# A Process for Engineer Domain Ontology: An Experience in Developing Business Analysis Ontology

Irena ATANASOVA
South-West University "Neofit Rilski" - Blagoevgrad
Department of Computer Science, Bulgaria
irenatm@swu.bg

*During the last years several works have been aimed to improve ontology technological aspects, like representation language and inference mechanisms. This paper presents a discussion on the process and product of an experience in developing ontology for the public sector whose organization requires a strong knowledge management. This process is applied to engineer and develop ontology for Business analysis domain.*
***Keywords:*** *Ontology, Ontology Engineering, Methodology, Protégé, Business Analysis*

## 1 Introduction

Business analysis involves a set of well-defined skills and knowledge that combine topics from traditional business and IT disciplines. Because the IT world changes so rapidly, many business analysts need to update their knowledge. Business analysis relates the business and the IT communities within an organization. It further investigates the emerging role of the business analyst as the primary and key liaison between business and information technology functions within an organization. Analysts gather business requirements from various individuals in the organization and translate them into functional system design specifications that can be successfully executed by IT development teams.

Over the last several years, the field of Business Analysis has experienced exponential growth. More and more, Business Analysts are joining project teams to manage the scope of the product and solution, leaving project managers with more time to manage the scope of the project itself.

Business analysts help organizations improve how they conduct their functions and activities in order to reduce overall costs provide more efficient use of scarce resources and better support the customer and the solutions needed to remain competitive in a global economy.

Since the academy and the industry have recognized the ontologies, there are many definitions about what ontology is. These definitions have been used for different purposes and came from different disciplines. In computer science, ontology can be seen as a dictionary of terms which designed to yield a lexical or taxonomical framework for knowledge-representation which can be shared by different information systems communities [8].

The term ontology can be used to describe models with different degrees of structure. Particularly, the ontology defined in this paper is a formal structure expressed in artificial formally defined languages.

There are three strategies for the creation of ontology: top-down, bottom-up, and a combination of the two. There is no best approach – it is depending of the Ontology creator's preferences and abilities.

We use the bottom-up approach to design, engineer and create our ontology – OntoBAn. There are several methodologies that have been proposed to structure ontology engineering process and thus to facilitate it. These methodologies have been demonstrated in a number of applications [1]. The ontology development has not been as expected in some areas. An example is the public sector area. This area is characterized by a wide range of task and work arrangements. Some process can be fully automated but its scope is limited to simple processes of registering, accounting and calculating – processes in which stakeholders participated (legal rules and knowledge play an important role [5]).

The making of decision in Business analysis area occurs at policy and organizational levels. It is characteristic of operative work of the leadership of corporation. (They must be able to access knowledge base and information to help their tasks.)
The purpose of this paper is to demonstrate the development process of ontology for Business analysis domain and to share it with the ontology community. This work is organized as follows: Section 2 shows the main ontology development methodologies and the methodology which we choice to engineer and define the Business analysis domain ontology. Section 3 and section 4 present a summary of results and the conclusions of this work.

## 2 Methodologies. How to engineer and build a Business analysis ontology

Different development methodologies were studied. Two main groups of methodologies can be identified. The first group is experience-based methodologies that belong to the enterprise modeler domain. The second one is the methodologies that propose evaluative prototypes models. They suggest a set of activities to develop ontologies based on its life cycle and the prototype refinement (iterative approach). There is not just one methodology or correct way for developing ontologies. It is common to merge different methodologies since each of them provides different design ideas that distinguish it from the others. Ontology construction must be supported by some building software engineering techniques. Thus, to support engineer and building of ontology we are going to use methods and tools from software engineering.
In general, the ontology development process can be divided into five main phases:

- Specification;
- Conceptualization;
- Integration;
- Implementation;
- Documentation.

Now, we describe the process of an experience in developing ontology for Business analysis domain.

