

Methodology Framework for Analysis and Design of Business Intelligence Systems

Martin Závodný

Department of Information Engineering
Faculty of Economics and Management
Czech University of Life Sciences in Prague
Kamýcká 129, 165 21 Praha 6–Suchbát, Czech Republic
zavodny@pef.czu.cz

Copyright © 2013 Martin Závodný. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Business Intelligence (BI) system enables its users to monitor performance indicators (measures) related to chosen enterprise agenda and supports decision making processes for all management levels. The contribution presents a general methodology framework for analysis and design of Business Intelligence systems and describes significant factors and risks of BI projects.

Mathematics Subject Classification: 68M01, 68U35

Keywords: Business Intelligence, methodology framework, analysis, design

1 Methodology framework

This presented methodology framework was designed on the basis of the analysis of specifics of BI projects, relevant literature and also the authors' experience. The methodology framework comprises of key activities, which should be fulfilled during analysis and design phases of BI systems and can be adapted and modified according to the specific needs of the company which decides to use this framework. The article is focused mainly on specialists from BI area; the basic knowledge of Business Intelligence systems is required.

1.1 Business Intelligence systems

Business Intelligence (BI) systems belong to the group of decision support systems. These systems process chiefly the data from internal enterprise information systems and provide it in a suitable analytic form to the end users, who make decisions of operational, tactical or strategic significance from this [2].

The basic technology, which is used in BI systems, is On-Line Analytical Processing (OLAP). OLAP is a technology based on the concept of multidimensional databases, which enables one to analyse data (so-called fact data) necessary for decision-making from multiple perspectives [4]. An example of fact data analysed by a BI system can be revenues of particular company. Suitable perspectives (called dimensions) could be time, products sold, client segments, retailers, etc. An illustration of the multidimensional analysis of revenue measure is depicted in Fig. 1.

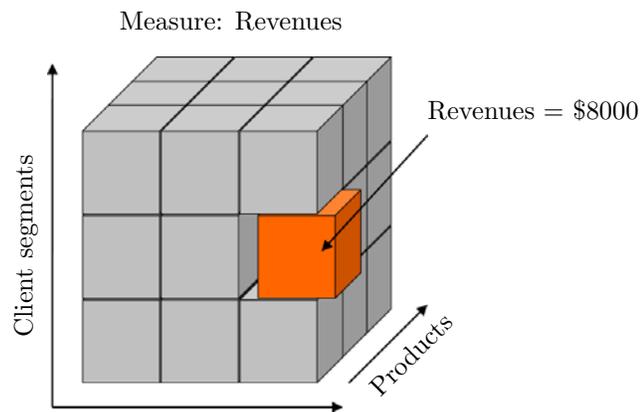


Figure 1: Multidimensional analysis – analysis of revenue measure over time, product and client segment dimensions

1.2 Concept of analysis and design of BI systems

Creation of BI system requires the usage of a suitable concept, which ensures that the designed solution (BI applications) will correspond to the initial user's requirements and that specific characteristics of the environment, in which the system should be implemented, will be taken into consideration.

Following paragraphs describe each particular phase of BI systems analysis and design, and identify risks factors determining successful implementation of the systems.

1.2.1 Analysis of information needs

Analysis of the users' information needs is the initial phase of BI system development. Its target is to identify data, which should be presented to the BI

system users in a suitable form (e.g. in the form of report sets and graphs) or used as a source for ad hoc analysis, data mining and other analytic outputs.

It is worthwhile to utilize existing industry domain models¹ as a prototype during the identification and giving precision to the users' information needs. The models can recommend usage of particular types of reports or performance indicators, which is suitable to monitor within BI systems. These models very often represent valuable know-how of specialised consultancies. The aim of BI projects is to ensure on-time delivery of measure values defined by such domain models to the management of the given organizations.

In the phase of analysis is necessary to define the content of BI applications, suitable scope of measures and collect requirements for the visual formatting of the analytic outputs. Regarding the optimum number of performance indicators, this number should be limited. Within one managerial area it is recommended to focus rather on a smaller number of indicators (e.g. up to 10). A larger amount of implemented indicators more likely to increase the cost of the project and is not fully utilized by the end users.

Part and parcel of this phase is the dimension analysis of measures. Within this activity it is determined whether a particular measure is required to be analysed over a particular dimension and what structure should the dimension have (e.g. measure revenues can be analysed over time dimension, product dimension, retailers etc.). A suitable tool for the analysis is usage of a dimension analysis table, which depicts whether the particular dimension is relevant to the specific measure or not. Hierarchy structure must be defined for each dimension over which the measure can be analysed (time dimension can have the hierarchy structure of a year, month or a day). Analysis of the measure from multiple perspectives is one of the key functionalities of BI systems.

Targeted value for particular measure can be set and according to it evaluated positive or negative trend of corresponding business target, which is linked with the measure. BI system can also automatically notify the responsible user of the last development of measure values.

Every measure should be assigned to a responsible owner and also a precise definition of the final value calculation. For example, if it is necessary to implement a revenue measure, it is necessary to calculate the measure value as a sum of the particular invoice selling prices on a data set determined by chosen dimensions categories (particular items of dimension).

