Techniques and Process Based Decision Support in Business Intelligence System

P.Pasatcha
Mahanakorn University of Technology, Thailand
parichat_noina@hotmail.com

Abstract

Business Intelligence (BI) has been viewed as sets of powerful tools and approaches to improving business executive decision-making, business operations, and increasing the value of the enterprise. The technology categories of BI mainly encompass Data Warehousing, OLAP, and Data Mining. This article reviews the concept of business process management to current business intelligence system and provides a survey, from a comprehensive point of view, on the BI technical framework, process, and architecture for business intelligence. In addition, the conclusions point out the business process management to current business intelligence system so as to provide the business intelligence system with the ability of process-driven decision making.

Key words: Business intelligence, business process management, business decision-marking, knowledge management

1. Introduction

As businesses continue to use computer systems for a growing number of functions in today's competitive, fast-evolving world, most companies face the challenges of processing and analysing huge amounts of data and turning it into profits. They have large volumes of detailed operational data, but key business analysts and decision makers still cannot get the answers they need to react quickly enough to changing conditions because the data are spread across many departments in the organisation or are locked in a sluggish technology environment. In these cases, Business Intelligence (BI) is resented, which are sets of tools, technologies and solutions designed for end users to efficiently extract useful business information from an ocean of data. [1]

Business Intelligence, a phrase coined by Howard Dresner of Gartner Group in 1996, which has become popular since the 90s. Nowadays, most business intelligence tools are based on data warehouse and data analysing software. As a kind of data-driven decision support system [3], business intelligence tools on the data layer and decision making is based on data analysis reports. When a new task comes, traditional business intelligence tools usually dig into the raw data stored in the data warehouse and then try to make a report through a time consuming process using OLAP, data mining and other data analysis software. Such kind of decision making is not efficient enough when solving new problems.

In order to improve the efficiency of the business intelligence system, they need to transform data into useful information and knowledge management techniques are then implemented to manage such information and to support decision making. On the other hand, today's companies are more process-oriented than in the past and process-driven decision support system is emerging in business operations. In this paper I introduce the concept of business process management to current business intelligence system so as to provide the business intelligence system with the ability of process-driven decision making. The processes stored in the process model base are flexible and reusable.

The remainder of this paper is structured as follows. Section 2, introduce the architecture of the business intelligence system. Section 3, describes a related literature for business intelligence. In section 4. there is a discussion followed by the conclusion in section 5.

2. Business intelligence technical architecture

A. Review Stage

The concept of Business Intelligence was first introduced by Garter Group in 1996 [2], and incipiently referred to the tools and technologies including data warehouses (or data mart), reporting query and analysis. According to Revelli [3], Business Intelligence "is the process of collection, treatment and diffusion of information that has as an objective, the reduction of uncertainty in the making of all strategic decisions". Today's BI evolves from EIS (Executive Information System), DSS (Decision Support System), query and reporting tools and multidimensional analysis which also known as OLAP (On-Line Analytical Processing). [4]

B. Technical Framework

Business Intelligence framework is never delivered by way of a single technology, product or vendor. The very nature of BI encourages business users to have more access to, and control over, the data. The successful application of BI in an enterprise should consider the following two points. The first is correct, valid, integrated, and in-time data, and another is the means which will transform the data into decision information. However, neither satisfied data nor effective means are easily acquired. BI technical framework is used to solve the two above questions.

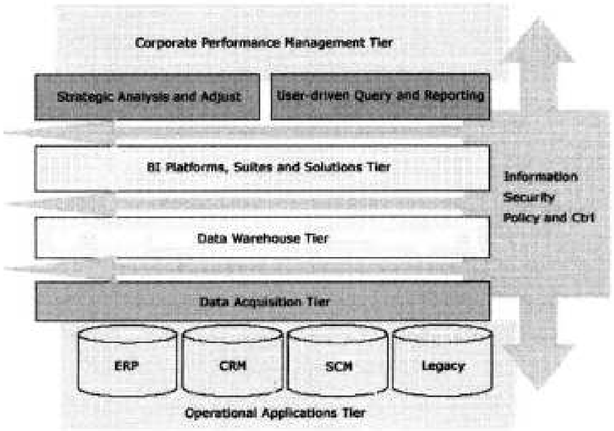

Figure 1. Generic BI Technical Framework. [5]

