Original Article

Ontology-supported database marketing

Received (in revised form): 10th March 2009

Filipe Mota Pinto
is Professor of Database Marketing at Polytechnic Institute of Leiria – Portugal.

Alzira Marques
is Professor of Relationship Marketing at Polytechnic Institute of Leiria – Portugal.

Manuel Filipe Santos
is Professor of Artificial Intelligence at University of Minho – Portugal.

ABSTRACT Database marketing (DBM) provides in -depth analysis of marketing databases. Knowledge discovery in database techniques is one of the most prominent approaches to supporting some of the DBM process phases. However, in many cases, the benefits of these tools are not fully exploited by marketers. Complexity and amount of data constitute two major factors limiting the application of knowledge discovery techniques in marketing activities. Currently, ontologies may here play an important role in the marketing discipline. Motivated by its success in the area of artificial intelligence, we propose an ontology-supported DBM approach. The approach aims to enhance DBM with ontology by providing detailed step-phase specific information. Our research work has its foundations in a double methodological approach using the Delphi and Action Research methodologies. First, we use Delphi to structure related DBM knowledge, and then we align our work to the Action Research methodology in order to systematise the knowledge extraction process and knowledge base creation. The issues raised in this paper both respond and contribute to calls for a DBM process improvement. Our work was evaluated in the relationship marketing domain focusing on a relational marketing programme database. The findings of this study not only advance the state of DBM research, but also shed light on future research directions.

Journal of Database Marketing & Customer Strategy Management (2009) **16,** 76–91.
doi:10.1057/dbm.2009.9

Keywords: database marketing; knowledge extraction; ontologies; marketing databases; knowledge base

Correspondence:
Filipe Mota Pinto
School of Technology and Management, Polytechnic Institute of Leiria, Campus 2, Morro do Lena – Alto do Vieiro, Apartado 4163, Leiria 2411-901, Portugal
E-mail: fpinto@estg.ipleiria.pt

INTRODUCTION

Database marketing (DBM) is a database-oriented process that explores database information in order to support marketing activities and/or decisions.

The Knowledge Discovery from Databases (KDD) process is well established in the scientific community as a three-phase process: data preparation, data mining and deployment/evaluation. This process is guided and controlled by both domain experts and database analysts. The KDD has been successfully applied in various domains, particularly in the marketing field.

Nevertheless, there seems to be a lack of knowledge regarding its application to different requirements and conditions, such as marketing objectives, available data, database types or even missing domain expertise.

Our work focuses on the integration of knowledge extraction techniques within the DBM discipline. Here, we introduce ontologies as a support to the knowledge structure and integration of both fields. In the context of knowledge-sharing, the term 'ontology' means a specification of a conceptualisation. That is, ontology is a description of the concepts and relationships that can exist for a single technological application or as a reference in a decision support system, and can be designed for the purpose of enabling knowledge-sharing and reuse.[1–3] In this paper, we provide an approach based on high-level abstraction using domain ontologies in order to construct a formal framework from data to marketing knowledge.

Current situation

Technology has provided marketers with huge amounts of data, and artificial intelligence researchers with high level processing rate machines. Isolated practical DBM samples have been developed in different research fields.[4–6] In addition, there are some artificial intelligence projects that focus on marketing problems, but their usage remains based on a single methodology, for example, algorithm performance analysis, data processing rates or data mining sample projects.[7,8] That is, excluding proprietary business projects (marketing databases are normally used in a confidential environment), many of the research tasks (for example, data preparation, data mining or evaluation phases) are focused on solving a specific problem without further inferences or information registration for other future cases or knowledge-sharing.

Problem statement

Any time marketers need to develop DBM projects, they almost always start from scratch – much of the previous knowledge is unavailable, or when available is in an unhelpful format.

Much of the research developed in both fields (marketing and knowledge extraction techniques) focuses on DBM process and its results. Knowledge reuse in the marketing field is an innovation that could solve many of the practitioner's problems when developing his/her database-based marketing activities.

Proposed solution

In computer science, ontologies provide a shared understanding of knowledge about a particular domain.[1] Marketing ontologies, although low in number, are starting to come to the light through some marketing or computer research centres.[9–14] Marketing ontologies are becoming more and more available, and are contributing to the understanding of the large amounts of data existent in the marketing field.

One of the promising possibilities for marketing ontologies is their use in guiding the process of knowledge extraction in DBM projects. A tool that gradually accumulates knowledge of the previous domain-developed processes is appropriate because of its iterative nature. Researchers often rework their data in order to optimise further interactions.[15] Integrating this knowledge with ontology extends the usefulness of ontology.

Therefore, the purpose of this work (Figure 1) is to focus on DBM as the intersection of two other disciplines (knowledge-extraction techniques and marketing). It intends to capture main DBM concepts through knowledge discovery in databases and relationship

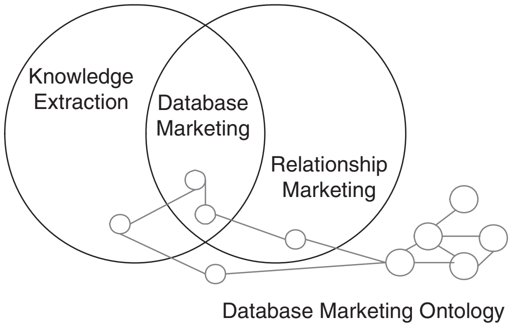


Figure 1: Database marketing ontology context.

marketing. The DBM Ontology (DBMO) should cover a semantic description of the supporting DBM process, comprising classified marketing objectives and activities, knowledge-extraction methods, objectives, and tasks.

The impact of this research is the future initiation of a shared DBM knowledge platform that will provide a trusted base among marketers, DBM practitioners and artificial intelligence researchers. Moreover, the ontology is intended to become the basis for future core ontology in the domain of DBM community.

This paper unfolds in the following manner: we start with the ontologies basis and knowledge issues in the marketing discipline, and then we outline the research approach. Research questions and research findings are presented in two subsequent sections. The results discussion is presented in the Discussion section, followed by conclusions and areas for potential further research.

