Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 221 (2023) 232–236

Procedia Computer Science

www.elsevier.com/locate/procedia

10th International Conference on Information Technology and Quantitative Management

The Data Quality in a Complex Web Based Decision Support System

Florin Boboșatu[a], Clara Maria Boboșatu[b], Maria Vișan[c] *

[a] *Credit Europe Bank, B-dul Timisoara, no. 26Z, Cladire Anchor Plaza, Sector 6, 061331, Bucharest, Romania*
[b] *Grigore Moisil National College, B-dul Timisoara, no. 33, Sector 6, 061309, Bucharest, Romania*
[c] *School of Advanced Studies of the Romanian Academy - Department of Engineering, Mechanics, Computers, Romanian Academy, 125 Calea Victoriei, Bucharest, Romania*

Abstract

The paper focuses on the implementation of new concepts and business models to increase data quality by defining clear rules for real-time error detection in an experimental decision support system (DSS), which helps financial decision-makers to analyse and process useful and business information from raw data, documents, personal knowledge and/or business models in order to identify and solve complex financial problems and to choose the best decisions.

Keywords: Big Data; DSS; Extract-Transform-Load; Knowledge Management.

1. Introduction

A decision support system is a specific class of computerized information system used to improve a company's decision-making capabilities and organizational decision-making activities [1]. A properly designed DSS analyses, processes large amounts of data and helps decision-makers to compile useful information from raw data, documents, personal knowledge, and/or business models, in order to identify and solve the business problems and to take effective decisions.

Decision support systems bring together data and knowledge from different business areas and sources to provide users with information beyond the usual reports and summaries. This is intended to help and guide the people to make the best decision [2].

A DSS is meant to improve the business decision-making ability of managers allowing them to take faster and better decisions within the constraints of cognitive, time and economic limits. It increases the business productivity of the decision-makers and supplements one or more of a decision maker's abilities [3]. In addition, it can facilitate one or more of the decision-making stages [4] (such as intelligence, design, choice and implementation), facilitate smother problem solving flows and can aid enables decision maker to address unstructured or semi-structured decisions [5]. Thus, the decision maker's knowledge management

* Corresponding author.
E-mail address: fbobosatu@yahoo.com

1877-0509
10.1016/j.procs.2023.07.032

competence is enhanced; his knowledge management (KM) skills are supplemented with the computer-based KM capabilities [6].

The rest of the paper is organized as it follows. In the first section, it is presented DSS general aspects section. Then it is proposed the general structure of the system. Then there are described the main components (front end and back end) of the system. The main advantages of the proposed solution are eventually presented. The financial and business information utilized in this Complex Web Based DSS are stored in the relational SQL databases.

The financial and business data arrives in the Databases using SQL scheduled jobs with daily or monthly frequency.

2. DSS general aspects

In the specialized literature, from a practically point of view, a significant number of authors agree [7] with the following DSS set of features [8]:

a) Facilitation. DSS helps and guide the managers to make good decisions for complex business problems. DSS contains facilities and supports specific decision-making activities and/or decision processes.

b) Interaction. DSS are computer-based systems designed for interactive use by decision makers or staff users who control the sequence of interaction and the operations performed.

c) Ancillary. DSS can support decision makers at any level in an organization and guide them to get the best solution. They are NOT intended to replace decision makers.

d) Repeated Use. DSS are intended for repeated use. A specific DSS may be used routinely or used as needed for ad hoc decision support tasks.

f) Task-oriented. DSS provide specific capabilities and facilities that support one or more tasks related to decision-making, including intelligence and business data analysis; identification and design of business alternatives; choice between business alternatives and taken decision implementation.

g) **Identifiable**. DSS could be an independent business system that collect or replicate data from other information business systems or subsystems of a larger, more integrated information system.

f) **Decision Impact**. DSS are intended to improve the data accuracy, data timeliness, data quality and overall effectiveness of a specific decision or a set of related decisions.

This paper aims at presenting a complex web-based decision support system, which meets most of those above characteristics.

