Semantic Approach for Finding Suitable Commercial Business Location

Felix Mata
Computing Mobile Laboratory
IPN-UPIITA
Mexico City, Mexico
mmatar@ipn.mx

Omar Juarez, Consuelo Garcia, Roberto Zagal
Systems Department
IPN-ESCOM
Mexico City, Mexico
{omarjg82@gmail.com, varinia400@hotmail.com, rzagalf@ipn.mx}

***Abstract*—Nowadays, in order to find the best suitable location for a commercial business, one needs to apply a market analysis in several steps. This research improves this process by applying spatial semantics. We define a methodology to find a list of candidates for the best commercial business location. Some parameters involved in spatial semantics and market analysis are: floating population, taxes cost, market competition, transportation, land use, geographical area, among many others. All of them should be evaluated with different methods. A list of semantic rules are defined to find possible location, ontology is built based on the geographic and market conditions. Each one of places found complies with law restrictions, administrative and geographic requirements defined as input parameters. Additionally, it is possible to relax the parameters' weight in order to obtain new results for possible locations. Proofs include results obtained in conjunction (semantic and market analysis) that shows better results that when these criteria are used independently.**

Keywords-Spatial Semantics; geographic domain conceptualization; axiomatic relation.

I. INTRODUCTION

Nowadays, finding the adequate location for establishing a business, is a task that includes review aspects of market (e.g., business market competition, economy, geographic conditions, etc.). This process is not done by just one person; it is made by the collaborative work of specialists from different disciplines (marketers, economists, geographers). Then, the possibility of automating this task is possible, if we consider that: 1) ontologies can represent the knowledge involved in this context (spatial and market conditions), 2) semantic analysis allows to explore ontologies and define rules based on certain constraints, 3) the processing of spatial semantics is used to find locations, based on semantic rules and geographic conditions. Thus, these three statements are integrated as a methodology to find a list of possible locations for a business. In addition, users can assign higher or lower relevance to each parameter (qualitative or quantitative configuration) in order to generate a new list of possible places for locating a business. For example, suppose the following query:

Q_1= "find a business location in a county without market competition in a radius of 5 km". It can be relaxed qualitatively by the user in parameter market competition. Then, the system generates new queries as follows:
Q_{1a} = "find a business location in a county with low market competition in a radius of 10 km".
Q_{1b}: "find a business location in a county without market competition of first impact in a radius of 5 km".

The case study focuses on three types of business: stationer' shops, drugstores and restaurants. The methodology includes definition of semantic processing algorithms, geographic queries implementation and the design and implementation of an ontology. The research is focused on assisting in making a decision about the possible locations for establishing a business. The implementation was done into a Web system. The rest of paper is organized as follows: Section II describes the related work; Section III explains the methodology definition; Section IV shows the obtained results and, finally, the conclusions and future work are outlined in Section V.

II. RELATED WORK

The use of an ontology as a resource of knowledge is well known from tasks such as extracting the semantics, the meaning, of natural texts to entity recognition (people, places, companies, and prices) [14]. Ontology is also employed as a method for identifying categories, concepts, relations, and rules [5], [6], [8]. In our work, we use this principle to find the best locations for a business. On the other hand, the use of ontology in market domain is not new. Researchers have suggested employing market mechanisms for the allocation of Web Services [16], while Lamparter and Schnizler [15] presented an architecture of an ontology-driven market for trading Semantic Web Services. Auction schema is enriched by a set of components enabling semantics based matching, as well as price-based allocations. Meersman [13] uses the market information into the knowledge system of a company in order to contribute to the process of product innovation, competence development and relevant social interaction. Nevertheless, these previous works do not consider the geographic parameters that are relevant in the market analysis.