ONTOLOGIES

Currently, ontologies are one of the most popular knowledge representation (KR) techniques. They have been proposed since the eighteenth century, and have been developed and deployed for sharable and reusable models. These ontologies aim to allow information modelling and knowledge management and reuse.

Ontology definition

Ontology is a description of conceptual knowledge organised in a computer-based representation.[16] In artificial intelligence literature, the most commonly quoted definition of ontology is *a formal, explicit specification of a shared conceptualization*.[1] A *conceptualization* refers to an abstract model of one factor that describes the semantics of the data. An *explicit specification* means that the concepts and relationships in the abstract model are given explicit names (terms) and definitions (specification of the meaning of the concept or relation) that can be communicated among people and across application systems. *Formal* refers to how the meaning specification is encoded in a language the formal properties of which are well understood – in practice, this usually means logic-based languages that have emerged from the KR community within the field of artificial intelligence. *Shared* means that the main purpose of ontology is generally to be used and reused across different applications and communities.

At a higher level, ontology specifies the classes of concepts that are relevant to the application domain and the relations that exist between these classes. Ontology captures the intrinsic conceptual structure of a domain. For any given domain, its ontology forms the heart of the KR. Here, we very briefly describe what entities are found in an ontology language.

- *Classes* or *concepts* are the main entities of ontology. They are interpreted as a set of individuals in the domain, for example, *data* or *algorithms*. It is possible to assign sub-classes to each class, like *datasource* or *datavaluetype* for the class *data*;
- *Instances* or *objects* are interpreted as a particular individual of a domain, for example, *age* is an instance of the sub-class *demographics*;
- *Relations* are the ideal notion of a relation independently to why it applies (for

example, the name relation in itself), they are interpreted as a subset of the products of the domain;

- *Properties* are the relations precisely applied to a class (for example, the gender of an individual); property instances are the relations applied to precise objects (the name of this individual);
- *Datatypes* are a particular part of the domain that specifies values (as opposed to individuals); values do not have identities.

Ontologies use a formal domain or KR agreed on consensus and shared by the entire community. There are several ways to represent such ontologies, and many languages have been defined to represent them. There is a wide range of languages that goes from first-order logic (for example, Ontology Web Language (OWL) or Resource Description Framework (RDF)) to frame-based languages implemented in ontology management systems (for example, Protégé or Ontolingua).

KNOWLEDGE ISSUES

Knowledge management involves the representation, organisation, acquisition, creation, use and evolution of knowledge in its many forms. In order to build effective technologies for knowledge management, we need to further our understanding of how individuals, groups and organisations use knowledge.[17,18] Currently, more and more knowledge is represented in computer-readable forms, stressing the need to build tools that can effectively search databases, files and websites to extract information, capture its meaning, organise and analyse it, and make it useful.

Ontologies are becoming more and more abundant in KR and management. Ontologies model the structure of data (classes and their properties or attributes), the semantics of data (in the form of axioms that express constraints such as inheritance relationships, or constraints on properties) and data instances (individuals). To integrate ontologies, we must understand the relationship between structures (classes and properties) and data (individuals) from different ontologies. Furthermore, we must be able to use the semantics of ontology to model these relationships, and create a coherent and consistent integrated ontology.[19]

Knowledge representation and ontologies

Knowledge representation has long been considered one of the principal elements of artificial intelligence, and a critical part of all problem-solving.[20] The sub-fields of KR range from the purely philosophical aspects of epistemology to the more practical problems of handling huge amounts of data.[21] This diversity is unified by the central problem of encoding human knowledge – in all its various forms – in such a way that the knowledge can be used.

Knowledge representation must unambiguously represent any interpretation of a sentence (logical adequacy), have a method for translating from natural language to that representation and must be reusable.

The central tenet of KR systems is a notation based on the specification of objects (concepts) and their relationships to each other. The main features of such a language are:[22]

(i) *Object-orientedness*. All the information about a specific concept is stored with that concept, as opposed, for example, to rule-based systems where information about one concept may be scattered throughout the rule base.

(ii) *Generalisation/Specialisation*. Long recognised as a key aspect of human cognition, KR provides a natural way to group concepts in hierarchies in which higher-level concepts represent more general, shared attributes of the concepts below.

(iii) *Reasoning*. The ability to state in a formal way that the existence of a piece of knowledge implies the existence of one other previously unknown piece of knowledge is important to KR.

(iv) *Classification*. Given an abstract description of a concept, most KR languages provide the ability to determine whether or not a concept fits that description. This is actually a common special form of reasoning.

KR systems have some limitations when dealing with procedural knowledge. An example of procedural knowledge[23] would be Newton's Law of Gravity – *the attraction between two masses is inversely proportional to the square of their distances from each other*. Given two bodies, with slots holding their positions and masses, the value of the gravitational attraction between them cannot be inferred declaratively using the standard reasoning mechanisms available in KR languages. A function or procedure in a programming language could, however, represent the mechanism for performing this 'inference' quite well. Ontologies can deal with this kind of knowledge by adding a procedural language to its representation. Therefore, the knowledge is not being represented in a declarative way; it is being represented as C or LISP (computer programming languages) code, which is accessed through a slot. This is an important distinction – there is knowledge being encoded in these computer programming functions that is not fully accessible. The system can reason with this knowledge, but not about it – here is the ontological role.

Ontologies are a key part of a broader range of semantics-based technologies that include the areas of KR and automated inference that arose within the artificial intelligence community.[2,24] Many different representation formalisms have been explored, and reasoning engines developed. In strict sense, ontologies may be considered a sub-area within KR,[25] as almost every knowledge base frequently has ontology as its main backbone. Ontologies capture the intrinsic conceptual structure of a domain.

The focus on knowledge-sharing and reuse constitute the major difference between ontologies and KR in general. Moreover, ontologies go beyond KR limits, as they are designed to allow reasoning activities.

ONTOLOGIES IN THE CONTEXT OF DATABASE MARKETING

This research on marketing ontologies is part of a larger project that deals with the extraction of marketing knowledge from large and heterogeneous marketing databases. Thus, we need a tool for KR, reasoning and decision support.