The framework consists of the Operational Applications Tier; Data Acquisition Tier; Data Warehouse Tier; BI Platforms, Suites and Solutions Tier and finally the Corporate Performance Tier. The Operational Applications Tier has systems such as Enterprise Resource Planning; Customer Relationship Management; Supply Chain Management. and Legacy. Extraction, Transformation and Loading obviously belong to the Data Acquisition Tier. Besides Data Warehouse, Data Warehouse Tier includes Data Marts and the Operational Data Store. Data Warehousing, OLAP [6], and Data Mining [7], are three of the most significant technologies in this tier, as there is a "very large" repository of historical data pertaining to an organization. OLAP refers to the techniques of performing complex analysis over the information stored in a data warehouse. The complexity of queries required to support OLAP applications makes it difficult to implement using standard relational database technology. Data mining is the process of identifying and interpreting patterns in data to solve a specific business problem. Data mining strategies for BI include classification, estimation, prediction, time series analysis, unsupervised clustering, and association analysis or market basket analysis, e.g. Apriorio , Decision Tree Induction, Support Vector Machine, Nearest Neighbor, Genetic Algorithms, Rough Set, Fuzzy Set [14], K-means, Case-Based Reasoning, etc. It could be said, that BI is technically a combination of several disciplines and techniques mentioned above instead of an independent, novel or original approach. [8]-[10]

C. Architecting principles for business Intelligence

The principal objectives of business intelligence can be summed up as follows: [11]

• To provide a “single version of the truth” across an entire organization.

• To provide a simplified system of implementation, deployment and administration.

• To deliver strategic, tactical and operational knowledge and actionable insight.

Due to the focus on information in business intelligence applications, the privileged point of view of the supporting architecture has to be the information view [8]. From this point of view, the most popular paradigms [9, 10, 11] are:

• The hub-and-spoke architecture with centralised data warehouse and dependant data marts

• The data-mart bus architecture with linked conformed dimensional data marts

• Independent non-integrated data marts

Figure 2 represents a typical layered view of architecture for business intelligence.

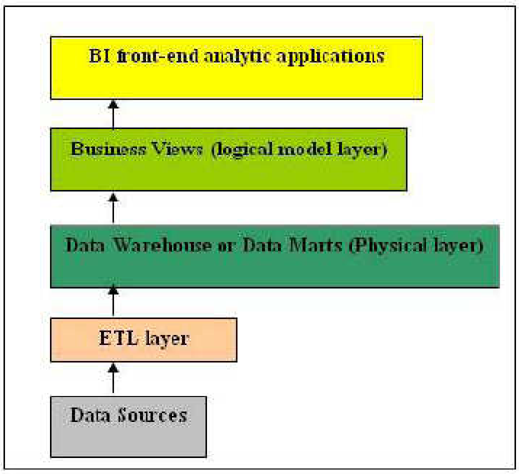

Figure 2: conceptual architecture for business intelligence. [15]

In today’s heterogeneous environments where many disparate systems and domains hold different parts of the necessary data, the most difficult challenges in achieving the above mentioned objectives are effective information delivery and technology integration.

3. Related Works

Conversely, several researchers have developed architecture for Business intelligence systems with the ability of processes-driven decision making.

Srivastava and Cooley [12] developed a general architecture for web-based business intelligence systems that can be applied to individuals or organisations. Their work led to the identification of the two major components of web-based business intelligence systems that guide this research. Namely (1) content acquisition. and (2) knowledge creation. Content acquisition refers to the automated acquisition of relevant, detailed, and reliable information from the web and knowledge creation refers to the generation of new knowledge through pattern discovery and prediction. The content acquisition component is by far the more problematic of these two functions. From a temporal perspective, content acquisition precedes knowledge generation and results in a structured, mineable data set. Whereas knowledge creation can be addressed by traditional data mining techniques, automated content acquisition is a far more challenging topic.

