



PERGAMON

Available at  
www.ComputerScienceWeb.com  
POWERED BY SCIENCE @ DIRECT®

Information Systems 30 (2005) 1–19



www.elsevier.com/locate/infosys

## Generating domain representations using a relationship model<sup>☆</sup>

Irene Díaz<sup>a,\*</sup>, Juan Llorens<sup>b</sup>, Gonzalo Genova<sup>b</sup>, J. Miguel Fuentes<sup>c</sup>

<sup>a</sup> Artificial Intelligence Center, Computer Science Department, Universidad de Oviedo, Carr. de Castiello de, Bernueces s/n 33203 Gijon, Spain

<sup>b</sup> Information Engineering Group, Computer Science Department, Universidad Carlos III de Madrid, Spain

<sup>c</sup> Research Department, dTinf S.L., Spain

Received 31 January 2002; received in revised form 29 July 2003; accepted 31 July 2003

### Abstract

Domain analysis (DA) techniques are more and more applied in information science, software engineering and even knowledge management, to help in the creation of a controlled vocabulary to represent, index and retrieve every kind of information. The main limitation of DA expansion is related with the enormous human resources investment these techniques demand. This paper presents a particular technology to perform computer based (semi) Automatic Domain Analysis (ADA) based on a relationships information representation model called RSHP. The proposed technology is presented, describing its different stages and algorithms. Finally, experimental results obtained when the methodology was applied to the “Software Reuse” domain are presented.

© 2003 Elsevier Ltd. All rights reserved.

*Keywords:* Classification; Domain analysis; Filtering

### 1. Introduction

A domain can be defined as a knowledge area for which a certain software system is developed. Domain analysis (DA) can then be defined as the process of eliciting, classifying and modeling domain information at a non-executable level [1]. The results of performing DA are domain models, formed by knowledge structures representing the commonalities and variability of all the possible

software applications that can be modeled within the selected domain [2]. Therefore, the main advantage of performing DA is that once a particular domain has been modeled (usually inside an organization), the development of information systems can be made based on high level reuse principles. Information systems developed in that way are usually cheaper and of best quality. An introduction to domain analysis and its application to information systems modeling can be found in [3].

Along the last 20 years, different methods for performing domain analysis have been developed. Neighbors in [4,5] presents the first approach to DA, named Draco. Draco is an interactive system that supplies mechanisms to enable the definition

<sup>☆</sup> Recommended by Maurizio Lenzerini.

\*Corresponding author. Tel.: +34-985-18-26-65;  
fax: +34-985-18-21-25.

E-mail addresses: sirene@correo.uniovi.es (I. Díaz),  
llorens@inf.uc3m.es (J. Llorens), ggenova@inf.uc3m.es  
(G. Genova), jmfuentes@dtinf.es (J. Miguel Fuentes).

of problem domains as special-purpose, high-level languages and manipulate statements in these languages into an executable form.

McCain in [6] creates a hierarchical DA process with three steps: market analysis, DA and specification and implementation of reusable resources.

Prieto-Díaz in [7] presents a technology based on a software library organized around a faceted classification scheme. The system supports search and retrieval of reusable components and librarian functions such as cataloging and classification. However, he demonstrated that it is not easy to make automatic faceted classification, as huge human resources are needed.

DARE [8,9] is a CASE tool that supports an analyst in extracting and recording domain information from documents and code, acquiring and recording domain knowledge from experts, analyzing domain knowledge, producing various domain models, and producing a repository of reusable assets for the domain.

Many other DA methods have been proposed, as the intelligent libraries of Simos [10], the Feature Oriented Domain Analysis (FODA) [11], the Synthesis project [12], the intelligent design of Lubars [13], the Rapid project [14], the KAPTUR tool [15], the Goma's domain analysis method [16] or the Organization Domain Modeling [17].

Although DA presents clear benefits in the development of information systems, one of the main drawbacks of starting a DA project within an organization, is that it needs huge human resources, most of them highly qualified (experts in the subjects) and therefore very expensive. Trying to find ways of automating DA is the main goal of the research presented in this paper. An Automatic Domain Analysis (ADA) methodology will allow organizations to perform DA for their expertise domains at a low cost level, and this will tremendously ease the development of their information systems. In this paper we present a methodology to perform Automatic DA based on a relationships representation model, called *RSHP*, used to represent the domain. Section 2 presents the general aspects of how automatic DA should be made. Section 3 describes *RSHP* domain representation model. The *RSHP-DA* methodol-

ogy is presented in Section 4. The results of applying *RSHP-DA* to a particular domain are shown in Section 5, and some conclusions are presented in Section 6.

## 2. Automatic domain analysis

ADA states for a set of techniques, methods and algorithms that allow a computer to create candidate domain representations gathered from electronic information sources.

The different DA techniques try to build a domain according to the steps described in Fig. 1 [18].

The first step is called plan of the project; this is a previous phase, not specific of the DA process where the domain has to be identified and limited.

In the second step relevant domain information is selected in order to generate a corpus of documents to be used as input information into computer-based algorithms. The documents corpus must be as complete as possible. It must cover the whole domain to be modeled.

The third step is composed of the following two activities: information analysis and classification.

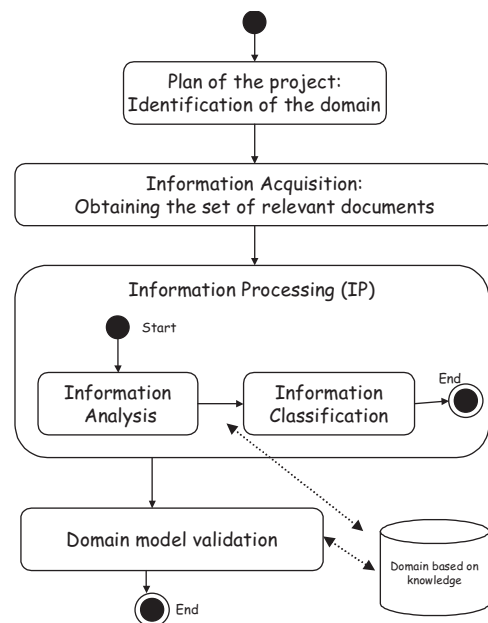


Fig. 1. Steps for building automatic domain representations.

Download English Version:

<https://daneshyari.com/en/article/10325319>

Download Persian Version:

<https://daneshyari.com/article/10325319>

[Daneshyari.com](https://daneshyari.com)