

Zenon Biniek

University of Economics and Human Sciences in Warsaw, Poland
e-mail: z.biniek@vizja.pl
ORCID: 0000-0003-0157-5351

APPLICATION OF THE META-MODELLING METHOD IN BUSINESS INTELLIGENCE

ZASTOSOWANIE METODY METAMODELOWANIA W BUSINESS INTELLIGENCE

DOI: 10.15611/ie.2020.4.04

JEL Classification: M10, M15

@ 2020 Zenon Biniek

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>

Quote as: Biniek, Z. (2020). Application of the meta-modelling method in Business Intelligence. *Informatyka Ekonomiczna. Business Informatics*, (4).

Abstract: The article deals with the issue of conceptual modelling of the MOLAP database structures used in multidimensional data analysis for BI. Management information for decision-making purposes was collected in the MOLAP database and delivered to the user via the management cockpit. Building OLAP databases is discussed from the database designer point of view. The meta model of the OLAP database was described in conceptual and logical terms. The work includes a critical analysis of the meta-model assumptions for the design of systems for multidimensional data analysis. A postulate to revise the classic meta-model was presented and a new type of relationship was proposed. The assumptions of the meta-model for modelling and simulating business decisions were presented, as well as the application of meta modeling to define business processes. Based on Gozinto graphs, the meta model of structural and technological developments was defined.

Keywords: Business Intelligence, Gozinto graphs, meta-modeling, MOLAP.

Streszczenie: W artykule podjęto zagadnienie modelowania konceptualnego struktur baz danych typu MOLAP stosowanych w wielowymiarowej analizie danych dla BI. Informacja zarządcza dla potrzeb podejmowania decyzji gromadzona jest w bazie MOLAP i dostarczana użytkownikowi przez kokpit menedżerski. Zagadnienie budowy baz OLAP jest omawiane z punktu widzenia projektanta baz danych. Opisany został meta model bazy OLAP w ujęciu konceptualnym i logicznym. Praca zawiera krytyczną analizę założeń metamodelu dla potrzeb projektowania systemów do wielowymiarowej analizy danych. Zaprezentowano postulat rewizji klasycznego metamodelu i zaproponowano dodanie nowego typu relacji. Przedstawiono założenia metamodelu dla modelowania i symulacji decyzji biznesowych. Zaprezentowano

zastosowanie metamodelowania do definiowania procesów biznesowych. Na podstawie grafów Gozinto zdefiniowano metamodel rozwinięć konstrukcyjno-technologicznych (RKT).

Słowa kluczowe: *Business Intelligence*, grafy Gozinto, metamodelowanie, MOLAP.

1. Introduction

The large-scale application of BI (Business Intelligence) applications for decision support has created a need for new information processing techniques, in particular for new technologies for storing data in databases. The technology of classic SQL databases is modified to meet the requirements of advanced business users. For IT management of management information it is necessary to meet the challenge of creating modern tools to support the construction of manager dashboards. New types of databases to support multi-dimensional data analysis have appeared, such as the so-called MOLAP database (Multidimensional On-line Analytical Processing). The new modified meaning is attributed to computer simulation technique. There are three main areas of data structure modelling in BI applications:

- ROLAP technology, where data are collected in two-dimensional tables, the database has a relational structure and is divided into tables of dimensions and facts,
- MOLAP technology, data are collected in OLAP cubes in multidimensional databases,
- simulation technique, data are collected in method and model databases, in addition to real data, simulation data are collected, models and methods are saved in the databases.

As a result of the dynamic development of BI applications, classic SQL databases are modernized. On the one hand, new technological solutions are emerging in the field of using SQL to support data warehouses (ROLAP), and on the other hand, work is underway on a new generation of databases to support multidimensional data analysis that stores data in the form of OLAP cubes (MOLAP). Simulation in management is no longer an independent discipline, but rather is implemented in a comprehensive system called the decision support system.

2. Business Intelligence applications

BI (Business Intelligence) is information technology focused on preparing information for making decisions. This technology is designed to enable the user to transform data into decision information.

BI is based on modern information technologies supporting the analysis and design of information resources for the purposes of decision support. Obtaining data from the resources of transaction systems is of key importance. This acquisition requires tools from the fields of Data Mining and ETL (Extraction, Transaction,

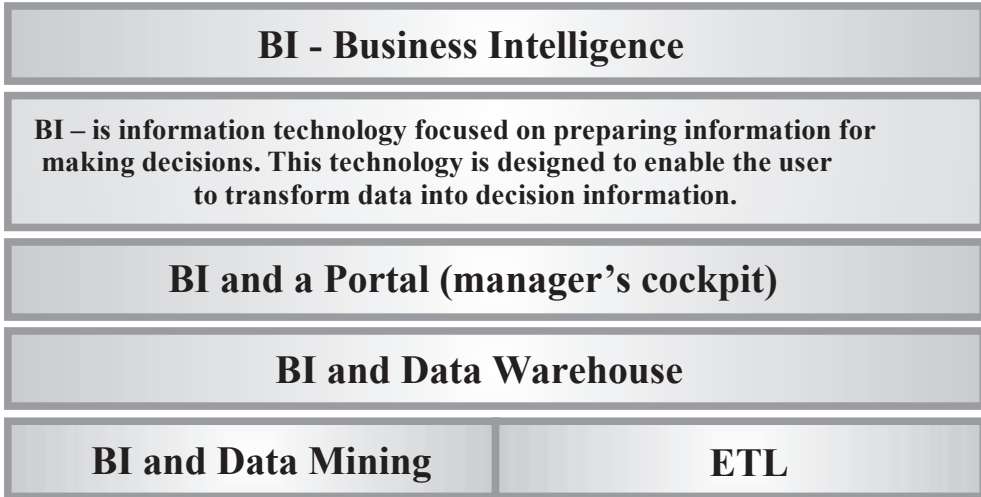


Fig. 1. Functional scope of the BI application

Source: own work.

