Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2011 Proceedings - All Submissions

8-5-2011

A capacity and value based model for data architectures adopting integration technologies

Carlo Batini
DISCo - University of Milano-Bicocca, batini@disco.unimib.it

Cinzia Cappiello
Politecnico di Milano, cappiello@elet.polimi.it

Chiara Francalanci
Politecnico di Milano, francalanci@elet.polimi.it

Andrea Maurino
DISCo - University of Milano-Bicocca, maurino@disco.unimib.it

Viscusi Gianluigi
DISCo - University of Milano-Bicocca, gianluigi.viscusi@epfl.ch

A capacity and value based model for data architectures adopting integration technologies

Carlo Batini
DISCo - University of Milano-Bicocca
Viale Sarca 336-U14, Milan, Italy
batini@disco.unimib.it

Cinzia Cappiello
Politecnico di Milano
Piazza L. da Vinci, 32 - Milan, Italy
cappiello@elet.polimi.it

Chiara Francalanci
Politecnico di Milano
Piazza L. da Vinci, 32 - Milan, Italy
francalanci@elet.polimi.it

Andrea Maurino
DISCo - University of Milano-Bicocca
Viale Sarca 336-U14, Milan, Italy
maurino@disco.unimib.it

Gianluigi Viscusi
DISCo - University of Milano-Bicocca
Viale Sarca 336-U14, Milan, Italy
viscusi@disco.unimib.it

ABSTRACT

The paper discusses two concepts that have been associated with various approaches to data and information, namely capacity and value, focusing on data base architectures, and on two types of technologies diffusely used in integration projects, namely data integration, in the area of Enterprise Information Integration, and publish & subscribe, in the area of Enterprise Application Integration. Furthermore, the paper proposes and discusses a unifying model for information capacity and value, that considers also quality constraints and run time costs of the data base architecture.

KEYWORDS

Information capacity, information value, data quality, data architecture.

INTRODUCTION

Many have been the efforts in the past to investigate information and data in terms of economic categories to measure or define their value and the paying off of the investments in information systems (Dedrick et al. 2003). The literature both in computer science and information systems area has shown a change in the focus of the analyses from a focus on information economics (Gilboa et al. 1991) rather than on technology value for a given organization (Glazer 1993) to information as economic asset for an organization (Boisot 1998; Moody et al. 1999) and IT business value in terms of return on investments on technology (Melville et al. 2004). In this paper we focus on two concepts that have been associated with various approaches to the concepts of data and information, namely capacity and value. In particular, the goal of the study is to provide answers to the following research questions:

- *How to model the concept of information capacity?*
- *Which is the maximal total increase in information capacity achieved using data integration technologies over a set of databases?*
- *How the information capacity is related to quality of data, such as accuracy and completeness, and to the value produced by data integration initiatives?*

Furthermore, the problems considered in this paper have practical relevance for business in strategic initiatives and operations such as e.g. mergers and acquisitions of organizations having an information asset characterized by very large data bases. In an attempt to provide first results toward a systematization of the whole issue, we first provide four different coordinates to categorize the different approaches, and analyze main contributions in the literature using such coordinates. Since integration is seen frequently as a means to enrich the data/information capacity/value, then we focus on a relevant issue in the above defined multidimensional space, namely the context of data base architectures, and on two types of technologies diffusely used in integration projects, namely data integration and publish & subscribe. We resume the approaches proposed and discuss several drawbacks, providing a more comprehensive model that considers several definitions of information capacity, a new cost model, the data quality issue, and moves finally from capacity to value. Future work concludes the paper.

BACKGROUND AND MOTIVATIONS

When considering the concepts of capacity and value as they are discussed in the literature, we have first of all to notice that other two terms with similar or specular meanings are introduced, namely *utility* and *cost*. Second, in order to compare the different meanings and usages of the two concepts, one has to deal, besides the capacity vs value coordinate, with other three relevant issues, namely a. data vs information, b. the role of quality of data and information, and c. the technological architectures considered. We discuss first the three issues separately and then all together in their relationship with capacity and value.

The term *data* is usually adopted when an underlying technology is considered, that allows to store, manipulate, query, exchange information represented by means of simple structured domains. Usually, the reference technology considered is the relational DBMS technology. So, data are usually anchored to a model, such as "data in the relational model". Within this interpretation, two characteristics are usually associated to data:

- a structure, from which the term *structured data*, and a model that defines such structure;
- a query language, such as SQL or relational algebra in the relational model, that allows to extract data from other data, depending on the query.