### 2.1 Specification of Business analysis domain

The specification phase is orientated to acquire informal knowledge about the domain.
This ontology considers the needs for creating an analytic business analysis with concepts related to reports, planning, forecasting and marketing and sales. It considers the concepts related to other stages as budgetary executing, accounting, payments, or fiscal year closure. It Include general concepts for the business analysis life cycle and specific concepts for the formulation.
In domain analysis, the application to formulating the business analysis and its related documentations were studied and revised. Meeting with a group of experts were carried out. This group was confirmed by software engineers who bring informatics supports for business analysis formulation process-tasks and experts in statistics area and finance.
*Goal of business analysis*
Business analysis wants to achieve the following outcomes:

- Reduce waste;
- Create solutions;
- Complete projects on time;
- Improve efficiency;
- Document the right requirements.

*The ontology goal and scope*
The scope limits the ontology, specifying what must be included and what must not. It is a very important step for minimizing the amount of data and concepts to be analyzed, especially for the extent and complexity of the business analysis semantics. This ontology considers the needs for creating an analytic business analysis with concepts related to planning and control, accounting, sales, forecasting and evaluations. It does not consider the concepts related to other stages as budgetary executing. Therefore, it includes general concepts for the business analysis life cycle and specific concepts for the formulation.
*Domain description*
In previous domain analysis, the application to formulating the business analysis and its related documentations were studied and revised. The group of experts was conformed by software engineer who brings informatics

supports for these tasks, statistician who summarizes statistical information and financier. The group of experts was very heterogeneous. This group was the support for knowledge acquisition during the ontology development.

Business analysis of a company is the discipline of identifying business needs and determining solutions to business problems. Business analysis focuses on identifying the changes to an organization that are required for it to achieve strategic goals. These changes include changes to strategies, structures, policies, processes, and information systems.

Business analysis includes:

- Enterprise analysis or company analysis - it focuses on understanding the needs of the business as a whole, its strategic direction, and identifying initiatives that will allow a business to meet those strategic goals.
- Requirements planning and management - involves planning the requirements development process, determining which requirements are the highest priority for implementation, and managing change.
- Requirements elicitation - describes techniques for collecting requirements from stakeholders in a project.
- Requirements analysis - describes how to develop and specify requirements in enough detail to allow them to be successfully implemented by a project team.
- Requirements communication - describes techniques for ensuring that stakeholders have a shared understanding of the requirements and how they will be implemented.
- Solution assessment and validation - describes how the business analyst can verify the correctness of a proposed solution, how to support the implementation of a solution, and how to assess possible shortcomings in the implementation.

**Table 1.** Steps of Business analysis life cycle

| |
|---|
| 1. Create financial statements |
| 2. Create a balance sheet |
| 3. Movement of cash flows |
| 4. Analysis of accounts and ratios |
| 5. Preparation of budget and planning cycle |
| 6. Forecasting and preparation of estimates |
| 7. Statistical control of determining the level of quality |
| 8. Examination of the business situation and the criteria for decision making for business analysis |
| 9. Planning gain |
| 10. Making investment decisions (under conditions of uncertainty) |
| 11. Recording of revenue and profits down |
| 12. Analysis of marginal revenue |
| 13. Close of business analysis for a given period |

Table 1 shows the business analysis life cycle steps.

The business analysis is prepared by different entities in different public areas.

There is common information for all business analysis life cycle stages.

There are two situations where the availability of semantic information associated to business analysis data is critical: business analysis formulation and approval tasks.

- In the first case, only company leadership with specific knowledge can be involved, concentrating a great responsibility on a few people;
- In the second case, semantics information is necessary for analyzing business analysis data. This is more complex because all legislators have to vote and most of them do not have specific knowledge.

*Ontology granularity and type*

According to the level of conceptualization and granularity [2], the ontology proposed here is domain ontology. In this case study, the ontology describes the business analysis domain of a company. The ontology objective is to facilitate communication among the members of the central administration staff that must deal with the local business analysis, bringing adequate terminology to non-expert users.

### 2.2 Conceptualization of Business analysis domain

The goal of the conceptualization phase is to organize, classify and structure this knowledge.

*Domain conceptual model*

Each concept definition includes:

- a name – short phrase or a word used by the ontology creators to refer to the concept;
- a description – a textual definition of each concept (in English).

Both attributes can be used in documenting the ontology, in clarifying the process of ontology creation.

A list of the most important term is elaborated in this step. To this aim, the bottom-up approach was used. With this approach:

- The set of basic terms is identified first;
- They are specified;
- They are generalized;
- Definition of key term list (with these concepts as reference).

