SOFTWARE TOOLS FOR BUSINESS INTELLIGENCE

Ilia Vecherkov[1], Nikolay Yanev[2], Yordanka Anastasova[3]

[1] University of Mining and Geology "St. Ivan Rilski", 1700 Sofia, e-mail: ilia.vecherkov@mgu.bg
[2] University of Mining and Geology "St. Ivan Rilski", 1700 Sofia, e-mail: nikolay.yanev@mgu.bg
[3] University of Mining and Geology "St. Ivan Rilski", 1700 Sofia, e-mail: yordanka.anastasova@mgu.bg

ABSRACT. Nowadays we are witnessing a new industrial revolution, driving forces of which are digital data, computing technologies/ computers and automation. The mining industry is no exception to this trend.
In recent years, large mining companies have massively implemented different information systems to manage individual processes or modules of ERP systems. With the development and integration of technologies, the volume of gathered information increases. Large and complex arrays of data appear. They are usually difficult to process and analyze.
Solution to these problems offers Business Intelligence (BI). BI is a collection of methods for extracting, processing, analyzing and visualizing of data. Its purpose is the retrieving of meaningful and business-friendly information.
An important feature of BI systems is their ability to process large amounts of data in real time, making them especially useful in managing the dynamics of conditions and processes in the mining industry.

Keywords: Business Intelligence, Software tools, Mining Industry

СОФТУЕРНИ ИНСТРУМЕНТИ ЗА БИЗНЕС ИНТЕЛИДЖЪНС

Илия Вечерков, Николай Янев [2], Йорданка Анастасова [3]
[1] Минно-геоложки университет „Св. Иван Рилски, 1700 София, e-mail: ilia.vecherkov@mgu.bg
[2] Минно-геоложки университет „Св. Иван Рилски, 1700 София, e-mail: nikolay.yanev@mgu.bg
[3] Минно-геоложки университет „Св. Иван Рилски, 1700 София, e-mail: yordanka.anastasova@mgu.bg

РЕЗЮМЕ: В днешно време сме свидетели на нова промишлена революция, чиито движещи сили са цифровите данни, изчислителните технологии и автоматизацията.
В последните години големите минни предприятия масово внедряват различни информационни системи за управление на отделни процеси или модули от ERP системи. С развитието и интегрирането на технологиите нараства и обемът на събраната информация. Получават се големи и сложни масиви от данни, които са трудни за обработване и анализ.
Решение на тези проблеми предлага Бизнес Интелигентността. BI представлява съвкупност от методи за придобиване, обработване, анализ и визуализация на данни с цел извличане на смислена и полезна за бизнеса информация.
Важна характеристика на BI системите е способността им да обработват големи количества данни в реално време, което ги прави особено полезни при управление динамиката на условия и процеси в минната индустрия.

Ключови думи: Бизнес интелиджънс, софтуерни инструменти, минна индустрия

Introduction

Nowadays we are witnessing a new industrial revolution, driving forces of which are digital data, computing technologies and automation. The human activity, the industrial processes and the scientific research are invariably related to data collection, data processing and data visualization. This stimulates the emergence of new products and services as well as new business processes and scientific methodologies. The trend is global and covers various areas of the human activity like healthcare, security, climate and resource efficiency, energy, intelligent transport systems and smart cities, digitization of public services and others.

This also applies to the mining industry. The mine enterprises are usually located on vast areas. Special intranet networks are built in order to make information resources accessible from each point of the enterprise. The information process is dynamic. The data obtained from one source is supplemented by new elements. The results are large volumes of data, which makes it difficult to use the classical analytical methods for solving many of the tasks in the modern mining industry. With the development of the IT and particularly data processing and data analyzing, now it is possible to resolve the task more quickly and easily. The main purpose of the modern data processing and analysis systems is to extract knowledge. This knowledge is used to support decision-making. The decisions are based on the provided analyses.