Daniel [13] developed the Automated Web-based Business Intelligence (AWBI). The AWBI architecture described builds on previous research from Srivastava and Cooley, but differs from and extends it in several important ways. Firstly, the framework described in this paper focuses specifically on automated web-based business intelligence systems; i.e. those web-based business intelligence systems in which content is acquired entirely from the web by an automated agent. AWBI systems by definition do not allow for manual data gathering from the web. Second, AWBI systems as operationalised herein are intended to provide a supplementary source of decision support information that can be integrated into an existing organizational decision making infrastructure. They are not intended to be standalone decision support systems. Third, the AWBI architecture is proposed only in support of organisational decision making for competitive advantage. AWBI systems do not fall within the realm of personal decision making agents such as those offered as services to individual web users in the financial, travel, and comparison shopping domains. The proposed AWBI architecture is depicted graphically in Figure 3.

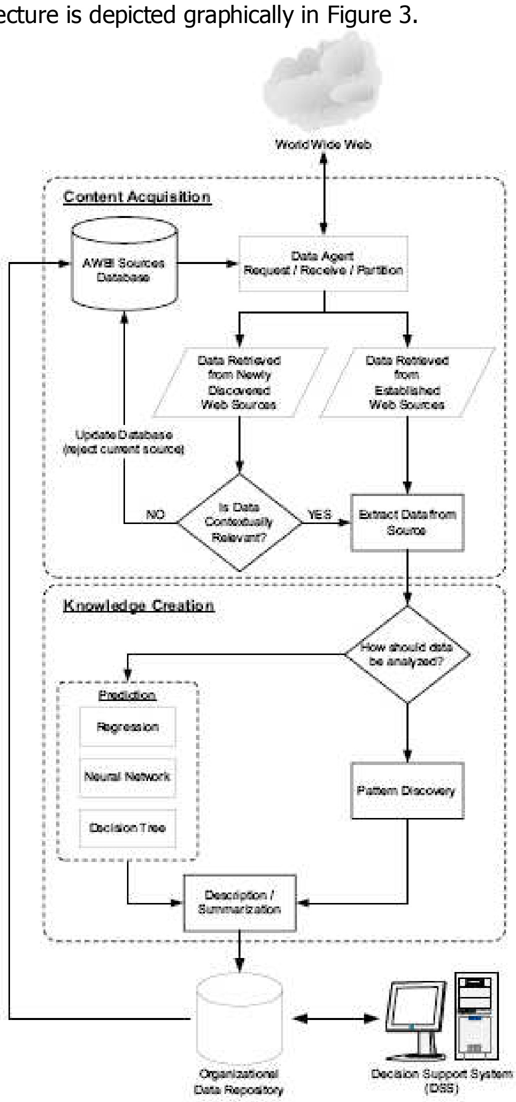

Figure 3. The AWBI architecture. [13]

Every AWBI system should contain an automated exploratory agent, the purpose of which is to seek out and identify new sources of web data that may potentially be relevant to the organizational decision making process. AWBI systems should not, however, rely solely on newly discovered web sources, but should establish semi-permanent relationships with web sources that are known to be dependable and trustworthy. These web sources can then be accessed at regular intervals by the AWBI system. A key component in the content acquisition process within the proposed AWBI framework is the data agent. The intent of the data agent is to breakdown the flow of information between the AWBI system and the web. It does this by negotiating with the AWBI sources database to request information from both known and unknown sources. When the requested information is retrieved from the web, the data agent partitions it based on prior knowledge regarding the relevance of the data source. If the source is established and is known to be dependable, then the data does not need to be examined for contextual relevance. Conversely, if the data is from a source unknown to the AWBI sources database, then the contextual relevance of the data must be determined.

The prototype AWBI system built on the architecture described in the previous section led directly to improved decision making performance, thereby providing evidentiary support in favor of the proposed AWBI framework for use as a supplementary source of decision-making data for organisational decision support systems.

The goal of AWBI systems is to automatically acquire contextually relevant information from the web regarding the environment outside the organisation, and use that information to inform the corporate decision making process, thereby yielding competitive advantage. Although the prototype described herein lends support to the notion that AWBI systems can serve as a valuable addendum to corporate decision support endeavors, most organizations have yet to undertake an AWBI initiative. As such, the AWBI architecture presented is both timely and desirable.

The value to organizational decision makers of having this sort of information available to them in real-time can hardly be overestimated. Ultimately, AWBI systems of this sort may have the potential to replace corporate espionage activities altogether. To that end, research related to these types of AWBI systems may indeed be very fruitful.