3. Ensuring data quality for a complex web-based decision support system

Data quality is critical in any decision support system, particularly in complicated web-based systems. To create relevant insights and support decision-making processes, a decision-support system relies on accurate, trustworthy, and timely data. Maintaining data quality becomes considerably more difficult in a complicated web-based decision support system when data is collected from numerous sources and integrated into a central platform. Here are some crucial factors [9] to consider when addressing large data quality in such a system:

a) **Scalability**: Big data systems handle massive amounts, speeds, and types of data. Checking the data quality processes and technologies can handle the enormous amounts of data being processed.
b) **Data Pre-processing**: Big data is frequently noisy, incomplete, or inconsistent, and requires increasing the data quality before analysis, by using pre-processing techniques such as data cleaning, transformation, and normalization.
c) **Data Sampling and Sampling Bias**: due to its dimension, sometimes big data is required to deal with samples rather than the complete dataset. Ensuring that the sample procedure is representative and that sampling bias is avoided, guarantees expected analytic results.
d) **Data Integration Difficulties**: to achieve accurate and relevant insights when a variety of data sources are used, including organized and unstructured data, data integration methods must accommodate the heterogeneity of data formats, schemas, and semantics.
e) **Monitoring and Metrics for Quality**: big data quality measures take into account elements such as data accuracy, completeness, consistency, and timeliness. The solution is to implement strong monitoring methods to verify data quality and spot anomalies constantly.
f) **Parallel Processing and Distributed Computing**: Processing and analysing enormous datasets, sometimes requires distributed computing frameworks, to ensure that data quality processes can be conducted in parallel and dispersed across the computing cluster.
g) **Data Privacy and Security**: sensitive information requires the implementation of strong security measures to preserve data

privacy and maintain regulatory compliance, so anonymization techniques and access controls to protect the quality and integrity of your data become usual.

h) **Machine Learning and AI**: machine learning and AI algorithms used frequently in big data systems for analysis and decision-making require addressing the data quality processes, such as feature engineering, model training, and evaluation, which should be in line with the needs of these algorithms.

i) **Encourage user feedback and validation**: based on their experience and domain knowledge, decision support system users can contribute insights into data accuracy, relevance, and reliability. To increase overall data quality, users are asked to detect potential data quality issues in their daily interaction with Big.

By taking these factors into account the data quality is increased and specific quality rules could be generated at the level of a quality framework (as a main component of the web-based decision support system). Continuous monitoring, proactive data management strategies, and user participation are critical for improving the quality framework and for preserving data quality over time for determining the relevance of data-driven decisions.

4. General structure

The general structure of the Experimental Web Based DSS is shown in Fig. 1. The hardware configuration platform is composed out of many machines that run software applications independent of each other. They are called generically: **Core Banking Server** machine, **Data Warehouse (DWH) Server** machine, **Reporting Server** machine, **Web Server** machine and **Decidend** machine that are independent of each other.
They are described in the sequel:

a) **The Core Banking Server machine** is always a power server, which allows users to connect and use the capabilities and the facilities of the banking application. The core banking application can be one of the following core banking applications: **ABSOLUT** ***Core Banking, Oracle Flex Cube Core Banking, bMaster Core Banking etc.*** All these banking applications store the data in one of the following strong relational databases: Oracle Database 12c, Microsoft SQL Server etc.

b) On **the Data Warehouse (DWH) machine** there is also installed a database Server (as Oracle Database 12c, Microsoft SQL Server, IBM DB2 etc.) to store the historical data with Big Data methods and technologies. To minimize the cost with licenses or to reduce the cost of training with employees is recommended to use the same type of database on the both machines. At the DWH database level there is a data quality module with SQL rules that verifies daily the successful transfer of data from the core banking application tables to the DWH tables. There are used also many jobs with daily or monthly frequency that checks the rules, move or update incrementally the data in the DWH tables.

