

# A Review on BI Tools and Data Warehouse

Vivek Haralkar<sup>1</sup>, Prof. Prathibha Adkar<sup>2</sup>

<sup>1,2</sup>MCA Department, Savitribai Phule Pune University/Modern college of Engineering, Pune, Maharashtra, India

**Abstract**-Business intelligence is the process of analyzing large amount of data to take different types of decisions. there are so many tools are used to analysis data such as, yellowfi BI, Clear Analytics, Sisense, Micro Strategy, Pentaho, Oracle Hyperion System . I briefly mention three tools BIRT , Pentaho , Spago BI.

**Keywords:** Architecture, Application, Tools.

## I. DEFINITION OF BUSINESS INTELLIGENCE

**B**usiness Intelligence (BI) is: “The processes, technologies and tools needed to turn data into information and information into knowledge and knowledge into plans that drive profitable business action. BI encompasses data warehousing, business analytics and knowledge management. It is also defined as, Business Intelligence is defined as "knowledge gained about a business through the use of various hardware/software technologies which enable organizations to turn data into information”.

## II. INTRODUCTION

Many companies directly contact with large numbers of customers, however, a growing number applications (e.g. e-commerce support, call centre support) create a new data management challenge: that is effective way to implement enterprise applications in real time. To learn from the past and predict the future, many companies are adopting Business Intelligence (BI) tools and systems. Every business is dynamic in nature and is affected by various external and internal things. These things include market conditions, competitors. Business intelligence is big term which is used to gather information about business and the market condition . This information is collected to improve efficiency of business, reduce costs, reduce time and bring many positive changes. Organisation believe that up-to-date, accurate and integrated information about their product supply chain, products and customers are critical Successful organizations improving the use of their data assets and human empowerment. Rapid advances in computer technology allow business intelligence (BI) systems to provide managers with access to a tremendous amount of data. To function these systems combine complex front-end software with ETL capabilities that extract enormous amounts of data. At the heart of these systems are huge enterprise data warehouses that can populate a possible infinite combination of advanced reports, OLAP cubes and datasets for data mining. The underlying belief is that technically advanced systems are the most important drivers of effective decision making. Based on this belief BI vendors focus on technologically advanced systems while paying relatively

little attention to whether these systems meet the needs of decision makers.

The drive towards technological sophistication and away from improving decision making is one of the reasons for the low success rates (50%), high costs and time overruns associated with BI projects. While vendors and information technology departments get the technology right they often fail to deliver a product that is useful to management. Executives and managers will reject BI systems that do not facilitate efforts to evaluate and improve strategic and operational effectiveness. What practitioners fail to understand is that technologically advanced software, by itself, does not allow management to perform these vital functions. Effective business intelligence systems must account for the goal oriented behavior of decision makers. One way to incorporate the purposive behavior of decision makers in BI systems is the use of the frame as the basic building block of a pervasive BI system architecture. It does this since a frame is a "scheme of interpretation" in which the particulars of events and activities are organized and made sensible (Goffman 1974). The frame organizes the world according to the goal oriented behavior of the frame owner. Specifically, the frame captures the two parts of decision makers' goal oriented behavior. The first are the objectives disseminated from the ESM in the form of scorecards. These provide the goals that guide the activities of the business unit. The other part is the experience of the business unit leader. It is their experience that provides the guidance for how to achieve the objectives. The goal oriented behavior of the frame owner is codified in a business unit strategy map based on the objectives that the unit must achieve along with the means of achieving those objectives. At the heart of this system are strategy maps and scorecards. We propose strategy maps can be used to organize the company as a system since each scorecard contains the necessary elements to create a control system. A control system requires a measuring unit, norm, comparator and correction unit. The analogs in a scorecard are the measures, budgets, reports and initiatives. The importance of the control system is that it allows the business unit leader to evaluate and correct issues that impede operational effectiveness. Once management is certain they are operating effectively they can report to executives.