Denilson Sell et al [14] developed an architecture in which conceptualisations for business analysis can be captured, represented and processed in order to offer the decision maker more tailored and flexible exploratory functionalities. This architecture developed for building semantic-based analytical tools. In this approach use IRS-III, a framework and implemented infrastructure for developing Semantic Web Services compliant with WSMO . The main meta-models of WSMO are Ontologies, Goals, Web Services and Mediators. Ontologies provide the basic glue for semantic interoperability. Goals represent the types of objectives that users would like to achieve via Web services. Web services descriptions describe the functional behavior of an actual Web service. Mediators provide the means to link two components together, defining mappings between them. In IRS-III a published web service may be selected during a selection process and then invoked for achieving a goal.

The architecture is composed of a set of loosely-coupled modules that are illustrated in Figure 4. The first describe the semantic infrastructure based on ontologies and Semantic Web Services and then they describe how the functional modules handle the integrated knowledge models

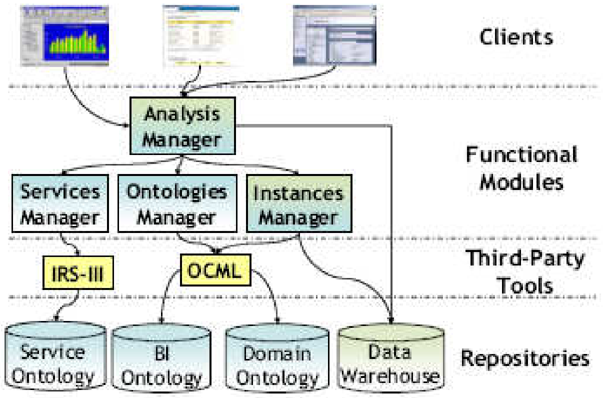

Figure 4. Illustration of the OntoDSS modules. [14]

In this approach, they use ontologies to capture business semantics and to define the necessary knowledge models for generating flexible exploratory functionalities in analytical tools. More specifically, they use OCML [16] for creating the business intelligence (BI) model, the service models and the application domain models. They are then able to define semantic functionalities including filters, relation navigation and Semantic Web Services.

This approach supporting query rewrite is similar to the one presented in [17], but they apply more extensively the business semantics described in the Domain Ontology to support the rewrite of query conditions and to combine OLAP features in this process. The rewrite process is guided by the users of analytical tools, who choose which semantic filter will be applied to extend the results of their queries. For instance, the user could be interested in finding products that are related to the same category of one particular product listed in the result of the analysis. The BI domain ontology models the main concepts related to business intelligence. This representation takes advantage of the structure of data sources in terms of dimensions and facts in the DW. An analysis formatted by a user is defined in this ontology as shown in listing 1. This definition includes dimensions, measures, filters, privileges and parameters. Parameters are used to bind values in a filter definition. Like dimensions and measures, parameters are instantiated with domain concepts defined in the Domain Ontology. Thus, the analysis concept can be used in any business domain.

OntoDSS architecture in order to illustrate some of the functionalities provided by a semantic web based analytical tool. By using OntoDSS, users browse the semantic definitions of their data sources in order to select the data items that they want to include in their analysis.

Listing 1. Partial OCML definition of the classes used in the Analysis definition.

```
(def-class Analysis () (
 (has_description :type string)
 (has_measures    :type Analysis_Measure)
 (has_dimension :type Analysis_Dimension)
 (has_filter    :type Analysis_Filter)
 (has_parameter :type Analysis_Parameter)
 (has_creator   :type User)
 (has_allowed_user    :type User)
 (has_allowed_role    :type Role)))

(def-class Analysis_Filter () (
 (has_attribute :type DB_Attribute)
 (has_operator :type DB_Operation)
```

Besides the traditional OLAP functionalities, OntoDSS provides a set of exploratory functionalities relying on the semantic descriptions defined in architecture repositories. Figure. 5 illustrates the results of an analysis defined by the user in OntoDSS in a specific business scenario. This application scenario includes information about universities and the correspondent number of students and researchers. Users can right-click over any of the columns in the analysis and see what options are recommended by the tool. OntoDSS automatically identifies which concepts were used in the analysis definition and uses these concepts to recommend options to the user.