The term *information* has a much wider scope, covering representations ranging from unstructured or semi-structured texts, maps, images, sounds, etc. Definitions of information in the literature are a wide number. Here we mention the Goguen's definition of *item of information* as "the interpretation of a configuration of signs for which members of some social group are accountable". A recent General Definition of Information (GDI) is in Floridi (2011) where "σ is an instance of information, understood as semantic content, if and only if: (GDI.1) σ consists of one or more data; (GDI.2) the data in σ are well-formed; (GDI.3) the well-formed data in σ are meaningful". So, data is considered to convey meaning only when such properties are achieved, producing a mutation into information. We do not want to go deeper on this foundational issue, we simply remark that the large majority of approaches in the literature combines the terms capacity and value with the concept of information. Information is increasingly being recognised as a key economic resource and the basis for achieving competitive advantage (Moody and Walsh 1999). However, the information value decreases if information contains errors, inconsistencies or out-of-date values. Therefore, high information quality levels can be considered an initial guarantee for the potential usefulness of the information objects. Data and information quality literature provides a thorough classification of quality dimensions. Analyzing the most relevant contributions (such as e.g. Naumann 2002) it is possible to define a common basic set of quality dimensions including *accuracy*, *completeness*, *currency*, *consistency* (Batini and Scannapieco 2006).
The correlation between the information value and accuracy has been analyzed in (Moody and Walsh 1999): the higher the accuracy of information, the higher its usefulness and value. Low accuracy levels can be very costly since they can cause both operational errors and incorrect decision making. (Moody and Walsh 1999) point out that the information value also depends on the age of the information. Information is often very dynamic and its validity decreases over time.

With the term *data architecture* we define the allocation of the data of interest to an organization among the (usually many) database management systems available in the organization's information system. Organizations tend to create databases of interest through a series of projects and realizations that result in a database architecture characterized by a set of anomalous behaviours. Among existing integration technologies in the market two are the solutions that are emerging (Bernstein and Haas, 2007), respectively Enterprise Information Integration (EII) and Enterprise Application Integration (EAI). EII is seen as a framework made of middleware and services to provide a single interface for viewing all the data within an organization to appear to users as a single, homogeneous data source. EII achieves integration using data technologies; among them, *data integration* (DI) allows users to read-only access data stored in heterogeneous data sources through the presentation of a unified view of these data. In the last few years, both industry and academia have investigated data integration solutions both from theoretical and practical view points, see Bergamaschi, Maurino (2009) for a survey. In *virtual data integration* the unified view called *global schema* is virtual, and data reside only at sources (Wiederhold, 1992). EAI achieves the integration through the usage of middleware solutions. Among them, *publish and subscribe* (P&S) (Eugster et.al. 2003) is a kind of message oriented middleware that realizes a many to many and anonymous interaction among a group of participants. The participants are divided in two groups (i) the publisher, namely the producer, message sender, and (ii) the subscribers or consumers that are interested in receiving specific typologies of update messages. P&S is used for the integration of updates.

THE FOUR ISSUES CONSIDERED TOGETHER

Among the four concepts previously introduced, relating data/information to capacity/value, the most investigated concept is information value. The analysis of the literature shows how different approaches to information value have been followed depending on the disciplinary area and focus. As anticipated in the introduction, the academic interest in information value has evolved in terms of focus for different disciplines. Due to the radical change at societal and enterprise level, related to the diffusion of internet and the (supposed) commoditization of information and communication technologies (Carr 2004), in the following we consider in our analysis as period of interest mainly the last twenty years. Early studies in information value have been carried out in the area of information economics. Economists have investigated information value mainly from a mathematical perspective, by focusing in particular on classical problems of asymmetric information and game theory (Gilboa et al. 1991).

