

VOLUME I | Issue 1 | 2013 | CIEO

Guest Editors: Marielba Zacarias and Paula Ventura
Editor: Teresa de Noronha

Journal of Spatial and Organizational Dynamics

ICT in Organizations

Crawling Big Data in a New Frontier for Socioeconomic Research: Testing with social tagging
Juan D. Borrero and Estrella Gualda

A Maturity Model for Higher Education Institutions
Duarte Duarte and Paula Ventura Martins

How to Model People Work Practices From Ontological Transactions
António Gonçalves, Marielba Zacarias and Pedro Sousa

TECHNICAL INFORMATION

Volume I, Issue 1, 2013

JOURNAL OF SPATIAL AND ORGANIZATIONAL DYNAMICS

ICT in Organizations

Authors:

Antonio Leonardo Gonçalves
Duarte Duarte
Estrella Gualda
Juan D. Borrero
Marielba Zacarias
Paula Ventura Martins
Pedro Sousa

Guest Editors: Marielba Zacarias and Paula Ventura Martins

Editor-in-chief: Teresa de Noronha

Publisher:

Research Centre for Spatial and Organizational Dynamics - CIEO
Campus de Gambelas, Faculdade de Economia, Edifício 9
8005-139, Faro
cieo@ualg.pt
www.cieo.pt

Editorial Board:

Teresa de Noronha, Faculty of Economics, University of Algarve, Portugal (mtvaz@ualg.pt)
André Torre, Institut National de la Recherche Agronomique, Agro Paris Tech, France (andre.torre@wanadoo.fr)
Charlie Karlsson, Jönköping International Business School, Jönköping University, Sweden (Charlie.Karlsson@ihh.hj.se)
Eric Vaz, Department of Geography, Ryerson University, Canada (evaz@GEOGRAPHY.Ryerson.Ca)
Helen Lawton Smith, Department of Management - Birkbeck, University of London, U.K (helen.lawtonsmith@ouce.ox.ac.uk)
Jafar Jafari, School of Hospitality Leadership, University of Wisconsin-Stout, USA (jafari@uwstout.edu)
Purificación Galindo, Department of Statistics, University of Salamanca, Spain (purivic@yahoo.com)
Rafael Alberto Peres, Universidad Complutense de Madrid, Spain (estrategia@rafaelalbertoperez.com)
Saul Neves de Jesus, Faculty of Human and Social Sciences, University of Algarve, Portugal (snjesus@ualg.pt)
Thomas Panagopoulos, Faculty of Sciences and Technology, University of Algarve, Portugal (tpanago@ualg.pt)

Open Access Policy:

This is a blind peer-reviewed journal.

Articles submitted to this journal should not have been published before in their current or substantially similar form.

The JSOD is diffused by all links related to the Research Center for Spatial and Organizational Dynamics and is continually online (http://www.cieo.pt/discussion_papers_editions.php)

This journal is supported by the Portuguese Foundation for Science and Technology (FCT).

Networking and Indexing:

Silvia Fernandes

Page Layout:

Marlene Fernandes

Design and Cover Concept:

Bloco D, Design e Comunicação

Quarterly Edition
ISSN: 2183-1912
CIEO, Faro, Portugal

TABLE OF CONTENTS

Crawling Big Data in a New Frontier for Socioeconomic Research: Testing with social tagging..03

Juan D. Borrero

Estrella Gualda

1. Introduction	03
2. Literature Review	05
3. Methodology	10
4. Results	13
5. Discussion	16
6. Conclusions and Future Research.....	17

A Maturity Model for Higher Education Institutions..... 25

Duarte Duarte

Paula Ventura Martins

1. Introduction	25
2. Academic Institutions and Process Improvement Needs.....	26
3. Process Maturity Models	28
4. Investigation Problems.....	41
5. Present and Future Work.....	42

How to Model People Work Practices from Ontological Transactions 46

António Gonçalves

Marielba Zacarias

Pedro Sousa

1. Introduction and Motivation.....	46
2. Theoretical Framework	48
3. Proposed Approach	53
4. Example.....	58
5. Conclusion and Future Work	62

CRAWLING BIG DATA IN A NEW FRONTIER FOR SOCIOECONOMIC RESEARCH: TESTING WITH SOCIAL TAGGING

Juan D. Borrero
Estrella Gualda

ABSTRACT

Tags and keywords, freely chosen by users for annotating resources, offer a new way for organizing and retrieving web resources that closely reflect the users' interests and preferences, as well as automatically generate folksonomies. Social tagging systems have gained increasing popularity as a method for annotating and categorizing a wide range of different web resources. They also attract researchers in social sciences because they offer a huge quantity of user-generated annotations that reveal the interests of millions of people. To date, the study using digital trace data methods continues to lack a theoretical framework, particularly in social science research. This paper presents a methodology to use big data from Web 2.0 in social research. At the same time, it applies a method to extract data from a particular social bookmarking site (*Delicious*) and shows the sort of results that this type of analysis can offer to social scientists. The illustration is made around the topic of globalization of agriculture. Using data crawled from a large social tagging system can have an important impact in the discovering of latent patterns, which is needed to provide effective recommendations to different actors. In this paper, a sample of 851 users, 526 URLs and 1,700 tags from the *Delicious* classification system on the subject of globalization were retrieved and analyzed. Through the analysis, main users and websites around globalization issues in *Delicious* emerged, along with discovering the most important tags that were applied by users to describe the globalization of agriculture. The implications of these methodology and findings for further research are discussed.