The most important task in the methodology is the definition of a Domain Conceptual Model. Then, it is important to assign all the necessary time to carry out a good conceptual analysis. The conceptual model resumes the knowledge acquired during the specification phase and it is the basis of conceptualization. This conceptualization has to be agreed on by domain experts.

Formulate Conceptual Model

The conceptual modeling process comprises the inter-related activities as acquire and analyze structure description, determine modeling objectives, determine object roles and level of detail, and determine model structure.

The conceptual modeling process usually starts with the acquisition of system descriptions. These types of descriptions are unstructured and ambiguous, which are result from consultations between the analyst and domain expert. The nature of descriptions is that they are partial. From this point of view, an important aspect of the conceptual modeling process is determining the adequacy and completeness of an existing description.

Table 2 does not include total overlapping of concepts, synonyms, properties, relations and attributes.

**Table 2.** Key terms

| | | |
|---|---|---|
| Analysis | Confidence intervals | Period |
| a part of | Forecasting and evaluations | Planning and control |
| Assessments | Investment | Regression analysis |
| Autoregresion | Investment decisions | Report |
| base on | Large amount of data | Sales and marketing |
| by | Moving average | Smoothing |
| Computer programs | of | Standard deviation |

To understand the conceptual aspects, a Unified Modeling Language (UML) diagram is elaborated to represent the main relations among defined concepts.

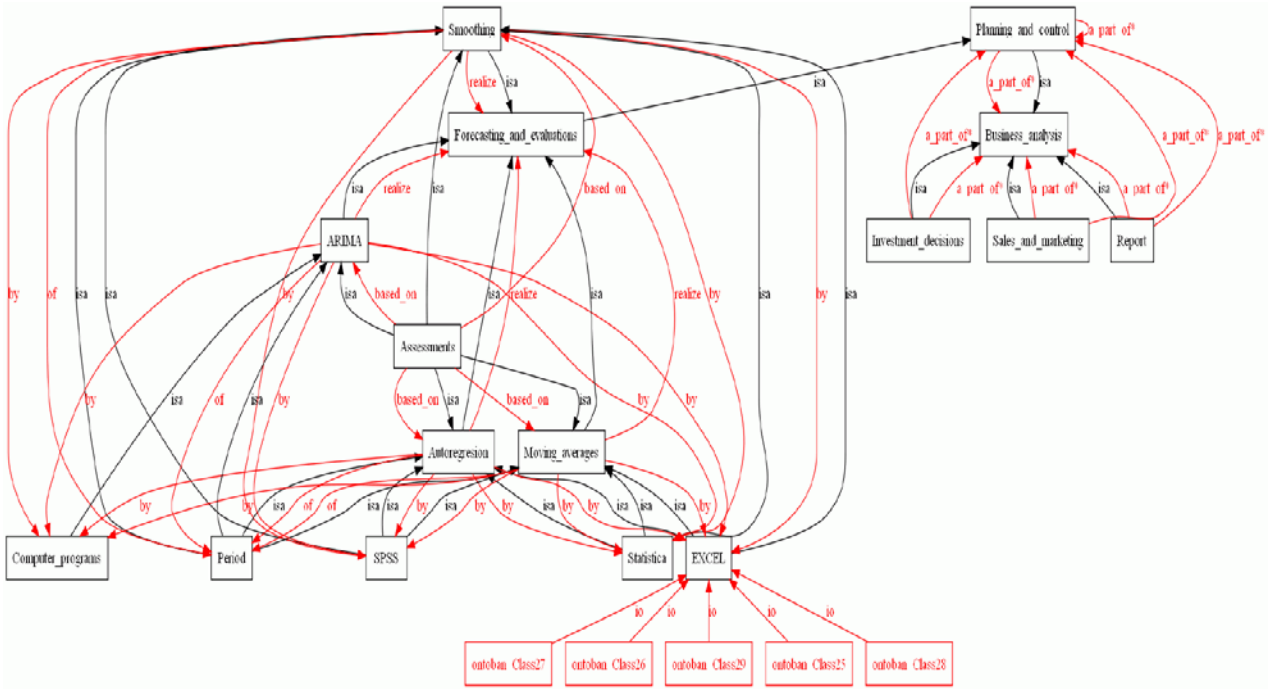


**Fig. 1.** A part of Domain conceptual model

The UML class diagram can be used to express concepts in terms of classes and relationships among them. On the other hand UML in its standard form is not suitable for semantic representation. The information modeled in the UML class diagram is the base for building the ontology term glossary. This UML model is useful to verify the ontology scope and to discover two granularity levels for business analysis domain concepts – Domain Ontology (contains the general concepts for the business analysis life cycle) and Formulation Ontology (contains the semantic specific for formulating a business analysis).