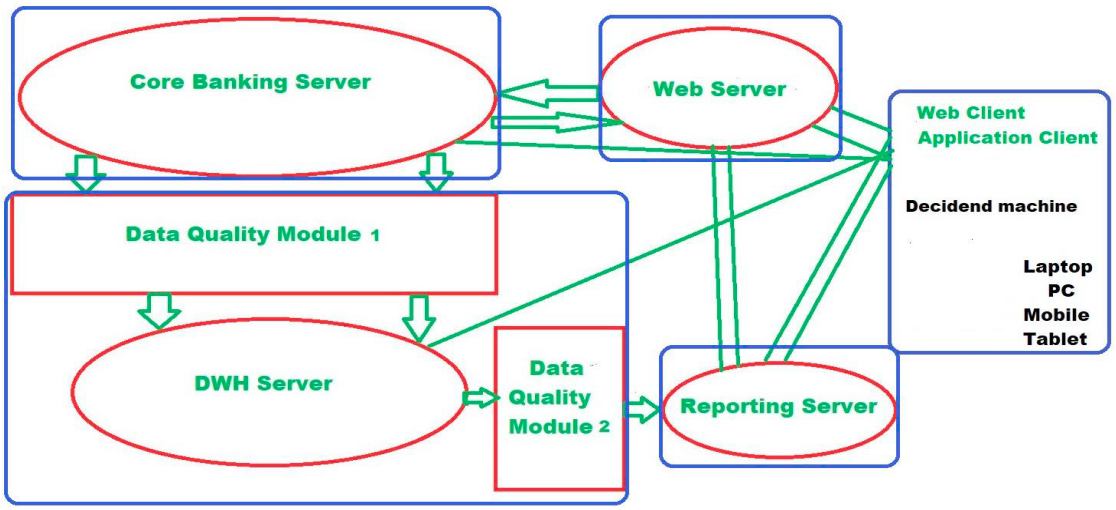

Figure 1. The general structure of the experimental web-based DSS

c) **On Web Server machine** there is installed a Web server and a Web Application that contains a nice interface that allowed the users to see the business reports. The Web server can be Internet Information Services (IIS), Apache HTTP Server or

another HTTP server. At the level of the Web server, a software driver allows the users to connect and make queries to the database.

d) **Reporting Server machine** can be a Repository Database, File System Server or Elite Reporting Centre where the reports are generated in Excel, CSV or PDF formats with daily and monthly frequency. Another possible solution is to use a mail server and have users receive the business reports in Excel, CSV or PDF formats by mail as an attachment.

e) **The Decidend machine** can be a laptop, a mobile phone, a PC or a tablet. There is a browser installed on these devices that can be allowed decision makers to connect to Web application and use the capabilities of the system.

5. Back-end components

In the case of our system, there are three back-end components:

a) **Transfer component.** It is composed of a large set of SQL jobs that that move or update incremental the data in the DWH tables. Most of all SQL jobs have a daily or monthly frequency. In the function of the size of the tables, the jobs add automatically partitions at every table of big size.

b) **Process Component.** It is composed of a large set of ETL (Extract, Transform and Load) jobs that prepare the data for *containing more years of historical data*. In the functions of the requests of the decision makers, the SQL tables have daily or monthly partitions to increase the speed of SQL queries used in the SQL views or tables.

ETL provides the foundation for data analytics and machine learning work streams. Through a series of business rules, ETL cleanses and organizes data in a way, which addresses specific business intelligence needs, like monthly, or daily reporting, but it can also tackle more advanced analytics, which can improve back-end processes or end user experiences. ETL capability is often used by an organization to extract data from legacy systems, to clean the business data, to improve data quality and establish data consistency and at the end to load data into a target relational database.

These SQL jobs contain a set of ETL, which process the raw data using decision-making models and business techniques, and at the end the result (the processed data) is stored into the SQL tables.

c) Reporting Component. It contains many business reports, which are getting from a large set of SQL jobs that run SQL stored procedures or views, which contain the business models for every business report. At the end, every report should solve the model of the business required by the decision maker.

6. Front end component

The decision makers interact with the Web Based Decision Support System by using any browser installed on a laptop, PC, mobile phone or Tablet. The authentication with the system is made using a user name and password or at the operating system level for the decision makers that are using an Intranet Explorer or Microsoft Edge, user account and they connect to the system from intranet network. The decision makers with special rights can view the following:

a) The aggregated data for all departments for every month or year to follow the increases or decreases of the cost amounts per employee.

b) The total exposure of the loans in original currency or in EUR equivalent at the NBR exchange rate [10], total number of loans for every month.

c) The total exposure of cards in original currency or in EUR equivalent at the NBR exchange rate, total number of cards for every month or year.

d) The total portfolio of deposits in original currency or in EUR equivalent at the BNR exchange rate, total number of deposits for every day, month or year.

e) The entire volume of transactions of cards in original currency or in EUR equivalent at the NBR exchange rate, total number of transactions for every day, month and year.

f) The entire volume of FX transactions for every currency calculated in original amount or conversion amount using NBR rate, total number of FX transactions for every currency in a day, a month or a year.

g) The total portfolio of new deposits in original currency or in EUR equivalent at the NBR exchange rate, total number of new deposits for every day, month or year.

The access rights can be limited for certain information at department level, group of employees or for particular decision makers.

7. Conclusions

Many organizations have tried to analyse a large set of business project proposals in order to choose project portfolios that maximize the performance, meet the resource constraints and the minimize the risk. The model of Web Based Decision Support System presented in this paper is intended to assist business managers to choose the best solution in the decision process.