Keywords: Information Retrieval; Social Network Analysis; Collaborative Tagging; Web 2.0

JEL Classification: C81, D85, F60

1. INTRODUCTION

The advent of the social web has significantly contributed to the explosion of web content and, as a side effect, to the consequent, explosive growth of the information overload. Thus, in recent years, there has been a substantial growth in social computing systems (Parameswaran and Whinston, 2007a; Kwai Fun and Wagner, 2008) that serves as intermediaries for social relations (Schuler, 1994) and are characterized by online community formation and user content creation (Parameswaran and Whinston, 2007b; Duan *et al.*, 2008). Some of the best known social computing systems are content sites such as Wikipedia, Flickr, YouTube, social networking sites such as Facebook, microblogging such as Twitter, and social bookmarking services such as *Delicious* (Marlow *et al.*, 2006; Parameswaran and Whinston, 2007a). Overall, these social computing systems are characterized by a high

heterogeneity of information sources and make large amounts of information available to their users (Schueler *et al.*, 2007; Vickery and Wunsch-Vincent, 2007). For some authors, certain data collection methods commonly applied in social studies like interviews or surveys have yielded inconclusive results, especially when it is in a web environment (Herring *et al.*, 2004; Nardi *et al.*, 2004; Yao, 2009). However, some studies are already deploying automatic data extraction techniques (Garrido and Halavais, 2003; Jones *et al.*, 2008; Shumate and Dewitt, 2008; Xu *et al.*, 2008; Carmel *et al.*, 2009; Wang and Jin, 2010; Ackland and O'Neil, 2011) from big data. These techniques are of interest to social researchers; but, to date, the study using digital trace data methods continues to lack a theoretical framework (Janetsko, 2009: 170). Some authors clearly point out the need for flexibility and adaptability of the methodology (quantitative and qualitative techniques) to the object (Domínguez *et al.*, 2010: 10), and also focus on discovering collaborative process methods and tools (Zacarias and Ventura, 2011: 45). Our first objective will be to present a methodology to use big data from Web 2.0 in social research, and then to show the sort of results that this type of analysis can offer to social scientists.

Currently, the extensive use of the social web is emphasizing the central role of users and their (cor)relations. The focus is on the profile, preferences, needs, feedbacks, reputation, relationships, and, last but not least, the personal way each user classifies the huge amount of information at her/his disposal in the form of tags (Golder and Huberman, 2006; Marlow *et al.*, 2006; Cattuto *et al.*, 2007a; Rattenbury *et al.*, 2007; Dattolo *et al.*, 2012). Tags are keywords freely chosen by users (e.g. “globalization”, “agriculture”, “trade”), employed to annotate various types of digital content including images, bookmarks, blogs, and videos (Golder and Huberman, 2006; Shneiderman *et al.*, 2006; Rattenbury *et al.*, 2007; Trant, 2009). The idea behind the concept of tagging is simple: a user enjoys a resource – e.g. an image or a website – and, according to her/his mental model, identifies those terms that better describe the information conveyed by that resource. The same resource can be annotated by several users: some of them will reuse the tags already assigned to that resource, while others will adopt new tags. Social tags produced by users are usually regarded as high quality descriptors of the web pages' topics and a good indicator of web users' interests and preferences. This process also allows building of a socially-constructed classification schema, called folksonomy (Vander Wal, 2007).

Social tagging systems have recently begun receiving increasing attention from the scientific community. The growing number of scientific publications concerning this issue and the development of real, social tagging systems, such as social networks (Twitter), social bookmarking applications (Delicious), sharing systems (Flickr), and in the e-commerce field (Amazon), confirm this tendency. The popularity of tagging is attributed, at least in part, to the benefits users gain from effectively organizing and sharing very large amounts of information (Cattuto *et al.*, 2007b) and users' interests (Golder and Huberman, 2006).

Some prominent examples of a tagging-intensive social computing system are social bookmarking sites such as Digg, StumbleUpon, Reddit, Pinterest, and Delicious¹. These services are an inestimable source of information for scholars, as they produce a huge amount of user annotations (tags) and reflect the interests of millions of users. The social aspects of these systems derive from the fact that the resources (mainly websites) are tagged by the community, a feature known as collaborative tagging, which provides important metadata for investigators and others practitioners.

Delicious (www.delicious.com) – formerly *del.icio.us* in 2003 – is a free, social bookmarking web service for storing, sharing, and discovering web bookmarks. Delicious uses a non-hierarchical classification system in which users can tag each of their bookmarks with freely chosen index terms. Its collective nature makes it possible to view bookmarks added by

¹ <http://www.ebizmba.com/articles/social-bookmarking-websites> (retrieved 10.09.2012).

other users. All bookmarks posted to Delicious are publicly viewable by default. Tagging in systems like Delicious is an important change in the way web bookmarks are organized and shared (Ames and Naaman, 2007).

Introducing folksonomies as the basis for social network analysis means that the usual binary relation between users and resources, which is largely employed by traditional Recommender Systems, changes into a ternary relation between users, resources, and tags which is more complex to manage. Nevertheless, very few works highlight how to employ folksonomies in the field of social research. This leads us to think that further researches, evaluation studies, and insights are needed. Hence, our second objective in this paper will be to use data crawled from a large social tagging system to discover latent patterns, which will form a basis in order to provide effective recommendations to different actors.

Due to the current lack of a theoretical framework in retrieving automatic data and analyzing digital data in social science research, this paper presents a methodology to use big data from Web 2.0 in this field. At the same time, it applies this method to automatically extract data from a particular social bookmarking site (Delicious) and to show the type of results that this kind of analysis can offer to social scientists.

