence, or the intelligent use of information, is a new vision and a fundamental shift in the way businesses use information. The Data Warehouse has enabled significant advances in the integration and use of information but its underlying architectural approach is now being questioned. Information Intelligence extends BI beyond the traditional Data Warehouse and query tools to include automated decision-making and real-time/event-driven technologies. In addition, Information Intelligence is about building smarter business processes and making BI more user friendly and flexible.

In the following paragraphs we will discuss one of the technologies we believe is capable of transforming future BI architectures, i.e. Enterprise Search Engines. Based on some recent integration projects for our clients we found that Enterprise Search Engines have the capacity to simplify and improve BI in large organizations.



### **Search-based BI or how to make BI user friendly**

In the BI space we see more and more evidence that Search Engines are revolutionizing traditional BI architectures based on Relational Database Management Systems (RDBMS) and Multi-Dimensional Database Management Systems (MDBMS).

#### **There are multiple reasons for this:**

1. Search Engines are very flexible in handling any format and type of information, be it structured, unstructured or external. This is in contrast to the more rigid single structured view of data within Data Warehouses.
2. Search Engines are able to cope with continuously evolving data structures. Indexing both existing and new data does not require extensive data-modeling. This in contrast with the modeling of the Data Warehouse which is time consuming not only when the model is created but each time new data is added to the Data Warehouse.
3. Search Engines enable content-driven dimensional navigation. At each step of navigation Search Engines propose different possibilities to filter results according to the content of the datasets that are being indexed and analyzed in near real-time.
4. Search Engines, unlike solutions based on RDBMS, are able to analyze data without the need to know the various data types: e.g. a Search Engine can easily look for any event that occurred at a specific point in time, whereas in traditional RDBMS data models, you can only search on precise data fields (e.g. on “invoice date” or “customer creation date”).
5. End users are now quite familiar with the “Google” interface. As a consequence, they are much more independent from IT departments if they can access decision support data through a Search Engine.
6. When external and unstructured data is needed to support decision making, traditional Data Warehouse architectures are limited and Search Engines can help to fill the gap.
7. Search Engines include functionality to automatically generate categories and clusters, hence improving the contextualization and meaning of data.
8. Most of today’s Search Engines include functionality to aggregate and analyze data. In addition, Search Engines enable end-users to expose relationships and to find patterns in data without the necessity of the perfectly formulated question or query [7].
9. Search Engines can work with existing information systems (e.g. Data Warehouses, Data Marts, Production Systems, etc.) and provide a federated view of data without compromising on performance. At the same time, the federated business view can encompass new data sources and provide cross-domain data navigation. This is in contrast to federation approaches based on RDBMS that failed to address performance requirements.

#### **Towards a Search-based BI**

Over the past couple of years Capgemini has implemented multiple Enterprise Search Engines and Search-based solutions for a growing number of customers. Based on technologies like Exalead, Autonomy, and Fast, billions of documents are being indexed on a daily basis from multiples sources like ECM (Enterprise Content Management), ERP (Enterprise Resource Planning), CRM (Customer Relationship Management), Data Warehouses, and other Legacy Systems. In most of

the cases information is being collected in near real-time and presented to end-users through highly friendly and helpful interfaces (fuzzy logic, spelling checker, categories, etc.).

Because of the nature of Enterprise Search Engines, i.e. one single product to extract, load, index, and present data, the time required to implement a Search-based BI solution is heavily reduced compared to the time that is needed to design and build a traditional BI system. Furthermore, performance is not an issue in Search-based BI, neither in terms of number of users nor in volume of data.

Future BI systems, integrating non-structured and external information, will benefit from the proven scalability features of Search Engines. Within Search environments volumes are measured in petabytes instead of gigabytes and terabytes. Take for instance Google which is indexing billions of documents daily on the Internet while providing access to a tremendous number of concurrent users. Search-based BI is leveraging investments in existing BI systems and is capable of getting the long-awaited business benefits out of the investments in existing Data Warehousing environments.

Will Search-based BI replace current BI systems? In the short-term, we don't think so. For the moment, Search-based applications are being used as a complement to cater for the shortcomings of existing BI systems; however, in the long run we will see Search-based solutions transforming the BI domain because of its inherent features. The combination of BI and Search-based solutions will preserve the strengths of both and mitigate the drawbacks of each. As discussed in the previous paragraph, Enterprise Search Engines have some really nice capabilities to make existing BI systems user friendly, agile and flexible. Also, in terms of cost and performance, Enterprise Search Engines are the way to go.

### Conclusion

In this paper we have described the evolution of BI in large organizations. We have argued that the existing model of Data Warehousing is getting too complex and expensive to maintain and that relational or multi-dimensional databases do not provide the appropriate functionality and flexibility for user friendly data exploration.

New technologies and concepts are needed to make BI user friendly, agile and flexible. Enterprise Search Engines are one of the exciting new technologies capable of transforming the BI industry. In addition to the scalability and performance features, Search Engines are easy to implement and user friendly. For the moment we see Enterprise Search Engines as an interesting alternative to complement traditional BI architectures.

### Literature

1. Rittman M. *An interview with Nigel Pendse*, 2010.
2. Rockart J. *Executive Information Support Systems*, 1980.
3. Inmon B. *Building the Data Warehouse*, 1991.
4. Kimball R. *The Data Warehouse Toolkit*, 1996.
5. Nicholls C. *BI 2.0: The Next Generation*, 2006.
6. Nelson G. *BI 2.0: Are we there yet?* 2010.
7. Caruso D. *The BI-Search Evolution*. 2010.



### About Capgemini and the Collaborative Business Experience™

Capgemini, one of the world's foremost providers of consulting, technology and outsourcing services, enables its clients to transform and perform through technologies.

Capgemini provides its clients with insights and capabilities that boost their freedom to achieve superior results through a unique way of working, the Collaborative Business Experience™. The Group relies on its global delivery

model called Rightshore®, which aims to get the right balance of the best talent from multiple locations, working as one team to create and deliver the optimum solution for clients.

Present in more than 30 countries, Capgemini reported 2009 global revenues of EUR 8.4 billion and employs 90,000 people worldwide.

More information is available at [www.capgemini.com](http://www.capgemini.com)

#### About the authors

**Christian Becht** is the Vice President in charge of the Business Information Management service line at Capgemini France. He has been involved in many BI projects since 1990 and is considered a global industry leader. For more information about Business Information Management or to ask for Christian as a speaker, contact him at [christian.becht@capgemini.com](mailto:christian.becht@capgemini.com)

**Marcel de Grauw** is a Managing Consultant at Capgemini in Paris. He has been active in BI since the early 1990s and is currently working on next generation BI architectures based on new technologies, like Search Engines and AI. For more information about Search-based BI or to ask for Marcel as a speaker, contact him at [marcel.de-grauw@capgemini.com](mailto:marcel.de-grauw@capgemini.com)

To find out more about Business Information Management, visit us at [www.capgemini.com/BIM](http://www.capgemini.com/BIM) or e-mail [BIM@capgemini.com](mailto:BIM@capgemini.com).