*Identification of Classes, Relations, Attributes, and Instances*

For determining classes, we identified those terms of independent existence from the key terms list and the glossary.

When the conceptual model of the ontology has been created, the first step is to define relevant instances inside an instance table. Each instance should be provided a definition of its name, the name of the concept it belongs to, and its attribute values if known. An excerpt of the Instance Table of the Business analysis Ontology is shown in Figure 2.

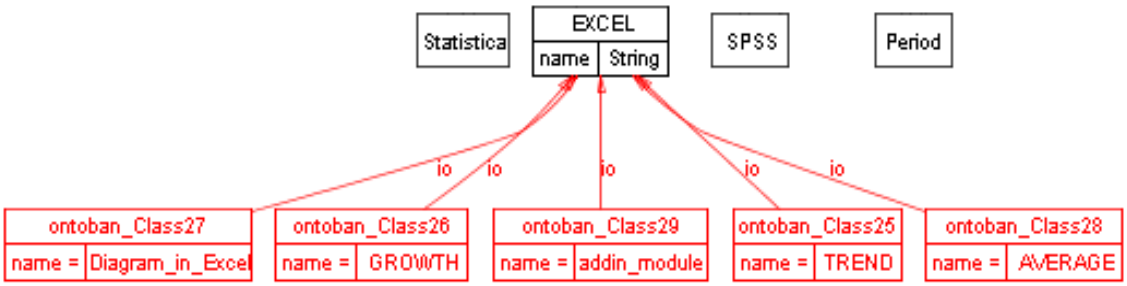


**Fig. 2.** An Excerpt of the Instances of the Business Analysis Ontology OntoBAn

**2.3 Integration of Business analysis domain ontology**

In integration phase we may integrate and reuse the existing and built ontology in ontology which we engineer.

Reusing and integrated ontologies are not straightforward. There is no guarantee that two systems with the same vocabulary have the same conceptualization. This problem is called ontology integration problem. To deal with this problem, some authors argue in flavor of mapping mechanisms between ontologies, while others propose using different kinds of ontologies. In the process of building and engineer of our ontology OntoBAn of business analysis domain we do not use or integrate any existing ontologies.

### 2.4 Implementation of Business analysis domain ontology

The goal of the implementation phase is to implement the ontology by using some tool from software engineering.

To compare the ontology implementation with its conceptualization, graphics using the OWL Viz and Onto Viz plug-ins were generated and compared with UML diagrams. On the one hand, OWL Viz enables the class hierarchies in OWL Ontology to be viewed allowing comparison of the asserted class hierarchy and the inferred class hierarchy. On the other hand, Onto Viz generates diverse combinations of graphics with all relations defined in the ontology, instances and attributes.

Analysis of tools for Ontology creation

*Building ontologies with Protégé*

Protégé is a popular open source ontology editor and knowledge base framework [7].

Protégé supports the export to standard ontology languages as OWL and RDF schema, but it is easily extensible through its plug-in interface. This makes it a flexible base for a rapid prototyping and application development.

We are going to describe the Protégé features in detail, and demonstrate how we used them for sub-ontology of Planning and control development.

Protégé is an ontology development environment with a large community of active users. The representation mechanism for ontologies and knowledge bases (Protégé's model) is based on a flexible meta-model which is comparable to object-orientated and frame-based systems. It basically can represent ontologies consisting of classes, properties (slots), property characteristics (facets and constraints), and instances.