We focus our study on the Delicious site, specifically, in its user community around the issue of globalization of agriculture. According to Stiglitz (2006), issues such as bilateral trade are impeding development in the world's poorest countries. The globalization of the agriculture system is at the centre of this debate, because so many poor people depend on agriculture as an income source and because they spend a large proportion of their resources on food. Given that the majority of the poor inhabit rural areas and earn a living as small farmers, the effects of globalization on employment and small-farm competitiveness are central to determining its impact on poverty.

This paper begins by reviewing the literature around collaborative tagging, paying specific attention to meta-knowledge and networks perspectives. Then we expound the methodology, laying out the empirical data and describing in detail the data extraction process applied. The next sections analyze and provide the results of a study that involved 851 users on 1,077 URLs and 1,700 tags, before concluding with a discussion of several implications for research and practice.

2. LITERATURE REVIEW

2.1. Web 2.0 and Collaborative tagging

The Web 2.0 concept was developed by Tim O'Reilly (O'Reilly, 2007). According to O'Reilly, "Web 2.0 is the business revolution in the computer industry caused by the move to the Internet as platform, and an attempt to understand the rules for success on that new platform²." Web 2.0 is combined with the programmable Web and was designed to be dynamic, peer to peer, and an online storage of knowledge.

Collaborative – or social – tagging is the activity in Web 2.0 of annotating digital resources with keywords, or so-called tags (Golder and Huberman, 2006; Trant, 2009). This process is easy to indulge in, because it does not require any professional background; all that is needed is to freely choose keywords from an individual's vocabulary to annotate a Web resource. This process of annotation has converted ordinary people into metadata generators. Hence, collaborative tagging is a form of a user-centric, social, and democratic method of indexing. The use of tags creates a collective classification scheme and provides a snapshot of the current mindset of the user.

² <http://radar.oreilly.com/2006/12/web-20-compact-definition-tryi.html> (retrieved 01.10.2012).

Collaborative tagging has two purposes: first, it can quickly generate personal categorizations for later information retrieval; second, the collective use of tags makes inferences about related resources and tags possible. A resource can be tagged with an unlimited number of tags. The collaborative tagging of websites allows for the organization and sharing of digital resources. These websites allow users to publicly tag available resources and share content; therefore, users can categorize information by themselves and browse or search for the information by using these tags (Golder and Huberman, 2006). In this sense, the collaborative tagging of websites works as a shared resource for a given community of actors that could be used for different motives and in different moments.

A collaborative tagging system is mainly composed of three interconnected components – users, tags, and resources (Smith, 2008) – which can be described as follows:

- Users: They employ a tagging system to create tags, and sometimes they add resources. Users – who have a variety of different interests, needs, goals, and motivations – try to share or label a resource so they can find it later.
- Resources: They are the items that users tag such as the Web pages in Delicious and the photos in Flickr.
- Tags: They are the keywords added by users. Tags are essentially metadata about the resource. Users can tag just about any kind of term to resources, and different users have different tagging patterns.

2.2. Tagging and Folksonomy

Social tagging systems aggregate the tags of all users and describe the resources in a so-called folksonomy (Vander Wal, 2004; Trant, 2009). The word “folksonomy” (Vander Wal, 2004) is a combination of “folks” and “taxonomy”. Recently, the use of folksonomies gained more attention because of their simplicity: using tags, users can freely model the information without the constraints of a predefined lexicon or hierarchy (Mathes, 2004). Folks are the common people of a society; taxonomy means a hierarchical structure of classification. However, the simplicity of the approach also has an important drawback: the information managed by folksonomies is modeled in a simple, syntactical way. Therefore, collaborative tagging systems suffer from the vague-meaning problem when users retrieve or present resources with keyword-based tags. The vague-meaning problem is created by the following causes (Kroski, 2005; Golder *et al.*, 2006; Hope *et al.*, 2007; Marchetti *et al.*, 2007):

- Synonyms: It is when multiple tags share the same meaning. For example, resources tagged as *Web site* and *Website*, or *global warming* and *climate change* could have the same meaning: the first ones are semantically similar, and in the second example the words are different. However, collaborative tagging systems do not understand it.
- Term variations: There is no standard for the structure of tags; for instance, a noun can be singular or plural, uppercase or lowercase. In collaborative tagging systems, we can also have simple, morphological variations. Moreover, mis-tagging due to spelling errors occurs often. Also, spacing is not allowed in a tag in most collaborative tagging systems; therefore, both the underscore and the hyphen are typically used to separate words by a single tag. Additionally, different possible spellings of the same word and tags using different languages generate term variations such as *globalization* and *globalisation*.
- Lack of relationships: Relationships between tags cannot be structured in existing, collaborative tagging systems. For instance, resources might be labeled with the tags *fast food* or *hamburger*, and there is no mechanism that might indicate that hamburger is a sub-class of fast food.

These drawbacks hinder the use of folksonomies for tasks more complex than the simple browsing of resources. In order to avoid these problems, in recent years many tools have

been developed to facilitate the user in the task of tagging, by also speeding up the tag convergence (Cattuto *et al.*, 2007b).