Loading). One of the central elements of the BI system is the data warehouse, where information processed in analytical mode (OLAP) is stored. The manager’s cockpit is responsible for displaying information in a form adapted to decision needs.

3. Conceptual modelling of data structures in BI

The BI application development process can be divided into several phases. The starting point in the design process is to conduct a requirement analysis, i.e. to determine the users’ information needs. Having a specific information scope, we can create a conceptual BI application model.

Conceptual modeling uses a set of basic concepts such as classes, objects, relationships, processes and events. In this arrangement, the use of the meta model is important for the construction of diagrams. The metamodel is a generalization for a given class of conceptual models, it contains semantic invariants for marking graphic symbols in diagrams intended for modeling data structures. Using these symbols, one can define a specific database model or business process.

The concept of a meta-model was introduced to the IT literature by Ferstl and Sinz (2001, p. 122), defining the meta-model of E/R data, which significantly structured the discussion on semantic modelling of relational database structures. Using the meta-model category makes it easier to define specific data models that fall into a particular category and must meet the general rules for a given meta-model, where the classic E/R notation also fulfills its important role here (Burmester & Goecken, 2008).

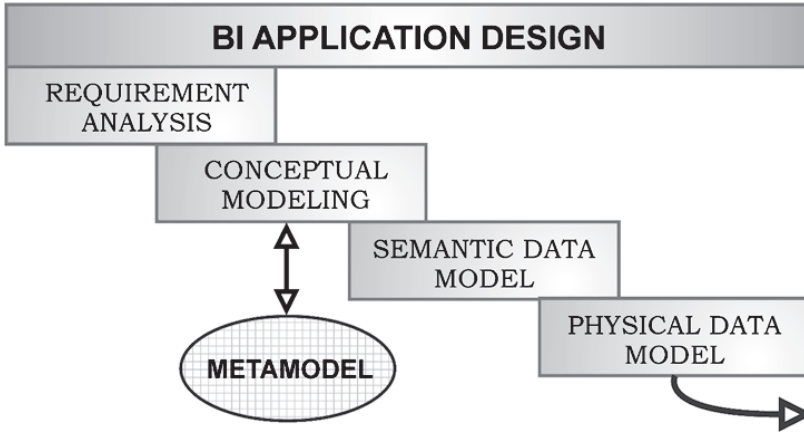


Fig. 2. BI application design phases

Source: own work.

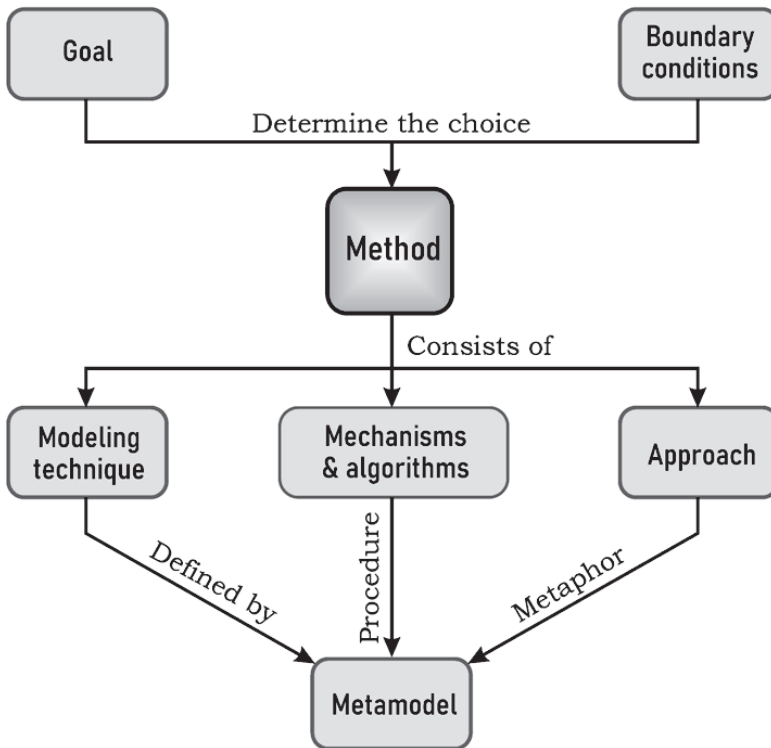


Fig 3. The idea of meta modeling

Source: own work.

Conceptual design is a set of actions aimed at transferring the specified requirements specification to the semantic scheme of the future database. At the stage of the requirements analysis, the information specification of the future BI system is determined, and the semantic data model presents the data structure where data is collected for preparing information reports.

4. Metamodelling of the OLAP database

The first phase of building OLAP systems was dominated by the approach focused on creating relational databases supporting data warehouses. From a technical point of view, the approach to data warehouse design was practically no different from the traditional approach to relational database design. The task of the database designer was, among others, to define a semantic data model containing definitions of data formats and relationships between individual data elements. Designers have distinguished, as is known, two basic types of semantic models to support data warehouses: snowflake or multi-stars (Biniek, 2009, p. 287). All operations on the relational database were carried out using the SQL language, initially not adapted for processing large data sets. It was only after some time that the SQL language was modified to adapt it to data warehouse services, for example in some language implementations Cube operator software support has been developed (Eickler & Kemper, 1997, p. 465).