It is also important to determine the data attribute according to it the fact data (measures) are assigned to the concrete dimension value. For example, revenues can be assigned to a time dimension on the basis of order date, invoice date, payment date, etc.

An important component of the analysis is the specification of requirements for BI application response times and timeliness of data provided to

¹e.g. Gartner Business Value Model [5]

the end users. A critical factor of this phase is the complete identification of key information needs and accurate definition of corresponding performance indicators.

1.2.2 Analysis of data sources

The analysis of information needs should be followed by the analysis of data source systems, since it is necessary to identify data sources, which will be used as a base data set for BI system. Mapping tables is recommended to use as a suitable tool, which depict coverage (mapping) of proposed measures and dimensions by source data entities from enterprise information systems and databases.

Analysis of data source systems determines further steps in solution design. Data source analysis helps find out if suitable data exists for a particular area, which should be monitored by BI system. Nonexistence of relevant data generates requirements for modification of source systems. These requirements are necessary to prioritize on the basis of cost benefit analysis and solve the designed changes within development projects of source systems. From that reason it is worthwhile to initially focus on requirements, which are covered by data and don't require excessive effort to obtain.

1.2.3 Solution design and its architecture

Data that is presented to BI users is, for an extensive solution, very suitable to be processed in several separate stages. Data from high-volume databases of enterprise information systems are loaded into temporary data storages and then loaded to a single-purpose analytic databases and also complex data warehouses. The aim is to have the data in such a form at the end of this chain that it can be flexibly aggregated and presented to the users or alternatively to allow the users to use the pre-processed data by other analytical tools (e.g. data mining tools). The source systems shouldn't be used for any data transformation, except loading data to temporary data stages, as this could cause an excessive load of source information systems and restrictions on their common operation. The preparation of proper data format in the data layer includes a number of transformation tasks. The preparation comprises aspects as a data quality, timeliness of data acquisition from source databases, metadata management, security etc.

The principle is that in the area of temporary data stages it is advisable to get all necessary data from source systems and then load and transform it into the subsequent data layers of the BI system, which is data warehouse or data mart, where the actual data and historical data is stored in a structure suitable for analysis. Also in the case of predictable future need for adding a particular data entity to the data warehouse or data mart it is advisable to

have the data prepared at least in the temporary data stages. The design of the most appropriate solution architecture is always depended on the particular conditions of the organisation. Architecture should take into account the amount of source systems, volumes of data transferred between components of the data layer, the requirements for the timeliness of data and also the capacity necessary for the storage of large volumes of historical data in data warehouse or data marts [1].

To support dimension analysis of defined measures, an appropriate data model should be designed. Its technical implementation is a star schema, snowflake schema, or the normalized data model [2].

When designing the BI architecture, considerations regarding future requirements for changes in BI content are needed. BI application similarly like other categories of applications goes through its life cycle and develops. This development entails primarily requirements for analytic output changes, changes in the data models and changes in source systems in particular, which must be reflected in the solution, so that the data load is not disrupted. For this reason, it is necessary for BI systems to keep metadata at all levels of the BI solution², so that an impact analysis could be carried out in the case of changes in source systems. The critical factor of this phase is the resolution of requirements ensuring the adequate performance of the BI infrastructure and also the design of quality data models. The top component of the designed solution is analytic output (in the form of reports, analysis, dashboards, scorecards and data mining outputs).

The analysis and design of BI systems is advisable to draw up a solution specification describing user requirements, containing a coherent view of the components of BI systems and data view addressing data layer and data transformation procedures.

2 Conclusion

The contribution addressed the methodology framework designed for analysis and design of Business Intelligence systems. The described framework covers the best practices and techniques that were successfully verified in practice and can be adapted according to specific conditions of a particular company. The presented methodology framework can be further elaborated in the area of BI technology infrastructures, data quality solutions, structure and content of analytic and design documents, aspects of effort estimations of BI projects, etc.

²Metadata of data source systems, data entities, data transformations and analytic outputs.

ACKNOWLEDGEMENTS. This contribution was created with the support of a grant Nr. 20121061 of the Internal Grant Agency of the Faculty of Economics and Management of the Czech University of Life Sciences in Prague.

References

- [1] H.W. Inmon, *Building the Data Warehouse*, Wiley Publishing, 2005, Indianapolis.
- [2] R. Kimball, M. Ross, *The Data Warehouse Toolkit – The Complete Guide To Dimensional Modeling*, Wiley Computer Publishing, New York, 2002.
- [3] H. Plattner, A. Zeier, *In-Memory Data Management: Technology and Applications*, 2nd ed., Springer, Berlin, 2012.
- [4] E. Thomsen, *OLAP Solutions: Building Multidimensional Information Systems*, 2nd ed., Wiley, New York, 2002.
- [5] The Gartner Business Value Model: A Framework for Measuring Business Performance, <http://www.gartner.com/id=492685> [10.12.2012]

Received: January, 2013