2.3. The collective knowledge inherent in social tags

Social tagging is the process by which many users add metadata in the form of keyword-based tags to shared resources. In social tagging systems, users can annotate a variety of digital resources with tags, for instance, bookmarks (e.g. www.delicious.com), pictures (e.g. www.flickr.com), or products (e.g. www.amazon.com). In most applications, users are free to choose any tags for describing their resources in order to structure, organize, and re-find their own stored Web material. The tags that are used will reflect individual associations with regard to resources, and they will describe a specific meaning or relevance for the respective users. The social aspect of social tagging systems lies in the opportunity to use other people's tags as navigation links for one's own search processes. The folksonomy process is developed in a bottom-up process of individual tagging in which the tags of many different users are aggregated and the resulting collective tag structure – such as tag cloud – depicts the collective knowledge of Web users (Cress *et al.*, 2012) – although, in some cases, such as by laziness, users could have aggregated tags suggested by the social tagging system. The individual users' tags establish a network of connections between resources and tags, and among those tags themselves. The more frequently tags are used for one resource, the stronger the connection becomes among them. Analogously, the more often two tags co-occur for one resource, the stronger they are related to each other. Social tagging systems can be considered shared, external knowledge structures of communities (Fu, 2008) and augment the collective structure of a community with the individual knowledge representations of individual users. When aggregating all tags from a community, a collective representation of the connections between related tags and their strengths of association will emerge. These associations are typically visualized by tag clouds, in which different font sizes represent the strength of association of tags to a related tag or a resource. Tag clouds externalize the community's associations between tags and the strengths of associations. In this way, social tags are able to provide visual representations of the conceptual structure of a domain that is built upon the knowledge of individuals who belong to a large Web community. One study (Kammerer *et al.*, 2009) showed that tags as *navigational signposts* are able to provide a kind of scaffold to learn new topics, leading to better understanding of a knowledge domain. Although a few studies have investigated the influence of tag clouds on visual attention, recognition, and tag selection (Rivadeneira *et al.*, 2007; Bateman *et al.*, 2008), this research addresses how users could use this collective knowledge representation.

Much research on social tagging has focused on the description of regularities in user activities (e.g. Golder and Huberman, 2006; Millen *et al.*, 2007). Research has also investigated the motivation of people to tag and the research question of how to design Web sites and platforms in order to motivate users to annotate content with social tags (Van Velsen and Melenhorst, 2009). However, surprisingly little is known about how these new technologies directly interact with individuals at the knowledge and cognitive level (Fu, 2008).

The first study that investigated the interplay between collective and individual knowledge was presented by Fu (2008). He introduced a rational model of social tagging in Delicious and provided evidence for the interaction between social and cognitive systems. Further studies addressed the emergence of stable tagging patterns (Kannampallil and Fu, 2009) or investigated how the use of social tags affects search performance (Fu *et al.*, 2010), tag choice, and the individual interpretation of documents through processes of imitation (Fu *et al.*, 2009; Kang *et al.*, 2009), users' behaviours (Kang and Fu, 2010), or innovation impacts in organizations (Parise and Iyer, 2011), with Delicious as well. In these studies,

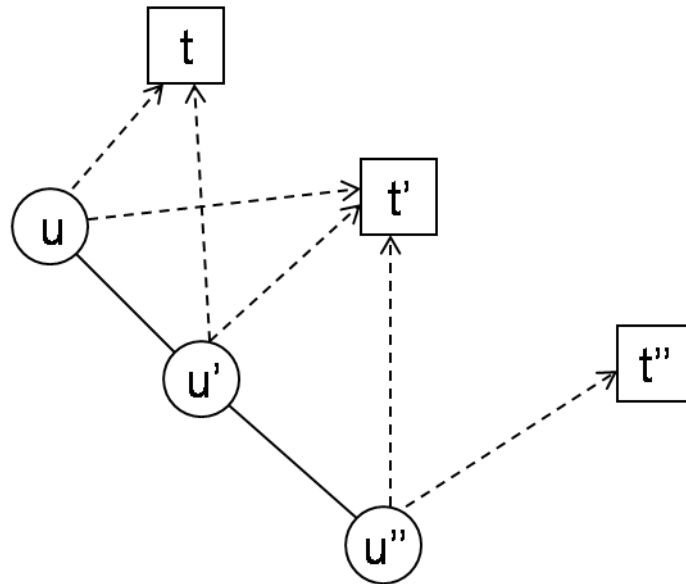
model simulations were used to demonstrate the exchange process between the collective knowledge that is inherent in social tags and the individual knowledge of users. These studies show the potential of social tagging systems.

2.4. Tagging and Social networks

The structure of Social tagging websites can be viewed as a network consisting of three parts, or a network of three different node types: the **U** users, the **R** resources (web sites – URLs), and the **T** tags that the **U** users deploy to tag the **R** web sites. A user can attach one or more tags to a URL. The network that emerges can be graphically illustrated by means of the *I* link between the *u* user and the *r* website that passes through the *t* tag.

A particular class of networks is the bipartite networks, whose nodes are divided into two sets (e.g. users and tags), and only the connection between two nodes in different sets is allowed, as illustrated in Figure 1.

Figure 1. A Bipartite Network made of three users $U=(u,u',u'')$, three tags $T=(t,t',t'')$ and two kinds of links: between users R^U (straight lines), and between users and tags R^T (dashed lines)



Source: Authors

Two kinds of bipartite networks are important because of their particular significance in social, economic, and information systems. One is the so-called *collaboration network*, which is generally defined as a network of actors connected by a common collaboration (Wasserman and Faust, 1994; Scott, 2000). Examples in the social systems are numerous, such as scientists connected by coauthoring a scientific paper or movie actors connected by co-starring in the same movie, or in other fields like on the technological collaboration of software and urban traffic systems. The other one is called the *opinion network* (Maslov and Zhang, 2001; Blattner *et al.*, 2007), in which users connect to the objects that they gather. For example, listeners are connected with the music groups they collected from a music-sharing library such as iTunes, web-users are connected with the webs they collected from a bookmark site such as Delicious, or customers are connected with the books they bought from a site such as Amazon.