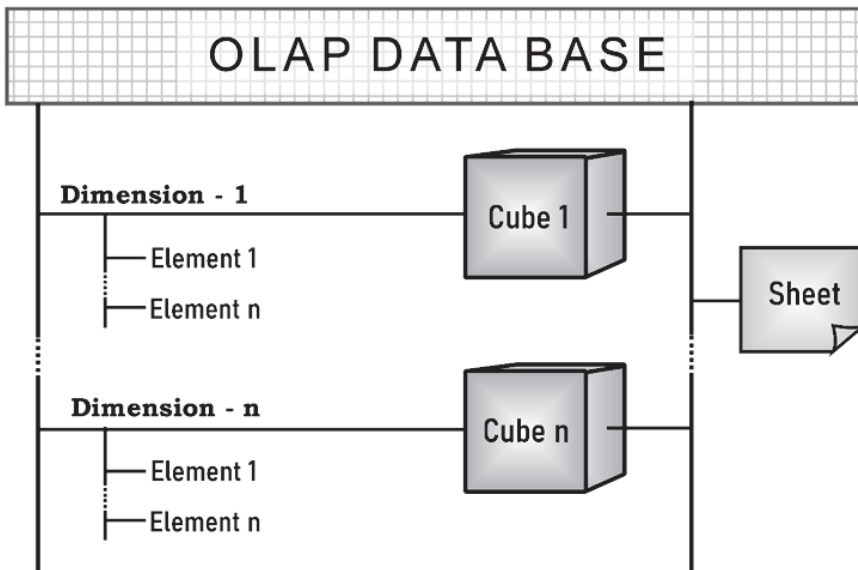


Fig. 4. Functional structure of the MOLAP database

Source: own work.

Another approach to creating OLAP systems is defining analytical databases that support spreadsheets. In this case, the OLAP database supports specialized multidimensional data analysis. Using the spreadsheet as before was assisted by an OLAP database containing appropriately aggregated data, adapted primarily to the needs of analytical summaries. Such a database is fundamentally different from the classic ROLAP database, where E / R diagrams can be used to model data. In the OLAP database, one does not use two-dimensional relational tables containing key columns and data columns. Figure 4 shows the functional diagram of an OLAP database that supports spreadsheets.

The MOLAP conceptual model does not have the traditional relational data model schemes i.e. stars and snowflakes. In this case, already during the construction of the conceptual model, the study defined only those data elements that are necessary to define a multidimensional OLAP cube, assuming that the defined cubes will be designed to support various analytical reports. As can be seen in the Figure 4., the designer defines the dimensions and their associated elements.

The MOLAP database consists of one or many multidimensional OLAP cubes. There can be many cubes in one base (Cube-1,..., Cube-n). The OLAP cube is structurally defined using predefined dimensions used to describe the components of data analysis that occur in a given decision support system, e.g. time, area and products. A specific dimension can occur in one or more cubes at once. There can be many dimensions in one database depending on the scope of data analysis. The dimension consists in data elements (Element-1,..., Element-n) always assigned to a specific dimension. An element is defined by giving it a unique name. Similarly, the dimension and the cube are defined. When defining elements, one first specifies the dimension, and then assigns elements to the dimension, which means that the elements do not exist independently in the database. The OLAP cube must contain at least one dimension. The sheet technique is adapted to support the input and output interface. The worksheet is used by the database user when entering data and preparing information sheets from the database.

Starting from the general premise of meta modelling, the author undertook the task of defining the meta-model of the MOLAP database, as shown in the Figure 5. This figure presents a graphic reflection of the relationship between the elements of the MOLAP database used in the software packages used so far, i.e. the so-called classic approach. Data objects represent business facts that occur in business operations. In the process of conceptual modeling, the fact is transferred to the database as an object of analytical data. The data object in the modeling process can take the form of an element, dimension or cube, where the data element consists of a homogeneous value defined in specific units of measure. The dimension consists of elements, and the OLAP cube of dimensions. Assigning elements to a dimension and a dimension to the cube is a sovereign decision of the OLAP database designer. Thus, the final form of a given data object is determined by the OLAP database designer, guided by the general modeling rules stored in the meta-model.

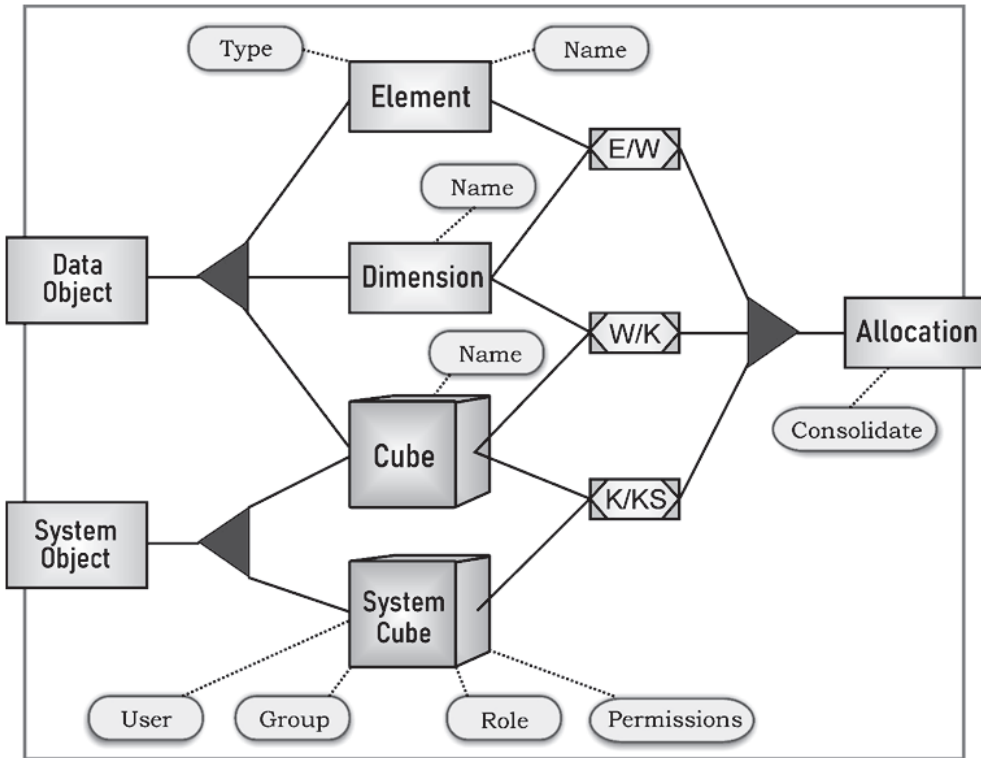


Fig. 5. MOLAP conceptual meta-model (classic)

Source: own work.