During ontology design, the most interesting reasoning capability from this type of tools is classification. Classification is used to infer specialization relationships between classes from their formal definitions. Basically, a classifier takes a class hierarchy including the logical expressions, and then returns a new class hierarchy, which is logically equivalent to the input hierarchy. Protégé can display the classification results graphically. After the user has clicked the classify button, the system display both the asserted and the inferred hierarchies, and highlights the differences between them. The visualization can utilize the interaction between knowledge engineers and domain specialists and thus should increase the effectiveness of the knowledge acquisition process.

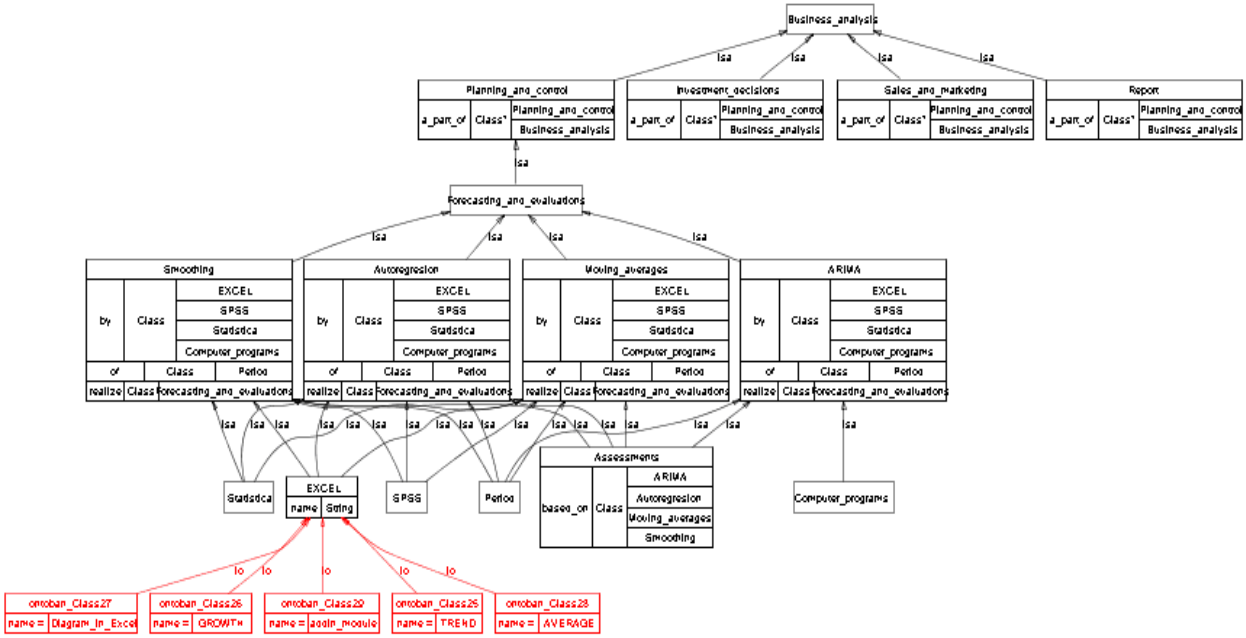


**Fig. 3.** Realization of a part of business analysis ontology

### 2.5 Documentation of Business analysis domain ontology

In documentation phase all documentation about the ontology must be prepared and viable.

Once the requirements are elicited, they are thoroughly analyzed by business analysts.

Business analysts document the Software Requirements Specification based on the results of the requirements analysis. Software requirement specifications serves as the base for the software design and development and can run into many pages depending on the complexity of the software application, size of the user groups and number of users.

A well-written software requirements specification not only lists out stakeholders' requirements but also communicates these requirements to the technical community for design and development of the system.

## 3 Business analysis ontology Onto-BAn

In this paper we propose a domain ontology for business analysis. In order to elaborate an ontology for that domain we have used terms proposed in [4] and [6] for the business analysis of a company.

**Onto-BAn ontology**

The main terms are organized in:

- Report - Nature and analysis of financial report ;
- Planning and control - Financial planning and control;
- Investment decisions;
- Sales and marketing.

