DESIGN WORKFLOW FOR CLOUD SERVICE INFORMATION SYSTEM FOR INTEGRATION AND KNOWLEDGE MANAGEMENT BASED IN RENEWABLE ENERGY

Associate Professor Adela BARA, Professor Ion LUNGU, PhD Simona Vasilica OPREA[1]

Associate Professor George CARUTASU, Professor Cornelia Paulina BOTEZATU, Associate Professor Cezar BOTEZATU, [2]

ABSTRACT:

The article presents the state of the art in decision support systems in energy field and further, the future work plan for design development of cloud service information system for integration and knowledge management based in renewable energy. The research is part of SIPAMER project, financed by NASR agency.

Keywords: renewable energy, decision support systems, data mining, business intelligence

1. INTRODUCTION

Decision-making process, on the organization level, generally is very complex, as is revealed by the proposed definition in [1], where the process is seen as all phases, processes which determine their objectives and embedded subsystems using a complex methods and techniques in order to achieve more efficient the reasons which led the establishment of such organization.

Decision-making involves making decisions that are classified according to [1] in terms of: functional content (organizational, managerial and planning decisions), decision-making level (strategic, tactical and operational decisions), objectives achievement certainty (certain, uncertain and risk decisions).

As is described in the book [2], making decision involves the following elements: establishing targets to be achieved, finding more alternative action plans, including in economic constraining factors decisional map (resources, time, work), substantiating the decision on scientific level and its formulation in terms that can be understood and applied. Organization management involves different activities types, and therefore, requires different types of information and to efficiently manage and process this information and give them for analyses into a synthetic form as it requires a decision support systems type (DSS) developed on entire organization level.

Nationally, in the renewable power plants, at the moment, renewable resource management isn't supported by a decision support system that could enable efficient monitoring and analysis of energy resources from these sources. In Europe, there are several countries that developed decision support systems used for more effective management of renewable resources (e.g. Germany, Spain), but the building cost of these systems are high and specific national energy potential makes it impossible application of these methods in

---

[1] The Bucharest University of Economic Study, Bucharest - Romania

[2] Romanian-American University, Bucharest - Romania

Romania. In terms of energy production forecasting systems, there are a series of software, e.g. production monitoring system based on the parameters of wind produced by GreenByte Sweden or forecasting services provided by the Fraunhofer Institute, Wind Energy and Energy System Technology. Problems related with the wide application of these solutions in Romania are related to the high costs related to forecasting services, but especially the errors recorded by these systems are caused primarily by the peculiarities of wind farms operation in areas with potential in Romania.

2. DECISION-MAKING PROCESS' REQUIREMENTS

The requirements of the decision-making process derive from the characteristics of the renewable production.

Thus, in wind power plants the factors are described in [3], [4], [5], but the most important factor consist in weather condition that affects the wind power generation. This leads to power fluctuations that require tertiary reserves.

The energy produced in photovoltaic power plants depends on technical parameters of photovoltaic panels, temperature variations, coefficients of power output and voltage at the maximum power of the photovoltaic panels and solar radiation level. The photovoltaic power plants energy production is more predictable, having determined how productive is on cycle day/night. But in this case the output is reliant on meteorological factors resulting power fluctuations.

In small hydro plants' case, the generated power is dependent on water flow, being more predictable than wind power generation, the fluctuations caused by water flow being lower. Those characteristics, operating conditions and solutions for monitoring activities are presented in studies and research projects [6].

The development of such system presents a number of issues and risks related to the following aspects:

1. **Heterogeneous systems and data sources** - in installed renewable power plants there are installed different monitoring systems (e.g. SCADA/EMS) that use a variety of algorithms for the allocation and management of equipment. From these plants originate information, forecasts and predictions with different errors from various types of applications and local devices. Based on the online transmission data system using EMS-SCADA, the TSO and NDC receive a series of notifications regarding the state of the elements and functioning of the generating groups. For obtaining the required information for the tactical and strategic decision making process, integration techniques and data processing, solutions for advanced analysis and data presentation in a friendly user interfaces are needed.
2. **Inaccurate predictions of the production** - currently, for wind power plants haven't implemented effective solutions for efficient forecasting and errors over 10-15% are being recorded. In order to assist the decision process, it is necessary to have accurate prediction that can lead to a solid base for financial plans and technical estimation. These lead to the need develop solutions to increase the accuracy of the forecasting methods.
3. **Analysis, simulation, and activities planning models' complexity** - data organized and integrated in the same manner must be processed and organized in order to be converted into information and useful knowledge for strategic decision making process. Such a requirement can only be met through the application of multidimensional analysis, simulation, forecasting and planning models for