Modelling organizational and functional dependencies uses system objects, namely: Cube, View, Dim, Subset, Function. These are objects existing in every MOLAP database, regardless of specific material entries (element, dimension, cube). In addition, the database designer must ensure that the base elements are assigned. The general rule is that there can be no element not assigned to a dimension in the database. Assignment is understood here as the logical connection of elements and dimensions (E / D), the logical connection of dimensions and cubes (D / C) and the logical connection of data cubes and system cube (C / SC). In the E / D assignment, some elements were consolidated, which necessitated defining consolidation rules. Element consolidation consists in the hierarchical linking of elements, i.e. determining their belonging to a certain defined parent group of elements. Each element of the base can have attributes defined, expressing some additional features of this element.

The relationship between logical units of the OLAP database is created through a unique element name. The element, in addition to the unique name, is assigned

a type (numeric, text). Special diagrams are used to model relationships between dimensions and elements, the so-called diagrams of facts (Lechtenbörger & Vossen, 2003, p. 190). Figure 6 shows an example of the diagram of facts.

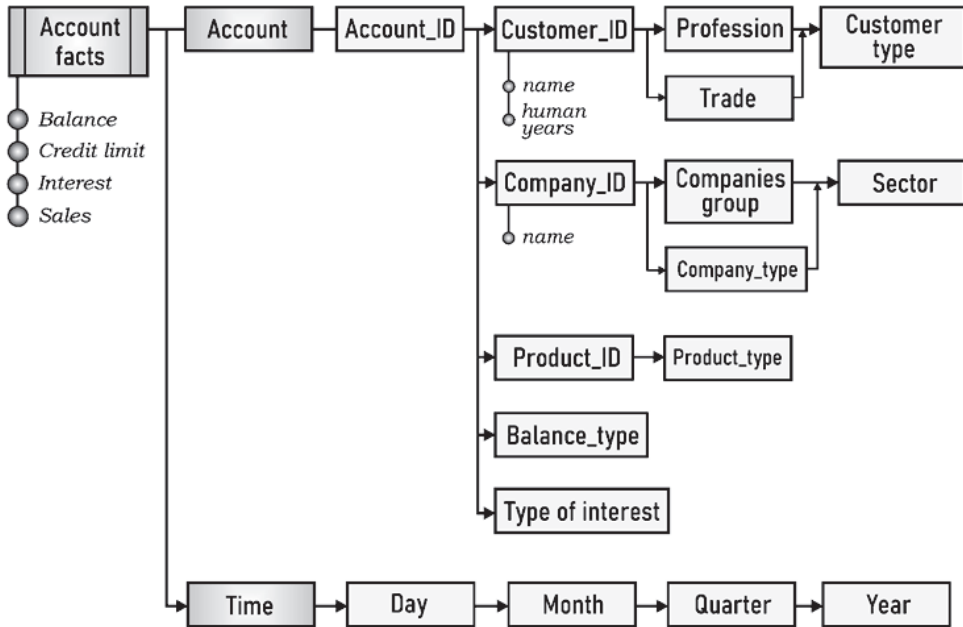


Fig. 6. Diagram of dimensions and facts of the OLAP database

Source: (Hüsemann, Lechtenbörger, & Vossen, 2000, p. 63).

Facts represent unitary elements of information (data) in a multidimensional analytical base. The fact describes quantitative values in the form of a measure (units of measure) and qualitative features that are defined in the base through various dimensions starting from the terminal dimension level. One can define several aggregation paths for the same dimension and the same baseline. Elements of a given dimension are functionally dependent on the base element. Each individual dimension contains a defined group of instances or elements. Dimensions are hierarchically shaped and have a specific aggregation path and each element of the base must belong to at least one specific dimension.

Using the facts diagram, it is possible to define the hierarchical structure of elements present in a specific MOLAP database. The central fact table (sales) was surrounded by various necessary hierarchical or linear dimensions.

The meta-model of the MOLAP database presented in Figure 5 is insufficient to fully represent objects and relationships in multidimensional data analysis. Classic MOLAP packages usually have the disadvantage that they allow defining only the

hierarchy of elements and dimensions in cubes, and they are unable to define the functional relationships of elements occurring in independent dimensions. The basic requirement of the MOLAP base is that the functional relationship occurs within many dimensions. Generally, in conceptual modeling of the MOLAP database, three types of relationships should be distinguished: functionality, aggregation and membership relations. Aggregation and membership relationships dominate in systems already in use, while functional relations are not present in the currently used MOLAP databases. Assigning one or more dependencies on other elements requires changes in the MOLAP meta-model base.

