

Memòria justificativa de recerca de les convocatòries BE, PIV, BCC, NANOS i BP

La memòria justificativa consta de les dues parts que venen a continuació:

- 1.- Dades bàsiques i resums
- 2.- Memòria del treball (informe científic)

Tots els camps són obligatoris

1.- Dades bàsiques i resums

Nom de la convocatòria

BE

Llegenda per a les convocatòries:

BCC	Convocatòria de beques per a joves membres de comunitats catalanes a l'exterior (BCC)
BE	Beques per a estades per a la recerca fora de Catalunya (BE)
BP	Convocatòria d'ajuts postdoctorals dins del programa Beatriu de Pinós (BP)
NANOS	Beques de recerca per a la formació en el camp de les nanotecnologies (NANOS)
PIV	Beques de recerca per a professors i investigadors visitants a Catalunya (PIV)

Títol del projecte: ha de sintetitzar la temàtica científica del vostre document.
Systematic Construction of Goal-Oriented COTS Taxonomies

Dades de l'investigador

Nom: Claudia Patricia Cognoms: Ayala Martínez

Correu electrònic:
cayala@lsi.upc.edu

Dades del centre d'origen

Universitat Politècnica de Catalunya

Número d'expedient

2006BE 00313

Paraules clau: cal que esmenteu cinc conceptes que defineixin el contingut de la vostra memòria.
Component-Based Development, Off-The-Shelf (OTS) components, components selection, reuse, components repository.

Data de presentació de la justificació

09-07-2007

Resum del projecte: cal adjuntar dos resums del document, l'un en anglès i l'altre en la llengua del document, on s'esmenti la durada de l'acció

Resum en la llengua del projecte (màxim 300 paraules)

El uso de componentes de software llamados Commercial-Off-The-Shelf (COTS) en el desarrollo de sistemas basados en componentes implica varios retos. Uno de ellos es la falta de información disponible y adecuada para dar soporte al proceso de selección de componentes a ser integrados.

Para lidiar con estos problemas, mi trabajo de tesis propone un método llamado GOTHIC (Goal-Oriented Taxonomy and reuse Infrastructure Construction). El método está orientado a construir una infraestructura de reuse para facilitar la búsqueda y reuse de componentes COTS.

La estancia de investigación reportada en este documento se realizó en el período de 4 meses (Septiembre-Diciembre 2006) y tuvo como objetivo primordial las mejora del método y la obtención de datos empíricos para darle soporte.

Algunos de los principales resultados fueron la obtención de datos empíricos fundamentando la utilización del método en ámbitos industriales de selección de componentes COTS, así como una nueva estrategia para conseguir de forma factible e incremental, la federación y reuso de los diferentes esfuerzos existentes para encontrar, seleccionar y mantener componentes COTS y Open Source (OSS) -comúnmente llamados componentes Off-The-Shelf (OTS) - en forma estructurada.

Resum en anglès(màxim 300 paraules)

The use of Commercial-Off-The-Shelf (COTS) software components in Component-Based Development implies many challenges. One of them is the lack of available and well-suited data to support selection of suitable COTS components. I

To deal with these problems, my thesis work proposes a method called GOTHIC (Goal-Oriented Taxonomy and reuse Infrastructure Construction). The method is aimed at building a reuse infrastructure for facilitating COTS components searching and reuse.

The research stay reported in this document was performed during 4 months (September-December 2006) and had as main objective to improve the method and obtain empirical data supporting it.

Some of the main results were the obtention of empirical data supporting the use of the method in industrial settings of COTS components selection, as well as a new strategy to reach a feasible and incremental way to federate and reuse the different efforts for finding, selecting, and maintaining COTS and Open Source Software (OSS) components (the so called Off-The-Shelf -OTS- components) in a structured way.

2.- Memòria del treball (informe científic sense limitació de paraules). Pot incloure altres fitxers de qualsevol mena, no més grans de 10 MB cadascun d'ells.