### 1) Business Intelligence and Reporting Tools (BIRT):

BIRT project is a flexible, open source, and 100% pure Java reporting tool for building and publishing reports against data sources ranging from typical business relational databases, to XML data sources, to in-memory Java objects. BIRT is

developed as a top-level project within the Eclipse Foundation and leverages the rich capabilities of the Eclipse platform and a very active open source community of users. Using BIRT, developers of all levels can incorporate powerful reporting into their Java, J2EE and Eclipse-based applications. BIRT originated from the open source Eclipse project, and was first released in 2004. BIRT is an open source technology platform used to create data visualizations and reports. Project sponsors include OpenText, IBM, and Innovent Solutions.

BIRT consists of several components. The main components include a report designer and the BIRT runtime, but BIRT also provides three extra components: a chart engine, chart designer, and viewer. With these components, you should be able to develop and publish reports as a standalone solution. BIRT is written in Java, and is licensed under the Eclipse Public License. Its latest release, which runs on Windows, Linux, and Mac. BIRT is a built on Eclipse open source reporting system for web applications, especially those based on Java™ and J2EE. You can access BIRT through the TADDM Domain Manager.

BIRT has the following two main components:

- A report designer built on Eclipse
- A runtime component that you can add to your application server.

More information and tutorials can be found at the BIRT Web site:

<http://www.eclipse.org/birt>

To generate a report, BIRT requires a report design file, which is an XML file that contains all the information about how the report looks, what data sources to contact, and which SQL queries to run. BIRT supports a variety of ways to get data. Some of the relevant ones are described in the following list. You can browse through the BIRT tutorials to learn more details about other data sources.

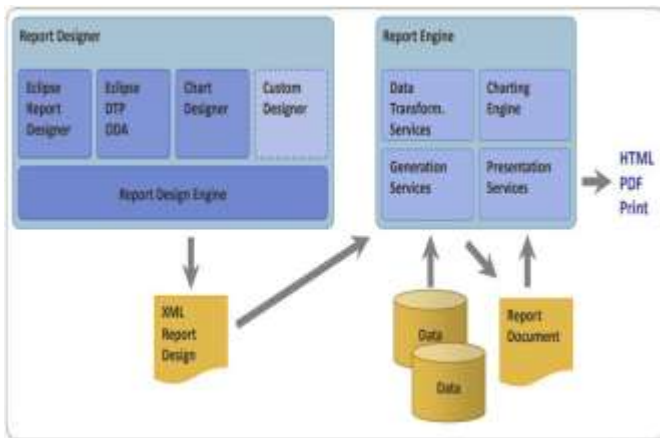


Figure 1: BIRT tool processing

Report Preview – Allows you to test your report at any time with real data. · Code Editor – Edits custom business logic in

Java or JavaScript to be executed during data access, report generation, or viewing.

Chart Editor – Defines and formats charts and their accompanying legends and axes. Outline – Provides a compact overview of your entire report structure.

Cheat Sheets – Offer step-by-step instruction on commonly performed tasks to shorten the learning curve. · Report Libraries – Archives of commonly used report items are shared amongst report developers to speed report design.

Report Templates – Common report types and layouts ease report design.

2) Spago BI tool:

SpagoBI supports the real-time monitoring, analysis and presentation of business data and processes. You can keep business processes under control by constantly monitoring their state. SpagoBI allows you to go further than this: you can detect inefficiencies and bottlenecks in your business processes, promptly react to events requiring quick decision making, as well as discover new business opportunities hidden in your own data.

It is the only platform Business Intelligence 100 % Open Source, covers and satisfies all requirements of BI, both in terms of analysis and data management, administration and security. Stratebi is the first and only certified partner in Spain. The analytical world offers solutions for reporting, multidimensional analysis (OLAP), data mining (Data Mining), dashboards (Dashboard) and ad-hoc queries. Add original for managing collaborative processes and analysis of geo-reference.

SpagoBI is an integration platform (and not a product platform), since it is not built around a predefined set of tools. It has a modular structure in which all modules are related to the core system, ensuring the harmony of the platform along with their evolving capacities.