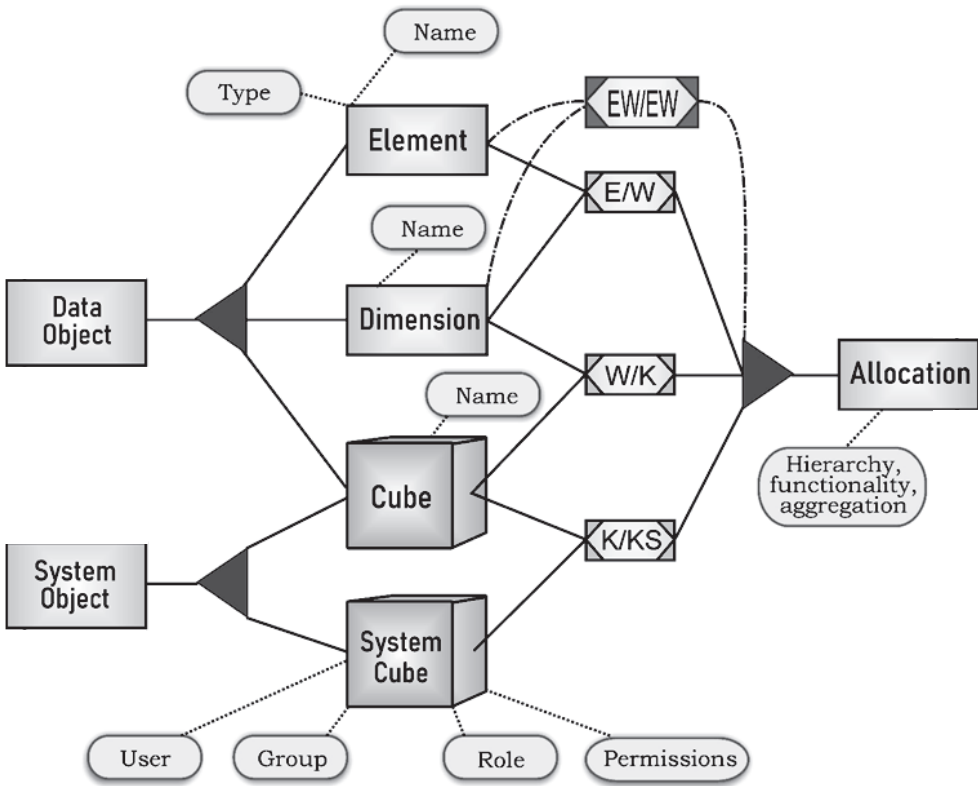


Fig. 7. Extended meta-model of the MOLAP database

Source: own work.

It is necessary for OLAP-based multidimensional data analysis packages to support all distinguishable relationships between elements, including functional

dependency (Vossen, 1999, p. 152). The many-to-many (n: m)¹ mapping between elements of different dimensions within the same cube is marked grey, e.g. employee (employee dimension element) supports one or several products (products dimension). To the three traditionally used associations (to support hierarchical relationships and memberships) there should be added the EW / EW type relationship supporting functional relationships. Therefore, in the EW / EW table the author placed four entries for one functional relation. The relationship indicates which Ex element of the Wx dimension is functionally related to the Ey element of the Wy dimension. In addition to entries indicating a functional relationship, one can also place entries containing the status of the relationship, e.g. active or outdated, date of entry, etc. In addition, related elements can be assigned a cardinality (multiple of the relationship), e.g. a 1: 1 relation or n: m relation, and a set of connections elements is a finite set of four related elements. Each element of a given set of relations belongs to a specific dimension. An assignment of this type significantly facilitates the use of base elements when preparing specific multi-sectional reports, and in particular preparing reports for dashboards.

During the requirements analysis, a conceptual model of a multidimensional OLAP database is created in a few steps:

- defining measures, functional dependence of the base, determining the relationship between dimensions and measures,
- definition of dimensions and their hierarchy,
- defining element consolidation.

So far, the study addressed functional and structural modeling. In the created diagrams, the attention of modellers was focused on the definition of objects and the connections of these highlighted objects. Correctly defined relationships between dimensions within a given OLAP cube are important. When preparing multidimensional data analyses, the author used tabular functions that enable downloading data from OLAP cubes and placing the data in prepared reports. Functions are system objects called by names intended primarily for manipulating data. There are both functions designed for filtering data and the function of operating on arrays. The database designer at the stage of creating its conceptual model can predict what functions will be used in a particular implementation.

In the conceptual modeling of databases to support BI applications, the simulation approach should also be highlighted. Computer simulation significantly contributes to the improvement of business decision making, including by simulating examining the potential effects of decisions. The simulation concerns strategic decisions, also medium and long-term operational decisions can be subject to simulation. The concept of computer simulation is presented in the Figure 8. The simulation model is inherently related to describing changes in variable values on the time axis.

¹ Properties typical of relationships found in relational databases cannot be directly transferred to OLAP databases.