Here, ontologies' role in DBM has particular significance, as they focus on a crossover of areas. That is, to develop DBM, both marketing and extraction techniques knowledge is needed. Thus, ontologies can play an important role in describing in a semantic form all concepts and techniques around the process. Moreover, with such a description it will also be possible, in a second phase, to introduce metrics in order to compare and therefore select and suggest the best approaches and methods in the context of a new project.

Ontologies should provide consensual knowledge about a certain domain or area interchangeable by the community. Such ontologies would allow common applications to be developed because of their compatible formats. In this study, we are not designing a global marketing ontology representing all the varied aspects of the marketing domain. We are proposing domain ontology as an integral part of a global marketing system. Our ontological proposal deals with some marketing knowledge and extraction process methods and tasks necessary to the DBM process and thereafter for marketing ontology. According to some researchers, this ontology is classified as application ontology [26] serving our main global project. As such, we focus only the study of DBM-related concepts.

Ontologies are also like conceptual schemata in database systems. A conceptual schema provides a logical description of shared data, allowing application programmes and databases to interoperate without having to share data structures. While a conceptual schema defines relations on data, ontology defines terms with which to represent knowledge.[13] For present purposes, one can think of data as that expressible in ground atomic facts, and knowledge as that expressible in logical sentences with existentially and universally quantified variables. Ontology defines the vocabulary used to compose complex expressions, such as those used to describe resource constraints in planning problems. From a finite, well-defined vocabulary one can compose a large number of coherent sentences. This is one reason why vocabulary, rather than form, is the focus of specifications of ontological commitments.

In computer science, ontologies have appeared in a variety of forms, ranging from lexicons to dictionaries and thesauri, or even first-order logical theories. Lexicons provide a standardised dictionary of terms for use during indexing or retrieval. Dictionaries can be organised according to specific relations to form hierarchies (taxonomies, meronomies and so on). Thesauri add related terms to any given term. In DBM, as in any of these forms, ontologies are useful because they encourage standardisation of the terms used to represent knowledge about a domain. When ontologies are formalised in first-order logic, they can also support inference mechanisms.[18] For a given collection of facts, these mechanisms can be used to derive new facts or check for consistency. Such computational aids are clearly useful for knowledge management, especially when one is dealing with complex problems or handling large amounts of knowledge.

The essence of our marketing ontology is a collection of DBM process relationships between marketing row data and marketing knowledge. The basic facts we need to represent are of the form that a specific classification of marketing activities, data used and knowledge extraction process techniques adopted. As example, ‘a married man with children buys beer and diapers during world football cup’. The challenge is to find a representation of this kind of knowledge in a convenient and economical way that fits into our DBMO framework.

RESEARCH APPROACH

Because of the nature of the research, we split the project into two sections, adopting different research approaches for each one: Delphi and Action Research approaches.

First, we focus our work on marketing knowledge, in order to structure its main concepts and systematise the overall organisation, namely marketing objectives and activities and main marketing database data types. Second, in order to design and improve overall DBM perspective, we focus the process on the semantic description of used procedures and methods in order to systematise the knowledge extraction process and knowledge base creation.

Our work ends by pointing to a possible framework that will lead future DBM projects supported by ontologies.

Delphi methodology

The Delphi method is normally used to structure a group communication process to deal with and to build consensus about a particular and complex topic. The method works based on an expert panel group (anonymous experts – no expert knows who else is on the panel), who answer proposed questions and formulate a set of hypotheses about them.[27,28] The method is then developed on the dialectical inquiry approach: the researcher introduces a set of questions in order to establish an opinion or view from the expert panel. The expert panel (individually) then answers, reporting a formulation (conflicting opinion or view). The researcher in charge has to generate a synthesis (a new agreement or consensus) and submit it again to the expert panel. This loop only ends when the researcher achieves a consensus with all expert panel members (Figure 2).

Currently, Delphi is considered a useful method for eliciting and aggregating expert opinion whenever there is a lack of viable or practical statistical techniques. It can be defined as a medium-term qualitative forecasting method that is based on building a consensus among a group of experts.[29] A Delphi-type study enables an exchange of information among experts over a number of rounds (iterations), and allows experts to react to the information gathered during each round, and to fine-tune their forecast by means of a feedback mechanism (controlled retroaction). Beyond these three main principles (anonymity – iteration – retroaction), the method's validity is first based on a rigorous selection of experts whose combined knowledge and expertise must reflect the full scope of the problem area.

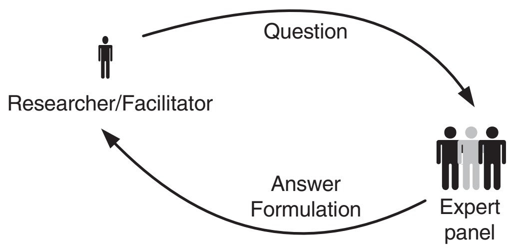


Figure 2: Delphi methodology process.

Some authors have suggested asking the persons involved to estimate their own degree of expertise, with others considering that the level of expertise does not necessarily need to be high.[30] Delphi's validity is also dependent on the size of the group of experts[31] (research suggests that the minimum threshold is 5–7 experts, and that a range of 8–10 offers the best precision/cost ratio. Beyond 12 experts, information contributions are marginal). The method's validity relies on a strict implementation of the process: three iterations are usually needed to obtain a satisfactory consensus.[29]

Action – Research methodology

This research study continued through the DBM process. We adopted an Action Research methodology for DBM development and deployment. This methodology incorporates the four-step process of planning, acting, observing and reflecting on results from a particular project or body of work.[32,33] The concept essentially involves a group of people who work together to improve their work processes.[34,35]

This choice of Action Research was based on two factors. First, because of the low number of scientific studies that have been conducted on supporting DBM process over intelligent structures like ontologies, the process through which this may be completed was unclear. Second, ontologies can play an important role in the knowledge development as long as they register past knowledge for future reuse (Figure 3). Thus, exploratory research was required, and Action Research provides this capability better than many other alternatives.[36]

Nevertheless, first we need to formulate, test, deploy and evaluate a complete DBM process interaction, and then annotate it in a semantic language like RDF, OWL or even in Semantic Web Language Rule (SWRL).

