

Automation System for Textile Industry Using BI

Mrs. Kanchan A. Khedikar^{#1}, Miss. Vaishnavi V. Yevale^{#2}, Miss. Payal M. Mehta^{#3}, Mr. Suyash S. Badnore^{#4}

Assistant Professor, Computer Science & Engineering Department, Walchand Institute of Technology, Solapur, India.

Student Computer Science & Engineering dept., Walchand Institute of Technology, Solapur, India

¹kanchan.khedikar@gmail.com, kakhedikar@witsolapur.org

²vaishnaviyevale@gmail.com

³payalmehta11071997@gmail.com

⁴suyashona0004@gmail.com

Abstract— Business intelligence is a technology and an application which can be used when it comes to automate the industry. Now a days industries need to process large amount of data. This data needs to be processed to give the proper information for making the business decisions. But often these data have so many challenges such as, many industries generate huge chunk of data in any format and in many unstructured files. These data must be combined, compared to work as a whole to get the proper decision. So to overcome these problems the solution is to use the concept of Data Integration which would allow data from different databases and formats to communicate with each other. Among the various available ETL tools in the market, Informatica is the market's leading integration platform.

This paper presents the proposed idea of implementing the system that will be capable of analyzing its environment and presenting business data to support better business decision making. The system consists of a website for an industry which will be used for data integration and will also keep their information secure and make their management flexible. It will integrate different modules such as designing database, data processing using ETL tool Informatica and at the end final reporting using the Pentaho reporting tool.

This paper explains the system architecture of proposed idea, various technologies used for this system and methodology used in the project.

Keywords— Intelligence, ETL, SQL, Informatica, Pentaho.

I. INTRODUCTION

In today's world, the data is growing rapidly in many formats and in unstructured files. This project's perspective is to high-tech the textile industry. The core purpose of designing this system is to process and manage Industry data. It is a system that will convert manual working to computerized system.

Advanced ETL tool interpret necessary information to manage data in an integrated enterprise approach. Enterprise information management brings Business Intelligence and Enterprise Content Management together. The website will contain all the information and different modules linked with each other which will make the management flexible. Usage is not limited to the particular industry and even useful for other textile industries. Collective information from different modules will be able to reduce workload and utilize their time efficiently. The friendly ETL tool called Informatica based on business intelligence is used in the project, which is gradually growing for the best synchronization. The main objective of this system is to automate the industry and make the management easier. Advanced ETL tools interpret necessary information to manage data in an integrated enterprise approach. Enterprise information management brings Business Intelligence and Enterprise Content Management together.

The main and important objective of this system is, it synchronizes the working of all the modules. The work is useful for saving valuable time and reduces the huge paper work.

Collective information from different modules will be able to reduce workload and utilize their time efficiently. The ETL tool called Informatica based on business intelligence is used in the project, which is gradually growing for the best synchronization between human and data.

II. LITERATURE REVIEW

“Application of ETL Tools in Business Intelligence” by Nitin Anand [1]. In this paper, the author explains, Business intelligence (BI) is considered to have a high impact on businesses. Research activity has risen in the last years. An important part of BI systems is a well performing implementation of the Extract, Transform, and Load (ETL) process.

“Extract, Transform, Load Data with ETL Tools”, Preeti handa and Neetu Sharma [2]. In this work the authors presents the use of ETL tools which have evolved, matured, and now present with productive environments for Big Data, Data warehouse, Business Intelligence, and analytics processing. Informatica is the best ETL tool in the marketplace It can extract data from numerous heterogeneous sources, transforming them as per business needs and loading to target tables. It's used in Data

migration and loading projects. The authors explain all the applications and advantages of using the Informatica. This solution serves as the foundation for all data warehousing and enterprise data warehousing projects. It accelerates their deployment, minimizing costs and risks, by ensuring that enterprise data warehouses are populated and maintained with trustworthy, actionable, and authoritative data.

“The research & application of Business Intelligence system in retail industry” [3]. In this research paper they have said, Business intelligence technologies emerge as the times require. Business intelligence is a very wide-ranging set of the collection, consolidation, analysis and information access capabilities for a solution, including ETL, data warehouse, data query and reporting, multidimensional data analysis, data mining and other technologies. The key technology of Business Intelligence is the establishment & application of Business Intelligence System in retail industry.

