Resum de Tesi Doctoral

Oficina de Doctorat
UNIVERSITAT POLITÈCNICA DE CATALUNYA

DNI/NIE/Passaport	53329728L
Nom i cognoms	Antonio Villegas Niño
Títol de la tesi	A Filtering Engine for Large Conceptual Schemas
Unitat estructural	Llenguatges i Sistemes Informàtics
Programa	Computació
Codis UNESCO	120317 / 120318 / 120608 /

(Mínim 1 i màxim 4, podeu veure els codis a http://doctorat.upc.edu/doc/impresos/impres_codunesco2.pdf)

Resum de la tesi de 4000 caràcters màxim (si supera els 4000 es tallarà automàticament)

Nowadays, the need for representation and conceptualization of real world information has dramatically increased. Organizations evolution and diversification require the management and maintenance of large amounts of knowledge from their domains of interest. That growth also has an impact in the size of conceptual schemas of information systems, making them larger. The sheer size of those schemas transforms them into very useful artifacts for the communities and organizations for which they are developed. However, the size of the schemas and their overall structure and organization make it difficult to manually extract knowledge from them, to understand their characteristics, and to change them.

There are many information system development activities in which people needs to get a piece of the knowledge contained in the conceptual schema. For example, a conceptual modeler needs to check with a domain expert that the knowledge is correct, a database designer needs to implement that knowledge into a relational database, a software tester needs to write tests checking that the knowledge has been correctly implemented in the system components, or a member of the maintenance team needs to change that knowledge.

Dealing with large conceptual schemas is one of the most challenging and long-standing goals in conceptual modeling. The purpose of this thesis is to formally define a new information filtering methodology to help users of very large conceptual schemas to understand the characteristics and knowledge these schemas contain. This thesis analyzes and describes the different phases of an information filtering engine, identifies and studies several properties of relevance for elements of large conceptual schemas, provides a catalog of specific filtering requests to explore a schema in several filtering scenarios, and implements and evaluates the efficiency and effectiveness of a filtering engine prototype with several real case studies.

The filtering methodology studies the characteristics of the knowledge contained within a large conceptual schema, and proposes ways to select and represent the user interest in order to specialize the results of a filtering engine. Consequently, a user focus on a fragment of the large schema of interest to her and our methodology automatically obtains a reduced conceptual schema extracted from the large schema and focused on the knowledge that has a closer relation with the focus of the user. Such filtered conceptual schema is a subset of the original one, and because of its reduced size it is more comprehensible to the user.

The filtering approach provides knowledge extraction techniques aligned with the user interest representation, and presents such knowledge in an appropriate way to simplify its understanding. Furthermore, using this filtering approach the large conceptual schemas are navigated more quickly, increasing their usability and reducing the user effort.

Lloc Barcelona Data 13 de Juny de 2012

Signatura