The growing relevance and adoption of enterprise systems and the consequent role in business of information and communication technologies (ICT) shifted the attention towards the value of information associated with value chain (Glazer, 1993). The adoption of ICT and, in particular of enterprise systems (Markus et al. 2000) leads scholars in information systems and knowledge management fields to investigate frameworks for the management of information at organizational level (Moody et al. 1999; Simpson et al. 1995; Skyrme 1994). Here it is worth noting a first association of information value to information quality dimensions such as e.g. accuracy, accessibility, completeness, currency, reliability, timeliness, and usability. Furthermore, the growing of investments in ICT poses the question of the ICT business value, mainly in the information systems field; in an accurate survey, Melville et al. (2004) show how ICT business value extent and dimensions are dependent upon internal and external factors (complementary organizational resources of the firm, partners, competitive and macro environment). In general, economy and knowledge oriented perspectives pay little attention to the type of technology mediating the information provision and its use/consumption, whereas in management of information systems approaches information value can be hardly disentangled from the technological and organizational resources and environment. As a consequence, a common interpretative framework for information value has to consider a set of characteristics which can provide dimensions suitable to evaluate information facets at different level of abstraction. As to technology, it is worth noting that it represents an independent variable or *weight* for each characteristics.

Concerning information capacity, the contributions are more limited, and the term is used with quite different meanings. In (Francalanci et al 2008), the information capacity of an organization, that is modelled as a set of cooperative processes, is the effort required to produce the quantity of information that the cooperative processes represented with a network can process in a time unit. Information is measured in terms of number of information units. So, technologies considered are generically cooperating processes that share information in an organization, and can be classified as pertaining to EAI. In (Miller et al. 1993) the information capacity of a schema S corresponds to its set of instances, and the information capacity preservation is investigated when integration transformations are performed on a set of schemas. In Batini et al. (2010) data is heterogeneously represented in terms of a set of databases that can be queried and updated with a DBMS language, and information is any query (at the intensional level) or result of a query (at the extensional level). The concept of information capacity is investigated within data integration architectures, seen as the increment in the number of queries that can be expressed over a set of databases integrated in a DI architecture, and that could not be performed querying databases locally.

In conclusion, we may say that the concept of capacity seems to evoke an intrinsic property of data and information, a potential that can be defined and evaluated independently from the usage, while value seems to evoke a property that can be intrinsic too, but more properly often depends on several factors, such as the context and the usage or the process that uses data. Furthermore, a comprehensive model of the concepts of information capacity and value is needed, also in their relationships with integration technologies and quality constraints.

A VALUE BASED MODEL OF DATA AND INFORMATION INTEGRATION TECHNOLOGIES UNDER QUALITY CONSTRAINTS

In this section we propose a model for characterizing the information capacity of a technological data architecture, namely a set of databases db_1, db_2, .., db_n federated through data integration (DI) and publish & subscribe (P&S) technologies, and considering also several quality dimensions of data and information.

We first provide a definition of a DI & P&S database architecture. A data integration (DI) architecture is usually defined as a triple $< G, S, M >$ where G is the global schema, S is a set of local schemas and M is the mapping between G and S. We also assume that the global schema contains all the attributes of the local sources.

The global schema G can be defined as a set of entities $E = \{e_1, e_2, \dots, e_j\}$ each one expressed in terms of a view on the set of local schemas $S = \{s_1, s_2, \dots, s_i\}$; this type of mapping is called the Global as View (GAV) mapping, whereas in the Local as View mapping the entities of local schemas are expressed as views on the global schema. The mapping M can be expressed in different ways depending on the architecture that is adopted, but in practice it defines, for each attribute A_{jk} contained in the j-th entity of the global schema, its relationships with the n-th attribute $A_{in}$of the i-th local schema.

In our approach a *DI & P&S architecture (DA)* is a set of data bases $DA_{DI\ \&\ P\&S}$ = [db_1, db_2, …, db_i] integrated through a DI technologies, and extended with the adoption of P&S technologies in order to coordinate the updates of entities common to two or more schemas. Databases are associated with their schema [s_1, s_2, …, s_i] that provide a map of represented objects together with their attributes and relationships. Due to the heterogeneity of the databases db_1, db_2, …, db_i and of update operations on entities that are common to at least two databases, several instance-level conflicts can arise on attributes of such entities, leading to the possibility of having different values for the same attributes. We assume that a policy is adopted to manage such conflicts among the different policies described in the literature, see (Batini, Scannapieco 2006).

In our model we consider four different quality dimensions, namely *global schema completeness*, *data accuracy*, *data completeness* and *data currency*. Global schema completeness states that the schema should contain all the relevant attributes that allow an appropriate representation of the local schema entities. According to our previous statement, this quality is achieved in our approach by definition. Syntactic accuracy of a table, in relational terms, can be measured in terms of the percentage of n-ple values that are contained in a look-up table of all valid values. Concerning data completeness, it means that the local sources have to include all the instances about the objects they represent. Currency provides a measure of the updateness of a retrieved instance.