So like this we have gone through multiple literatures and we come to the conclusion that, every company have to process large sets of data from varied sources. This data needs to be processed to give insightful information for making business decisions. But, quite often such data have following challenges.

1. Large industries generate lots of data and such huge chunk of data can be in any format. They would be available in multiple databases and many unstructured files.
2. This data must be collated, combined, compared, and made to work as a seamless whole. But the different databases don't communicate well.

To overcome these challenges, the best possible solution is by using the concepts of Data Integration which would allow data from different databases and formats to communicate with each other.

Among the various available ETL tools available in the market, Informatica PowerCenter is the market's leading data integration platform.

The existing environment in industry is a manual process so it is very time consuming and of huge paperwork. Though it is manageable but it is limited when it comes to store large amount of data. Thus our system aims at automation of the industry reducing its manual work and making the system easier. The development of the new system contains the following activities, which try to automate the entire process keeping in view of the database integration approach. There is no risk of data mismanagement at any level while the project development is under process. User friendliness is provided in the software with various controls. It provides high level of security using BI.

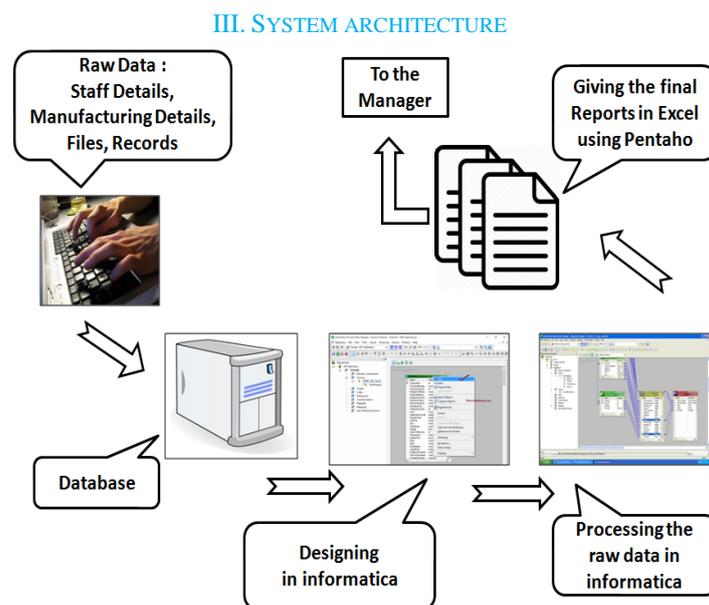


Fig 1: System Architecture

A. Row Data

Industries these days contains huge amount of raw data in unstructured files and in many formats. The raw data will be taken from the workers about manufacturing details, production details, total material, staff details, other files etc. The input will be given by the manager and will be stored into the database which will be used further for the production details.

B. Designing Database

The important part of the system is to design the database for different modules like weaving, stitching, production, delivery. These modules will be interlinked to each other. The main Master module will be used in all the other modules as the input data.

C. Data Processing

After designing the database, processing the data is also an important and difficult task to be performed. The raw data will be processed and made into required or appropriate format so that all the details will be accessible to the manager easily.

D. Final Report

The final reports of the processed data will be given in the required format such as in excel, pdf, text or printed reports using the reporting tool Pentaho to the manager as Pentaho Reporting's development is driven by the goal to create a flexible yet simple to use reporting engine.

IV. TECHNOLOGY USED

A. Tools

1. Informatica: The processing of raw data into required format using informatica will be profitable as it is a data integration tool and provides various services. Informatica uses the Extract, Transform & Load (ETL) architecture which is the most popular architecture to perform data integration. Informatica PowerCenter is a premium data integration solution available today. The reason it provides the best solution in large enterprises is because it is database neutral and hence can communicate with any database and it is the most powerful data transformations tool. It converts one application's data to another's format. PowerCenter reads data, row by row, from a table (or group of related tables) in a database, or from a file. This database or file is referred to as the source. The structure of the source is contained in a source definition object. Informatica PowerCenter converts the rows into a format the second (target) system will be able to use. The logic for this conversion is defined in transformation objects. Informatica PowerCenter writes data, row by row, to a table (or group of related tables) in a database, or to a file. This database or file is referred to as the target. The structure of the target is contained in a target definition object.