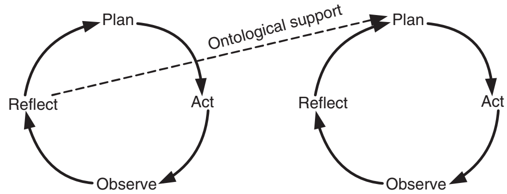


Figure 3: Action research methodology process adopted.

Because of its ontological characteristic, this stage of the project turned out to be emancipatory Action Research,[37] rather than merely technical or practical.[38] The relationship between the team leader and other participants (marketers from participant organisations) was collaborative. The implementation of this integrated system within the Action Research organisation transformed some business models from an ineffectual relationship to developing a more effective relationship with individual customers based on its individual profile: differentiating, interacting, personalising and also learning from each interaction between customer and organisation.

RESEARCH QUESTIONS

The framework for this project was conceived from different research area literature reviews: relationship marketing,[39,40] DBM,[5,41,42] ontologies[9,10,13,43] and knowledge discovery in databases.[44–47]

To define the expert panel, we focused on the individual's reputation and recognition in academic and business circles. To avoid any type of collusion or friendship side effects, we did not ask for experts' names, but devised a questionnaire on practitioners and researchers. We then sent the questionnaire to each one of them. We aimed to ascertain their opinion from the answers to the questions.

Based on Delphi methodology we've defined a marketing knowledge tree focusing main marketing objectives, action programmes and related activities.

According to these first-stage objectives, we proposed the following questions for discussion by our expert panel:

(i) *Regarding the relationship marketing context, what are the main marketing activities that use the DBM approach?*
(ii) *Regarding the relationship marketing context, what are the main DBM objectives?*
(iii) *What is the main type of data used in DBM projects?*

After constructing the marketing knowledge structure tree, we proceeded with Action Research methodology that led us to the answers to the following main concerns:

(i) *Principal marketing database data type information;*
(ii) *Main DBM steps from data to customer knowledge;*
(iii) *Operational DBM matrix aligning knowledge extractions methods and marketing activities and objectives.*

Research cycles from both methodologies, in combination with the reconnaissance of the expert panel (first phase) and professional marketers (second phase), led to the development of the final framework of DBM process, supported by ontologies and knowledge discovery in databases. The proposed framework has the capacity to suggest solutions from previous knowledge registered in the knowledge base.

RESEARCH FINDINGS

As referred to previously, this first-phase work was developed according to Delphi methodology. We sent the questionnaire to each expert. From each one of them we aimed to ascertain his/her opinion from his/her answer to the question. This interaction took place over five cycles. That is, there were four iterations before we considered (common agreement about the subject) all the answers to the proposed questions to be stable.

Our findings at this stage are summarised in Table 1.

Following the previous Delphi methodology research, which has given us a marketing knowledge concepts structure tree, we proceed with Action Research methodology. At this point, we aim to address the following main concerns:

(i) Principal marketing database data type information;
(ii) Main DBM steps from data to customer knowledge;
(iii) DBM matrix: marketing activities objectives, knowledge discovery type models and marketing data type connection.

We have developed Action Research at two simultaneous theoretical and practical levels, and therefore two working focus groups:

(i) practice over a real relationship marketing programme database,
(ii) literature-oriented field research (an expert panel had explored scientific literature and achieved a set of possible tracks to each of the research focuses).

Data description was collected through both focus groups. Because the research phenomenon is contemporary and no prior research has been conducted or was known at the time this paper was written, both of them had collected DBM process descriptions, as well as knowledge discovery approaches. These focus groups were interesting because they generated insights [35] from both DBM practice and process knowledge (namely at data preparation and pre-processing levels).

In combination with the focus groups, convergent work was used to further test and refine the aimed theoretical framework. Convergent work involves, for example, transposing from reviewed literature approaches or suggestions for the practical domain. Each interaction was then registered in terms of type of data, data analysis algorithm used and results achieved with them. Convergent work also involves conducting a series of in-depth working groups in order to explore other insights that were not previously registered. That is, the process is very structured and ends only when no new information remains uncovered or unregistered (Table 2).

After ending the Action Research, a practical and functional analysis was made towards a possible conceptual semantic map. Turning to analytic generalisation, we can then build a theoretical framework [49] linked to extant literature that shows how the DBM process is developed, how associated marketing knowledge can be structured and which knowledge discovery approaches may be used. Our research allows the idenfication of three main components of the DBM process (Figure 4): inputs (marketing objectives, marketing activities and marketing data), tasks (data handling and data modelling) and outputs (evaluation, deployment and business value).

The ontological commitment is a form of matrix evaluation, whereas data loaded, tasks and methods taken and results obtained are evaluated and registered in a knowledge base.

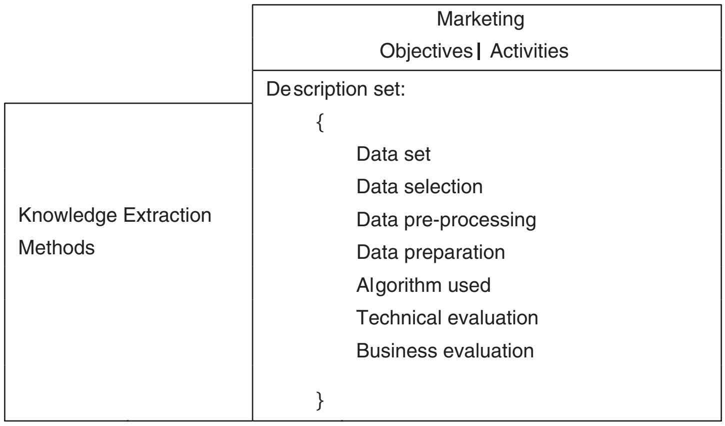
Knowledge base:
$\{\text{Results} = \{\text{DBM}_i\text{models}[\{\text{input}\}\{\text{tasks}\}]\}\}$

In order to test and verify the knowledge consistency and therefore the knowledge structure, we have collected a large amount of relationship marketing data from a multinational distribution company. Our database contains, at an individual level different kinds of marketing information, such as demographics, psychographics, lifestyle and transactional information. Moreover, some external data are presented as example market or as financial information.

