COMPETITIVE INTELLIGENCE AS A USEFUL DECISION SUPPORT TOOL

Konkurenční zpravodajství jako užitečný nástroj pro podporu rozhodování

Roman Chladný[1]

[1] *Slezská univerzita v Opavě, Obchodně podnikatelská fakulta v Karviné, Univerzitní nám. 1934/3,73340 Karviná*
Email:chladny@opf.slu.cz

Abstract: This paper deals with Competitive Intelligence as a highly efficient tool for decision support process. At the beginning the typical features of Competitive Intelligence are described and its place in company's hierarchy is presented. The core of this contribution is aimed at the Competitive Intelligence cycle and its particular phases are described in more details. The most widespread methods of Competitive Intelligence are presented because they determine what data will subsequently be collected. Within data gathering an emphasis is put on the Web Mining technique, as a modern tool for data collecting from the Internet. The paper pays an attention also to the subsequent storage of the collected data and an application of Data Warehouse, built on a multidimensional principle is suggested as a suitable solution. The next part of this contribution presents Data Mining technique as an effective tool for data analysis with the aim of transforming the collected data and information into the strategic knowledge. The final part of the paper outlines an idea of an expert system projection using a knowledge base obtained within data mining process.

Abstrakt*:* Předkládaný příspěvek se zabývá problematikou konkurenčního zpravodajství, jako účinného nástroje pro podporu rozhodování. V úvodu jsou popsány charakteristické rysy konkurenčního zpravodajství a jeho místo v podnikové hierarchii. Jádro příspěvku blíže popisuje cyklus konkurenčního zpravodajství a jeho jednotlivé fáze: plánování, sběr dat, analýzu dat a distribuci výsledků. Jsou prezentovány nejčastěji využívané metody konkurenčního zpravodajství, jejichž volba determinuje, jaké data budou následně shromažďovány. V rámci fáze sběru dat je prezentována technika Web Miningu jako moderního nástroje sběru relevantních dat z Internetu. Práce věnuje značnou pozornost také následnému uskladnění nasbíraných dat, kde jako vhodné řešení navrhuje aplikaci datového skladu na multidimenzionálním principu. Další část příspěvku popisuje techniku dolování z dat, která se jeví jako efektivní prostředek pro analýzu uskladněných dat s cílem transformace dat na strategické znalosti. Závěrečná část práce nastiňuje myšlenku projekce expertního systému pro konkurenční zpravodajství s využitím báze znalostí získané v předešlém procesu dolování z dat.

Keywords: competitive intelligence, data mining, data warehouse, expert system

Klíčová slova: datový sklad, dolování z dat, expertní systém, konkurenční zpravodajství

JEL classification: F23, M10, M31

Introduction

In today's dynamic, turbulent and highly competitive environment it is not easy for companies to survive. Very often early and precise information can help significantly. As a result, a new discipline has arisen – Competitive Intelligence (CI). This relatively new and useful tool develops very quickly and becomes a substantial part of a management process in each bigger company. Competitive Intelligence can be characterised by the following typical features:

- continuous monitoring and evaluating of competitive environment with the aim of discovering competitors' strengths and weaknesses,
- exploitation of legal available information,
- analysis and synthesis of the collected information, which are subsequently transformed into the strategic knowledge during a Competitive Intelligence Process (CIP). The strategic knowledge obtained within CIP are further used in the decision making process as a powerful tool for decreasing business risk and getting a competitive advantage.

Competitive Intelligence can be defined as a collection of legal and ethical activities, which aim at getting necessary information for decision-making process. It is a business strategy and an interdisciplinary area, which draws knowledge from marketing, strategic management, economics and informatics. CI should provide insight into what may happen in a near future. It is a powerful tool of early warning, which signalises the external threats and opportunities. By analysing rivals' moves, CI allows companies to anticipate market development – rather than merely react on it (Govoreanu 2004).