We now proceed to define the concept of information capacity. Intuitively, the information capacity of a data architecture $DA_{DI\ \&\ P\&S}$ is the incremental data that can be extracted from the architecture $DA_{DI\ \&\ P\&S}$ and that cannot be extracted from the set of local schemas LS = [s_1, s_2, …, s_i]; this concept can be defined at three levels:

1. intensional level,
2. extensional level ignoring the data quality issue, and
3. extensional level considering the data quality issue,

leading to three different definitions that we discuss in the following. Besides information capacity, seen as the benefit arising from the adoption of DI&P&S technologies, we will also discuss the cost issue, providing in such a way a complete model for the whole problem. The *intensional information capacity* of a schema s is defined as the number of all the possible queries Q than can be expressed on s. We observe that the distribution of queries relating n entities through join paths in a real life application load decreases rapidly with n; due to this issue, we consider only queries on the schema with maximum length equal to a fixed value n.

To calculate the intensional information capacity, we extend the approach presented in Batini et al. (2010). Let be G a graph where E are the entities of the schema, and edges R are paths among entities. Let AM be the adjacency matrix of the graph G. We first calculate all paths of length N, by elevating the adjacency matrix to N, but, instead of just counting the paths, we lists them, so that the (i,j) entry of the matrix is a list of relationship paths from i to j; then, instead of multiplying and adding cell values in the matrix, the multiplication algorithm concatenates cell values. This way the final matrix represents the query paths.

On the basis of this definition, the *intensional information capacity* IIC of a local schema s_i can be defined as the number of all the possible queries Q that are calculated by the previous algorithm:

$$IIC(s_i) = |Q_i|$$

And the information capacity of the set LS of local schemas is

$$\mathrm{IIC}(LS) = \sum_i IIC(s_i)$$

The *incremental intensional information capacity* $IIC(DA_{DI\&P\&S})$ of a data architecture is the number of queries that can be expressed on the global schema and that cannot be expressed on the set of local schemas. To calculate the incremental intensional information capacity, we extend the previous algorithm as follows. Let C the set of entities that are common to at least two local schemas, G a graph where E are the entities of the global schema, and edges R are paths among entities. Let AM be the adjacency matrix of the graph G. We first calculate all paths of length N, by elevating the adjacency matrix to N, but, instead of just counting the paths, we lists them, so that the (i,j) entry of the matrix is a list of relationship paths from i to j; this way the final matrix AM^N represents the query paths. Finally, all relationship paths not including entities of C are removed.

As to the incremental extensional information capacity of the architecture DA, we are not interested to queries, rather we focus on instances retrieved by queries. We define I_k the set of instances retrieved by the k-th query q_k listed in the previous AM^N matrix. Therefore, the extensional information capacity for the data architecture DA can be defined as:

$$EIC(DA) = \sum_k I_k$$

The definition of extensional information capacity introduced above does not consider the fact that in real life databases, data are characterized by errors of various types, and the number of retrieved instances is influenced by the quality of data.

Using the relational terminology, assume we have two databases made both of a unique table DB1 = [T1(A,B)] and DB2 = [T1(B,C)] and assume that the accuracy of values of the attribute B in the two tables is 0.9, meaning that only 90% of values are correct. In this case a join will not link the 100% of the n-ples, but only a subset depending on the distribution of errors in values for the attribute(s) B, and the percentage of linked n-ples decreases as far as the incorrect n-ples increase. The same phenomenon occurs in the case the tables are incomplete, namely some of the instances of the domain of B are not represented in n-ples of T1 and/or T2: n-ples generated are less than n-ples generated when completeness (B) = 100% in both tables. Besides accuracy and completeness, the nature and quality of instances retrieved is influenced also by currency. If we have two local schemas made of the same unique table Employee (Emp#, Address), when an employee with Emp# = "0318" represented in both databases changes her/his address, if the update is performed only on one table the two addresses are inconsistent. A query on the global schema looking for the address of the employee with Emp# = "0318" retrieves two different addresses. Depending on the policy adopted for conflict resolution the instance resulting from the query has an intrinsic correctness that is always smaller or equal to the correctness of the instance resulting from the query, in case the currency and coherence of all copies is ensured by a P&S technology. We define therefore a *quality constrained information capacity* as the number of *correct* instances retrieved by all queries, in a formula:

$$QCEIC\ (DA_{DI\&P\&S}) = \alpha\ EIC\ (DA_{DI\&P\&S})$$

Where α is the probability that an instance is correct and α=f(accuracy, completeness, currency). Note that $QCEIC\ (DA_{DI\&P\&S})$ is always less than or equal to $EIC\ (DA_{DI\&P\&S})$. The parameter α can be determined by suitably extending methods for the evaluation of the quality of instances obtained as a result of queries, when it is known the quality of the instances in input to the query, see (Batini & Scannapieco 2006).