Table 1: Delphi method findings

Research issue	*Findings about the research issues*
Regarding the relationship marketing context, what are the main marketing activities that use the DBM approach?	Same marketing activities may be developed under different marketing disciplines, for example, customer identification, can be developed both in relationship marketing programme as well as in direct marketing. That is, there is a non-exclusive set of possible marketing activities available where DBM projects took place. Aligning with relationship marketing objectives we have organised as follows[48]: *To identify* • Customer knowledge or identification • Customer needs • Customer wants *To differentiate* • Customer segmentation • Customer categorisation • Customer profiling *To interact* • Cross and up-selling • Cross marketing • One-to-one marketing • Customer reactivation *To customise* • Customer loyalty acquisition • Customer fidelisation • Customer affiliation
Regarding the relationship marketing context, what are the main DBM objectives?	DBM process is aligned with the marketing activity that holds its context. Therefore among the proposed DBM objectives we have organised the following as main objectives: • Segmentation • Classification or clustering • Market basket analysis • Prediction future behaviour • Description • Churn • Reactivation
What is the main type of data used in DBM projects?	Both literature and expert panel suggest that the information gathered in marketing databases is mainly organised or well defined as the following data types (some examples of each one are presented): Psychographics: Personal data that can easily be changed. • Monthly income • Professional occupation • Scholarship Demographics: Physical and personal data that is almost definitive and almost never changes. • Gender • Marital status • Birth date • Children • Race Transactional: Consumer based information regarding its commercial activity • Monthly consumption • Number transactions/month • Number items/month • Shops visited • Promotional acceptance Lifestyle or behaviour: Consumer or social related information. • Hobbies • Car type • Holidays • Club membership In addition to the above customer-oriented data types there are two other groups of data: Market data: Environmental market data • Financial (for example, inflation tax rate) • Market (for example, market or product share) • Social (for example, national birth, death or other census) Trigger events data: • Consumer (for example, married status change or children number) • Life related (for example, new car or new house) • Others (for example, accident, prison, tax penalties)

Table 2: Action research findings

Research issue	*Findings about the research issues*
Principal marketing database data type information	From some literature review and supported by previous Delphi methodology • Psychographics • Demographics • Lifestyle • Transactions
Main DBM steps:	Based on both practice and literature review we have considered the following steps as a stable DBM process framework: • Marketing objectives definition • Data selection • Data preparation • Data pre-processing • Modelling • Deployment • Evaluation
DBM matrix	

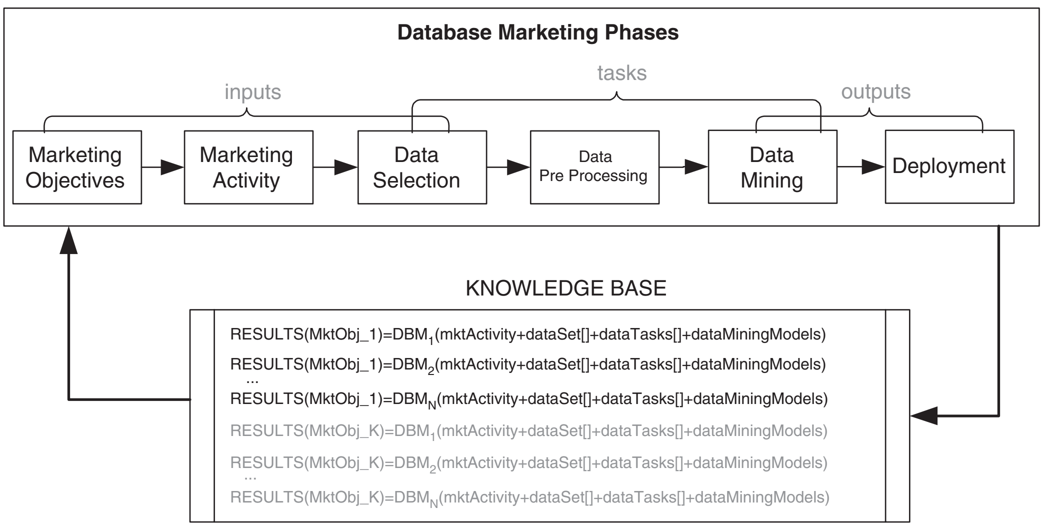


Figure 4: DBMO architecture.

Table 3: DBM process example

Case 1	
Marketing objective: customer profile	
Data: Personal-psychographics	
	birthDate
	gender
	children
	incomePerCapita
Personal-Demographics	
	maritalStatus
	houseHoldDimension
Personal-Transaction	
	customer id;
	productConsumption_1
	productConsumption_2
	…
	productConsumption_128
	supermarketMonthlyConsumption
Individuals:	613 000
Cleaned records:	64 000
Data preparation tasks used:	missing values;
	duplicationSelector;
	unitDeviations;
	outliers.
Data transform tasks:	matrizTranspose;
	discretization.
Data Mining Method:	Classification
Algorithms:	SOM
	C 5.0
Evaluation	pccConfusionMatrix

Table 4: Knowledge base table record example

```
{
marketing objectives;
marketing activity;
data used [{demographics}, {psychographics},
{life style},{ transactional}];
data quality[{outliers},{missing values}, … ]
data procedures [{selection},{preparation},
{pre-processing}]
algorithms used [{clusterers}, classifiers},
neuralNetworks}, geneticAlgorithms}, statistical} … ]
evaluation method [{auc}, {pcc} … ]
}
```

We have processed the data using WEKA,[50] free data mining software, and we have found different results according to different data and algorithms used (Table 3). Therefore, we extracted information and organised it according to an individual perspective.

All information regarding each developed DBM project has been registered in a knowledge base table that has information as in Table 4.