Following figure is the workflow designer of Informatica which shows the mapping of modules for our project. Here it is showing the transformation applied to the Master module.

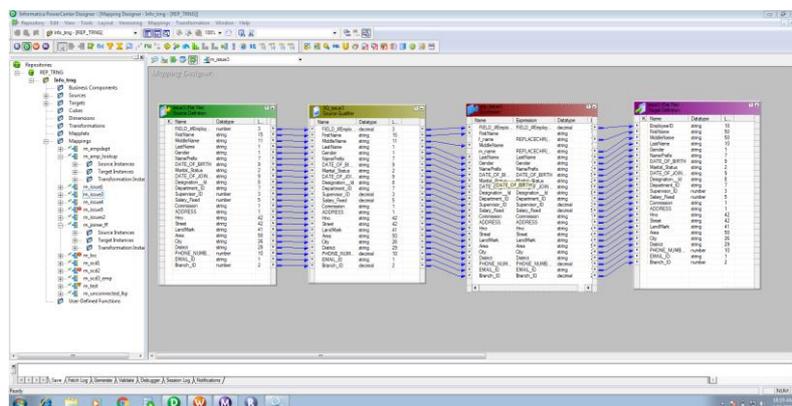


Fig 2: Mapping in Informatica

2. Pentaho: The important task for any industry is generating the final reports. Pentaho gives a uniform consistent format that is accessible and relevant to end users. Pentaho Reporting is a suite (collection of tools) for creating relational and analytical reports. It can be used to transform data into meaningful information. Pentaho allows generating reports in HTML, Excel, PDF, Text, CSV, and xml. Pentaho Report Designer generates professional reports. Pentaho can accept data from different data sources including SQL databases, OLAP data sources, and even the Pentaho Data Integration ETL tool. The following figure shows the visual representation of data.

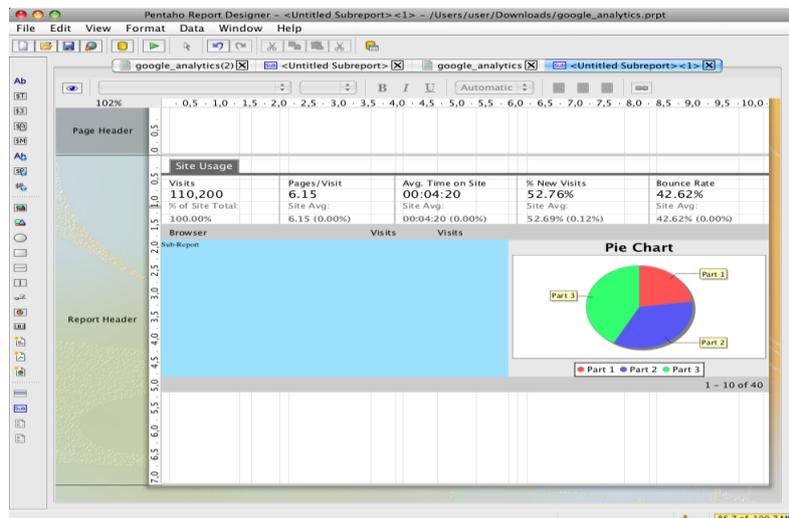


Fig 3: Pentaho Report Designer [11]

B. Web Technologies:

1. **HTML5 and CSS3:** These languages will be used for designing the webpage.
2. **PHP(5.5.6):** This scripting language will be used for data integration for making the webpage.
3. **MySQL:** The raw data will be stored and then the designing of database according to the requirement will be done using SQL.

V. DATA FLOW DIAGRAM

A. Level 0 DFD

The level-0 DFD gives the basic idea of taking the raw data from industry and storing. After that, integrating the data and processing it. Finally, after analysis giving the reports to the manager.

