

Real-estate ontology construction for market trend detection and user behavior analysis on the Israeli real-estate web sites / Yigal Maman

Abstract

Today, the internet is a central part of human life. It is being used for purposes like E-mail, surfing different websites, pay bills, bank account management and more. Also, a major influence is the development of E-Commerce.

E-Commerce is divided into five major categories (Schneider, 2010):

1. B2C - Business to Customer, a transaction that occurs between a private customer and a business entity. For this purpose, the business entity establishes a Web site used as a virtual store
2. B2B - Business to Business, transaction occurs between a business entity and another business entity. Trading is done through private infrastructure or infrastructure supplied by a third party
3. Business processes, processes used to support activities of purchase and sale. Companies and organizations share information for the identification and assessment of customers, suppliers and employees.
4. Consumer to Consumer - C2C, a transaction occurs between a private customer and another private customer. For example eBay, an auction site where transactions are carried out where both buyer and seller are individuals.
5. Business to Government - B2G, a transaction that occurs between a government entity to a business entity.

Studies have shown that E-Commerce reduces difficulties in the acquisition process, improves distribution and marketing and causes information to be more available (Porter, 2001). The implications of this are reflected in the development of the economy: developed markets, new customers and greater competition (Daniel & Grimshaw, 2002).

In the mid-90s of the 20th century, E-Commerce was still under development as a new way of doing business. This development was very rapid until 2000, when the explosion of "dot-com bubble" (dot-com boom) occurred. At this time, and for several years there was almost total cessation in the field. However, in 2003, companies that survived the crisis began to show growth in their sales and record

profits. In addition, with the development of the economy the E-Commerce also developed, but at a faster pace. As a result, E-Commerce took a larger share of the economy. Another interesting point is that in the global crisis that began in 2008 E-Commerce suffered less than most other areas of the economy (Schneider, 2010).

One of the services have had great success with the transition into online trading is bulletin boards of real estate. Unlike most online transactions, the business model and trade is directly between users (the seller and the buyer.) However, even companies (like brokerage firms and entrepreneurial firms) present their estate in these boards. Boards are usually free and allow to advertise a real estate property for sale by anyone. The service also includes tools used as a mechanism for sophisticated searches. In three major sites (Yad 2, WINWIN and Homeless), there are now over 150,000 ads selling real estate: about 82 000 at Yad 2, 70,000 at WINWIN and 3,500 at Homeless. These three sites were selected to serve as sources of information to a system that was designed and built in order to track the ads posted in them. Initially we carried out an analysis of these boards so we could build real estate ontology. The term ontology" was brought into computer science from philosophy where ontology represents the study of being and what is (Simperl, 2009). In computer science, an ontology describes a formal representation of entities using a dictionary containing vocabulary, definitions of terms in it, relations between these terms, as well as rules for proper use of these terms (Campbell & Shapiro, 1995). Ontologies are used for a wide range of areas such as knowledge management, E-Commerce, natural language processing, semantic web and more. In Semantic Web ontologies are actually components of knowledge represented in a formal way which allows them to be reused.

The ontology we have built is based on the ontology model of schema.org and makes use of existing classes in it. In addition, we created classes we found are required. To build the optimal ontology we used the method described in "Review of the Literature". After the development of the ontology model, we built a system for tracking the ads published in the boards. The system recorded various assets in 10 cities across the country: Ashdod, Ashkelon, Beersheba, Haifa, Jerusalem, Netanya, Petah Tikva, Kiryat Shmona, Rishon Lezion and Tel Aviv.