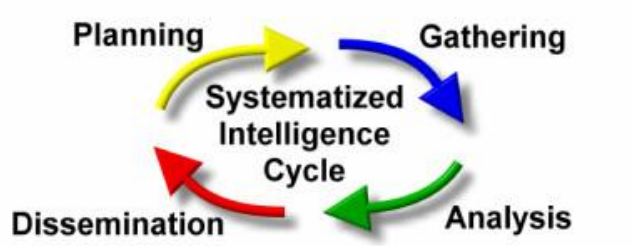
It is also desirable to define CI's place within company's information technology infrastructure. While the traditional components of company's IT infrastructure such as Enterprise Resource Planning (ERP) and Business Intelligence (BI) work with internal information, Competitive Intelligence (CI) uses especially external data and information. Therefore, it could be very effective to combine together these two types of information, which enables to evaluate the present state of the company in connection with the sector development and the competitive environment. The cross analysis of information obtained from BI and CI can be subsequently incorporated into the SWOT matrix or BCG matrix.

1 Competitive Intelligence Process

Competitive Intelligence represents a systematized cycle of the following activities:

1. Competitive Intelligence Planning
2. Data gathering
3. Data analysis
4. Dissemination of the results

Figure 1: Competitive Intelligence Cycle



Source: Govoreanu, A., Mora, A. and Serban, A. 2004.

1.1 Competitive Intelligence Planning

CI Planning defines the aim of CI and specifies what area will be monitored, what kind of data will be collected and what methods will be used to perform the task. It must be clear what is the purpose of the whole CI project. The most widespread CI models include (Colibasanu 2009):

- Porter's model of five competitive forces (present competitors, potential competitors, suppliers, customers, providers of substitute products),
- Value Net Model, which extends the five forces framework by examining the role of complementors, i.e. companies from which customers buy complementary products and their effect, which is the mirror image that of competitors,
- Product Life Cycle analysis indicates, where the company stands with regard to the developing phase of their products and shows a direction of the further company's focus – on market, on customers or on competitors,
- SWOT analysis helps to discover the competitive advantage of the company with regard to its market position. This approach analyses the strengths and weaknesses of the company in connection with the external threats and opportunities.
- BCG matrix - this approach, invented by the Boston Consulting Group, tries to classify company's products according to their market share and the market development speed,
- Supply Chain Analysis analyses relationships among producers, distributors, vendors and customers with regards to potential threats,
- Customer Segmentation Analysis helps the company to focus on the customer's segment, which shows the highest profit or market increase,
- Strategic Business Unit analysis is performed if a company operates in various sectors and needs to decide whether and in which sector to invest,
- PESTEL analysis gathers data from the political, economic, social, technological, ecological and legislation environment in order to evaluate how these external environments influence the business activities of the company,
- ADL matrix helps to understand how the company's strategy is being influenced by the state of the sector and company's competitive position in this sector,
- SPACE matrix defines the strategic position of the organisation analysing two internal factors: financial power of the company + competitive its advantage, and two external factors: power of the sector + stability of the environment,
- Monte Carlo Simulation tries to predict the future indicators on the basis of the historical and present data,
- Game Theory tries to anticipate the competitors' future steps on the basis of the mathematical models.

1.2 Data gathering

The selected method of CI determines which data are going to be collected. Among the most frequent data used for Competitive Intelligence project belong:

1. data concerning the external environment (inflation rate, gross domestic product, unemployment rate, balance of payments of the country, foreign exchange rate development, tax system, etc.),
2. data concerning sector's (competitors') environment (name of the competitor, its site, year of its establishing, legal form, number of employees, offered products, market share, offered vacancies, turnover, profit, total assets, payables, receivables, etc.),
3. data concerning company's internal environment (available sources, financial indicators, market position, management quality, etc.).

Information sources can be generally divided into 3 groups:

1. published sources can be found in a printed form or on-line (competitors' annual reports, competitors' websites, competitors' job offers, financial reports, books, magazines, newspapers, television, radio, stock exchange news, public registers, social networks, database centres etc.),
2. unpublished sources can be obtained by special methods, most often by a primary research when data are being newly gathered by own means,
3. semi published sources are hard to obtain because they are not published by classical means. These include especially research reports, special analysis performed by various institutions, technical reports etc.