For companies, the strategic decisions are directly related to achieving significant results in the future. They concern the management policy, the business areas, the financial management, the products and the services. The companies need flexibility in decision-making. The flexible decisions allow

dynamic change in the business processes of a company. They must be also effective decisions. For today's dynamic world improving efficiency is crucial for the organizations. They must function properly and remain competitive on the market. To achieve a good performance for the business means to achieve better results investing less funds. This could be achieved by optimizing the business processes.

Here comes the Business Intelligence or BI. BI includes analytical models, technologies, information and practices for making better and safer management decisions. The need for timely analysis of the generated data is at the core of the BI systems. By obtaining information and regularities, the BI system is capable of evaluating the received information and based on it to make forecasts for the company's development.

Business Intelligence

BI is a collection of methods for extracting, processing, analyzing and visualizing of data. Its purpose is the retrieving of meaningful and business-friendly information. The most important feature of the BI systems is the ability to process large amounts of data in real time. This enables rapid analysis and dynamic reporting to meet the business needs, helping to build more effective strategies, optimize and manage more accurately the production processes, reduce company costs and improve planning and forecasting. By extracting regularities and causal links from the data processing, the BI system evaluates and analyzes the retrieved information, which improves planning and makes forecasts for the company's development. The main objective of Business Intelligence is to provide useful and high quality information for the managers, when necessary undertake effective management actions. BI helps companies discover new opportunities for business growth.

Globally the mining industry today faces many operational and regulatory challenges. With the price of precious metals on the rise, these companies must position themselves accordingly to be ready for a rapid expansion into new regions, while at the same time they must find new and innovative ways of keeping their costs low. Along with the rapid growth, the mining companies also face a complex and diverse set of regulatory and legal standards that they must agree with. In order to overcome these industry challenges and stay ahead, many mining executives have already been leveraging the power of BI tools. BI tools provide them with efficiency, as well as, the right insights and agility to maintain performance while reducing costs. BI systems help mining executives make better and faster business decisions, when it comes to HR planning, legal compliance, industry standards, trend analysis, logistics and strategic planning. According to Jaques Du Preez (MD and founder of BI-Blue Consulting), BI provides new insight to the business: „With this tool, you can transform your view of the business using sophisticated visual representations of your performance and processes." [Miningweekly.com]

The BI solutions enable companies to find new business directions. This happened to Pulse Mining Systems, an ERP provider for the mining industry. Since the ERP back-end database used relatively old technology, it was a challenge for the company to retrieve and visualize data. Using Tableau BI they successfully created clearer and better visualizations of the extracted data. Rob Parvin, Manager of Visualization and Analytics claims that thanks to Business Intelligence, Pulse Mining Systems managed to create a new business line for the organization. Using Tableau BI they provided the CIO with a quick win, allowing delivery of a single reporting solution for the water quality across the organization. This stimulates the emergence of further initiatives for which the company is being funded.

BI architecture

The basic functionality of the BI system involves the availability of structured data like relational databases, NoSQL, Data Warehouse, Big Data and so on.

The trends in using databases make the BI developers find new ways to accelerate the analytical processes through a more efficient collaboration between hardware and software. Today the relational databases are not flexible enough to meet the Big Data's challenge. The new architectures are also associated with relational databases, but they are designed to serve large volumes of data. There are also BI manufacturers, who offer database management systems that do not rely on relational models.

The BI solution's architecture would have the following structure:

- At the bottom of the corporate IT there are standard ERP, CRM and other transaction systems;
- On the next level the integration of the collected data is performed by ETL (Extract, Transform and Load) tools, which clear the data and maintain the data quality;
- The next level is occupied by data storage systems – Data Warehouse, Big Data and so on;
- Next are the SQL applications, which provide access to the data;
- The following level is occupied by BI tools for statistics, data mining and operational analysis;
- At the top there are the platforms for collaboration and sharing of the received information.