Ontology is also used to inform designers of data models and information systems to make them better equipped for handling geographic concepts [1], [3], [7], while Smith and Mark [5] report the results of a series of experiments designed to establish how non-expert subjects conceptualize geospatial phenomena. Subjects were asked to give examples of geographical categories in response to a series of differently phrased elicitations. The results yield an ontology of geographical categories, which consists of a catalog of the prime geospatial concepts and categories. When combined with query languages, domain ontologies favor the design and development of domain-based search engines and their application to different areas such as in transportation [9]. A categorization of GIS (Geographic Information System) tasks and GIR (Geographic Information Retrieval) queries has been suggested in [10], the approach being applied to heterogeneous sources, including multimedia sources. Moreover, Biletskiy and Ranganathan [17] explain an ontology and rule-based framework for the development of business domain applications, which includes semantic processing of externalized business rules and, to some extent, externalization of application logic. The framework also includes a rule learning system to semi-automate the generation of information extraction rules from source documents. Summarizing, the use of semantic processing, and the ontology exploration, have been used in other domains, but not in combination with the geographic aspects in market analysis. Thus, in this paper, we show a methodology to be applied in order to get a useful application for real life.

III. Methodology for Querying and Semantic Processing

The methodology is composed of three phases: 1) Ontology modeling, 2) Ontology building, and 3) Semantic processing. Ontology modeling consists of capturing knowledge regarding market context for a set of specific business and their involved geographical parameters. For that, one is required to get a set of official data sources with: 1a) geographic data of places to analyze, 1b) statistics regarding to economy, soil use, population, among others, 1c) mechanisms used in the market to get the type and behavior of it. These components (1a, 1b, 1c) are transformed into semantic rules and concepts. The ontology building consists of reordering the concepts into a taxonomy and establishing the context for each concept following the methodology described in [2]. Finally, semantic processing is responsible to apply the semantic rules and measure the relevance of query parameters. This means, assign a lesser or greater importance to one parameter such as: the price of rent or the proximity of a communication route. This is achieved using the mechanisms of the market in previous step 1c). Basically, these mechanisms are surveys and quests made in others studies.

III.A Modeling Ontology

The modeling consists of taking the knowledge from one conceptual domain to a logical domain. The knowledge is acquired and abstracted, then is translated to logical model to obtain a design (classes and properties). To identify the domain of ontology, we classified the analysis based on semantics and market analysis for the location of a business in: 1) market analysis, 2) proximity analysis and 3) analysis of buildings infrastructure.

1. Market Analysis. It refers to the analysis of potential customers and demand generators. It is composed of three steps:

1a) Compute the size of local market. It is obtained by processing the total population living in the area, the floating population (people who do not live in the area but work or study in it), and income and expense levels by social stratum.

1b) Analyze the market competition. It consists of calculating the number of businesses in the local market to determine if the market is saturated or free. There may be commercial firms who are sacrificing their profit margin just to have a presence in the area (this is a market strategy).

1c) Detect demand generators: identifying corporate office buildings, schools, hospitals, recreation and leisure, and so on.

2. Proximity analysis. It refers to the analysis of the proximity of a building with relevant points (cultural, historical places, downtown, etc.). We establish minimum and maximum distance from 1 km to 5 km for walking distance, and from 5 km to 10 km driving distance.

2a) Locate the areas where most businesses are concentrated. Identify if in the area there are several centers and commercial corridors (including historical centers).

2b) Identify the areas with high vehicular and pedestrian traffic.

2c) Identify the roads surrounding the area, road type, high or low level of vehicular and pedestrian traffic.

3. Analysis of buildings. It refers to an analysis of the characteristics of the buildings: number of levels, construction material (concrete, wood, etc.), parking included, stairs or elevator, number of doors, windows.

III.B Ontology

The ontology was built based on data obtained from INEGI, an official Mexican organism that produces and manage the geographical data in Mexico [4] and documents from other sources such as [11], [12] and with an approach similar to [14] although it is possible to use Open Street Map for the mapping data, but the tabular data of economics parameters should be acquired from other sources as was described at the beginning of Section III. We identify the relevant concepts in a business by following the procedures to measure the behavior of a market in Mexico, (SNIIM in Spanish) [18].