Nowadays the Internet (the World Wide Web) represents one of the biggest sources of data and information for everyone. From this reason it can be very useful for Competitive Intelligence as well. A previous survey showed that about 90% of CI can be acquired from the World Wide Web (Thompson 2001). Information from the World Wide Web are mostly free of charge and it requires no special software to access them. In order to extract relevant information, however, it is necessary to check and analyse its content, structure and the possibility of the further exploitation. For this purpose, two suitable approaches can be used:

1. Search engines (e.g. Google, Yahoo, AltaVista, Lycos, Teoma, Bing, Ask Jeeves etc.),
2. Web Mining.

Search engines represent a classical and the most common method of collecting data from the Internet nowadays. They receive users' queries, usually in a form of key words, and provide a set of websites or documents, which should satisfy user's query. The search engines are generally composed of the following components (Boncella 2003):

- Web Crawlers or Web Spiders, which collect websites with the help of a graphical search technique,
- Indexers, which are used to index the gathered websites and to store the indices into a database,
- Retrieval and ranking methods, which retrieve search results and help to present them to users,
- User interface, which enables users to query into the database and to customize their searches.

Search engine approach, although still very popular, may provide too many irrelevant responses. On the other hand, Web Mining, as a newer searching technique, represents the methods for controlling the search in order to retrieve just the pages that are relevant for the search. This modern approach can be categorized into 3 concepts (Boncella 2003):

- Web Content Mining represents a process of searching the relevant data from Internet sources, which contain information usable for generating necessary knowledge. This concept is more precise than a standard searching by search engines and refines the initial found results with the help of Text Mining algorithms and techniques.
- Web Structure Mining tries to discover a model, which determines web hyperlink structure. If a website has more links from the others websites, this indicates that this website is relevant and important. Its rank will be increased in the hierarchy of the found websites.

- Web Usage Mining represents a process of searching users' behaviour patterns by data mining from various users' data warehouses – e.g. access web logs, users' profiles, registration data, cookies, users queries, click streams etc.

A Decision Support System based on a Web Mining technique consists of three layers generally (Chen 2002):

- data layer saves and reextracts information from all ERP databases and pass them on to the logical layer for processing.
- logical layer collects data, information and knowledge from ERP using Data mining algorithms, performs calculations, makes evaluation and finally transforms gathered data to the useful information. It also collects data, information and knowledge from the Internet using Web Mining techniques and provides them to the data layer.
- decision layer receives user's requirements and calls the logical layer in order to clarify requirements after their translation. After obtaining results from the logical layer, the decision layer transports these data into the format, which is required by the user.

As the analytical tools for Web Mining, the following products can be used: Google Analytics, Adobe Analytics, Universal Analytics, Site Catalyst, Omniture, Webtrends, Coremetrics, etc.

1.2.1 Storage of the collected data

The gathered data must be cleaned, scrubbed and filtered in accordance with their relevance and subsequently stored into appropriate data storage. Due to the fact that the collected data can be in various formats, a simple MS Excel sheet is not a sufficient tool for data storage. On the other hand, a database approach offers a much better solution how to store the gathered data effectively. Basically, there are two essential database concepts, which are used in practice nowadays:

1. relational database model,
2. multidimensional database model.

Tables, which are connected by relations, represent a basis of the relational databases. Data are stored in these tables in particular formats and the tables are linked according to their relationships. This approach is effectively applied in the transactions systems, where the manipulation with the data is performed on daily basis.

A modern competitive environment, however, requires the techniques, which enable to extract knowledge from the stored data. These procedures include the analytical tools of Business Intelligence such as Data warehouses, OLAP analysis and Data Mining. Due to the lack of these instruments, the relational database model is probably not very suitable for Competitive Intelligence purposes. On the other hand, a multidimensional approach appears to be an effective solution of this problem. This model organizes the data into the multidimensional structures. The database tables are not normalised and represent tables of dimensions. A multidimensional model of data is characterised by creating data aggregations and data grouping. A multidimensional data structure in a form of a cube, so called OLAP cube, is the result of this process. For creating an OLAP cube a lot of calculations and aggregations must be performed, which are running in a real time. A typical example of a standard OLAP cube is a solution with 3 dimensions: product, time and place. The subsequent analyses provide then results according to the product, time period or region (Lacko 2003).