Through the Web interface the proposed system can present graphical information to a multitude of collaborating users and it can be easily integrated with an expert system or other artificial intelligence-based tools (AI). It can also be intended for both business executives, top management decision makers and another group of knowledge workers.

The main benefits of the proposed architecture are:

a) Once the data is available on the Reporting Server, the values from the columns are automatically calculated and the decision makers will have a better speed performance in viewing the information;
b) The interface of the web system is very easy to use and very intuitive for the decision makers.
c) All data from the core banking database are stored for several years in one place in the DWH database using Big Data methods and technologies already calculated and available to the user with access rights;
d) Any decision maker can use a simple browser installed on his device to see the business reports or can use Microsoft Excel, Open Office or Acrobat Reader to view the business report;
e) The user with appropriate access right can access the system from virtually any internet location.
f) The relational databases and the software used are independent of each other.

The Web-based SSD concept described in this article can be easily implemented in any bank where the efficiency and automation of processes is desired so that decision-makers can choose the best solutions and minimize possible risks.

Further efforts are intended at building a more performing prototype to use digital cognitive systems and to explore your new ideas and show the intention behind a feature with a more correct placement of interface elements, increasing its usability [11], making the resources more attractive to decision-makers [12].

Consequently, the result of future business developments in ICT infrastructure and their impact on human well-being are not too easy to predict [13].

References

[1] Filip, Florin Gheorghe, Constantin Bala Zamfirescu, and Cristian Ciurea. (2017) *Computer- Supported Collaborative Decision-Making.* Springer International Publishing (http://link.springer.com/book/10.1007/978-3-319-47221-8).
[2] Filip, Florin Gheorghe. (2021) "Automation, computers, and their contribution to human well-being and resilience". *Studies in Informatics and Control* 30 (4), 5-18, https://doi.org/10.24846/v30i4y202101.
[3] Siddike, Md. Abul Kalam., Jim Spohrer, Haluk Demirkan, and Youji Kohda. (2018) "People's interactions with cognitive assistants for enhanced performances. in 51st Hawaii International Conference on System Sciences, HICSS 2018, Hilton Waikoloa Village, Hawaii, USA, January 3-6. pages 1-9, AIS Electronic Library (AISeL), http://hdl.handle.net/10125/50092.
[4] Boboșatu, Florin. (2008) *Advanced Decision Support Systems*, Doctoral thesis, Polytechnic University of Bucharest.
[5] Filip, Florin Gheorghe. (2008) "Decision support and control for large-scale complex systems ". *Annual Reviews in Control*, 32(1), 61-70
[6] Boboșatu, Florin, Alexandru Șerbănescu. (2007) "An Experimental Web-Based Decision Support System ", *Studies in Informatics and Control,* 265-270 https://rria.ici.ro/sistem-experimental-de-asistare-a-deciziilor/
[7] Boboșatu Florin, Luminita Duță, and Florin Gheorghe Filip. (2010) An Experimental Web - Based Decision Support System. IFAC Proceedings Volumes, Volume 43, Issue 8, 2010, Pag.43-46. https://www.sciencedirect.com/science/article/pii/S147466701533366 .
[8] Power Daniel (2001), "Supporting Decision-Makers: An Expanded Framework", University of Northern Iowa, https://proceedings.informingscience.org/IS2001Proceedings/pdf/PowerEBKSupp.pdf (last date found May 2023).
[9] Ehrlinger, Lisa, Wöß Wolfram. (2022) "A Survey of Data Quality Measurement and Monitoring Tools. Front Big Data ". https://doi.org/10.3389/fdata.2022.850611 (last date found May 2023).
[10] Isărescu, Mugur. (2016) "A quarter of a century since the reopening of the foreign exchange market in Romania ", https://www.bnro.ro/Home.aspx .
[11] Boboșatu, Florin, Alexandru Șerbănescu. (2007) *"*An Experimental Decision Support System *"*, *Romanian Journal of Information Technology and Automatic Control),* 55-62.
[12] Power, J Daniel. (2002), "Decision Support Systems: Concepts and Resources for Managers". Faculty Book Gallery. 67.https://scholarworks.uni.edu/facbook/67.
[13] Filip, Florin Gheorghe, (2022). Collaborative Decision-Making: Concepts and Supporting, Information and Communication Technology Tools and Systems, International Journal of Computers Communications & Control Online ISSN 1841-9844, ISSN-L 1841-9836, https://doi.org/10.15837/ijccc.2022.2.4732