The process of building software systems by assembling and integrating pre-packaged solutions in the form of Commercial-Off-The-Shelf (COTS) software components is becoming a strategic need in a wide variety of application areas [Meyers-Oberndorf02]. In general, COTS components are software components that provide a specific functionality, available in the market to be purchased, interfaced and integrated into other software systems. The potential benefits of this technology are mainly its reduced costs and shorter development time, while maintaining the quality [Oberndorf-Brownsword97]. Nevertheless, many challenges ranging from technical to legal issues must be faced for adapting the traditional software engineering activities in order to exploit these benefits.

Nowadays there is an increasingly huge marketplace of COTS components; therefore, one of the most critical activities in COTS-based development is the selection of the components to be integrated into the system under development [Basili-Boehm01]. Selection is basically composed of two main processes, namely: searching of candidates from the marketplace and their evaluation with respect to the system requirements. Unfortunately, most of the different existing methods for COTS selection focus their efforts on evaluation, letting aside the problem of searching components in the marketplace. Searching candidate COTS is not an easy task, having to cope with some challenging marketplace characteristics related to its evolvable and growing nature; and the lack of available and well-suited information to obtain a quality-assured search. Indeed, traditional reuse approaches also lack of appropriate solutions to reuse COTS components and the knowledge gained in each selection process. This lack of proposals is a serious drawback that makes the whole selection process highly risky, and often expensive and inefficient.

To deal with these problems, my thesis work is proposing a method called GOTHIC (Goal-Oriented Taxonomy and reuse Infrastructure Construction) [Ayala06]. It is aimed at building a reuse infrastructure for facilitating COTS components searching and reuse.

The main goal of the research stay reported in this document was the summative validation of the GOTHIC method in industrial settings. Such stay was performed at the Norwegian University of Science and Technology (NTNU), IDI department (Department of Computer and Information Science) was performed, specifically in the Software Engineering group (SU) led by Professor Reidar Conradi. They have a recognized worldwide expertise in empirical research and industrial validation issues. Therefore, their support has been crucial in the development of this stage of the thesis.

To reach our goal, several empirical validation approaches were analyzed and studied. From our study, we concluded that although the summative validation of the GOTHIC method may imply several alternatives; in terms of the nature of the method and the time restrictions of this thesis dissertation, we first had to tackle a short/medium term validation approach able to provide us forceful data supporting the claims of this thesis. However, a long term and further validation approach is also envisaged as an European financed project to be performed during the next three years.

In the following sections we summarize and discuss the main validation approaches taken, as well as the main results of the research stay. While theoretically different in research approach taken, each of these efforts seek to validate the method.

1.1 Evaluation of the success of the reuse infrastructure obtained from GOThIC

To correctly measure the success of a repository is not an easy task. It can be measured in many ways. Many of them come from different fields as knowledge-based system and economic models for software reuse.

In [Nick-etal01], the theoretically correct value of an experience-based repository (in terms of money and effort saved) is defined as the value of all query results minus the cost of all queries, as well as build-up and maintenance activities. The value of a query result is the value of the retrieved cases in terms of money or effort saved, and includes any future use of the retrieved cases that happen without querying the repository. Thus, we found that to obtain the theoretically correct value for evaluating the entire GOThIC reuse infrastructure would imply to collect empirical data along several life cycle stages of the repository, from its construction to its evolution and maintenance. But, it implied a costing and long-term analysis that is out of the time expected to report the thesis document (but not out of our objectives).

Other practical approaches as [Jedlitschka-Nick03] and [Cooper97] take into account the common restrictions related to cost and time for evaluating repository approaches and propose the use of an straightforward monitoring of indicators as the “personal utility” of the delivered information to the user (i.e., utility as perceived by stakeholders) and usability to identify significant trends regarding the value of the repository. Thus, based on these studies, we addressed the evaluation of the success of the GOThIC repository by considering the utility perceived by COTS selectors.

This study was performed in the context of the European ITEA project, Norwegian COSI (Co-development using inner & Open Source in Software Intensive products) [COSI] which aims to enable the Norwegian IT sector to fully exploit the benefits and advantages of COTS and Open Source Software (OSS) components, the so called Off-The-Shelf (OTS) components.