A central problem closely related to the opinion network is how to extract the hidden information. The exponential growth of the Internet confronts people with an information

overload. Two landmarks for social research are the use of digital trace data methods and the data analysis in the context of social networks analysis. This paper is an approach to these references.

2.5. Social Web and its impact on Information Retrieval (IR) and Recommender Systems (RS)

During the last few years, the advent of the Social Web has greatly changed the role of Web users, providing them with the opportunity to become key actors and to share knowledge, opinions and tastes thanks to the interaction through online media.

2.5.1. Recommender Systems and Social Web

Introducing folksonomies as the basis for recommendations means that the usual binary relation between users and resources, which is largely employed by traditional RS, changes into a ternary relation between users, resources, and tags, thus becoming more complex to manage.

Different surveys (Dattolo et al., 2012; Kumar and Thambidurai, 2010) analyze the use of social tagging activities for recommendations, focusing their attention particularly on the following aspects:

- RS improvement due to tags: an interesting overview on social tagging systems and their impact on RS is presented in Milicevic *et al.* (2010), while a methodology to improve RS due to Web 2.0 systems, and particularly to social bookmarking platforms, is offered by Siersdorfer and Sizov (2009); moreover, Xia (2010) provides a recommender system model based on tags.
- Role of tag recommendation: the system presented in Rendle and Lars (2010) exploits a factorization model to propose personalized tag recommendations, while Niwa *et al.* (2006) illustrate a strategy used in a Web page recommender system exploiting affinities between users and tags. In addition to these affinities, Durao and Dolog (2009) propose a recommender system exploiting tag popularity and representativeness to recommend web pages.
- Tags and User modeling: since RS rely on a user model to generally personalize recommendations, Wetzker *et al.* (2010) propose an original way to enhance modeling to improve tag recommendation. In a general context, Carmagnola *et al.* (2007) and Simpson and Butler (2009) also illustrate how tag activity can improve user modeling.

2.5.2. Information Retrieval and the Social Web

From a Social IR point of view, i.e. IR that uses folksonomies, tags and particularly the relations between tags have been studied as a novel knowledge base related to information exploited in the IR process:

- As a pull approach, users retrieving information need to understand what information is available to identify which one is relevant to their need. A tag cloud has been used in this context to offer an original and improved visual IR interface (Hassan-Montero and Herrero-Solana, 2006; Bar-Ilan *et al.*, 2010) which allows user browsing information. A more powerful visualization based on tag clusters (Knautz *et al.*, 2010) is considered as better than a tag cloud.
- FolkRank (Hotho *et al.*, 2006) is a new search algorithm for folksonomies. It can also be used to identify communities within the folksonomy that are used to adapt information ranking. This algorithm is inspired from the famous PageRank model from Google. Information ranking (scoring) has also been studied according to query (Liu *et al.*, 2009). Another document ranking based on relationships extracted from the different node types - user, tag, and resource - is illustrated in Bender *et al.* (2008).

- IR have also been improved thanks to folksonomies and two original measures (Bao *et al.*, 2007): SocialPageRank, which computes the popularity of web pages, and SocialSimRank, which calculates the similarity between tags and queries.
- Query expansion based on tag co-occurrence has been studied in Wang and Davison (2008), Biancalana and Micarelli (2009), and Jin *et al.* (2009). Results show that such an approach consistently improves retrieval performance.

In summary, this paper aims to exhibit a methodology to retrieve big data from Web 2.0 and use social network analysis in order to represent the main users and websites around the globalization of agriculture issue in a particular social bookmarking site – Delicious –, along with the most important tags that were employed by users around this topic. An additional aim is examining if it is possible to discover latent pattern links to the activity of collaborative tagging, which could be key in order to provide effective recommendations to different actors.

3. METHODOLOGY

The setting chosen for this study is Delicious (www.delicious.com). Delicious is a prominent example of a social bookmarking system whose content is created, annotated and viewed by its users. Delicious uses a non-hierarchical classification system in which users can tag each of their bookmarks on the Delicious website, and it provides knowledge about the URL marked (Golder and Huberman, 2006; Marlow *et al.*, 2006). Its collective nature makes it possible to view bookmarks added by other users. Delicious also allows users to organize existing tags into groups, called tag bundles. In addition, a Delicious user can follow the latest discoveries from people who share their interests. Hence, we believe that Delicious would be a good setting to investigate how to discover latent structures by using data crawled from a large, social tagging system.

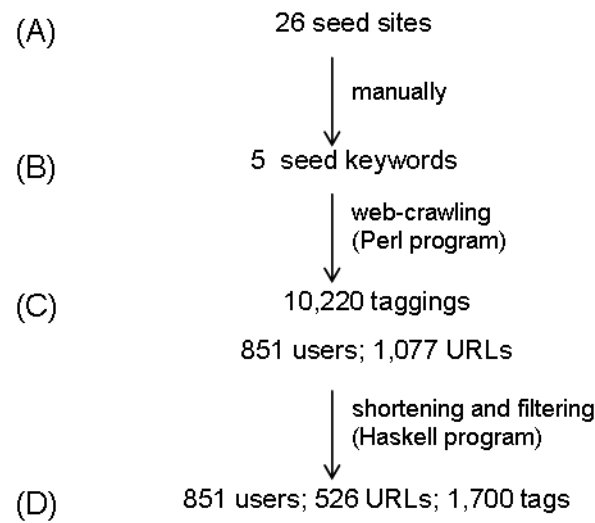
3.1. Data Collection procedure

In Social Bookmarking Services, an annotation typically consists of at least four parts. The link to the resource (e.g. to the website), one or more tags, the user who makes the annotation, and the moment when the annotation is made: user, resource, tag, and time. A user labels a resource with a specific tag at a given moment. This paper is less interested in when the annotation took place than in the co-occurrence of users, resources and tags (*user, resource, tag*). The dataset collected is written as: $\mathbf{U} = \{u1; u2; : : : ; uK\}$, $\mathbf{R} = \{r1; r2; : : : ; rM\}$, and $\mathbf{T} = \{t1; t2; : : : ; tN\}$ as the set of K users, M URLs, and N tags, respectively.