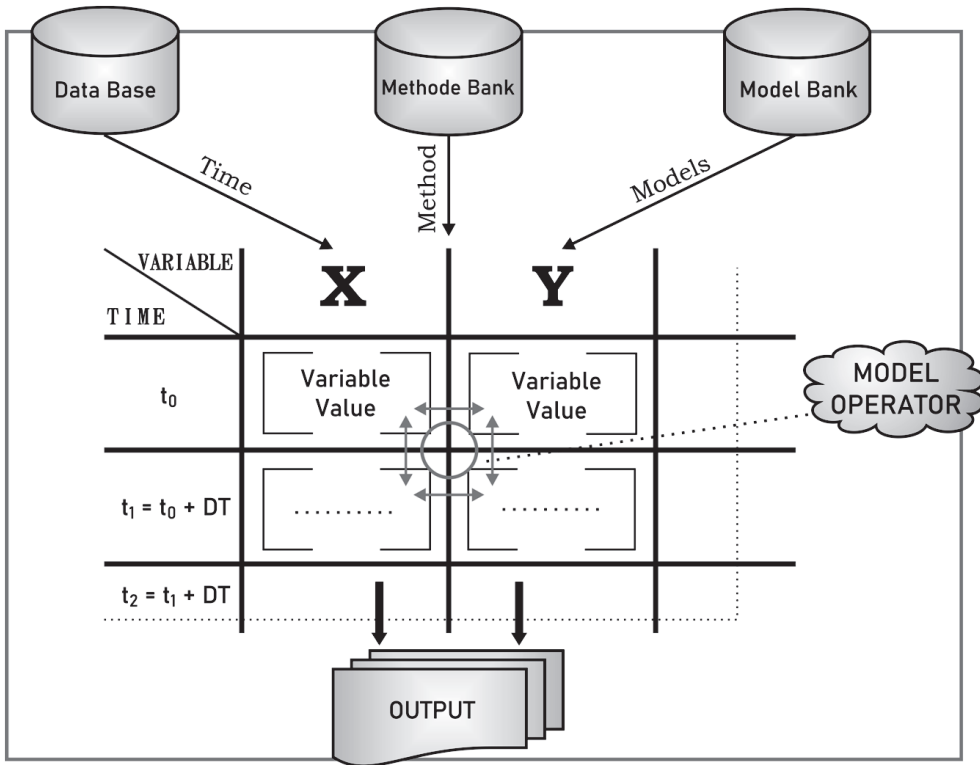


Fig. 8. Functional assumptions of computer simulation

Source: own work.

Time as a variable can be modeled in a continuous or discrete way. The model variables can be described algebraically or by functional dependencies. The simulation model has an internally built-in operationalization mechanism. There is a model operator for each model. In general, three types of simulation models are distinguished in computer simulation:

- what-if models (Golfarelli et al., 2006),
- goal seeking models,
- steady state models.

The above-mentioned types of models are characterized by a common plane of operationalization. The simulation takes the form of a system of equations (algebraic, differential, differential), and the control takes place either on the time axis or in the form of fixed steps, regardless of the time course.

Assumptions of the meta-simulation model occur in the conceptual modeling of computer simulation. Petri networks were adopted as the basis for creating simulation models.

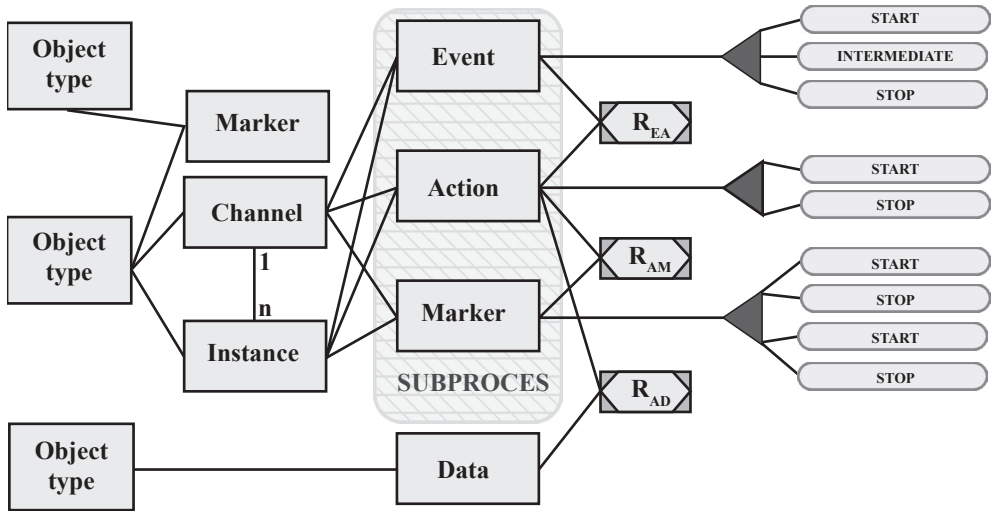


Fig. 9. A meta-simulation model in decision making

Source: own work.

The idea of a meta-simulation model is derived from the general assumptions of object-oriented meta-modelling defined by Ferstl and Sinz (2001, p. 199). Currently, the dominant notation in the construction of simulation models are algebraic equation and cybernetic models. For the semantic modelling of simulation model structures, among others, Petri networks were used (Biniek, 2009, p. 233).

5. Business facts diagram modelling typology

It should be noted that the use of the conceptual meta-modelling method of data structures allows for multiple benefits in the design of databases supporting business processes. By creating a universal model for the purpose of defining specific implementation models, the study thus brings semantic ordering of data structures, provide semantic completeness of specific notations, appropriate semantic accuracy, as well as the ability to supplement the notation with new elements available in a given meta-model. Meta-modelling greatly facilitates the creation of new database implementations to support BI applications. Based on the class definitions in the meta-model, one can create appropriate model libraries from which to retrieve elements to create specific implementations of business processes.

Modelling business processes constitute a separate issue in the design of BI applications. BPMN (Business Process Modelling Notation) notation, developed and maintained by BPMI (Business Process Modelling Initiative) organizations, are used to design applications that support the e-business sphere (Business Process Modeling Notation, 2003).

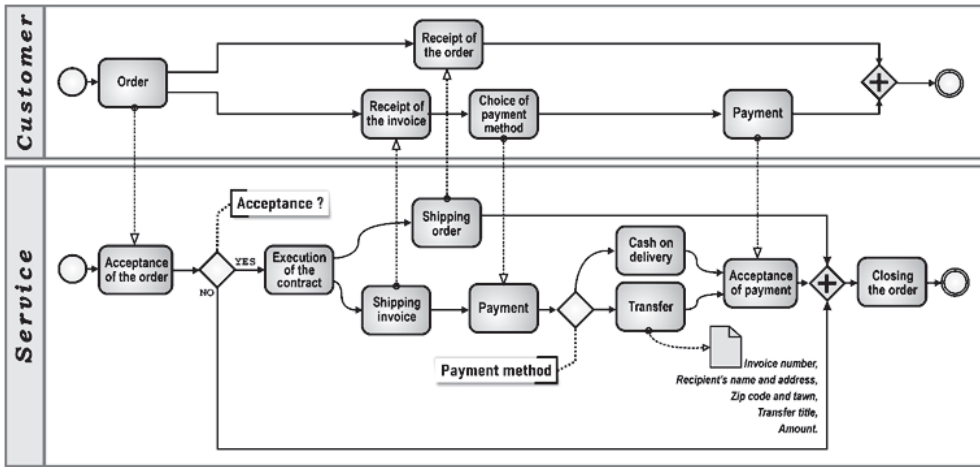


Fig. 10. BPD diagram online store service

Source: own work.