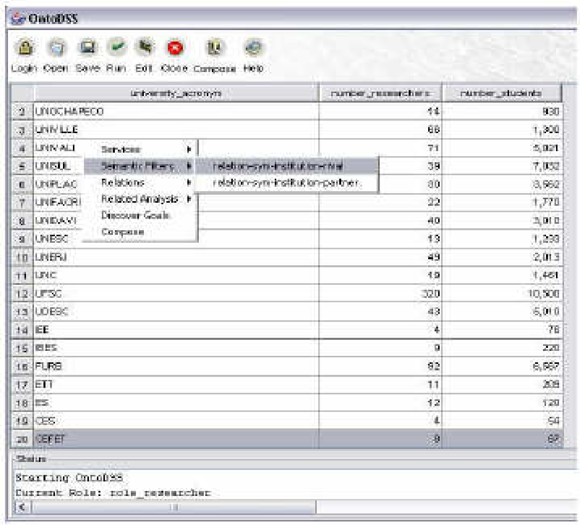

Figure 5. Illustration of OntoDSS functionalities. [14]

Relations also support the generation of exploratory operations over the results of an analysis without necessarily relying on the relationships between dimensions and fact tables. One can define a relation such as “product is supplied by company” and use this relation to analyse the prices offered by different suppliers, or the volume of sales by supplier.

Current analytical tools do not offer scalable ways for the generation or extension of functionalities. Frequently, users have to develop the required extension from the scratch, even when there is a code already implemented that could be used. Existing codes found in the organisation or on the Web can be semantically described and easily integrated to analytical tools through their architecture. Such semantic description supports the discovery, composition and invocation of services related to the users' analysis. In addition, definitions of business concepts and logic are done directly in the Domain Ontology, so one can modify or include definitions and change the behavior of analytical tools at any time. Therefore, modifications no longer depend on hours of coding, and the maintenance of the system can be done remotely.

Luan Ou and Hong Peng [15] introduce the concept of business process management to current business intelligence system so it provides the business intelligence system with the ability of process-driven decision making. The processes stored in process model base are flexible and reusable. Through using case based reasoning technology, the most similar process will be retrieved and show to the decision maker when solving a new problem.

The KBBIS system (Knowledge Based Business Intelligence System) is composed of five main components, as shown in figure.5 : the inference engine, the model base management system, the data storage management system, the knowledge base management system, the workflow management system. The inference engine combining case-based reasoning engine and rule-based reasoning engine is responsible for the control of reasoning mechanism, knowledge interaction and management. The data storage includes database, data warehouse and other data sources. The model base is composed of a process model base and a mathematical model base. Business process models are stored in the process model base. Mathematical models for analysis, evaluation, or forecasting are stored in the mathematical model base. Model generation, configuration and management are the responsibility of model based management system. The knowledge base stores business rules and related knowledge of many different domains.

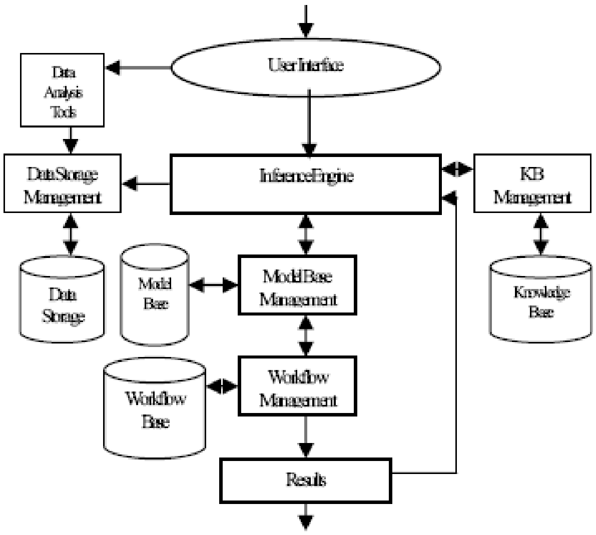

Figure 5. System architecture of KBBIS [15]

The workflow management system is in charge of workflow generation, configuration and execution. Process models from the model base are converted into executable workflows by the workflow management system. When a new task comes, it is first identified by the reference engine. If the task is a question query, domain or problem specific knowledge will be extracted from knowledge base to answer the question. If the task is to seek solutions for a certain complex problem, the most similar process will be retrieved from the process model base. The user can also make his own process with the aid of the model base management system and reference engine. The system is developed base on J2EE architecture.