The classification of the degree of success of a DBM project is very subjective. Nevertheless, according to our approach, we can perform, register and implement some analytical procedures that will lead to some DBM evaluation. Within this research, we assume that data mining evaluation models like area under curve (AUC), confusion matrix or principal components analysis are used. For each model, we also evaluate which kind of data were used, and related quality in terms of completeness, outliers and missing values. Regarding each data set used, we have registered all data tasks performed, like data cleaning, data transformation or data reduction. Related to the modelling phase, a table was created in order to register not only which algorithms were performed, but also which data from loaded data sets were used.

The model deployment is performed based on two factors: (i) analytical deployment: focusing the algorithm performance; (ii) business perspective: regarding its practical application, that is, there are models with high accuracy but with low interest (for example, a rule like all women buy female products), and others with low rating but with high impact regarding business value (for example, customers aged under 50 years, two children, married, high-level occupation have a 50 per cent probability of buying your product).

DISCUSSION

One of the promising interests of DBMO is its use for guiding the knowledge extraction process from marketing databases. This idea seems to be much more realistic now that semantic web advances have given rise to common standards and technologies for expressing and sharing ontologies.[51,52]

In this way, DBM can take advantage of domain knowledge embedded in DBMO:

(i) based on marketing activity definition, ontology can indicate a global perspective according the available resources (for example, data quality or completeness)

(ii) from a DBM objectives point of view, ontology may suggest or select the most appropriate approaches to treat the available data;

(iii) during the data-preparation step, DBMO can facilitate the integration of heterogeneous data, and guide the selection of relevant data to be used;

(iv) in the modelling phase (for example, data mining), domain knowledge allows specification of constraints to guide data mining algorithms by narrowing search space;

(v) during the interpretation step, domain knowledge helps experts to visualise and validate extracted units.

Therefore, using a general framework, it is possible to illustrate a general perspective of how the system works (Figure 5). We have considered a three-layer architectural approach:

— The physical layer holds the process development tasks, namely data handling (selection, preparation, pre-processing and transformation) and modelling.

— The ontological layer acts like a guide to the data analyst and as a reference to the marketer expert.

— The presentation or user layer plays the interaction role among above layers and users.

DBMO divides the DBM process into four main phases: marketing activity objectives, knowledge extraction, evaluation and business decision.

With this research, we suggest some general roles for the ontology in each DBM phase:

— *Marketing activity definition*: The role of ontologies in business understanding is not specific to the marketing discipline. Domain ontologies are an important vehicle to inspect a domain before committing to a particular task. Semi-formal ontologies can help a newcomer to become familiar with most important concepts and relationships, whereas formal ontologies allow the identification of conflicting assumptions that might not be obvious at first sight.

— *Knowledge extraction*: For improved data exploration, elements of ontology have to be (presumably manually) mapped onto elements of the data scheme and

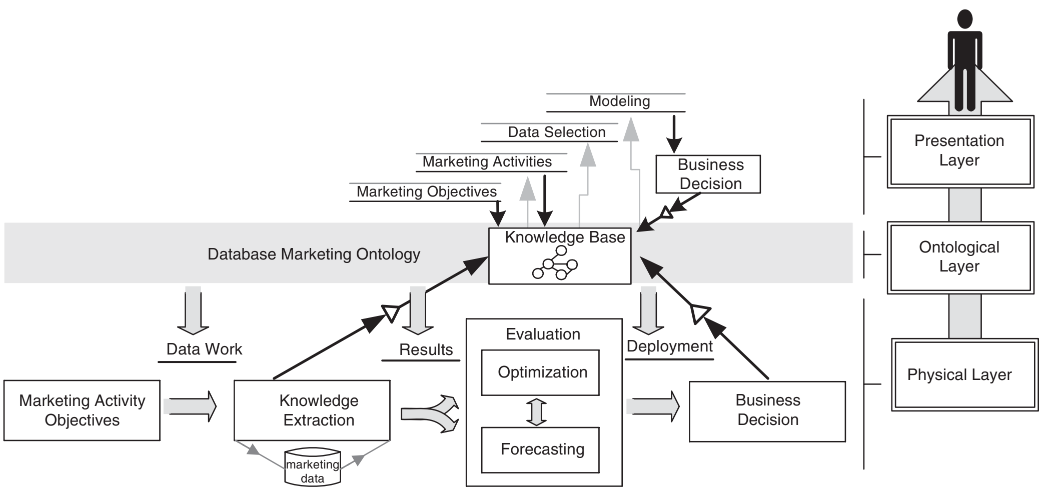


Figure 5: DBMO general framework.

vice versa. This will typically lead to selecting a relevant part of ontology (or multiple ontologies) only. Another relevant issue is the connection between the Data Preparation phase and the subsequent Modelling phase. Concrete use of domain ontology depends partially on the chosen mining tool/s. Ontology may characteristically help by identifying multiple groups of attributes and/or values according to semantic criteria. In the Modelling phase, ontologies might help to design the individual mining sessions. In particular, for large data sets, it might be worthwhile to introduce some ontological bias, for example, to skip the quantitative examination of hypotheses that would not make sense from the ontological point of view, or, on the other hand, of two obvious hypotheses.

— *Evaluation phase*: The discovered model(s) has the character of structured knowledge built around the concepts (previously mapped on data attributes), and can be interpreted in terms of ontology and associated background knowledge.
— *In the Business Decision phase*: Extracted knowledge is fed back to the business environment. Provided that we previously modelled the business using ontological means, the integration of new knowledge can again be mediated by the business ontology. Furthermore, if the mining results are to be distributed across multiple organisations (say, using the semantic web infrastructure), mapping onto a shared ontology is inevitable.

CONCLUSIONS

The extent, degree and speed of communication enabled by the ontology makes it a synergistic component of DBM strategy. Our proposed DBMO, an ontological DBM approach solution, appears promising for both marketers and computer scientists.

The results of this research have implications for both theory and practice. Related to practice, the very first implication relates to the possible feedback among different DBM projects depicted in a table with all used resources registered. This enables the construction of a knowledge base containing suggestions or work profile capability. According to the previous registered experiments, the knowledge base will be capable of indicating for each marketing objective which marketing activities, data and tasks should be carried out.