technical and business processes. In decision support systems, there are several solutions already implemented in various areas, especially in the financial domain. There are many software companies worldwide that offer various systems and applications to assist decision-making, but these solutions are very expensive, their implementation usually lasts 3 years and have many limitations in terms of flexibility and adaptation to specific requirements. To assist decision-making in the production companies, researches funded by companies are conducted, but actually there aren't yet any implemented systems. Nationally there isn't a unified solution that can be successfully applied to the renewable power plants. In this field, the authors conducted a series of analyses into the research projects [7], [8], but also in other scientific papers [9], [10], [11].

4. **Infrastructure and technologies –** for decision process support it is necessary to conceive a system's architecture that contains data integration technologies from heterogeneous sources, forecasting and analysis technologies and advanced presentation technologies. The data extraction, processing and analysis requirements are high in terms of volume and structure, but also in terms of the response time required for the analytical reports. Therefore, the necessary hardware resources involve powerful servers, networks and access points, next-generation devices for presenting information. Also, technologies, systems and software packages that are used involve a series of high costs related to licensing and configuration. The risks in this case are related to: the impossibility of the renewable power plants to allocate significant financial resources for proper calibration of the system; the failure to recover investments made in infrastructure due to weaknesses in the implementation or use of the developed information system.

3. SOLUTION AND APPROACH

To properly address of issues listed above, it is envisioned a complex set of research instruments and technical approaches. To face the issues, a set of solutions are proposed:

1. **ETL process.** A solution of data processing and integration from heterogeneous systems and data sources might be to develop an ETL process (Extract, Transform and Load) in order to load data into a centralized database and then into a data warehouse to achieve advanced analysis and simulations.
2. **Data mining algorithms.** Fundamental research conducted by the authors in this area have shown that a series of data mining algorithms for wind energy production recorded accurate results mainly through power thresholds; the estimate is much closer to actually measured values.
3. **Technical and business processes monitoring and analysis model**. The authors envisage to design and develop a model for technical and business processes monitoring and analysis, in order to optimize the information flows and to increase economic efficiency of the production companies. To achieve the objective, business processes will be detailed, models and calculations of efficiency by defining performance indicators will be develop. For geospatial information representation is used Geographic Information Systems (GIS), which allows creation of maps, scenarios viewing etc. Implementing of such

model provides a general view over the organization's management, enables coherent activities planning, optimize decision making process and resource consumption. Such a model designed for technical and business processes analysis will enable planning support, control and monitoring activities, corrective action, audits and analysis. The system can be supported by a document management component that will allow the end-users to efficiently store the content in a centralized system, control access permissions to the content, and submit the content to different workflows of approval. The implementation of a document management solution reduces risk and lowers costs associated with the regulatory and legal compliance processes in the field of renewable energy.

4. **Business Intelligence and Cloud Computing*:*** Currently there are Business Intelligence (BI) technologies which can be used in building this architecture. A solution is given by the architecture where computing power, databases, storage and applications coexist in a complex network of servers that allows users to access the information as services, accessible via the Internet using personal computers and mobile devices. The concept is fairly recently defined (see [12]), but has evolved fulminate, many companies offering such services. Advantages of Cloud Computing platform are related to significant cost reduction for renewable energy producing companies related to a minimum implementation infrastructure, high speed of data processing and analysis, the information system being offered as a service in hybrid cloud architecture. For analysis and reporting access, authors take into account to develop the solution into an online platform containing all sub-modules and components providing centralized access to components for management, analysis, prediction, collaboration, etc.

These technologies can be combined in a DSS's architecture (figure 1).