Different studies were tackled in order to gather valuable qualitative information supporting the effectiveness of the method in industrial settings:

- ▶ Two Industrial seminars addressed to NTNU researchers and the industrial COSI participants respectively. The intentions of these seminars were twofold:
 - First, we would like to explain the GOThIC method and report our results in several COTS selection processes. During our talk, participants were able to interrupt and ask questions
 - As most data we had showed academic cases and some post-mortem industrial results, we were interested in discussing and getting feedback about the perception of the utility and usability of GOThIC from academics in the first seminar, and from industrial experts in the second one. We therefore had in both cases time for discussion and feedbacks.

Participatory observation technique was used to gather data and annotations, as well as to process the data obtained from the seminars.

The first seminar was held at NTNU in May 2006. 11 researchers from the IDI department were attended. Our discussions and their feedbacks were significantly useful

to realize and improve several technical approaches forming part of the method as well as the industrial strategy to put forward the GOTHIC repository.

The second seminar was held in the context of a COSI project meeting at NTNU in October 2006. 12 industrial participants from the COSI project attended. Their industrial background helped us to mature our industrial conception of the method usability. After the seminar, some individual meetings with some interested participants were held. They were able to provide us some of their own COTS selection projects information to perform further studies.

- From the previous seminars, we realized the need of further analyze and get empirical evidence to support the method. One of our main concerns was: how COTS selection processes are dealt in industry in order to better synchronize the GOTHIC method activities, as well as understand the existing practices for improving the method processes. Thus, taking advantage of the availability of students taking the course TDT4735 Software engineering at NTNU we asked two of them to perform the study with our guidance. In the design stage of the study, we realized the need of focus not only in COTS, but also to include OSS. It means we focus our study in Off-The-Shelf (OTS) components. It was because industrial processes are actually addressing both paradigms.

The goal of the study was to discover the actual processes used in industry when it comes to selection and evaluation of OTS components, mainly to ask the next questions: Where and how OTS components are found? How are they evaluated? How are they learnt? How to take care of knowledge about the chosen component? How to make best use them?

A pre-study was performed and several semi-structured interviews were applied to software engineers in some companies participating in the COSI project as Keymind Computing, SINTEF ICT, FAST ASA, Statoil ASA, eZ Systems, SUN, and Linpro. It was reported in [Gerea06]. This former study was the basis for the main qualitative descriptive study reported in [Gerea07]. It includes improved semi-structured interviews to software engineers in the following Norwegian companies: Visma, Sirius IT, TietoEnator, WebOn, Abeo, DKDigital, Commitment, Grieg Multimedia, and Riventy.

1.2 Intended GOTHIC Populating and Maintenance Strategy

From the industrial formative evaluation of GOTHIC we found some concerns regarding its population and maintenance: heavy upstart cost (i.e., small and medium enterprises will not be able in the general case to adopt it), and difficulty to maintain complete and up-to-date information due to the highly changing nature of the COTS components marketplace.

To overcome the issues found for populating and maintaining the GOTHIC knowledge base, together with the people from the Norwegian University Of Science and Technology (NTNU), we designed a population and maintenance strategy that makes use of the creative and productive potential of “open-source collaboration”. These results were published in an important conference in the area [Ayala-etal07].

By using our proposed strategy, the COTS and OSS components consumers or (re)users (i.e., individuals, organizations, academic researchers, industrials) can be harnessed to work as a community dedicated to incrementally build and maintain an open OTS knowledge base built

with GOTHIC. This provides all the GOTHIC benefits to all OTS consumers whilst ensuring smooth start-up and maintenance cost, as well as highly reliable information.