Another implication relates to the benefits of a global view of marketing databases' roles in marketing objectives. There is only one way to have a successful DBM project: it must have appropriate data type and quality.

The research findings and contributions have several implications for the theory about ontologies and DBM, as well as for the integration of research methodologies such as Delphi and Action Research. This research provides new insights into DBM theory in two ways:

(i) It appears to provide the first global investigation about the intersection of ontologies and DBM in organisations, and how it may be achieved. This research contributes to the theory-deficient area of the integration of ontologies and DBM, providing the first approach to a theoretical framework for such a phenomenon.
(ii) There is little literature dedicated to marketing ontologies, and thus this research appears to be the first academic investigation of this phenomenon.

The DBMO model further emphasises the importance of the marketing knowledge being structured in order to allow resource reuse or even to achieve synergies in marketing activities development. Thus, managers and marketers should be aware of this issue, because there is a loop through which performance of DBM process can effectively be improved.

This research showed that the most important output of the ontological approach is an enabling of effective DBM assistance without in-depth expertise in data mining tools. Supported by the knowledge base, ontology is capable of suggesting the pathway from data to desired knowledge.

REFERENCES

1 Gruber, T. R. (1993) A translation approach to portable ontology specifications. *Knowledge Acquisition* 5: 199–220. http://tomgruber.org/writing/ontolingua-kaj-1993.pdf.

2 Jasper, R. and Uschold, M. (1999) A framework for understanding and classifying ontology applications. In: V. R. Benjamins, B. Chandrasekaran, A. Gomez-Perez, N. Guarino and M. Uschold (eds.) IJCAI 99 Ontology Workshop, pp. 16–21, http://citeseerx.ist.psu.edu/viewdoc/summary? doi=10.1.1.30.1927.

3 Zhou, L. (2007) Ontology learning: State of the art and open issues. *Information Technology and Management* 8: 241–252.

4 Payne, A. and Frow, P. (2005) A strategic framework for customer relationship management. *Journal of Marketing* 69(4): 167–176.

5 Ozimek, J. (2004) Case studies: The 2003 information management project awards. *Journal of Database Marketing & Customer Strategy Management* 12(1): 55.

6 Kamakura, W. A., Wedel, M., de Rosa, F. and Mazzon, J. A. (2003) Cross-selling through database marketing: A mixed data factor analyzer for data augmentation and prediction. *International Journal of Research in Marketing* 20: 45–65.

7 Watada, J. and Yamashiro, K. (2006) A data mining approach to consumer behavior. In: Proceedings of the First International Conference on Innovative Computing, Information and Control (ICICIC'06), http://ieeexplore.ieee.org/iel5/11141/35657/01692071.pdf?.

8 Jans, M., Lybaert, N. and Vanhoof, K. (2008) Data mining as a methodology for internal fraud risk reduction. *Journal of Information Systems*. 32: 1–29.

9 Dabholkar, P. A. and Neeley, S. M. (1998) Managing interdependency: A taxonomy for business-to-business relationships. *Journal of Business and Industrial Marketing* 13: 439–460.

10 Grassl, W. (1999) The reality of brands: Towards an ontology of marketing. *American Journal of Economics and Sociology* 58(2): 313–319.

11 Bouquet, P., Dona, A., Serafini, L. and Zanobini, S. (2002) ConTeXtualized local ontology specification via CTXML. In: P. Bouquet, F. Harmelen, F. Giunchiglia, D. McGuinness and M. Warglien (eds.) for Artificial Intelligence AA. MeaN-02 AAAI Workshop on Meaning Negotiation. Edmonton, Alberta, Canada: AAAI, http://sra.itc.it/people/serafini/distribution/aaai-ws-ctxml.pdf.

12 Zairate, L. E., Nogueira, B. M., Santos, T. R. A. and Song, M. A. J. (2006) Techniques for missing value recovering in imbalanced databases: Application in a marketing database with massive missing data. In: IEEE International Conference on Systems, Man, and Cybernetics. Taiwan: IEEE, pp. 2658–2664.

13 Zhou, X., Geller, J. and Halper, Y. P. M. (2006) Theoretical computer science. In: S. B. Heidelberg (ed.) *An Application Intersection Marketing Ontology. Lecture Notes in Computer Science*. Berlin/Heidelberg, Germany: Springer, pp. 143–163.

14 Nogueira, B. M., Santos, T. R. A. and Zarate, L. E. (2007) Comparison of classifiers efficiency on missing values recovering: Application in a marketing database with massive missing data. In: Proceedings of the 2007 IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007).

15 Siqueira, S. W., Silva, D., Uchoa, E., Braz, H. and Melo, R. (2001) An architecture for database marketing systems. *Lecture Notes in Computer Science*: 2113.

16 Nédellec, C. and Nazarenko, A. (2005) Ontologies and information extraction. In: P. Buitelaar, P. Cimiano and B. Magnini (eds.) LIPN Internal Report. LIPN, http://arxiv.org/ftp/cs/papers/0609/0609137.pdf.

17 Jurisica, I., Mylopoulos, J. and Yu, E. (1999) *Ontologies for Knowledge Management: An Information Systems Perspective*, In: for Information Sciences AS. Proceedings of the Annual Conference of the American Society for Information Sciences (ASIS'99). American Society for Information Sciences, ftp://ftp.cs.toronto.edu/pub/eric/KAIS-IJ030114.pdf.

18 Mylopoulos, J., Jurisica, I. and Yu, E. (2004) Ontologies for knowledge management. *Knowledge and Information Systems* 3: 380–401.

19 Udrea, O., Getoor, L. and Miller, R. J. (2007) *Leveraging Data and Structure in Ontology Integration*, In: C. Y. Chan, B. C. Ooi and A. Zhou (eds.) 2007 ACM SIGMOD international conference on Management of data P. International Conference on Management of Data; pp. 449–460.

20 Newell, A. (1982) The knowledge level. *Artificial Intelligence* 18: 87–127, citeseer.ist.psu.edu/newell02artificial.html.