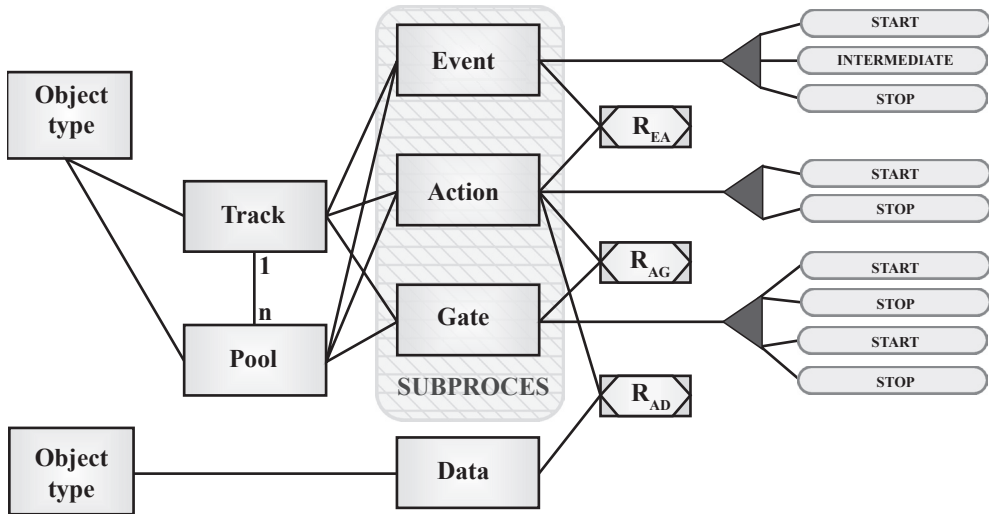


Fig. 11. Model meta diagram for modelling business processes

Source: own work.

BPD consists of a set of graphic elements used to define specific diagrams. These elements allow to create simple charts that are easily accepted by most business process analysts. The elements were designed to be distinguishable and their shapes easy to remember for most people linked with business modelling. For example,

actions are rectangles and decisions are diamond-shaped. The author illustrated the issue presenting an example of a simple BPD diagram showing activities related to the operation of an online store.

One of the main assumptions when designing BPMN diagrams was to create a simple mechanism for creating business process models, while also meeting the possibilities of creating complex business processes. Combining these two conflicting requirements affects the allocation of graphic elements of notation to specific departments. Finally, a small number of categories were created so that the BPD user could easily recognize the basic types and understand the charts. The most accessible way to present diagrams for modeling business processes is the meta modelling method. Figure 11 presents the assumptions of the meta model containing the elements used when defining BPMN diagrams.

The use of the meta modelling method requires the definition of two basic elements that make up the modelling process: the metaphor (context) and the meta model diagram. To illustrate the modelling principle, the metaphor of the swimming pool (a pool with swimming lanes/tracks) was adopted. Figure 11 shows the meta model containing the system of concepts and elements used in business process diagrams. A specific business process diagram describes a process or subprocess. According to the metaphor, the diagram is placed in the lane (track) and the track in the swimming pool (marked pool). At least one track is placed in the diagram, and there can be n tracks in the bases (1, n) Two types of objects are used to define diagrams (data, process). There are four basic categories of elements in diagrams:

- Flow Objects,
- Connecting Objects,
- Swimming Lanes (tracks),
- Artifacts.

Recently, streaming databases have become an important element of information technologies used in e-business. Unlike relational databases operating on transactions and records, data streams describe the continuous flow of data between the end device and the database server.

6. Metamodel STD (Structural and Technological Development)

Gozinto graphs are the most common structure and production modelling technology. Creating the Gozinto metamodel is oriented towards structures, not processes. Creating business models involves such concepts as: Metamodel, m2m, model of models.

Figure 12 shows the functional structure of the Gozinto graph. Product P has a specific functional structure, which consists of selected sub-elements, such as: Intermediates, Elements of parts list, Structure (relationships), Variants, Products.

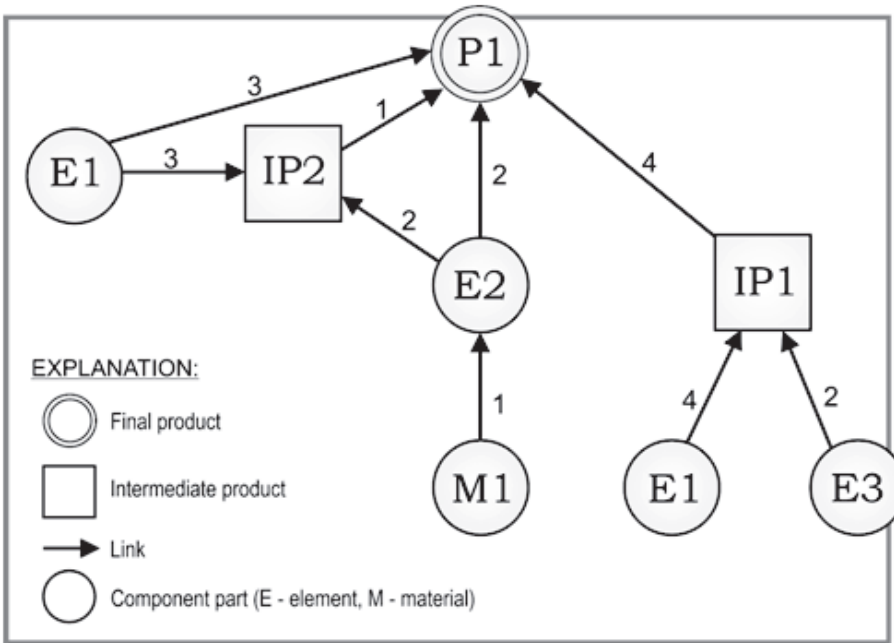


Fig. 12. Functional diagram of the Gozinto graph

Source: own study.