We move now to the concept of *value of a DI &P&S architecture*. First, we associate a value to single queries in all of the above definitions of capacity. Each query q_k has the potential to produce useful information to the users that submit it. The value can be associated with a single query but also with the single instances retrieved by the query. In general, we can assume that a function V = value(obj) exists, where the object *obj* is either a query or an instance, and the domain of V can be either a-dimensional or else a monetary domain. We can distinguish between intensional value and extensional value. The *intensional value of a DI&P&S architecture* is

$$IV(DA_{DI\&P\&S}) = \sum_k V(q_k)$$

Similarly we can introduce the *extensional value of a DI&P&S architecture*. We consider the value associated with each t-th instance retrieved with the k-th query:

$$EV(DB_{DI\&P\&S}) = \sum_k \sum_t V(i_{kt})$$

In the above definitions the critical issue is how to determine a. the function [value(obj)] → V. Two possible methods suitable to determine the function are:

1. by means of questionnaires that gather user evaluations.

2. considering business processes that make use of queries, and analyzing the increase in efficiency and effectiveness of processes resulting from the exploitation of the query from the business process.

A detailed discussion on the above two issues is outside the scope of the paper.

Combining heterogeneous data sources using DI and P&S technologies is a complex and costly operation. In fact, DI and P&S architectures raise *design costs* and *execution time costs.* The design costs are divided in four groups:

– C_S: design cost for the source wrapping. The set of local sources has to be analyzed and registered in the data architecture.
– C_E: design cost for the definition of the global schema.
– C_M: design cost for mapping local schemas to the corresponding entities e_j of the global schema. These data are reconstructed in a mapping table that lists the mapping of the entities in the global schema in terms of the entities in the local schemas. As to the publish and subscribe architecture, the mapping table supports the propagation of the update events in a source in the other sources that contain similar data.

  We focus in the following on the design-time costs associated with constructing the mapping table. For each entity e_j, we model this cost as a function of the number of distinct queries involving e_j. For each query involving e_j, designers must understand the instance set involved in answering the query and store the corresponding information in the mapping table, accordingly. These costs are indicated as C_{MD} and are expressed as:

$$C_M(e_j) = d \cdot |Q_j|$$

  where d represents the average cost of analyzing a query and storing the corresponding information in the mapping table, while $|Q_j|$ indicates the number of queries involving entity e_j.
  Thus, the total mapping cost for DI and P&S architecture is:

$$C_M = \sum_j C_M(e_j)$$

– C_{DQ}: possible costs at design time for procedures to clean and reconcile data

The total design cost can be defined as:

$$C_D = \sum_i C_{s_i} + C_E + C_M + C_{DQ}$$

Concerning run time costs, each entity e_j involves a runtime cost, c_{exe}, that depends on the frequency with which the entity is queried and updated called respectively $q(e_j)$ and $u(e_j)$:

$$C_{exe}(e_j) = p_j \cdot cq(e_j) + r_j \cdot cu(e_j)$$

where p_j and r_j represent the average cost of querying and updating entity e_j.

The total runtime cost of the global schema E is the sum of the runtime cost of all entities $e_j \in E$ that is:

$$C_{exe}(E) = \sum_j C_{exe}(e_j)$$

CONCLUSION AND FUTURE WORK

In this paper we have proposed a model which aims at unifying the concepts of information value and capacity in data architectures. The considered architectures make use of data integration and publish & subscribe integration technologies. Furthermore, we have assumed data may have quality problems with respect to accuracy, completeness and currency quality dimensions. The proposed model is based on literature analysis and formalization of authors experiences; as a consequence, it requires application and testing on real cases. Due to these issues, the model needs future work in at least three directions.