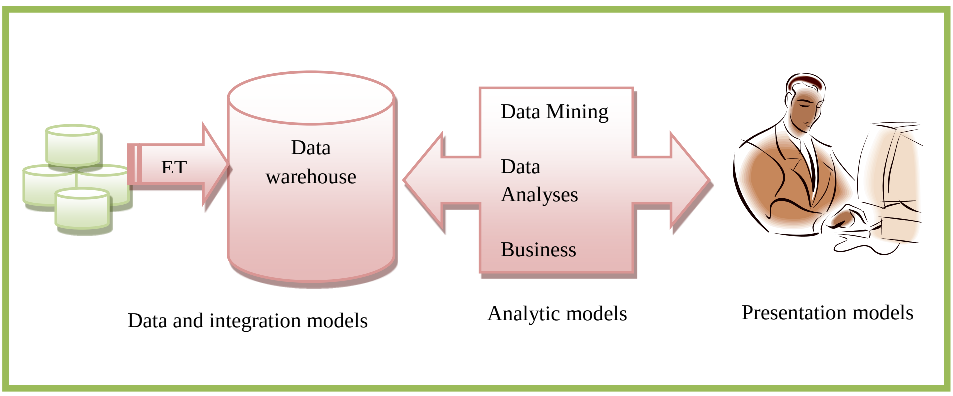

Figure 1 – The DSS's architecture

The architecture will be refined and we'll propose several components for a DSS prototype.

4. WORKPLAN AND OUTCOMES

In the following is described the major development phases for cloud service information system for integration and knowledge management based in renewable energy, structured according to the decision support systems development cycle (figure 2):

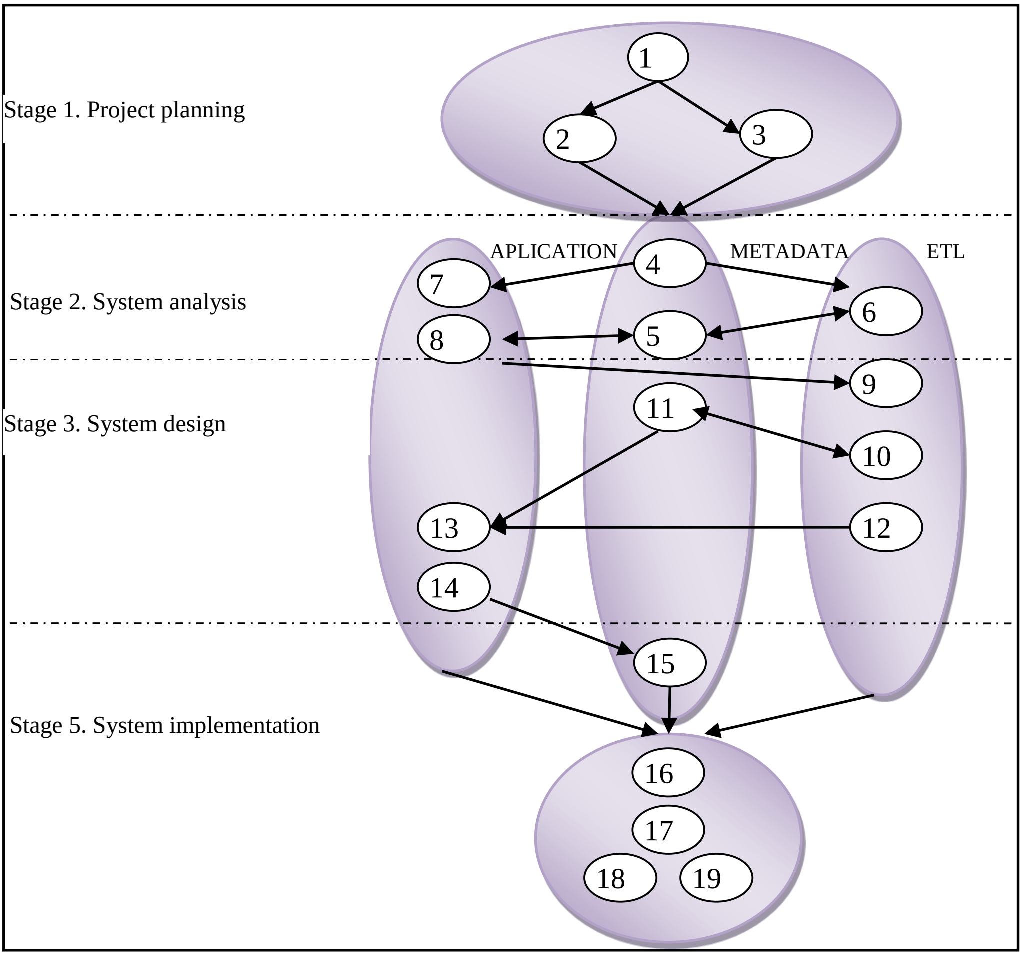

Figure 2 – The development lifecycle of the proposed decision support system (source: adapted form [3])

Stage 1: Planning

Step 1: Preliminary evaluation - the general requirements of the project are discussed and evaluated.