21 Guarino, N. (1995) Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human and Computer Studies* 43: 625–640.

22 Welty, C. and Murdock, J. W. (2006) Towards knowledge acquisition from information extraction. In: I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold and L. Aroyo (eds.) Springer. Proceedings of ISWC-2006. Athens: The Semantic Web – ISWC 2006.

23 Welty, C. A. (1996) An integrated representation for software development and discovery. Rensselaer Polytechnic Institute, http://www.cs.vassar.edu/weltyc/papers/phd/HTML/dissertation-1.html.

24 Uschold, M. and King, M. (1995) Towards a methodology for building ontologies. AIAI-TR Workshop on Basic Ontological Issues in Knowledge Sharing, 1. http://www.aiai.ed.ac.uk/project/pub/documents/1995/92-ont-ijcai95-ont-ijcai95-ont-method.ps, http://citeseer.ist.psu.edu/uschold95toward.html.

25 Uschold, M. and Gruninger, M. (2004) Ontologies and semantics for seamless connectivity. *SIGMOD Rec* 33(4): 58–64.

26 Sowa, J. F. (2000) *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Pacific Grove, CA: Brooks Cole Publishing, http://www.jfsowa.com/krbook/index.htm.

27 Linstone, H. A. and Turoff, M. (eds.) (2002) *The Delphi Method – Techniques and Applications*. Newark: New Jersey Institute of Technology.

28 Chu, H. C. and Hwang, G. J. (2008) A Delphi-based approach to developing expert systems with the cooperation of multiple experts. *Expert Systems with Applications* 34: 2826–2840.

29 Armstrong, J. S. (2006) Findings from evidence-based forecasting: Methods for reducing forecast error. *International Journal of Forecasting* 22(3): 583–598.

30 Rowe, G. and Wright, G. (2001) *Expert Opinions in Forecasting: The Role of the Delphi Technique. Principles of Forecasting*. Boston: Kluwer Academic Publishers, pp. 125–144.

31 Vernette, E. (1997) Evaluation de la validation predictive de la methode Delphi-leader. In: Congress International de l AFM; pp. 988–1010.

32 Zuber-Skerritt, O. (ed.) (2000) *New Directions in Action Research*. London: Falmer Press.

33 O'Brien, R. (2002) An overview of the methodological approach of action research. In: R. Richardson (ed.) *Theory and Practice of Action Research*. Toronto, Canada: Universidade Federal da Paraiba, http://www.web.ca/robrien/papers/arfinal.html.

34 Baskerville, R. L. (1999) Investigating information systems with action research. *Communications of AIS* 2: 19–51, Article 19.

35 Carson, D., Guilmor, A., Perry, C. and Gronhaug, K. (2001) *Qualitative Marketing Research*. London: Sage.

36 Dick, B. (2008) Postgraduate programs using action research. In: *Action Learning, Action Research and Process Management: Theory, Practice, Praxis*. Faculty of Education, Toowong, Australia: Griffith University, http://www.scu.edu.au/schools/gcm/ar/arp/ppar.html.

37 Leary, C. O., Rao, S. and Perry, C. (2004) Improving customer relationship management through database/internet marketing. *European Journal of Marketing* 38(3/4): 338–354.

38 Zuber-Skerrit, P. C. (2000) Action research in graduate management theses. *Action Learning, Action Research and Process Management: Theory, Practice, Praxis* 1: 84.

39 Coviello, N. and Brodie, J. (1998) From transaction to relationship marketing: An investigation of managerial perceptions and practices. *Journal of Strategic Marketing* 6(3): 171–186.

40 Coviello, N., Winklhofer, H. and Hamilton, K. (2006) Marketing practices and performance of small service firms – An examination in the tourism accommodation sector. *Journal of Service Research* 9(1): 38–58.

41 Coviello, N., Milley, R. and Marcolin, B. (2004) Understanding IT-enabled interactivity in contemporary marketing. *Journal of Interactive Marketing* 15(4): 18–33.

42 Wehmeyer, K. (2005) Aligning IT and marketing – The impact of database marketing and CRM. *Journal of Database Marketing & Customer Strategy Management* 12(2): 243.

43 Diamantini, C., Potena, D. and Smari, W. (2006) Collaborative knowledge discovery in databases: A knowledge exchange perspective. In: W. W. Smari and B. McQuay (eds.) Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition. AAAI, Arlington, VA: IEEE Computer Society, pp. 24–31, http://boole.diiga.univpm.it/paper/2.pdf.

44 Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) From data mining to knowledge discovery in databases. In: *AI Magazine*, Vol. 17 Waltham, MA: American Association for Artificial Intelligence, pp. 37–54, citeseer.ist.psu.edu/fayyad96from.html.

45 Phillips, J. and Buchanan, B. G. (2001) Ontology-guided knowledge discovery in databases. In: ACM International Conference on Knowledge Capture 1st International Conference on Knowledge Capture, pp. 123–130.

46 Buckinx, W. and den Poel, D. V. (2005) Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research* 164(1): 252–268.

47 Buckinx, W., Verstraeten, G. and den Poel, D. V. (2007) Predicting customer loyalty using the internal transactional database. *Expert Systems with Applications* 32: 125–134.

48 Peppers, D. and Rogers, M. (1999) *One To One Future: Building Relationships One Customer at a Time*. New York: Doubleday.

49 Yin, R. K. (2003) *Applications of Case Study Research. Vol. 34 of Applied Social Research Methods Series*, 2nd edn. Thousand Oaks, CA: Sage Publications.

50 Witten, I. H. and Frank, E. (2000) Data mining: Practical machine learning tools and technique, 2nd edn. In: M. R. Jim Gray (ed.) *The Morgan Kaufmann Series in Data Management Systems*. San Francisco, CA.

51 Coulet, A., Smail-Tabbone, M., Benlian, P., Napoli, A. and Devignes, M. D. (2008) Ontology-guided data preparation for discovering genotype-phenotype relationships. *Journal of BMC Bioinformatics* 9: 1–9.

52 Smith, M. K., Welty, C. and McGuinness, D. L. (2008) OWL web ontology language guide, http://www.w3.org/TR/owl-features/.