The concepts and terms are abstracted based on definitions of Royal Academy of the Spanish language. Some examples are:

• Offer: Set of property or goods being offered on the market at one price and at a specific time.
• Demand: Overall amount of purchases of goods and services performed or provided by a community.
• Market: Set of consumers able to purchase a product or service.
• Product: Profit.

Moreover, we identify properties and relationships of each one of the concepts. Then, we obtain the taxonomy and classes shown in Fig. 1.

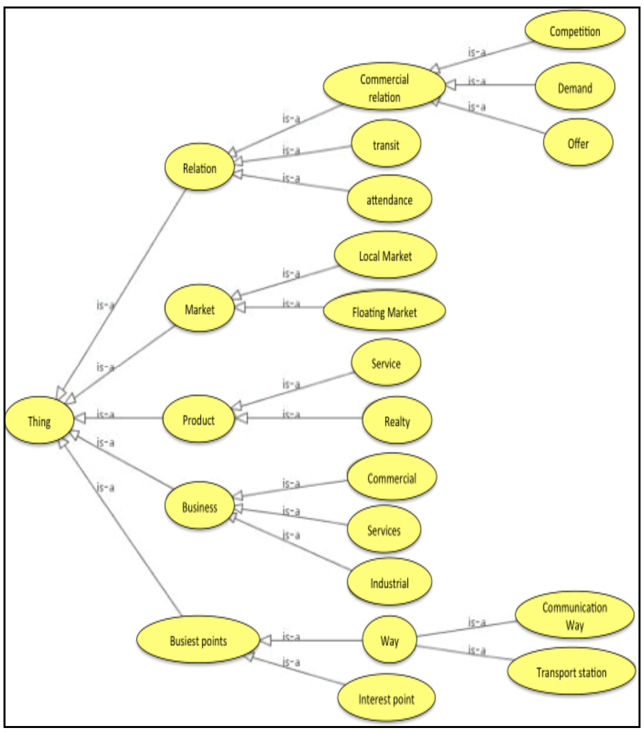

Figure 1. Taxonomy of Ontology.

For the definition of concepts and final model of ontology, the knowledge acquired and surveys were used. Some rules and terms used in market context are as follows:

a) Check type and amount of products offered by the business. b) Check floating market, demand products of business and the importance of offer and demand. c) Check the places that people of floating market attend and how often they attend. d) Check the ways and stations where the floating market is often attended. f) Evaluate the influence of interest points and attractions in the area.

Regarding the surveys conducted during the research, we define a range of values to qualify characteristics of a building or property (see Table 1).

TABLE I. TABLE RELATION QUALIFYING- WEIGHT VALUES

Qualifying	Weight value
Without importance	0
Without importance but required	25
desirable	50
necessary	75
required, essential	100

We used reasoner FaCT ++, for instance creation and classification, and we explored the ontology using OWL (Ontology Web language) Java API (Application Program Interface).

III.C Semantic Processing

The query is received and the processing starts: the query is analyzed and contextualized through the exploration of the ontology. Next, the rules of the context are applied (market analysis constraints). This results in the parameters that define a geographic query in market domain. Then, the ontology is explored to find a matching between query elements and the concepts stored into ontology. Each element of a query is searched within the ontology (e.g., by label names). When a concept is found into the ontology, then its context is extracted. A context is formed by the neighboring concepts of a matched concept and its semantic relations are extracted and stored into a vector. The process works according to the following algorithm depicted in pseudo code (see Fig. 2):

Begin

Identification and parsing of query elements.

Semantic processing

1. Search (by label name) the concept corresponding to the elements in the query

If (match)

2. Context extraction using the semantic relations of the ontology

    Extraction of relations and properties
    Set context, relations and properties into the context vector

3. Define the market search criteria data using the geographic parameters contained into the context vector

Spatial exploration

4. Define the spatial search criteria using the geographic parameters contained into the context vector relation

Query execution

5. Execution of the query to the database
6. Execution of the function to find market concepts
7. Display results

End

Figure 2. Ontology exploration algorithm.