In the knowledge based process model management system (KBPMS), they use hybrid technology of casebased reasoning (CBR) and rule-based reasoning (RBR) to facilitate reuse and management of process models.

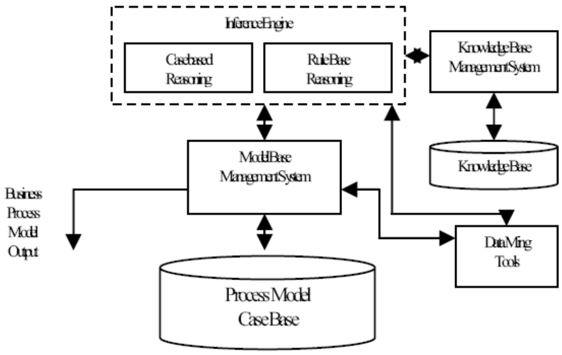

Figure 6. Framework of KBPMS. [15]

They development of the architecture of knowledge based process model management system (KBPMS) is shown in figure.6, a framework for facilitating process model reuse implementing hybrid of rule-based reasoning (RBR) and case-based reasoning (CBR) techniques. The innovative features proposed are: ontology based business process models representation which is flexible and can be easily extended and inherited; combination of rule-based reasoning and case-based reasoning that can overcome the limitations of these two methods and get more satisfactory results; strategy for knowledge management in business intelligence system which is the central part of business intelligence information accumulation and application. Currently, they are implementing the KBPMS on a prototype of CRM (Customer Relationship Management). The results of application show that decision mechanism based on knowledge and the process model can help the business intelligence system improve the efficiency of problem solving.

Liya Wu [16] developed a service-oriented architecture of IT Performance management system for business intelligence which makes possible a seamless integration of technologies into a coherent business intelligence environment, thus enabling simplified data delivery and low-latency analytics. The first step in re-architecting legacy system is the break down the legacy components into service-oriented reusable components able to communicate through open standard messaging protocols, based on XML, WS-* and SOAP. The resulting service-oriented architecture of IT performance management system (SOA-ITPA) is described by figure 7.

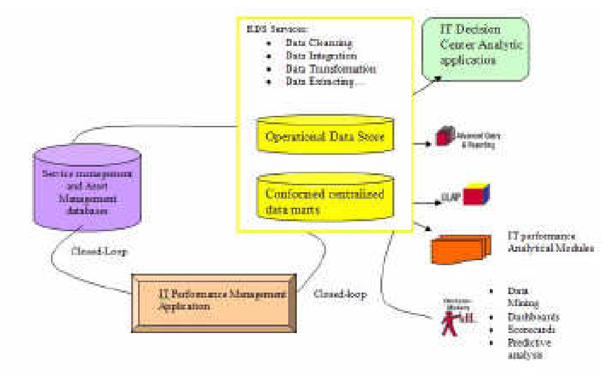

Figure 7. IT performance management system [16]

The SOA-ITPA architecture is illustrated in figure 8. The essential part of SOA-ITPA is the centralized integrated reporting data store (RDS). RDS consists of historic, current and predictive data. A number of component services are built to populate data from sources into RDS. The SOA-ITPA system was implemented in multiple phases with agile development principle. The focus of the first phase was the centralized reporting data store (RDS) schema management, to support metadata management of the data warehouse, creation and customization of the RDS schema. Building on the first phase, the second phase concentrates on extract, transform and load (ETL) services, implementing a number of services to do with data cleansing, mapping, loading and corresponding auditing services. Following on that, the third phase sees the implementation of RDS publish services, where data from the RDS can be made available to be consumed as analytical information by external systems. The focus of the fourth phase is on the measurements of IT key performance indicators from the information made available in the previous phase and the presentation of such measurements through scorecards that allow a guided analysis of IT performance. Finally in a fifth phase a fully fledged business intelligent solution was implemented that orchestrated the ETL services and processes to provide data marts for service management, asset management and other aspect of IT service management performance.