From the above-mentioned facts we can come to the conclusion that a multidimensional approach can be an effective solution for Competitive Intelligence as well. The gathered data will be stored in a Data Warehouse built on the multidimensional principle. Basically, a multidimensional model stores the data in two ways:

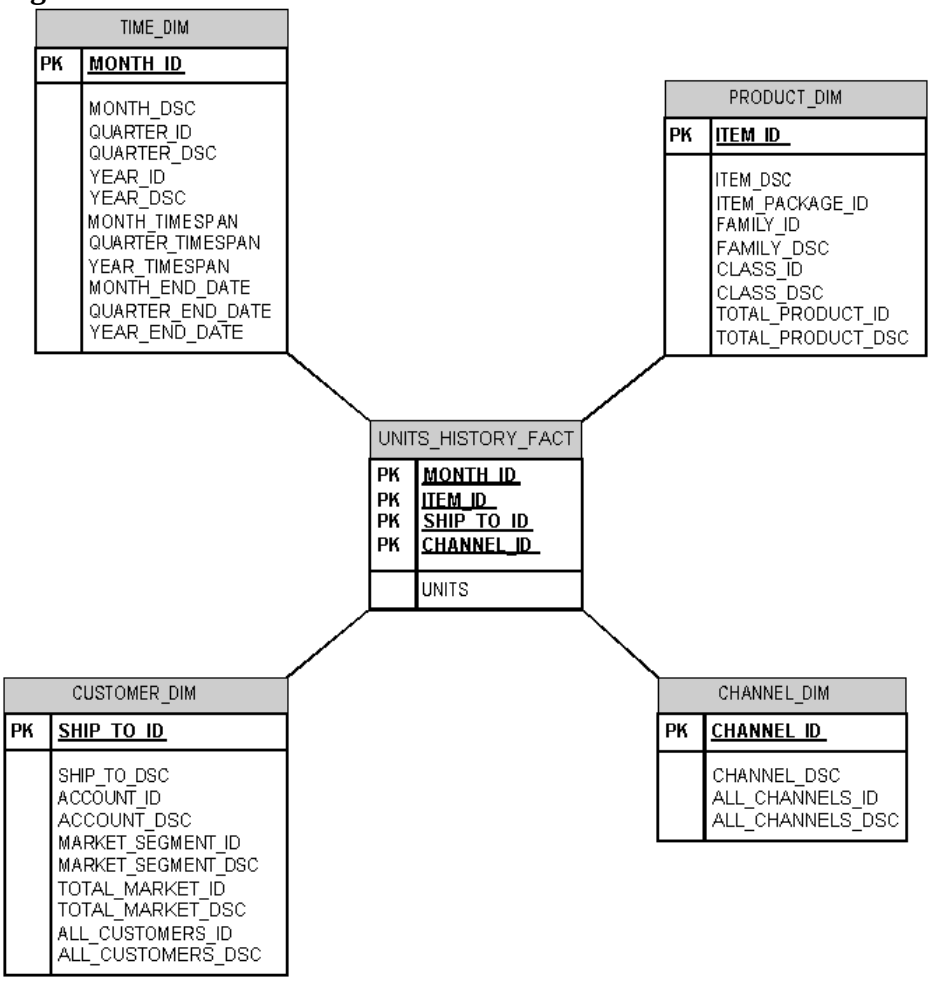
1. a star scheme
2. a snowflake scheme

In both cases we distinguish two types of tables:

- a table of the facts, which is located in the middle of the scheme. This is generally the largest table in a database with a huge amount of data. This table serves for a monitoring of the selected indicators and contains measurable records according to the time development.
- tables of dimensions

A star scheme includes one table of facts and foreign keys. These keys build links to the primary keys in tables of dimensions. The next picture provides an example of a database model based on a star scheme.