In general, BI relies on interactive dashboards, which provide good graphical interfaces. The dashboards are easy to use and quickly meet the business needs. They provide general information online. The BI dashboards also provide detailed graphical visualization of the data, which reveal the data regularities. They also ensure a high level (context-sensitive) monitoring of the large amounts of data. Thus, when deviation from key performance metrics (KPIs) happens, BI allows going deeper into the corresponding BI report (so called *drill down*). In this way, only the data, whose timing parameters are currently relevant, is observed. One of the most specific features of Business Intelligence tools is the ability to shift the data categories in order to group data by different dimensions and convert the report according to the business needs.

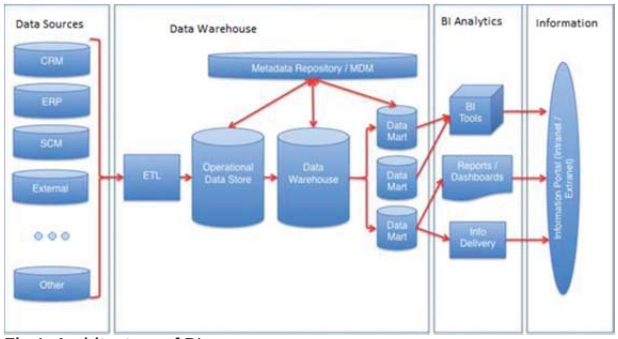

Fig 1. Architecture of BI

BI software

Gartner is a world leading research and consulting company in the field of IT. Each year, Gartner evaluates 24 of the leading Business Intelligence software solutions on multiple metrics such as system infrastructure, data management, analytics and visualization, information sharing and collaboration, and others. The Gartner's Magic Quadrant focuses on products that are currently popular on the market and meet the criteria of today's BI platform.

The BI platforms are evaluated on fifteen critical capabilities across five main sections.

I. Infrastructure.

1. BI Platform Administration, Security and Architecture. Capabilities that enable platform security, administering users, auditing platform access and utilization, optimizing performance and ensuring high availability and disaster recovery.
2. Cloud BI. Platform-as-a-service and analytic-application-as-a-service capabilities for building, deploying and managing analytics and analytic applications in the cloud, based on data both in the cloud and on-premises.
3. Data Source Connectivity. Capabilities that allow users to connect to data contained within various types of storage platforms, both on-premises and in the cloud.

II. Data Management.

4. Metadata Management.
5. Self-Contained Extraction, Transformation and Loading (ETL) and Data Storage.
6. Self-Service Data Preparation. capabilities include user-driven data combination of different sources, and the creation of analytic models such as user-defined measures, sets, groups and hierarchies. Advanced capabilities include machine-learning-enabled semantic autodiscovery, intelligent joins, hierarchy generation, data blending on varied data sources, including multistructured data.

III. Analysis and Content Creation.

7. Embedded Advanced Analytics. Enables users to easily access advanced analytics capabilities that are self-contained within the platform itself or through the import and integration of externally developed models.
8. Analytic Dashboards.
9. Interactive Visual Exploration.
10. Smart Data Discovery: Automatically finds, visualizes and narrates important findings such as correlations, exceptions, clusters, links and predictions in data that are relevant to users without requiring them to build models or write algorithms. Users explore data via visualizations, natural-language-generated narration, search and NLQ technologies.
11. Mobile Exploration and Authoring.

IV. Sharing.

12. Embedding Analytic Content.
13. Publish, Share and Collaborate on Analytic Content.

V. Overall platform capabilities.

14. Platform Capabilities and Workflow.
15. Ease of use and interface.

The Magic Quadrant helps in choosing the right product. It classifies the BI solutions in four smaller quadrants – Niche Players, Challengers, Visionaries и Leaders (Fig. 2).

Fig. 2. Gartner's Magic Quadrant

Niche Players do well in a specific segment of the BI and analytic platform. They have a limited capability to innovate or outperform other vendors. They may focus on a specific domain or aspect of BI, but are likely to lack depth of functionality elsewhere. They also have a limited implementation and support capabilities. In addition, they may not have achieved the necessary scale to solidify their market positions yet.