In SOA-ITPA, this is much simplified, as the only steps to be undertaken are adding new XML metadata model for the new source and creating or modifying reports, metrics and dashboards. Other implementations are already available because of existing ETL and publish component services. Not only are the overall development steps reduced by more than 70%, but the effort of extension is much less as well. And the system impact and risk are reduced dramatically.

However, the gain in flexibility comes at a cost in complexity of a service-oriented architecture such as SOA-ITPA presented here with respect to traditional BI systems. This is the trade-off consideration that they need to make. For simpler BI systems than the one they considered here, it could be possible to build viable solutions with the traditional approach. But the SOA approach appears to be the best way to reduce the total development and maintenance cost, and to minimize the risk and impact across an entire enterprise when introducing business intelligence solutions. Another advantage of the approach is that there exist SOA development guidelines and many integrated development environment (IDE) tools that can simplify the development of complex SOA applications.

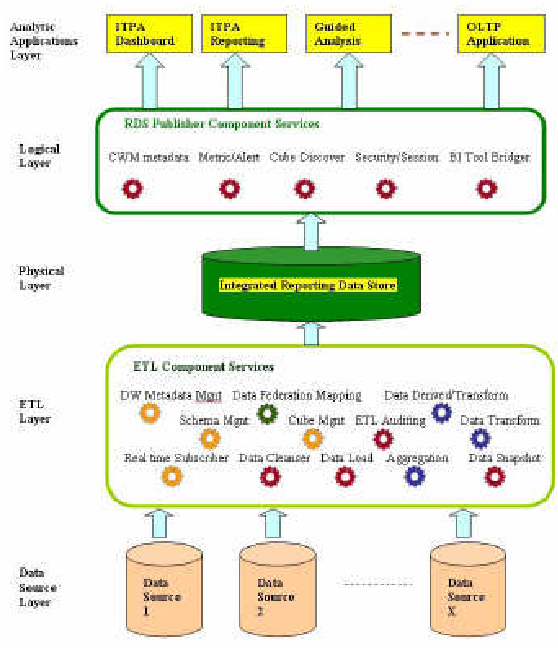

Figure 8. Service Oriented Architecture for IT Performance Analytic (SOA-ITPA). [16]

4. Discussion

Both the concept of business process management to current business intelligence system with the ability of process-driven decision making should be characterized by certain canons. In case of BI systems particular attention ought to be paid to the following issues :

- BI system should be rapidly implemented, which is quite difficult because such systems are specific for each enterprise. Although basing on standard components shortens time required to build BI, each implementation necessitates adjusting of a particular system to specific requirements of an enterprise. While choosing ready to use BI solutions, it is necessary to be very careful;
- BI solutions ought to be flexible. As soon as business changes, organizations should adjust their BI systems to new conditions;
- BI systems ought to be independent of their hardware and software platforms. Hence, it is recommended that a system of multidimensional analists should co-operate with different bases and work in already tested and commonly applied operation system. Such solutions will allow for better adjusting of the system in question to information technology related infrastructure of an enterprise;

- BI solutions have to be scaleable. Flexibility and open architecture allow for easy expansion of the system. It is necessary in a situation when there are new informational needs or when an amount of information to be processed remarkably increases; and
- BI system should be based on modern technologies. It is necessary to pay much attention to solutions provided by household names of the computer industry. Only then, it is possible to expect stability and reliability of purchased technologies.

BI systems pose a chance for the effective management of an enterprise. However, they require analysts, designers and users of high business, information and organizational culture. Skills to identify, model (in the processes and organization structures) and share knowledge constitute only some factors that determine a correct development of the BI systems.

5. Conclusion

This paper reviews the concept of business process management to current business intelligence system and provides a survey, from a comprehensive point of view, on the BI technical framework, process, and architecture for business intelligence. In addition, the conclusion points out the business process management to current business intelligence system so as to provide the business intelligence system with the ability of process-driven decision making.