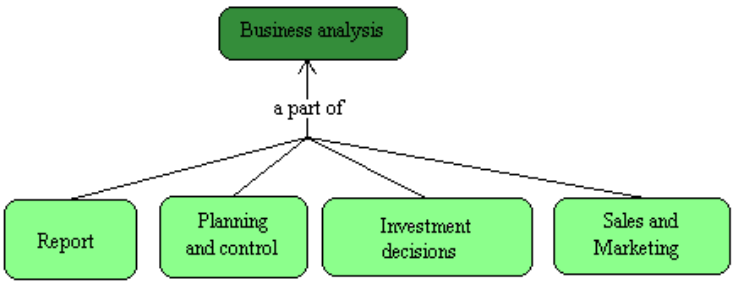


**Fig. 4.** General diagram of the ontology Onto-BAn (super class and classes)

Each term is defined in properties and relations, generating a complex network of classes, subclasses, instance and slots. The ontology Onto-BAn is designed as a reflection of the views of specialists in the areas of business analysis, statistics, and knowledge engineering.

We use the model of sub-ontology of Investment decisions to represent the mechanism and instrumentality for modeling of the Business analysis domain and how the business analysis ontology can be elaborated.

For the ontology graphical presentation we use objects with different shapes and colors and arrows with different colors.

In the following we describe the class "Investment decisions" of the super class "Business analysis" briefly.

*Description of class "Investment decisions"*

Investigation of business situations (investments) is the one of the most important activity, which the financial department can do to contribute the leadership (managers) to make decisions. The business situation has different form according to submitted question for discussion and the decision which can be made. She should confirm the logical assumptions, reporting the results of sensitivity analysis to changes in similar data. Business risk is the uncertainty that exists in every business activity.

The investment decision is one of the fundamental decisions of business management: Managers or leadership determine the investment value of the assets that a business enterprise has within its control or possession. These assets may be physical, intangible, or financial. Assets are used to produce streams of revenue that often are associated with particular costs or outflows. The manager must determine whether the net present value of the investment to the enterprise is positive that is associated with the particular area of business.

Investment comes with the risk of the loss of the principal sum. The investment that has not been thoroughly analyzed can be highly risky with respect to the investment owner because the possibility of losing money is not within the owner's control. The difference between speculation and investment can be subtle. It depends on the investment owner's mind whether the purpose is for lending the resource to someone else for economic purpose or not.

In the case of investment, rather than store the good produced or its money equivalent, the investor chooses to use that good either to create a durable consumer or producer good, or to lend the original saved good to another in exchange for either interest or a share of the profits. The consumer obtains a durable asset or investment, and accounts for that asset by recording an equivalent liability. As time passes, and both prices and interest rates change, the value of the asset and liability also change.

The concept "investment" is used differently in economics and in finance. Economists refer to a real investment (such as a machine or a house), while financial economists refer to a financial asset, such as money that is put into a bank or the market, which may then be used to buy a real asset.

Often, the decisions must be made without an access to reliable information. Some important concepts play fundamental role in the process of decisions making:

- Standard deviation – it is a widely used measure of the variability or dispersion. It shows how much variation there is from the "average" (mean). In finance, standard deviation is a representation of the risk associated with a given security, or the risk of a portfolio of securities. Risk is an important factor in determining how to efficiently manage a portfolio of investments because it determines the variation in returns on the asset and/or portfolio and gives investors a mathematical basis for investment decisions (known as mean-variance optimization). The overall concept of risk is that as it increases, the expected return on the asset will increase as a result of the risk premium earned – in other words, investors should expect a higher return on an investment when said investment carries a higher level of risk, or uncertainty of that return. When evaluating investments, investors should es-

timate both the expected return and the uncertainty of future returns. Standard deviation provides a quantified estimate of the uncertainty of future returns.
- Confidence intervals – With their help, it is possible to evaluate a given probability of unknown value of the general population. The boundaries are called confidence.
- Regression analysis - powerful tool that is used in the analysis of large amounts of data. This method is the basis for the study and understanding of dependencies between variables. In interpreting of the results of regression analysis, it is important to understand that the regression shows associative relationship between variables.