We built the network of globalization of agriculture using a combination of search techniques proposed for researching “issue networks” (Rogers and Zelman, 2002): associative reasoning, whereby educated guesses are made about relevant issues and related websites; public trust logics, finding groups commonly linked to by players trusted to be important in the debate; media stories, following links from an authoritative news source; and search engine crawls of key words.

The process to retrieve the data and of representing the *Delicious* community as a network follows a procedure that we present in Figure 2:

Figure 2. Data Collection Procedure



Source: Authors

Firstly (A), following links from an authoritative news source, we identify the search attributes on the basis of an original sample of a set of 26 web pages (Appendix 1), according to the Wikipedia definition of “critics of globalization³.” We could have randomly selected them from other sites or sources, but we focus on this page because it is one of the most popular Web 2.0 pages, and because it has a high reputation⁴. On the other hand, we could have chosen another starting point, and it may have changed the keywords, but that was not relevant at the time of this study. The important thing in this phase was to have an authoritative news source as baseline to find, as a first step, keywords connected to globalization, and, as a second step, the idea of ‘globalization of agriculture’ as the main issue for the present illustration. We propose future research considering other starting points.

Based on a detailed study of site content, we selected main concepts from external links to these webpages (B). The search attributes were extracted manually from the website homepages and from the tag clouds or the topics that appear on the homepage. Following Rogers and Zelman (2002), we decided to identify these keywords through associative reasoning, whereby we made educated guesses about relevant issues and finding key concepts commonly linked to all seed websites. Finally, we found a set of attributes related with agriculture – agriculture, food, organic, and GMO – that had been grouped along with the word globalization under the denomination of “globalization of agriculture”. Other different concepts were rejected at this step as they were not directly associated to agriculture, though could be linked to globalization. The decision to proceed at this stage with the manual extraction of these keywords, as opposed to using another automatic selection, was taken due to the importance that we give to the researcher in this stage, due to his/her expertise.

In a third stage (C), we gather the raw data sample of all the users’ records, URLs and tags available for the four tag pairs around the globalization main tag – globalization+agriculture; globalization+food; globalization+organic; globalization+GMO – identified by crawling through the social bookmarking website Delicious using a Perl-developed⁵ web crawler⁶. The data-gathering process from the four attributes covered the period between 22 April 2011

³ http://es.wikipedia.org/wiki/Categor%C3%ADa:Cr%C3%ADticos_de_la_globalizaci%C3%B3n (retrieved 02.04.2011).

⁴ <http://www.alexa.com/topsites> (consulted 13.09.2012).

⁵ José Carpio, Intelligent Systems and Data Mining research group from University of Huelva, Spain (TIC-198).

⁶ A web crawler is a program that automatically traverses a web site (e.g. Delicious) by retrieving all users, URLs and tags that match the search criteria.

and 21 May 2011 (one completed month), and produced 10,220 taggings that involved 851 users on 1,077 URLs and 1,720 tags.

Finally, we developed a program in Haskell⁷ to reduce the amount of data (D) by cutting the URLs and using key words, including the identification of synonyms, and eliminating words with capital letters and derivatives such as words in plural. Both software programs, Perl and Haskell, are free software and they are in line with Web 2.0 philosophy. The definitive data constituted 851 users, 526 URLs and 1,700 tags.

Table 1 shows the key words and the frequency with which they occurred around the topic of globalization of agriculture.

Table 1. Keywords Used in the topic “Globalization of agriculture”

Search attributes used	Number of resulting tags (I+II)	More frequent Tags / Main Tags
Globalization (I) + agriculture (II)	1,116	Food (268), economics (176), environment (145), politics (85), trade (81), sustainability (70)
Globalization (I) + food (II)	1,682	Economy (180), economics (171), environment (122), sustainability (78), politics (60)
Globalization (I) + organic (II)	22	Business (3), fair-trade (3)
Globalization (I) + GMO (II)	54	Food (13), agriculture (12)

Source: Authors, from Delicious dataset (from 22-4-2011 to 21-5-2011)

Note: Each user can label each URL with a different number of tags

3.2. Analysis procedure

We are interested in computing the proportion of links preferentially created towards some kind of agents, relative to the proportion of these agents in the whole network (Barabási *et al.*, 2002).

Node centrality, or the identification of the nodes that are more “central” than others, is a fundamental part of network analysis (Freeman, 1979; Bonacich, 1987; Borgatti, 2005; Borgatti *et al.*, 2006). It is a network level property which gives a rough idea of the node’s social power based on how well it “connects” to the network.

The literature on social networks conceptualizes centrality in many different ways (Freeman, 1979). The degree of a node is the number of ties it has; specifically, the number of direct connections individuals have with others in the group, which reflects the level of activity. The node with the highest degree exerts influence (or authority). The directed networks differentiate between *In-degree* and *Out-degree*. *In-degree* is the number of incoming ties that reflect the popularity of a website. As a result, the prominent, well-connected members (those with a high degree of centrality) are usually the opinion leaders. *Out-degree* is the number of outgoing ties which determine if a particular user is an active or passive participant within the network.