Step 2: Evaluation and design of infrastructure - this step estimates the equipment necessary to carry out the project by analyzing the following: infrastructure, components, devices, network equipment and any technical requirements. In this step the technical infrastructure of the project is built with secure access for all sides involved.

Step 3: Project Planning – in this step the project plan is detailed progressively, each stage.

Stage 2: System Analysis

Step 4: Analysis of the data source and data flows - meetings are organized with consultants, business analysts, experts and decision makers from the pilot companies where the project will be implemented, after which the data flow requirements necessary for tactical and strategic decision making are identified and defined. Based on this, there will be identified and analyzed the types of data sources needed and will examine how to integrate them into a centralized database at organizational level. Overall the database schema and data integration flows will be defined, by designing the entity-relationships diagram with significant attributes and constraints. An initial solution is proposed, discussed and adopted.

Step 5: Analysis of organizational data warehouse metadata – the system must use multidimensional models and allow some good predictions and getting key performance indicators. It establishes the general framework appropriate for the data integration system. In order to analyze both data and metadata, CASE tools and specific UML diagrams will be used.

Step 6: Analysis of multidimensional models - necessary analysis models will be identified and we will analyze the way to obtain a set of key performance indicators and patterns of information necessary for strategic and tactical decision making processes.

Step 7: Analysis models for good predictions - based on the characteristics of renewable energy, data mining algorithms will be analyzed.

Step 8: Specify the system architecture components - three levels of achievement will be considered: data level, analysis models and presentation level. For each level there will be specified components and interconnections between them.

Stage 3: Designing the system

Step 9: Designing a model for data processing and integration - starting from data sources and business processes the design of a centralized database will be taken into consideration. As a result the detailed schema will be obtained for the centralized database in the cloud environment for prediction and analysis.

Step 10: Designing the data warehouse and the ETL process (extract, transform and load) – the data warehouse will be built in a cloud environment, data sources being represented by the centralized database. ETL design includes data mappings and transformations of the centralized database and the methods to load them in the data warehouse and data-marts.

Step 11: Experimental data model development - in this step, logical level elements from the previous steps are physically designed. This involve developing patterns for data extraction and integration workflows using filtering tools, cleaners and operators, mapping techniques and schemas. The schema is detailed for the central database and the result will be its physical schema. To validate the experimental data model, previous implemented procedures are launched on a dataset of sample data available from the pilot companies. Following, data warehouse objects are created, mappings with data sources and database items will be prepared to load partial data so that we can move to the next step, the design and development of analytical models. The infrastructure of the system is integrated into a Private Cloud model.

Step 12: Designing models for data analysis – designing the indicators system, based on the requirements identified in the first analysis stage and data model specified in step 2, and design analytical models to obtain these indicators.

Step 13: Designing prediction models - this step is dedicated to the design of the prediction model based on the analysis of the algorithms performed at step 7. Using the data model, data mining algorithms is designed and optimized to be executed on cloud architectures. This predictive model will be implemented and validated on a test data set originating from pilot companies.

Step 14: Development and validation of the experimental model for data analysis - functionalities for the analysis model will be partially implemented, using the centralized database and data warehouse partially validated in step 2. This partial validation functionality is to reduce risks in the implementation phase of step 4 and will result in an experimental analysis model partially validated.

Stage 4: Implementation

Step 15: Data processing and integration model in the cloud computing architecture – the other modules of the data model will be implemented based on the specifications described in step 2. Procedures for extraction and processing of data from heterogeneous sources, database schema and data warehouse will be refined and implemented. Also, data access interfaces will be built and cloud environment will be configured for data access.

Step 16: Implementing data analysis model - this step includes the implementation of the framework for analyzing data in a test environment within the target group and assumes full implementation of analytical models to obtain performance indicators. Security features specific to cloud environment will be implemented.