To put forward this strategy we exploit the potential offered by a Wiki-based portal. A Wiki (from the Hawaiian Wikiwiki meaning “fast”) is a collaboratively created and iteratively improved set of web pages [Wagner04]. It is considered a powerful knowledge management tool that enables the creation of an incrementally growing system containing the shared knowledge of multiple sources in a centralized infrastructure/repository (i.e. a database server, an application server that runs the Wiki software, and a web server that serves the pages and facilitates the web-based interaction). Thus, exploiting some particular Wiki characteristics (based on the principles described by Wagner [Wagner04]) we have designed an OTS-Wiki portal [Aaslund-Larsen07]. The intended main goals of the project are:

- Fostering a OTS Community and Incremental Population of Content.** The OTS-Wiki provides the web-based infrastructure for enabling OTS technology (re)users to collaborate as a community in an open-source-like environment, see Fig. 1.1. Thus, OTS Community users are able, and even encouraged to share knowledge (e.g., experiences, components information, and vendor comments). Therefore, the incremental population of content in the portal based on the OTS Community participation is expected. We have designed proper templates and guidelines for editing and use in order to share the information in a structured way (as demonstrated in the Wikipedia, an on-line encyclopedia implemented as a Wiki).

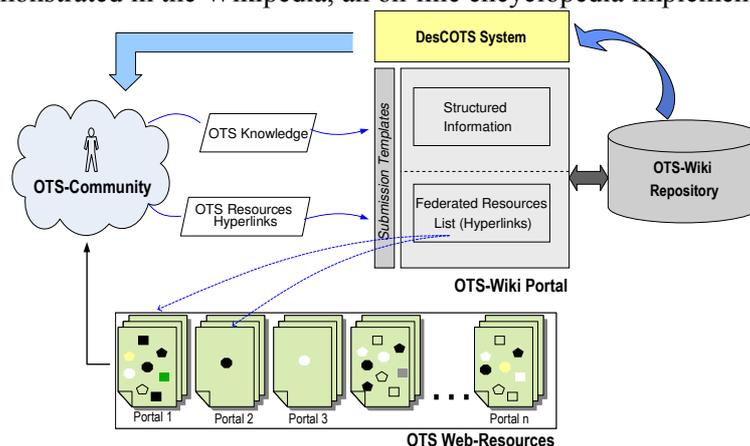


Fig. 1.1 OTS-Wiki Portal Main Interactions

- Federating Actual Efforts for Locating and Selecting OTS Components.** In this collaborative environment, OTS Community users are encouraged to add (as hyperlinks) and comment the helpfulness of existing web-resources for locating OTS components. This is a way of having an up-to date federated list of actual web-resources that the OTS Community users can exploit. Besides the obvious advantage of using hyperlinks for allowing users to make connections and to drill down into detailed knowledge, hyperlinks are also a potential quality assurance mechanism and relevance indicator. Pages with many links to them indicate a highly useful page. This factor fosters the OTS-Wiki portal to act as a meta-portal for promoting the progressive homogenization of the information contained in different OTS web resources. This is because such resources have an interest of being perceived as highly useful by the OTS Community users.
- Enabling Systematic Support for Selecting and Evaluating OTS Components.** Having structured OTS information as claimed by GOTHIC enables systematic support for evaluating and choosing components. We are integrating the DesCOTS system [Grau-etal04] into the *OTS-Wiki*, as stated in Fig. 1.1. It includes a set of tools that interoperate to support the whole OTS selection process. Nevertheless, some other existing or new tools

can be developed or designed for using the structured OTS component information from the *OTS-Wiki* portal.

In this scenario, any OTS Community user can use the OTS-Wiki portal as a meta-portal for providing support to:

- a) Searching OTS and information about them supported by a well-defined and dynamic taxonomy
- b) Recording component information in a structured way.
- c) Maintaining and reusing such information by the use of suitable and evolvable models capturing all the OTS informational dimensions proposed by GOTHIC [Ayala-Franch06].
- d) Getting tool support for performing selection processes (e.g., DesCOTS).

In [Ayala-etal07] this strategy is detailed and expected usage scenarios are provided. Such paper is attached.

It is worth to mention that the further exploration of this approach is considered as future work and is intended to be an European project involving several academy and industrial partners as the Norwegian University of Science and Technology, Politecnico di Torino, Centre de Recherche Public Henri Tudor and Hewlett Packard among others. Our aim is to use the GOTHIC approach and its metamodel to augment the OTS-Wiki functionalities with web-intelligence technologies. Web intelligence will be applied to analyze user logs of web-search queries, query responses, component choices, and all kinds of solicited comments and reviews – and from these build up, up-date and maintain revised and pragmatic taxonomies (ontologies) in an incremental way among other functionalities.