First, experiments have to be performed that show how our capacity and value measures are influenced by the number of schemas, the shared entities among the different schemas, the number of entity instances, the quality level of instances, the distribution of conflicts among instances, the policies adopted for conflict resolution. A tool has been developed that allows the casual generation of a set of local schemas that fit given values for the above dimensional indicators.

Further, we need a better formalization, investigation and experiments on the two methods sketched above for the expression of the function that associates a value to an architecture. Finally, the model needs to consider coopetitive systems, where the players are available to share entities only in case the advantage they achieve, expressed in terms of the capacity and value concepts introduced, exceeds the economic loss resulting from the availability of the entity by competitors in the system.

REFERENCES

1. Batini, C., Grega, S., Maurino, A. (2010): Optimal enterprise data architecture using publish and subscribe. HPDC 2010: 541-547
2. Bergamaschi S. and Maurino A.. (2009) Toward a unified view of data and services. In G. Vossen, D. D. E. Long, and J. X. Yu, editors, *WISE*, volume 5802 of *Lecture Notes in Computer Science*, pages 11–12. Springer.
3. Bernstein P. A. and Haas L. M. (2008) Information integration in the enterprise. *Commun. ACM*, 51(9):72–79.
4. Blanton, J.E., Watson, H.J., and Moody, J. (1992) Toward a Better Understanding of Information Technology Organization - a Comparative Case-Study, *MISQ Quarterly* (16:4), Dec 1992, pp 531-555.
5. Boisot, M. (1995) *Information space: a framework for learning in organizations, institutions, and culture* Routledge, London ; New York.
6. Boisot, M. (1998) *Knowledge assets : securing competitive advantage in the information economy* Oxford University Press, New York.
7. Carr, N.G. (2004) *Does IT matter? : information technology and the corrosion of competitive advantage* Harvard Business School Press, Boston.
8. Dedrick, J., Gurbaxani, V., and Kraemer, K. (2003) Information technology and economic performance: A critical review of the empirical evidence," *Acm Computing Surveys*), MAR 2003, pp 1-28.
9. Eugster P., Felder P. Guerraoui R., and Kerrmarreck A. (2003) The many faces of publish and subscribe. ACM Computing Surveys, Vol. 35, No. 2, June 2003, pp. 114–131.
10. Floridi, L. (2011) Semantic Conceptions of Information, in: *The Stanford Encyclopedia of Philosophy,* E.N. Zalta (ed.).
11. Francalanci, C., and Morabito, V. (2008) IS integration and business performance: The mediation effect of organizational absorptive capacity in SMEs, *J Inf technol* (23:4) 2008, pp 297-312.
12. Gilboa, I., and Lehrer, E. (1991) The value of information – an axiomatic approach, *Journal of Mathematical Economics*) 1991, pp 443-459.
13. Glazer, R. (1993) Measuring the value of information: The information-intensive organization, *IBM SYSTEMS JOURNAL* (32:1).
14. Goguen J. A. (1997) Towards a Social, Ethical Theory of Information, In Social Science Research, Technical Systems and Cooperative Work: Beyond the Great Divide, Georey Bowker, Les Gasser, Leigh Star and William Turner, Erlbaum, 1997.
15. Markus, M.L., and Cornelius, T. (2000) The enterprise systems experience— From adoption to success, in: *Framing the Domains of IT Research: Glimpsing the Future Through the Past,* R.W. Zmud (ed.), Pinnaflex Educational Resources, Cincinnati, OH.
16. Melville, N., Kraemer, K., and Gurbaxani, V. (2004) Review: Information Technology and Organizational Performance: An Integrative Model of IT Business Value, *MISQ* (28:2).
17. Miller R.J., Yoannidis Y.E, Ramakrishnan R. (1993) The use of information capacity in schema integration and trnslation, Proceedings of the 19th VLDB Conference, Dublin.
18. Moody, D., and Walsh, P. (2002) Measuring the value of Information: An asset valuation approach, ECIS 1999.
19. Naumann F. (2002) Quality-Driven Query Answering for Integrated Information Systems, LNCS 2261.
20. Sajko, M., Rabuzin, K., and Bača, M. (2006) How to calculate information value for effective security risk assessment, *Journal of Information and Organizational Sciences* (30:2).
21. Skyrme, D. (1994) Ten Ways to Add Value to Your Business, *Managing Information* (1:3) 1994, pp 20-25.
22. Wiederhold. G. (1992) Mediators in the architecture of future information systems. IEEE Computer, 25(3):38–49.