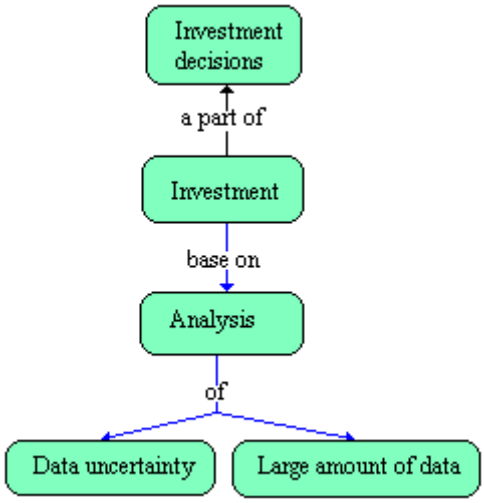


**Fig. 5.** A part of the class "Investment decisions"

The class "Investment decisions" (see Figure 5) is designed to describe the basic methods for analyzing of large amount data and data uncertainty. For the purpose of that investigation the concepts, connected to data, are grouped into two classes: *data uncertainty* and *large amount of data*.
Since investment decisions are very important for a company they are connected to other classes and subclasses of the ontology Onto-BAn – Planning and control, Forecasting and evaluations, Assessments, Period [3], Sales and Marketing. The class "Investment decisions" corresponds to the major issues directly, which makes them interesting for analysts – what is planned, what, how and where it is transacted, which are the results and the effects, what is the feedback to the corporate leadership.
The relations in class "Investment decisions" are of the type "a part of", "of", "base on", and "by".

## 4 Discussion

In order to develop the ontology presented in this paper, the methodology from part 2 has been followed. This methodology was divided into five phases: specification, conceptualization, integration, implementation, and documentation. The most important task in the methodology is the definition of a Domain Conceptual Model. The conceptual model resumes the knowledge acquired during the specification phase and it is the basis of conceptualization. This conceptualization has to be agreed on by domain experts. Graphical representation is essential in order to facilitate communication between ontology engineers and experts.
To making the ontology more flexible and allowing extensibility and reuse, it is important to modularize the ontology if possible. This modularization can be made through relations and attributes observation of conceptual aspects involved.

## 5 Conclusions

Business analysis is a crucial first step in the software consulting process because it identifies enterprise needs and helps build IT solutions aligned to these needs. Business analysis also gives management business intelligence, and allows them to study business performance, explain shortfalls in performance and devise strategies for improvement.

Ontology creation is iterative process of modelling the given domain, by choosing the\most important concepts and identifying the most relevant relations between them.

Building domain ontologies is not a simple task when domain experts have no background knowledge on engineering techniques and/or they have not much time to invest in domain conceptualization.

In this paper, we have shown how ontologists could develop Domain Ontology for Business analysis. The good practice on ontology building can be useful for the whole community. Then, the contribution of this paper is the implementation of domain ontologies.

## References

[1] O. Corcho, M. Fernandez-Lopez, A. Gomez-Perez, A. Lopez-Cima, "Building legal ontologies with METHONTOLOGY and WebODE. Law and the Semantic Web". *Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications*, 2005.

[2] A. Gomez-Perez, M. Fernandez Lopez, and O. Corcho, *Ontological Engineering with examples from the areas of knowledge management, e-commerce and semantic web*, London: Springer, 2004.

[3] E. Karashtranova, I. Atanasova, "A model for implementation of the dependence between random events in knowledge bases", *Economics and management*, Blagoevgrad, 2010.

[4] K. Karlberg. *Business analysis with Microsoft Excel*, SoftPress Ltd., 2003.

[5 ] R. Klichewski, K. Lenk, "Understanding and Modelling Flexibility in Administrative Processes". in: *Proc. EGOV 2002*, Ed. R. Traunmuller and K Lenk, 2002, pp. 129-136.

[6] P. Newbold, *Statistics for business and economics*, Prentice-Hall, New Jersey, 1984.

[7] Protégé (2008) *Protégé* Available online at http://protege.stanford.edu (Accessed January 2010).

[8] B. Smith, "Ontology", in L. Floridi (ed.), *Blackwell Guide to the Philosophy of Computing and Information*, Oxford: Blackwell, 2003, pp.155-166.

**Irena ATANASOVA** has graduated the Faculty of Mathematic and nature science of SWU – Blagoevgrad, Bulgaria in 1995. She holds Ms diploma in Mathematics, Informatics and Management and she joined the staff of the South-West University - assistant professor. She is the author of more than 10 journal articles in the field of modal logic and artificial intellect. Her work focuses on the process of engineer and building of ontologies.