Our aim is also to describe, in a simple manner, the resulting tag structure as a tag cloud that depicts the interests of Web users. In this way, social tags are able to provide visual representations of the conceptual structure of an issue, which is built upon the knowledge of individuals who belong to a large Web community.

⁷ Antonio Regidor, Agricultural Economics research group from University of Huelva, Spain (SEJ-110).

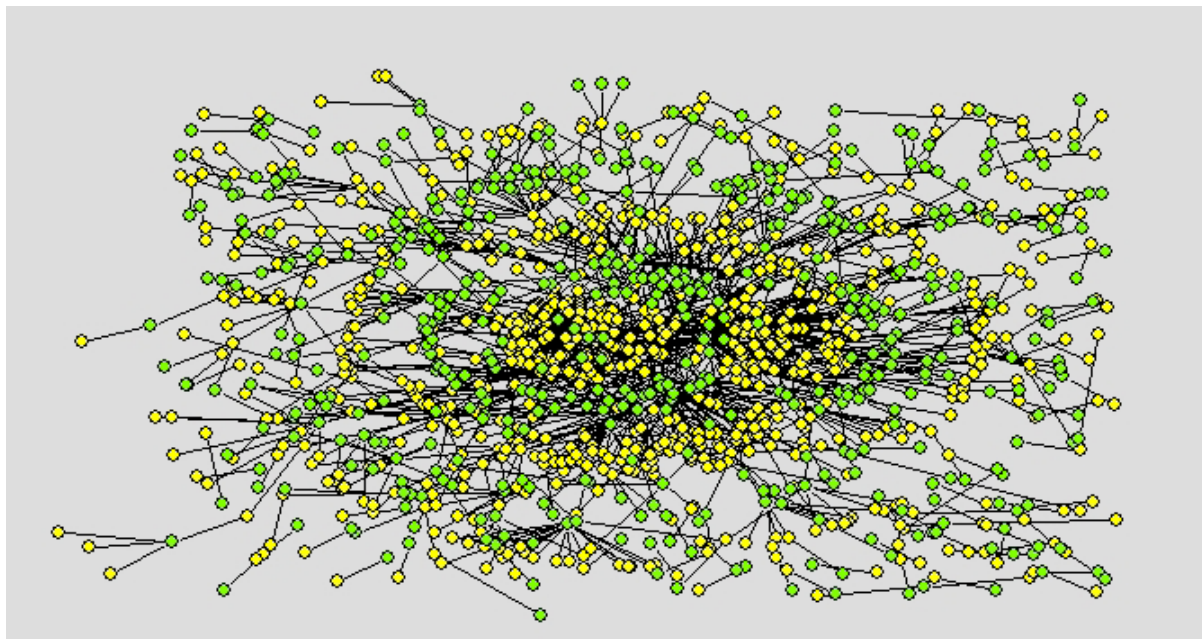
4. RESULTS

In this section, we present some empirical analysis of the network for the globalization of agriculture and a first approach to the tags associated to URLs found after the retrieval of information process. The next section reviews that analysis with reference to the three elements that compose the network: users, URLs, and tags. We use Social Network Analysis techniques (with the help of the software Pajek, which is better for big series of data than others such as Ucinet) to build the network⁸ that we have called “globalization of agriculture”⁹. Through the connections among three key elements (Users write Tags to characterise URLs), different calculations were made. The following pages focus on two different approaches that allow us to find visible and invisible patterns when a Delicious bookmarking system’s user is simply using Delicious. In that sense, we discovered latent structures. Firstly, we pay attention to power that emerges from the network – main users and websites. In the second section, we focus on concrete tags that were elaborated by users describing URLs and their importance.

4.1. Centralization: Authority

Centralization is a network-level property that broadly measures the distribution of power or prominence amongst actors in a given network (Hanneman, 2005). We calculate centralization by first computing a particular node-level degree centrality. Each time a user labels a particular URL, the intersection between user and URL was coded by 1. In the “useràURL” directed network (Figure 3), we calculate the indegree from each URL as the sum of total inbound links, and, in the same way, the outdegree from each user as the sum of the outbound links.

Figure 3. Hyperlink Network Energy Kamada-Kawai Map. Bipartite Network useràurl



Source: Authors by Pajek
Note: Users in Yellow color; URLs in Green color

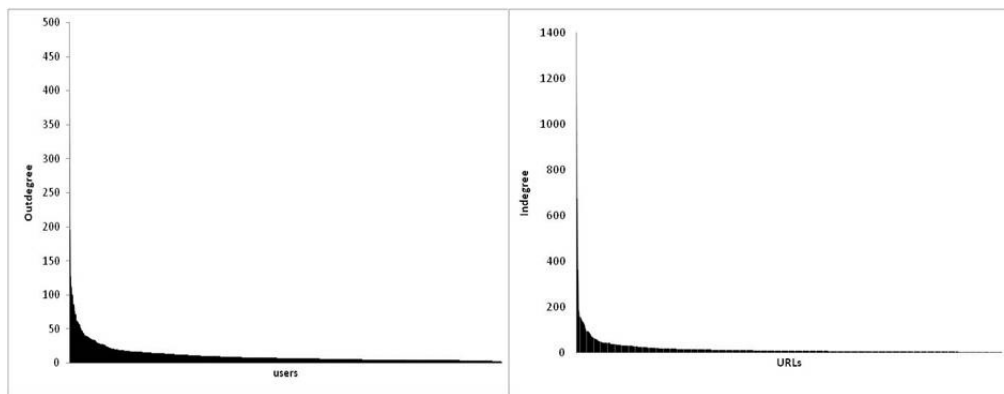
Most important Users and URLs are placed in the middle of the Figure 3, where density of connections is higher. Figure 4 shows the degree of variability in the website and user

⁸ In this work, we use Social Network Analysis for calculating some measures, but we do not show the visualization of the network.