Figure 2: Star Scheme

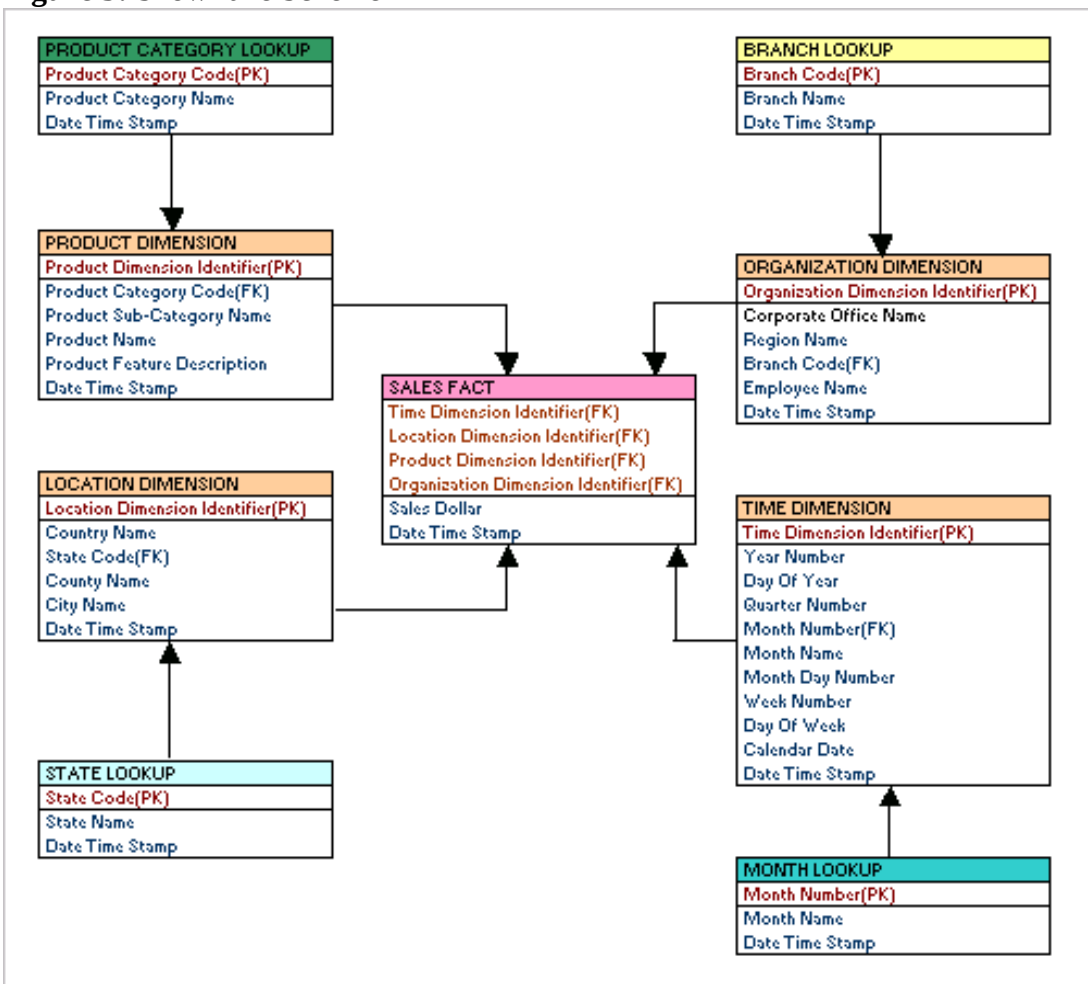


Source: Oracle Business Intelligence Concepts Guide, 2005

A table of facts is normalised, the tables of dimensions, however, not. A particular dimension is always represented by one particular table, which contains more attributes. This model offers a high inquiry performance and therefore is used very frequently (Lacko 2003).

A snowflake scheme includes a table of facts, which is connected with the tables of dimensions. These tables of dimensions are normalised into the further tables. The next picture illustrates an example of a database model based on a snowflake scheme.

Figure 3: Snowflake Scheme



Source: Oracle Business Intelligence Concepts Guide, 2005

This model offers a faster data implementation, an easy maintenance and lesser requirements on a disc capacity. However, a lower inquiry performance and the resulted decreased data analysis efficiency are a serious drawback of this model. As a result, this model is not used as often as the previous one (Lacko 2003).