 ISBN: 978-1-61208-293-6

In Fig. 2, the algorithm illustrates how the search parameters are used to find relevant information according to the steps involved in market analysis. A specific Web service has been developed in order to match a concept name with the term name commonly used in market analysis. The matching results are used as a parameter for other queries that are generated.

Moreover, the semantic processing includes an interpretation for each parameter. For example, in the case of stores, it is desirable that they are located in a corner (it is a defined rule), while for a book store, for example, it is more important that it is located near any school.

Table 2 shows the possible semantic values for some parameters in a market context.

TABLE II. SEMANTIC VALUES FOR MARKET PARAMETERS

Parameter	Possible semantic values
Influence of the zone	Low, medium, high.
Market competition	First impact, high impact, small impact, based on percentage.
Transportation	Fast, without traffic.
Population	Floating, fixed.

In Table 2, the possible values or range of values for each parameter is showed. For example, when a business requires a form of transportation, the evaluation is done in two ways: fast, in terms of time and distance, considering the traffic during the transportation.

IV. TEST AND RESULTS

The tests are made by the methodology implementation on a Web system, considering the three steps of methodology: semantic modeling, semantic building and semantic processing. A list of best suitable locations for a particular business on specific geographic area is generated. Then, users can set these restrictions and display the location of the commercial business according to requirements and constraints discussed in previous sections. The business types considered are: stationer' shops, drugstores and restaurants.

For testing the interface, the user sets the query through capture controls and options. The web interface is in Spanish, in a local system (see Fig. 3).

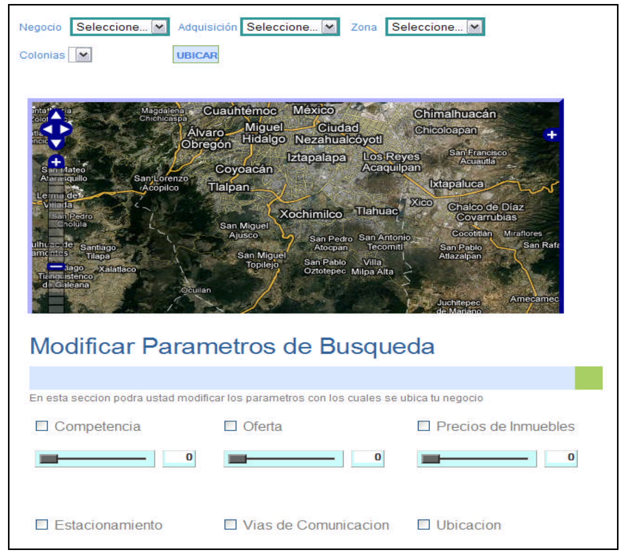

Figure 3. Input parameters (interface Web system).

As Fig. 3 shows, business requirements are defined through preset options: commercial business type, population size, neighborhood, type of acquisition, among others.

We consider the following query:

Q_{1a} = "find a business location in Mexico City with market competition in a radius of 10 km located in corner, or into the street, rent or buy". The result is the list of properties that meet the criteria, one of which is shown in Fig. 4.

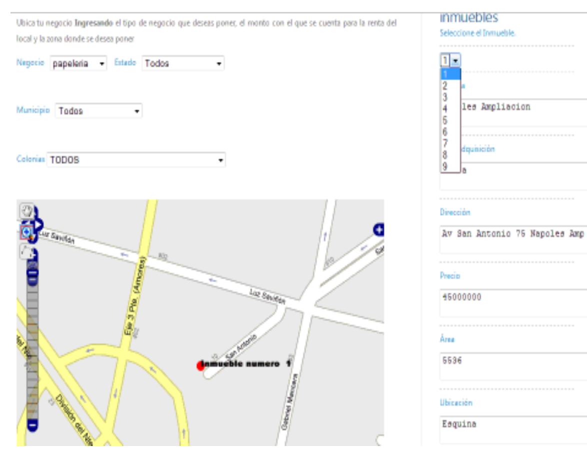

Figure 4. Best Rated Property.