Challengers are well-positioned to succeed in the market. However, they may be limited to specific use cases, technical environments or application domains. Their vision may be hampered by the lack of a coordinated strategy across the various products in their platform portfolios, or they may lack the marketing efforts and awareness of the vendors in the Leaders quadrant.

Visionaries have a strong and unique vision for delivering a modern BI and analytics platform. They offer depth of functionality in the areas they address. Visionaries are thought

leaders and innovators, but they may be more expensive and are more complex to maintain.

Leaders are vendors that demonstrate a solid understanding of the product capabilities and commitment to customer success, extend their products' capabilities to meet the customer needs, and offer an attractive pricing model. The platforms they offer are easy to use and easy to buy. Their BI solutions offer powerful analytics and detailed visualizations. They are characterized by being considerably easier to use. In some cases, IT specialist intervention is not even necessary.

Tableau

Tableau Software is a US company based in Seattle, Washington, USA. It offers interactive data visualization products in the BI sphere. The company was established at Stanford University's Department of Computer Science between 1997 and 2002. Tableau Business Analytics is one of the leaders in Business Intelligence business, according to this year's Gartner Magic Quadrant. In fact, the product is a market leader for the fifth consecutive year.
Tableau is also one of the most dynamically evolving BI solutions. The company's ambition to improve existing and add new functionalities in recent years has led to a significant growth in the number of customers, therefore, to its leadership position in this fast-growing technology sub-segment.

The company offers the following BI products: Tableau Desktop, Tableau Server и Tableau Online.

Tableau Desktop is a data visualization software. It provides a fast creation of interactive dashboards, beautiful graphics and reports.

Tableau Server is a platform for Business Intelligence that offers powerful tools for analyzing and visualizing data. It is easily scalable and supports multiprocessor processing. It also provides high availability and security. Tableau Server can work on both physical and virtual machines. Tableau Server also provides an online multiuser platform designed to share and collaborate with reports and visualizations created in Tableau Desktop.

Tableau Online is a Cloud-based BI solution. It provides flexibility and ease to the powerful data visualization application. As a cloud-based platform, there is no need for special server hardware, expensive software and IT support. Tableau Online can be expanded to suit the users' needs.

Tableau has an extremely scalable client-server architecture (Figure 3) that handles mobile, web and desktop clients.

This architecture supports fast and flexible expansion. It could be divided into several layers.

Data Layer - One of the main features of Tableau is that it works with a variety of data from different sources. For example, in a mining enterprise, the data environment is heterogeneous. The data is stored in Data Warehouse, Big Data structures, Relational Databases, Spreadsheets (Excel), and more. To optimize data processing and visualization, Tableau uses a specially developed In Memory Data Engine technology.

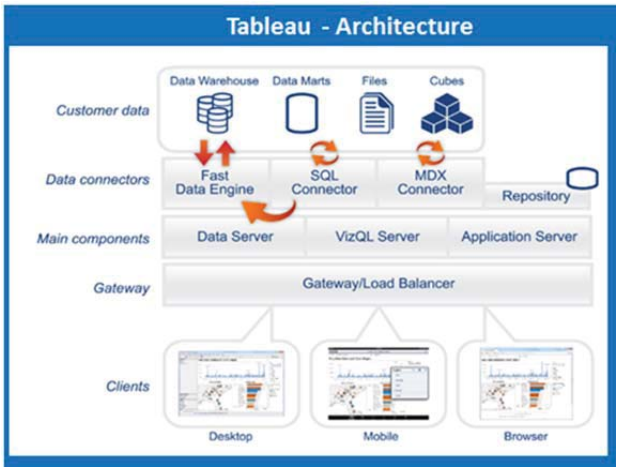

Fig 3. Tableau Architecture

Data Connectors - Tableau has multiple optimized data connectors for connections to various data sources such as Microsoft SQL Server, Microsoft Excel, Hadoop, MySQL, ORACLE and others. The system also has a general ODBC connector for data source systems that do not have one.