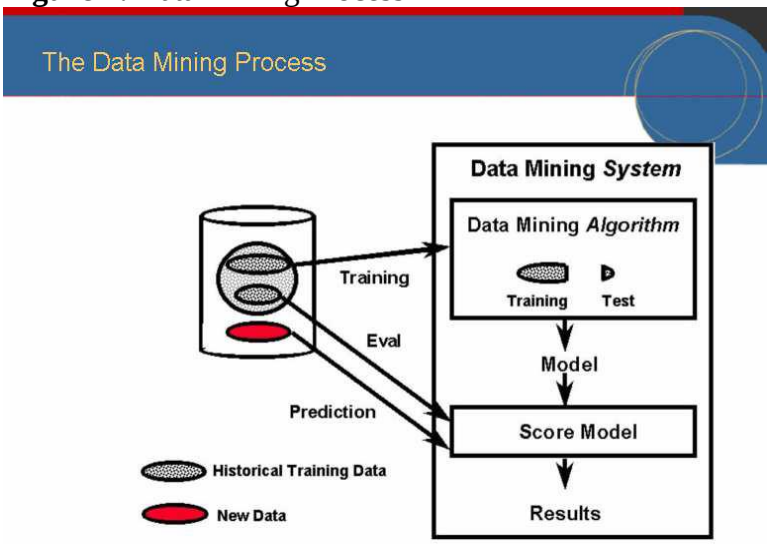
The above-mentioned schemes demonstrate how to store the data within a Competitive Intelligence Project. As for software products, which enable to build a Data Warehouse we can mention MS SQL Server, Oracle, My SQL, Cognos, Pentaho Business Intelligence, Jasper Soft Suite etc. Having stored the gathered data and information in an appropriate way, we can start to analyse them in order to transform them into the strategic knowledge.

1.3 Analysis of the gathered data

Analysis of the collected data is the key part of any Competitive Intelligence Process. There are various techniques how to approach the gathered data. Nowadays Data Mining represents the most widespread and the fastest developing technique of data analysis. Data Mining can be characterised as a process of discovering new, hidden correlations, patterns and trends among the collected data, which were unknown before, with the help of researching huge volumes of data saved in data warehouses using pattern recognition techniques and statistical or mathematical methods. Data Mining process should transform information into knowledge, i.e. knowledge is the final product of Data mining. A deep analysis process is based on the inductive methods of learning, with the aim to derive the general relationships and rules from a set of samples, which are composed of the past observations registered in one or more databases. A set of hypothesis is the result of this process. The following tools belong to the most widespread Data Mining techniques (Lacko 2003):

- Correlation analysis examines an intensity rate of dependence among the quantitative characteristics.
- Regression analysis examines a dependence of a monitored quantitative characteristic on one or more other quantitative attributes.
- Logistic regression is based on a similar concept as the classical regression, in this case, however, the dependent variable is discrete. This technique works with tools such as classification table or ROC curve.
- Discrimination analysis measures the importance of the factors, which determinate the categorization.
- Neuron networks are used for the creation of predictive models. They simulate the human brain structure. A neuron is the basic component of this model. It represents a cell for data input. Neuron networks gain information from a set of these inputs and on their basis they adjust the model parameters with the regard to the new knowledge.
- Genetic algorithms simulate a biologic evolution in solving assigned tasks. They assume the evolution process influence, when more models are being compared and adjusted by hybridising, mutation, cloning, random exchanges of values, marks and functions. This method demands a huge information technology capacity.
- Clustering and classification analysis groups and divides data on the basis of clusters into the groups with similar characteristics.
- Decision trees portray data in a form of a tree, where each root determines a criterion for the following separation into the particular leafs.
- Trends prediction is based on an analysis of the previous periods data and tries to model the future development of the monitored indicators.

Figure 4: Data Mining Process



Source: Thearling, K. *An Introduction to Data Mining* [online]

Software like MS Excel, SPSS, Matlab, Cognos DecisionStream, Data Analyser, Statistica Data Miner, Clementine, Oracle Data Mining, Kepler and Rapid Miner represent the typical analytical tools for Data mining purposes.