- Spago BI tool is written in java language.
- Spago BI is cross Plat form work on any operating system
- License Mozilla Public License
- It can be download from official website [www.spagobi.org](http://www.spagobi.org)



Fig 2: SpagoBI Architecture

SpagoBI Server, It includes analytic tools and features act as a core for the suite.

SpagoBI Studio, The integrated development– environment

SpagoBI SDK (Software Development Kit), It is the– integration layer allowing to use SpagoBI with external tools.

SpagoBI Applications, Using it a collection of vertical– analytic models are developed.

Data can be accessed from the MySQL, Oracle, Postgre or Ingres SQL databases. In addition to the BI product, Spago World includes apps for SOA development, enterprise Java, middle ware and quality measurement.

3) Pentaho tool:

Pentaho addresses the barriers that block your organization’s ability to get value from all your data. Our platform simplifies preparing and blending any data and includes a spectrum of tools to easily analyze, visualize, explore, report and predict. Open, embeddable and extensible, Pentaho is architected to ensure that each member of your team — from developers to business users can easily translate data into value. Pentaho is a business intelligence (BI) software company that offers open source products which provide data integration, OLAP services, reporting, information dashboards, data mining and extract, transform, load (ETL) capabilities.

Features of Pentaho:

Pentaho Reporting primarily includes a Reporting Engine, a Report Designer, a Business Intelligence (BI) Server. It comes loaded with the following features –

Report Designer – Used for creating pixel perfect report.

Metadata Editor – Allows to add user-friendly metadata domain to a data source.

Report Designer and Design Studio – Used for fine-tuning of reports and ad-hoc reporting.

Pentaho user console web interface – Used for easily managing reports and analyzing views.

Ad-Hoc reporting interface – Offers a step-by-step wizard for designing simple reports. Output formats include PDF, RTF, HTML, and XLS.

A complex scheduling sub-system – Allows users to execute reports at given intervals.

Mailing – Users can email a published report to other users.

Connectivity – Connectivity between the reporting tools and the BI server, which allows to publish the content directly to the BI server.

III. LITERATURE SURVEY

Data warehouse architecture

Layers of data warehouse

1) Tools : In this layer All tools are included such as BIRT, OLAP, Spago BI, Pentaho. These tools are used to analysis the data according to tools

2) Extraction : In this layer data extraction is done. Data is extracted from various sources like RDBMS, Sql Server, Flat files etc.

3) Staging: In this layer all data gets converted into one single format. Data gets clined, filter, split, join.

4) Loading : Loading is the process in which datagets loaded into data warehouse.

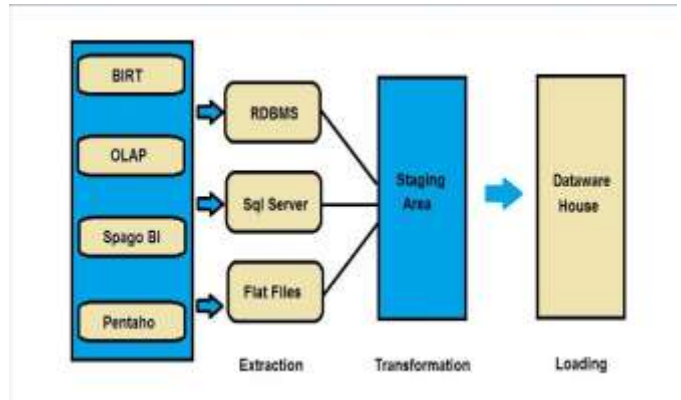


Figure 3 : Data warehouse architecture

There are two main components to building a data warehouse- an interface design from operational systems and the individual data warehouse design. Thus, the construction of DWH depends on the business requirements, where one development stage depends on the results of previously developed phase. The structure of a DWH can be understood better through its layered model, which lists the main components of the data warehousing architecture. You can see that it is nothing but the movement of data from source to staging area and then finally to conformed data marts through ETL (Extract, Transform and Load) technology.