There are two methods for linking and retrieving data - Live and In-memory. Users decide which method to use. Live connection is used when data is retrieved from databases. The data itself remains in the output system, and only the summary results of the request are sent to Tableau. In practice, Tableau can efficiently handle unlimited amounts of data. The performance of the information system and the server that manages the database have a great importance for the work of Tableau. According to the manufacturers Tableau is a front-end solution for analyzing some of the world's largest databases.

In-memory is a technology that optimizes the analysis and visualization of data. Once a connection has been established between the BI and the data source, the data can be retrieved in the Tableau memory. The technology uses disk, RAM and Cache memory. In the first case, the data is extracted into files with a special extension called Data Extracts. The Data Engine module is designed to make full utilization of system resources to provide fast processing and visualization of large volumes of data.

Power BI

Power BI is a cloud-based Microsoft product that provides business analytics services. According to Garner's Magic Quadrant, along with Tableau, Power BI is a leader in the BI sphere. The product is a collection of powerful intuitive tools and functionalities for statistical analysis and visualization of large data sets. Power BI service is a part of Azure, which is Microsoft's cloud service platform. The Power BI architecture is based on two clusters - the WFE (Web Front End) Cluster and

the Back End Cluster. WFE takes care of the initial connection and authentication, and then the Back End processes all subsequent user interactions. Power BI uses Azure Active Directory (AAD) to manage users as well as manage the data storage and metadata process via Azure BLOB and Azure SQL Database.

The WFE cluster manages the initial connection and authentication process through Azure Active Directory. Users use their email address to create their own Power BI Account. This email account serves as a credential name that provides access to the other resources. The Authentication Name is converted into a User Principal Name that associates with a corresponding Windows Account (Windows Active Directory) domain. UPN is a format that defines the "Internet" name. For example, if we have the following email UserName@example.com, UPN splits the entire string as follows: username UserName; Separator @ and UPN suffix or domain example.com. For greater security, AAD provides additional settings.

Power BI uses Azure Traffic Management to direct user traffic to the nearest datacenter determined by the DNS record of the client attempting to connect, for the authentication process and to download static content and files. Power BI uses the Azure Content Delivery Network (CDN) to efficiently distribute the necessary static content and files to users based on geographical locale.

The Back End cluster is how authenticated clients interact with the Power BI service. The Back End cluster manages visualizations, user dashboards, datasets, reports, data storage, data connections, data refresh, and other aspects of interacting with the Power BI service. The Gateway Role acts as a gateway between user requests and the Power BI service and is handled by Azure API Management. Only Azure API Management (APIM) and Gateway (GW) roles are accessible through the public Internet. They provide authentication, authorization, DDoS protection, Throttling, Load Balancing, Routing, and other capabilities.

Power BI uses two primary repositories for storing and managing data. Data that is uploaded from users is typically sent to Azure BLOB storage, and all metadata as well as system data are stored in Azure SQL Database.

When an authenticated user connects to the Power BI Service, the connection and any request by the client is accepted and managed by the Gateway Role, which then interacts on the user's behalf with the rest of the Power BI services. For example, when a client attempts to view a dashboard, the Gateway Role accepts that request then separately sends a request to the Presentation Role to retrieve the data needed by the browser to render the dashboard.

It is important to note that users are responsible for the data they share: if a user connects to data sources using her credentials, then shares a report (or dashboard, or dataset) based on that data, users, with whom the dashboard is shared, are not authenticated against the original data source, and will be granted access to the report

Using the On-premises Data Gateway resolves this problem. The dashboards are cached in Power BI, but access to underlying reports or datasets initiate authentication for the user attempting to access the report (or dataset). The access will only be granted if the user has sufficient credentials to access the data..

Power BI also offers a version suitable for embedding - Power BI Embedded. It makes it possible to integrate dynamic reports into web systems or mobile applications.