Now, if the query Q_1 is modified on the relevance of the parameters: parking and roads near. Then, the new Q_{1b} query is generated. The result is now a new list of locations.

Q_{1b}: "find a business location in a county without market competition of first impact in a radius of 5 km, located in corner, or into the street, rent or buy"".

Note, that the most appropriate business location from the previous test, was located now in eighth place (see Fig. 5).

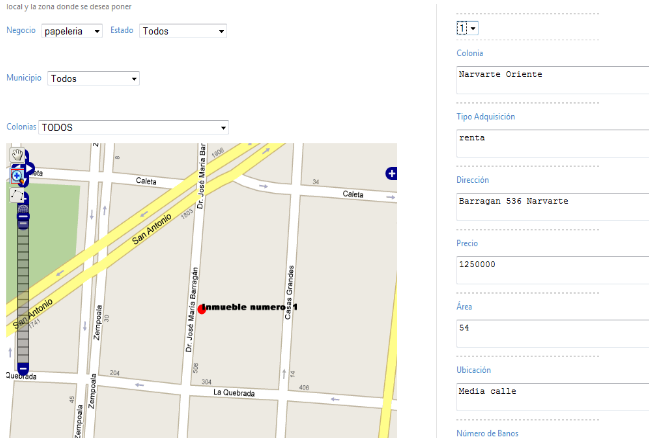

Figure 5. The best location for commercial business from Q_{1A}.

It is possible to modify / relax the parameter values to find new locations, changing the relevance of the parameters: market competition, offer, property rates, parking, roads, and location. This is achieved at the interface through sliders. For example, consider the query Q_{1A}, with the following changes: the weight factor assigned is set 50% to market competition. It generates new results shown in Fig. 6.

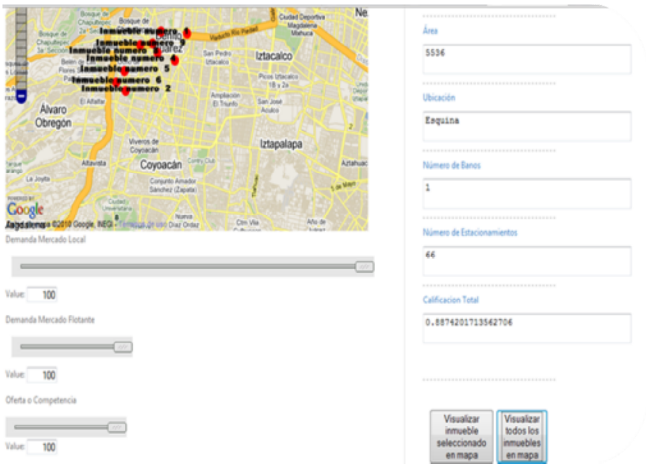

Figure 6. Results for new generated query Q_{1A}.

As we can see in Fig. 6, it is possible to relax the values for any parameter (change or configure to not be taken into account in the analysis process). Next, a new map will be displayed with the new results. Another test was conducted based on the query Q_3: “find a location for a restaurant in a corner and available for purchase, for any price, with fast transportation available”. The result of this query Q_3 is showed in Fig. 7.

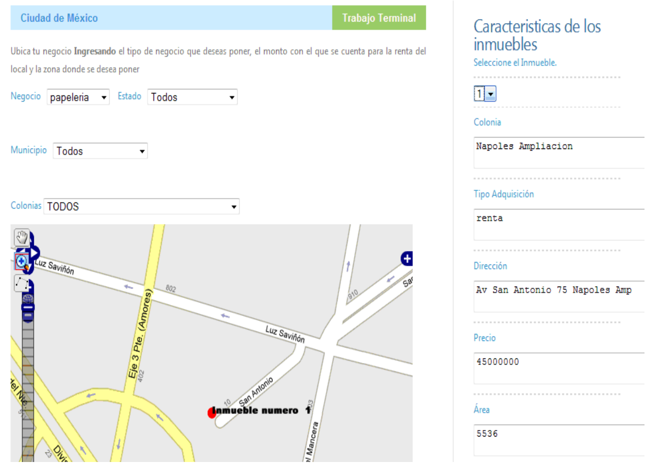

Figure 7. Best commercial locations for query Q_3.