Step 17: Testing analysis and data models - a series of formal tests will be performed to check whether the system meets the requirements documented in the analysis phase and refined during the project implementation. The test will be based on a comprehensive test plan, prepared during implementation. Possible problems encountered and proposed solutions will be documented.

Step 18: Evaluation of the solutions for industrial implementation of the prototype - following the successful completion of prototype testing, the large-scale implementation will take place, targeting renewable energy units in Romania.

Step 19: Patenting developed framework - the project team will submit required documentation for product patenting.

5. CONCLUSIONS

The general issues of producing renewable energy and overall energy system is depicted by authors, focusing on decision-making process and specific requirements. The article presents the design and development workflow applied for cloud service information system for integration and knowledge management based in renewable energy. The main issues that this system addresses are: heterogeneous systems and data sources, inaccurate predictions of the production, analysis, simulation, and activities planning models' complexity and infrastructure and technologies.

Solution provided by authors, regarding above presented issues enhance a complex approach, using specific meanings: ETL process, data mining algorithms, Technical and business processes monitoring and analysis model and Business Intelligence and Cloud Computing.

Moreover, the authors presents a detail plan from design to patent, covering the whole chain of development, starting from stating specific requirements of energy field to final implementation and system delivery. The design workflow may adapted to other fields where a Decision Support Systems is foreseen.

6. ACKNOWLEDGEMENT

This paper presents some results of the research project: *Sistem inteligent pentru predicția, analiza și monitorizarea indicatorilor de performanță a proceselor tehnologice și de afaceri în domeniul energiilor regenerabile (SIPAMER),* research project, PNII – Parteneriate în domeniile prioritare, PCCA 2013, code 0996, no. 49/2014 funded by NASR.

REFERENCES

[1] O. Nicolaescu, I. Verboncu, *Fundamentele managementului organizatiei*, Tribuna Economică Publishing House, 2001, ISBN 973-934-890-4;
[2] I. Lungu, A. Bara, *Sisteme informatice executive*, ASE Publishing House, 2007;
[3] T. Ackerman, *Wind Power*, John Wiley & Sons, 2005, 742 pagini;
[4] T. Burton, D. Sharpe, *Wind Energy Handbook*, John Wiley & Sons, 2001, 642 pagini;
[5] L. Landberg, G. Giebel, H.Aa. Nielsen, T.S. Nielsen, H. Madsen. *Short-term prediction - An overview*, Wind Energy 6(3), pp.273-280, 2003;
[6] PNCDI –CNMP - *SIMULATOR VIRTUAL PENTRU EXPLOATAREA UNEI MICROHIDROCENTRALE – MICROSIM*, Contract nr: 12-132 / 01.10.2008;
[7] G. Carutasu, C. Botezatu, CP Botezatu - *PNCDI II / Parteneriate nr.2387/2007 Dezvoltarea şi implementarea sistemelor integrate de management în domeniul energiei* DI-SIM, 2008-2010;
[8] I. Lungu, M. Velicanu, A. Bâra, V. Diaconiţa, I. Botha - *Soluţii informatice de asistare a procesului decizional şi pentru dezvoltarea managementului bazat pe cunoştinţe în instituţiile publice* GRANT CNCSIS, Ideas Program, PN II, code CNCSIS 820, 2007-2010;
[9]A. Bâra, A. Velicanu, I. Lungu, I. Botha, *Using Geographic Information System for Wind Parks' Software Solutions*, International Journal of Computers, nr. 2, vol. 5, 2011, pp. 149-156, ISSN 1998-4308
[10] G Carutasu, C Botezatu, S Ionescu, CP Botezatu, A Joita – *Sisteme integrate informatice pentru management*, Universitară Publishing House, Bucharest,. 2011, ISBN 978-606-591-106-2, 211 pag;
[11] G. Carutasu, C. Botezatu, CP Botezatu, - *Knowledge management System for Implementing Integrated Management System in Romanaian Industry*, Proceedings of the 19th International DAAAM Symposium, 2008, ISSN 1726-9679, ISBN 978-3-901509-68-1;
[12] D.S. Linthicum, *Cloud Computing and SOA Convergence in Your Enterprise: A Step-by-Step Guide*, Addison-Wesley Information Technology Series, 2009, ISBN 978-013-600-922-1;