The first layer is the **Data Source layer**, which refers to various data stores in multiple formats like relational database, Excel file and others. These stores can consists of different types of data – Operational data including business data like Sales, Customer, Finance, Product and others, web server logs, Internet research data and data relating to third party like census, survey. The next step is **Extract**, where the data from data sources is extracted and put into the warehouse staging area. The extracted data is minimally cleaned with no major transformations.

Then comes the **Staging area**, which is divided into two stages – data cleaning and data ordering. As the name suggests, this layer takes care of data processing methods, i.e. cleaning (removing data redundancy, filtering bad data) and ordering (allowing proper integration) of data. Overall, this stage allows application of business intelligent logic to transform transactional data into analytical data. It is indeed the most time consuming phase in the whole DWH

architecture and is the chief process between data source and presentation layer of DWH. Finally, we have the **Data Presentation** layer, which is the target data warehouse – the place where the successfully cleaned, integrated, transformed and ordered data is stored in a multi-dimensional environment. Now, the data is available for analysis and query purposes. The information is also available to end-users in the form of data marts.

#### IV. APPLICATION OF BI TOOLS

##### 1. Sales intelligence

Customer negotiation is a crucial skill that every organization's sales department should bolster. Sometimes it can be hard to convince potential clients to buy. However, with the business intelligence tools available, the process is becoming smoother and more predictable.

##### 2. Visualization

Business Intelligence software utilizes a range of data analytic tools that are designed to analyze and manage data related to your business operations. The data presented in the form of visualizations allows the organization to monitor logistics, sales, productivity, and much more. By presenting the data in intuitive visuals and easy to comprehend formats, the BI applications enable users to make giant steps in taking appropriate actions. An organization stands in a better position in optimizing data interpretation with the right BI application.

##### 3. Reporting

Business Intelligence tools continuously collect and study data. Collected data is used to generate a range of reports. The reporting includes staffing, customer services, expenses, sales, and other operational processes. The BI reports contain relevant information that is easy to understand. The organization, therefore, acts accordingly based on the compiled report and the use of customized reports is also

possible to adhere to the unique BI reporting needs of any organization.

##### 4. Performance management

With BI applications, organizations can input data with set goals such that the progress is monitored based on pre-defined and customizable timeframes. The data-driven goals may include target delivery time, projects completion deadlines or sales goals. The goals are tracked to ensure that the business gets real-time updates on the progress and knows what gaps may remain.

#### V. CONCLUSION

This review paper provides detailed information about BI tools and its application. In BI OLAP, BIRT, Spago BI, Pentaho provides a multidimensional view of the data, extending far beyond the ability of the relational database management system or spreadsheet. These tools provide a simple method to integrate, interrogate, browse and display data of various level.

#### REFERENCES

- [1]. Tvrđikova, M. (2007), 'Support of Decision Making by Business Intelligence Tools', Computer Information Systems and Industrial Management Applications, 2007. CISIM '07. 6th International Conference, pp. 368.
- [2]. DEJAN Nguyen Tho Manh, Schiefer Josef and Min Tjoa, A. (2005) 'Data warehouse design 2: Sense & response service architecture (SARESA): an approach towards a real-time business intelligence solution and its use for a fraud detection application', Proceedings of the 8th ACM international workshop on Data warehousing and OLAP, DOLAP '05, ACM Press
- [3]. Sc ha ff ha u s e r, D. Florida State U Transforms Reporting with Business Intelligence Campus Technology. 2010 (Last visited on 15 December 2014). <http://www.fsu.edu/~scshuff/>
- [4]. Sc ha ff ha u s e r, D. Florida State U Transforms Reporting with Business Intelligence Campus Technology. 2010 (Last visited on 15 December 2014).
- [4]. JISC (2011). Business Intelligence: Monitoring Performance and Planning Improvement. (Last visited on 16 December 2014).
- [5]. Infotool Data AB & Dimensional Insight Netherlands. 2014. A University BI-Solution: Better Insight and Control Over Operations, Easy Access for All Users, Reliable and Accessible Information. (Last visited on 18 December 2014).