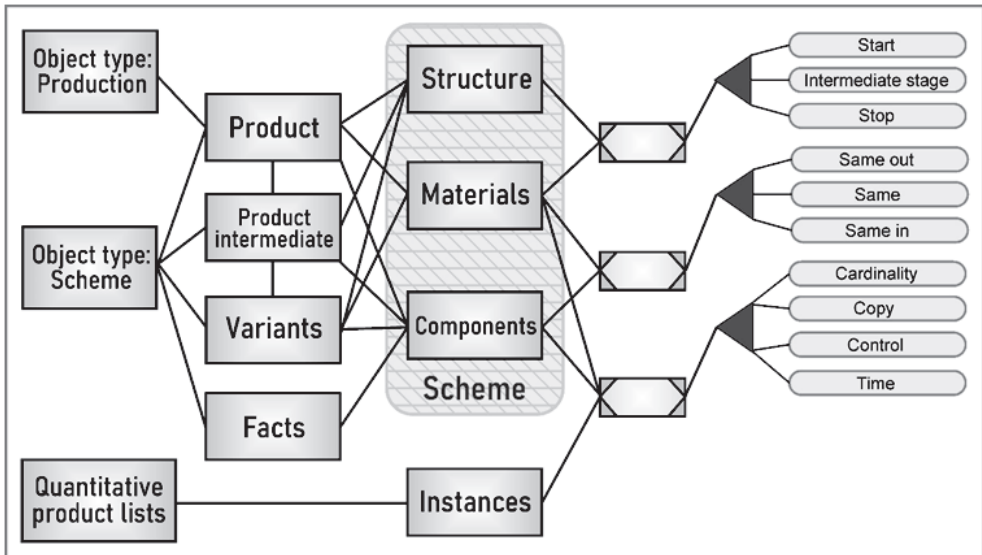


Fig. 13. Semantic Meta model of the STD Gozinto graph

Source: own work.

The goal of creating a Gozinto graph is to specify a quantitative list of products, including the structure of the product and the number of occurrences with connections within the previously defined structure. Instances include the number of parent and child instances. The Gozinto graph model allows us to define the parts list, alternative list, and parts usage list.

The structure of the P1 product (see Figure 12) includes links between the product, its intermediates and variants. The Gozinto graph contains a graphic model of the product structure, and also allows variants of technological processes. Building the semantic model is based on the theory of metamodelling of Gozinto graphs.

7. Conclusion

The paper presents the main assumptions for applying the metamodelling method to manage the structure of product production. Models of IT systems of small and medium enterprises on individual customer order enables differentiation of production orders, using models of B2B, BPMN, BI and Gozinto graph processes.

The Gozinto graph uses a hierarchical presentation, which means that it starts with raw materials (German: Rohstoffe = R) and inside contains intermediate products (German: Zwischenprodukte = Z), and top products are in the highest place (German: Endprodukte = E). Since it only shows the total quantity and does not focus on the base of production and on individual products, it can be compared with the quantity specification.

Thanks to Gozinto graphs, one can present structural relationships in production. Its basis is a structure specification that shows many (or all) production stages. Gozinto graphs use a diagram to show what quantities of raw materials are included in the semi-finished and final products. The amounts contained in these presentations are shown in Gozinto graphs by arrows.

The Gozinto graph is, in this order:

- a help in presenting the product structure to determine the demand
- an element supporting the planning of the implementation deadline and machine occupancy
- a visualization form positioned between the structure and quantity specifications.

The use of Gozinto graphs to map manufacturing processes is important in this case. Mapping is carried out using production process diagrams and manufacturing processes.

References

- Abello, A. (2001). *Understanding analysis dimensions in a multidimensional object-oriented model* (Proceedings of the International Workshop of Design and Management of Data Warehouses, Interlaken: DMDW'2001).

- Biniek, Z. (2009). *Informatyka w zarządzaniu*. Warszawa: Vizja Press&IT.
- Burmester L. and Goecken M. (2008). Relating system dynamics and multidimensional data models. A metamodel based language mapping, *Business Economic*. Retrieved from http://www.google.pl/search?sourceid=navclient&ie=UTF8&rlz=1T4SKPB_enPL306PL306&q=metamodel+system+dynamics
- Burmester L., & Goecken, M. (2008). Relating system dynamics and multidimensional data models, a metamodel based language mapping. Retrieved from *Business Economic*.
- Business Process Modeling Notation. (2003). *Working Draft 1.0*. Retrieved from www.BPMI.org
- Eickler, A., & Kemper, A. (1997). *Datenbanksysteme*. München: Oldenbourg Verlag.
- Erb, H., & Held, B. (2006). *Advanced controlling mit excel*. Poing: Franzis.
- Ferstl, K., & Sinz, E. (2001). *Grundlagen der Wirtschaftsinformatik*. München: Oldenbourg Verlag.
- Golfarelli, M. et al. (2006). *Designing what-if analysis: Toward a methodology* (Proceedings of the 9th ACM International Workshop on Data Warehousing, DOLAP).
- Hüsemann, B., Lechtenbörger, J., & Vossen, G. (2000). *Conceptual data warehouse design* (Proceedings of the International Workshop on Design and Management of Data Warehouses).
- Lechtenbörger, J., & Vossen, G. (2003). Qualitätsorientierter schemaentwurf für datenlager. *IT-Information Technology*, (45).
- Vossen, G. (1999). *Datenbankmodelle, datenbanksprachen und datenbankmanagement-systeme*. München: Oldenbourg Verlag.