Comparison between Tableau и Power BI

Tableau and Power BI provide powerful tools for extracting, processing, and visualizing data. Both products offer a wide variety of connectors to different data sources such as relational databases (MS SQL Server, MySQL, PostgreSQL, Oracle, etc.), spreadsheets (MS Excel), Web analytics services, data warehouses, Big Data Solutions. Loading fresh data can take place automatically at a time when servers and information systems are not heavily loaded to avoid delays or interruptions in the workflow. In their latest versions, both products add the automatic email notification option when certain data or manually defined calculations reach specific values.

Tableau and Power BI have predefined aggregate functions that make calculations very fast. These are SUM (), COUNT (), MAX (), MIN (), AVG (), median and trend finding, percentage, and so on. When extracted data is loaded, BI systems automatically determine the type of data and structure them in fields, if necessary (for example, when using JSON files as data sources). Both systems enable the user to convert the data type before it is rendered. The STRING fields can be trimmed.

Both products support programming languages by which users can create their own graphics or apply specific methods for deeper statistical analysis of data in order to retrieve business-friendly information. These are the R and the Python (Tableau only) languages.

Tableau Desktop and Power BI interfaces are intuitive and easy to use. And through cloud services, users can access their reports from different devices at any time.

They are available for Android OS, iOS, Windows OS.

Advantages of Tableau:

- More convenient interface;
- On-premises service that significantly reduces the possibility of leakage of valuable information;
- Can handle extremely large data;
- Provides better visuals;
- Story mode that draws attention to critical data values or anomalies;
- It has many and varied connectors for extracting data.

Advantages of Power BI:

- It has a free version that has limitations on the amount of data extracted, sharing and collaboration, automatic data update, and more. It is suitable for small companies with a limited budget;
- Great compatibility with other Microsoft products.
- Integrated in Microsoft Dynamics 365 (integrated in ERP Navision and CRM) business management, customer and marketing platform.

Conclusion

Nowadays, the information technologies are dynamically developing. New and new sources of raw data are constantly emerging increasing the need this data to be processed. Mining companies integrate high-tech systems for managing, allocating resources, and providing high security standards. All of these systems generate huge amounts of data.

BI gives us the opportunity to understand what data is "talking to us" through beautiful, detailed and interactive visualizations. The information we receive helps us in making decisions related to production, finance, optimal allocation of resources, and more. At a glance, we can compare past and present results with the past as well as make predictions about the future.

The choice of BI product depends on the specifics of work and the security policies. If the analyzed data is strictly confidential, then the user should choose Tableau's BI solution. Although the initial costs are relatively high because Tableau Server and Tableau Desktop are sold with separate licenses, the data is stored locally in the company and the security level is higher. The reports and dynamic boards can be displayed via Tableau Reader.

On the other hand, if the company uses Microsoft products and platforms, the more optimal choice would be Power BI. As it became clear, one of the key benefits of Power BI is that it integrates easily into Microsoft's ERP and CRM systems. This enables users to see real-time dynamic reports and boards with fresh data inside the system itself. This is extremely useful for directors, managers and financiers. It assists in making quick decisions that can be critical to business.

References

Inmon, W. H., “What is a Data Warehouse?”, Prism Tech Topic, 1995

Pieter M. Kroonenberg, Applied Multiway Data Analysis, Wiley 2008

BI and Big Data - accurate analyzes for reasoned decisions, www.cio.bg

Business intelligence can help mining overcome pressures, www.miningweekly.com

How digital innovation can improve mining productivity, www.mckinsey.com

Power BI Security and Power BI Architecture, www.powerbi.microsoft.com

Productivity in mining operations: Reversing the downward trend, www.mckinsey.com

Pulse Mining uses Tableau to create analytics solutions quickly, www.tableau.com

Tableau Knowledge Base, www.tableau.com

Tableau is the leader among Gartner Magic Quadrant business analysis systems, www.team-vision.bg

Using Tableau to Leverage Hadoop, www.tableau.com

The article is reviewed by Assoc. Prof. Dr. Mariana Trifonova and Assoc. Prof. Dr. Veselin Hristov.