1.4 Dissemination of the results

The output from the whole Competitive Intelligence Project should be delivered to the internal or external customer in an appropriate form and in the desirable time schedule, including conclusions and recommendations. A feedback from the customer, i.e. an evaluation of the project can help significantly at the next similar order.

2 Building an Expert System for Competitive Intelligence

Having obtained the strategic knowledge from a Data Mining process, it suggests itself to build an expert system for Competitive Intelligence, based on this knowledge base. In general, any expert system consists of the following basic components (Górecki 2014):

1. knowledge base
2. inference mechanism
3. working memory
4. explanatory module

A knowledge base will be provided by Data Mining process. Its implementation into an expert system is very often based on a system of rules. The rules represent the most widespread way how to represent the knowledge. At the same time, they are very understandable for human thinking. The rules are generally formed in the IF THEN structure, i.e. situation → action. A rule is called out if a condition emerges in its situation part.

An inference mechanism is a programme module, which:

- takes the necessary knowledge from the knowledge base,
- interprets the knowledge,
- derives new and modifies existing knowledge,
- manages a communication with a user.

A working memory is a temporary working database in which all external or derived data are kept during the solution process. The knowledge are applied on the data in the working memory with the help of the inference mechanism.

An explanatory module enables to provide a necessary explanation to a user concerning a particular case. Furthermore, it reasons the results and shows how the expert system derived them.

An Expert System can be the final product of the Competitive Intelligence project. It can represent an important decision support tool, under an assumption that knowledge based is continuously updated with the help of never ending Competitive Intelligence process.

Conclusion

In the end, we can say that Competitive Intelligence appears to be a very useful tool for supporting a decision making process in any company. From the above presented text it is clear that Competitive Intelligence is a very large, complex and never ending project, which incorporates the elements of various disciplines: strategic management, marketing and informatics. Strategic management is present especially in the first phase of CI, which defines the objectives and methods of CI in the line with company's strategy. Marketing elements are being utilised in the second phase of CI in a form of data collecting. Informatics, however, seems to have a dominant place in the whole CI process. Firstly, Web Mining could be a very efficient tool for data gathering. Secondly, concept of Data Warehouses represents an effective solution how to store a huge amount of collected data. Next, Data Mining techniques provide a wide range of useful tools for data analysis and finally, artificial intelligence area knowledge can be applied when projecting an expert system for Competitive Intelligence.

References

[1] BONCELLA, R. J. 2003. *Competitive Intelligence and the Web*. Communication of the Association for Information Systems. Volume 12. ISSN: 1529-3181.

[2] COLIBASANU, O. A. 2008. *Between Intelligence and Espionage in the Contemporary Business Environment*. Marketing and Management Sciences. Proceedings of the International Conference on ICMMS 2008. Imperial College Press. ISBN 101-84816-509-9.

[3] GOVOREANU, A., MORA, A. and SERBAN, A. 2004. *Competitive Intelligence*. Studies in Business and Economics. Volume 14.

[4] GÓRECKI, J. 2014. *Expertní systémy pro prezenční formu studia*. Karviná. Projekt OP VK č. CZ.1.07/2.2.00/28.0017 "Inovace studijních programů na Slezské univerzitě, Obchodně podnikatelské fakultě v Karviné".

[5] CHEN, H., CHAU, D. and D. ZENG. 2002. *CI Spider: A Tool for Competitive Intelligence on the Web*. Decision Support Systems, Volume 34.

[6] LACKO, L. 2003. *Datové sklady, analýza OLAP a dolování dat*. Brno. Computer Press. ISBN 80-7226-969-0.

[7] Oracle Business Intelligence Concepts Guide. Technical Fundamentals for Implementation. 2005. [online] [15th Sep. 2016]. Available from: <http://www.oracle.com>.

[8] THEARLING, K. *An Introduction to Data Mining*. [online] [15th Sep. 2016]. Available from: <http://www.thearling.com/dmintro/dmintro-frame.htm>.

[9] THOMPSON, S. and WING, C. Y. 2001. *Assessing the Impact of Using the Internet for Competitive Intelligence*. Information & Management, Volume 39 (1), pp. 67-83.