Business intelligence technologies include traditional data warehousing technology such as reporting, ad-hoc querying, online analytical processing (OLAP). More advanced business intelligence tools – such as HP Openview DecisionCenter – they also include data-mining, predictive analysis using rule-based simulations, web services and advanced visualization capabilities.BI systems have been created with focus on the back-end, which is usually powered by technologies for data warehousing. Lately, architects for business intelligence have evolved towards distributed multi-tier enterprise analytic applications and a service-oriented architecture for business intelligence that makes possible a seamless integration of technologies into a coherent business intelligence environment, thus enabling simplified data delivery and low-latency analytics.

However, The stage of exploration and discovery of new informational needs is of critical importance for the whole cycle of building and BI. Implemented BI environment casts new light on the role of information and competencies in an enterprise and on business relations and interdependencies. At this stage, new informational needs to emerge and new methods of information management are created. Using notions of iterative designing and rapid application development (RAD), it is possible to observe that the BI environment obtained is a prototype that should serve as the basis for evaluation and commencement of a new cycle of building a BI application. The exploration and discovery stage requires remarkable co-operation of representatives of IT departments with a centre for knowledge management along with end users whose all innovativeness and willingness to experiment with data are of very important.

6. References

[1] Li. Zeng et al., Techniques, Pocess, and Enterprise Solutions of Business Intelligence. *IEEE International Conference on Systems, Man and Cyberneties*, pp. 4722-4726. October 8-11, 2006.

[2] Luan Ou and Hong Peng, "Knowledge and Process Based Decision Support in Business Intelligence System," in *Proc. Conf. Computer and Computational Sciences,* IEEE 2006.

[3] D. J. Power, "A brief history of decision support systems," DSS Resources COM, World Wide Web, http://DSSResources.COM/history/dsshistory.html, version 2.8, May 31, 2003.

[4] K. Baina, S. Tata, and K. Benali, "A model for process service interaction," in Proceedings 1st Conference on Business Process Management, EindHoven, The Netherlands, 2003.

[5] C.Revelli. *Intelligence strategique sur Internet*. Paris. 1998: Dunod.

[6] S. Aissi, P. Malu, and K. Srinivasan, "E-business process modeling: the next big step," *IEEE Computer*, 35(5), pp. 55-62, 2002.

[7] Gartner, *How Secure Is Youir Business Intelligence Environment*. 2002. pp. 1-6.

[8] E.F. Codd, S.B. Codd, C.T. Salley, *Providing OLAP(on-line analytical processing) to user-analyst. an IT mandate*, E.F. Codd and Associates, Editor. 1993.

[9] J.W.Han, M.Kamber, *Data Ming: Concepts and Techniques*. 2001: Morgan Kaufmann. pp5-9.

[10] R. Agrawal, T. Imielinski, and A. Swami. Mining association rules between sets ofitems in large database. *In Proc. 1993 ACM-SIGMOD Int. Conf. Management of Data* (SIGMOD'93), pp:207-216, Washington, DC, May 1993.

[11] V. Vapnik. Statistical Learning Theory. John Wiley&Sons, 1998

[12] L.J. Cao, Francis E.H. Tay, Support Vector Machine With Adaptive Parameters in Financial Time Series Forecasting. *IEEE Transactions on neural networks*, 2003. 14(6): pp. 1506-1519.

[13] H.Yu, J. Yang, J.W. Han. Classifying Large Data Sets Using SVMs with Hierarchical Clusters. *in:Proc the ninth ACMSIGKDD international conference on Knowledge discovery and data mining*. Washington, D.C. August 24 - 27, 2005.

[14] Daniel S. Soper, "A Framework for Automated Web Business Intelligence System," in *Proc. Conf. Computer and Computational Sciences,* IEEE 2006.

[15] WSMO Working Group (2006). D2v1.0. Web Service Modeling Ontology (WSMO). WSMO Working Draft. http://www.wsmo.org/2004/d2/v1.0/.

[16] Denilson Sell at el., "Adding Semantics to Business Intelligence," in *Proc. Conf. Computer and Computational Sciences,* IEEE 2005.

[17] Daniel S. Soper, "A Framework for Automated Web Business Intelligence System," in *Proc. Conf. Database and Expert Systems Applications,* IEEE 2007.

[18] Dresner, H. J., Buytendijk, F., Linden, A., Friedman, T., Strange, K. H., Knox, M., & Camn, M. (2007). *The business intelligence center: An essential business strategy*. Gartner Research.