Bibliography

[Aauslund-Larsen07] Aaslund, K., Larsen S.: "OTS-Wiki: A Web Community for Fostering Evaluation and Selection of Off-The-Shelf Software Components" Master Thesis. Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU). Spring 2007. <http://www.idi.ntnu.no/grupper/su/su-diploma-2007/dipl07-larsen-aaslund.pdf>

[Ayala06] Ayala, C.: "Systematic Construction of Goal-Oriented COTS Taxonomies" In Proceedings of the 3rd Doctoral Consortium at the 18th Conference on Advanced Information Systems Engineering (CAISE 2006).5-9 June 2006, Luxembourg.

[Ayala-Franch06] Ayala, C.; Franch, X.: "Domain Analysis for Supporting Commercial Off-The-Shelf Components Selection". In Proceedings of the 25th International Conference on Conceptual Modelling (ER 2006). Tucson, Arizona, USA. Lecture Notes in Computer Science. Volumen: 4215/2006. Pages: 354-370.

[Ayala-etal07] Ayala, C., Sørensen, C.F., Conradi, R., Franch, X., Li, J.: "Open Source Collaboration for Fostering Off-The-Shelf Components Selection". In IFIP International Federation for Information Processing, Volume 234, Open Source Development, Adoption and Innovation. (OSS 2007). June 2007, pp. 17-30.

[Basili-Boehm01] Basili, V.R., Boehm, B.: "COTS-Based Systems Top 10 List", *IEEE Computer*, Vol. 34, No. 5, May 2001.

[Cooper97] Cooper, W.S.: On selecting a measure of retrieval effectiveness. In K. Jones and P. Willet (eds.), *Readings in Information Retrieval*, pages 191–204. Morgan Kaufmann Publishers, 1997.

[COSI] COSI: Co-development using inner & Open Source in Software Intensive products. European ITEA project.

[Gerea06] Gerea, M.: "Selection and Evaluation of Open Source Components", 15th Dec. 2006, 81 p. part of course TDT4735 Depth Project in Software Engineering, Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU). <http://www.idi.ntnu.no/grupper/su/fordypningsprosjekt-2006/gerea-fordyp06.pdf>

[Gerea07] Gerea, M.: "Selection of Open Source Components: A Qualitative Survey in Norwegian IT Industry". Master Thesis. Department of Computer and Information Science, Norwegian University of Science and Technology (NTNU). Spring 2007. <http://www.idi.ntnu.no/grupper/su/su-diploma-2007/dipl07-gerea.pdf>

[Grau-etal04] Grau, G., Carvallo, J.P., Franch, X., Quer, C. "DesCOTS: A Software System for Selecting COTS Components". In Proceedings of the 30th EUROMICRO Conference, IEEE Computer Society, 2004. The current version of the tool is available at <http://www.lsi.upc.es/~gessi/QMTool/QMTool.html>

[Jedlitschka-Nick03] Jedlitschka, A., Nick, M.: "Software Engineering Knowledge Repositories" Experiences from ESERNET 2003. LNCS Volume 2765/2003 pp. 55-80.

[Meyers-Oberndorf02] Meyers, B.C., Oberndorf, P.: "Managing Software Acquisition". SEI Series in Software Engineering, 2002

[Nick-etal01] Nick, M., Althoff, K.D., Tautz, C.: "Systematic Maintenance of Corporate Experience Repositories" *Computational Intelligence* 17(2): 364-386 (2001).

[Oberndorf-Brownsword97] Oberndorf, P., Brownsword, L. "Are You Ready for COTS?" Software Institute Engineering. August 1997.

[Wagner04] Wagner, C.: "Wiki: A Technology for Conversational Knowledge Management and Group Collaboration". Communications of the Association for Information Systems, vol. 13, article 19, pp. 256-289.