As shown in Fig. 7, the result is only one location that complied with the values of parameters indicated.

Another test conducted for the query Q_4: "Find location for a stationer’ shop, with a low floating population, located to 3 km from a zone influenced commercially." The results obtained show 10 possible locations, the most relevant being located in: Calle 15 No 100. This is shown in Fig. 8.

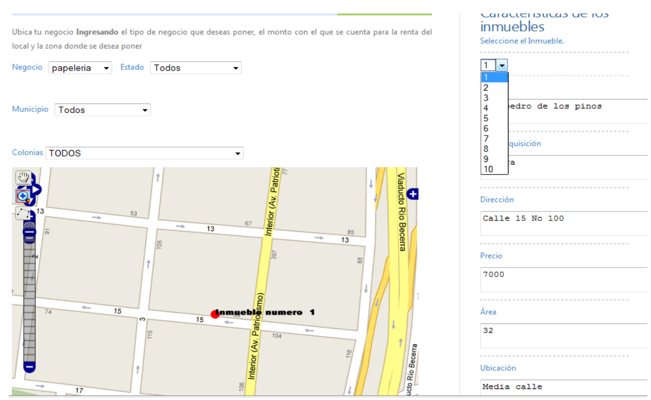

Figure 8. Rated building or property for query Q_4.

In Fig. 8, the best rated property for query Q_4 is shown; the relevancy is calculated based on the parameters values, if a property has the value requested, then it is considered relevant, but if not, then the value is evaluated, if it is lesser or higher, the relevance of it is assigned based on evaluation.

For example, the query Q_4 the influenced commercial zone around the all suitable locations is shown in Fig. 9.

 ISBN: 978-1-61208-293-6

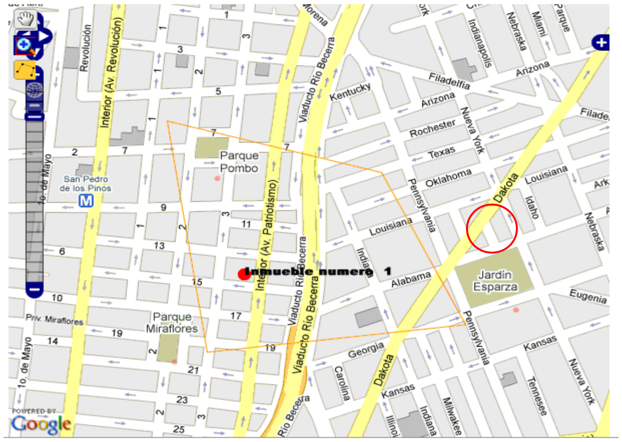

Figure 9. Zone of influence for the best location of query Q_4.

Fig. 9 shows a polygon (in orange) that represents the commercial zone that influences the locations found and evaluated as good candidates for query Q_4.

V. Conclusion

The research presented in this paper introduces an approach to perform geo-market analysis in order to find best possible places to locate a business. The methodology can be applied in other places using the corresponding geographic and statistical data of each country and procedures with applicable laws in a particular market. This can support people skilled in the branch to the location of a business and people without experience.

Semantic processing is driven by a domain ontology. The queries used are contextualized in order to closely relate information to market analysis and business location.

The approach is based on a method that matches user-defined queries with ontological concepts according to the market domains. The approach has been experimented on an illustrative case study applied in business from Mexico City.

Further work will consider larger information spaces and document collections like Web, and an integration of several distributed databases as well as additional semantic analysis. The domain ontology can also be enriched by considering additional spatial relations and by application-based conceptualizations.