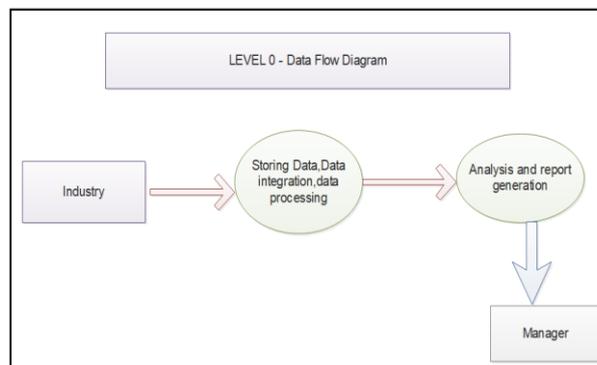


Fig. 4: Level-0 DFD

B. Level 1 DFD

The level-1 DFD briefly explains that the raw data taken from the industry will be stored into SQL server and then importing the files from database to the source of the Informatica. After importing, processing the data by different transformations and giving the daily, weekly, monthly reports to the manager in the required format.

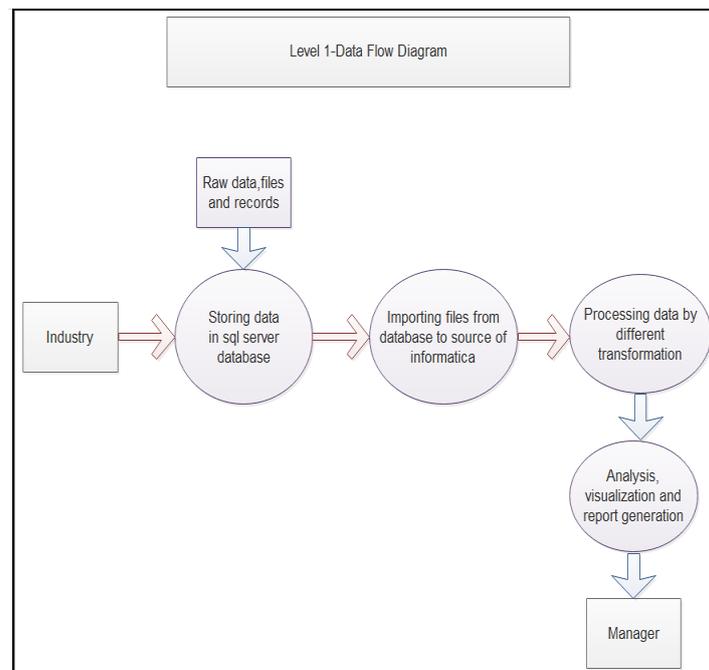


Fig. 5: Level-1 DFD

VI. FUTURE SCOPE

If we talk about the future scope then it depends on how innovative one could be to enhance the use of this project. The project is designed keeping in view the day to day problems faced by the industry. The usage is not limited to the particular industry and even useful for other textile industry.

VII. CONCLUSION

In this paper, the system deals with the issues related to textile industry. The purpose of this system is to design a webpage for automation of the industry. That should improve efficiency of the industry and reduce their workload. The webpage provides all the details in an appropriate format to the user using business intelligence. This system helps to improve the performance of workers. It saves time, can access data in anytime and anywhere. Collective information from different modules will be able to reduce workload and utilize their time efficiently.

REFERENCES

- [1] Nitin Anand, "Application of ETL Tools in Business Intelligence", International Journal of Scientific and Research Publications, Volume 2, Issue 11, November 2012
- [2] Preeti Dhanda and Neetu Sharma, "Extract Transform Load Data with ETL Tools", International Journal of Advanced Research in Computer Science, Volume 7, No. 3, May-June 2016
- [3] Tong Gang, Cui Kai, Song Bei, "The research & application of Business Intelligence scheme in retail industry", in IEEE International Conference on Automation and Logistics, Sept-2008
- [4] <https://tutorialspoint.com>>pentaho
- [5] <https://network.informatica.com>
- [6] <https://www.w3schools.com>
- [7] <https://www.edureka.co/blog/informatica-etl/>
- [8] <https://www.mysql.com/>
- [9] <https://en.wikipedia.org/wiki/MySQL/>
- [10] <https://www.on-reporting.com/blog/pentaho-report-designer-better-default/>