Acknowledgment

The authors of this paper wish to thank the Mobile Computing Laboratory-UPIITA, Instituto Politécnico Nacional in Mexico, IPN-UPIITA, IPN-ESCOM, COFAA-IPN, IPN-SIP and ICyTDF for their support.

References

[1] Camara, G., Monteiro, A., Paiva, J. and Souza, R., 2000. Action-driven ontologies of the geographical space: beyond the field-object debate. In: M.J. Egenhofer and D.M. Mark, eds. Proceedings of the first international conference on geographic information science, October 2000, GIScience, Savannah, GA. Georgia: AAG, 52–54.

[2] Quintero R., Guzmán G., Menchaca R., Torres M., and Moreno M. 2012. An ontology-driven approach for the extraction and description of geographic objects contained in raster spatial data. *Expert Syst. Appl.* 39, 10 (August 2012), 9008-9020.

[3] Fonseca, F., et al., 2002. Using ontologies for integrated geographic information systems. Transactions in Geographical Infomation Science, 6 (3), 231–257.

[4] INEGI, 1996, Diccionario de datos topograficos 1:50 000 (Vectorial). Agnascalientes: Instituto Nacional de Estadistica Geografia e Informatica. Technical report. ISO/IEC, 2005. ISO/IEC 18025:2005(E) Information technology - environmental data coding speci- fication (EDCS) [online]. Available from: http://standards.sedris.org/ [retrieved: March 2013].

[5] Smith, B. and Mark, D., 2001. Geographical categories: an ontological investigation. International Journal of Geographical Information Science, 15 (7), 591–612.

[6] Sorrows, M. and Hirtle, S., 1999. The nature of landmarks for real and electronic spaces. Lecture Notesin Computer Science, 1661, 37–50.

[7] Timpf, S., 2002. Ontologies of Wayfinding: a traveler's perspective. Networks and Spatial Economics, 2 (1), 9–33.

[8] Tversky, B. and Hemenway, K., 1984. Objects, parts, and categories. Journal of Experimental Psychology: General, 113 (2), 169–193.

[9] Huang, B., Claramunt, C., Spatiotemporal data model and query language for tracking land use change. Transportation Research Record, 107-113 (2005)

[10] Larson, R.R.: Geographic Information Retrieval and Spatial Browsing, University of California, Berkeley. Available from: https://sherlock.sims.berkeley.edu/geo_ir/PART1.html [retrieved: March 2010].

[11] MexicoTop. articulo locales comerciales [in spanish]. Available from: http://www.mexicotop.com/article/Locales+comerciales [retrieved: February 2013].

[12] MINDMAPS. Geomarketing Estudios de Mercado. Available from: http://www.mindmaps.com/ [retrieved: March 2013].

[13] Davor M. 2007. Market driven product ontologies. In *Proceedings of the 2007 OTM confederated international conference on On the move to meaningful internet systems - Volume Part I* (OTM'07), Robert Meersman, Zahir Tari, and Pilar Herrero (Eds.), Vol. Part I. Springer-Verlag, Berlin, Heidelberg, 275-283.

[14] Grefenstette G. 2010. Use of semantics in real life applications. In *Proceedings of the 19th ACM international conference on Information and knowledge management* (CIKM '10). ACM, New York, NY, USA, 5-6.

[15] Lamparter S. and Schnizler B.. 2006. Trading services in ontology-driven markets. In*Proceedings of the 2006 ACM symposium on Applied computing* (SAC '06). ACM, New York, NY, USA, 1679-1683.

[16] Li Z, Zhao H. and Ramanathan S. Pricing web services for optimizing resource allocation an implementation scheme. In *Proc. of the Web2003, Seattle, WA,* 2003.

[17] Biletskiy Y. and Ranganathan G. 2010. A semantic approach to a framework for business domain software systems. *Comput. Ind.* 61, 8 (October 2010), 750-759.

[18] SIIM (in spanish). Available from: http://www.economia-sniim.gob.mx/nuevo/ [Accessed: March 19, 2013].