⁹ This network is based on the original extraction of data from Delicious that took into account search attributes connected to globalization and agriculture (see Table 1).

centrality scores according to indegree and outdegree. As expected, the network is highly centralized within a few nodes because only 10 URLs from 526 (1.90%) account for 32.29% of the links to URLs¹⁰ and only 10 users from 851 (1.17%) account for 14.05%¹¹. This imbalance is not unusual given what we know about the long-tail distributions and the scale-free properties of the web. The power law is a defining characteristic of large-scale networks such as the web (e.g. Barabási and Albert, 1999), which implies a high degree of network centralization; it also proffers the empirical starting point for the question: Why?/ How come a few users and websites are better connected than the majority?

Figure 4. Hyperlink Network. 851 users arranged in rank order by number of outbound links and 1,077 URLs arranged in rank order by number of inbound links



Source: Authors, from data retrieved from Delicious dataset

Table 2 shows the 10 most centralized websites, and we can see that nine are media-based (online newspapers such as The New York Times, BBC, The Guardian, Washington Post, Financial Times, Reason, The Nation, Spiegel and The Economist).

Table 2. Top Authoritative Sites in the hyperlink network

	Indegree		Outdegree	
1	1203	http://www.nytimes.com/	433	/mritiunjoy
2	674	http://news.bbc.co.uk/	195	/laura208
3	365	http://www.guardian.co.uk/	127	/rd108
4	186	http://www.washingtonpost.com/	112	/amaah
5	158	http://www.ft.com/	111	/thepouncer
6	154	http://www.reason.com/	100	/anilius
7	147	http://www.thenation.com/	100	/emmarlyb
8	137	http://www.spiegel.de/	87	/adorngeography
9	136	http://www.foodfirst.org/	86	/pagolnari
10	130	http://www.economist.com/	85	/freemanlc

Source: Authors, from data retrieved from Delicious dataset

¹⁰ These ten URLs got 3,290 inbound links from a total of 10,190.

¹¹ These ten Users got 1,436 outbound links from a total of 10,219.

This table also shows the users with a greater degree of centrality. We observe that the user, *mritiunjoy*, plays a very important role in the network. We could take it a step further to know more about this central user and his possible connection to the links. For example, we discover on a Delicious web page that *mritiunjoy* joined Delicious on 12 March 2007, and, to date, he has 10,020 links and is following 38 users. However, on the internet, we also discover that *mritiunjoy* – Mritiunjoy Mohanty – is a professor at the Indian Institute of Management Calcutta, India, and his Research Interests are the Political Economy of growth and development.

Remembering Diani (2003), the analysis has identified valuable nodes (websites). Its value is not only due to the links that they receive (its instrumental nature) but also to the profile of these organizations (newspapers that channel big quantities of resources – information), due to the quality of the links. This last particularity (quality of links), added to the first (instrumentality), determines that these URLs are central and have some authority. As a consequence, they could be relevant to produce currents of opinion. Most URLs bookmarked are singulars, because they could create or modify opinions.

In addition, the results suggest that the most centralized users (those with the highest number of links) do that because they have other interests than simply bookmarking, sharing, or labelling a resource.

4.2. Node Tags: Users producing Tags

In this section we explore the collective tag structure (excluding the key search words, such as globalization, agriculture, food and organic, and GMO) in an attempt to identify topics around our main theme. A natural approach to identifying the topical groupings in a tag network is to use tag clouds. Thus, Figure 5 shows a selection of highly descriptive keywords for the globalization of agriculture system in Delicious. Cluster keywords were automatically identified.

The clouds were produced with Wordle (Viégas *et al.*, 2009), where the sizes of the terms in the tag clouds are proportional to the weights, with the top 25 highest weighted tags included. The resulting key topics were *economics* and the *environment*, which were the main keywords used by users to describe or characterise in Delicious the topic ‘globalization of agriculture’.

Figure 5. Tag Cloud for Agriculture Globalization Network Identified on the Delicious Data Set



Source: Authors

After this brief description, we want to clarify that these are not the unique results that could have been exposed here, after the complex process described for the retrieval of information. We have chosen them as good examples that give sense to the operation of crawling big data, as first points of departure for knowing a bit more about the topic of globalization of agriculture, and for demonstrating the way that people describe and share

websites about this issue through a modern and collaborative process of tagging. In the next section, a discussion about the results will be presented in order to know more about some alternative analyses, reflections, etc.

5. DISCUSSION

5.1. Centrality and Power

Hanneman (2005) reminds us: “a very simple, but often very effective measure of an actor’s centrality and power potential is their degree”. In our case, as indegree concern URLs and as these represent some kind of collective actor, the determination of centrality measures make sense. Higher indegrees mean that the URLs are chosen by more users (they received more links). It is evident that the New York Times, in this network of globalization of agriculture in Delicious, greatly surpasses other URLs (with 1,203 inbound links, followed by the BBC website with 674 links). Most cited, recommended, or considered websites with regards to a topic occupy a central place and have an important role in the process of dissemination of news, events, trending topics, ideology, culture, etc. Knowing this previously hidden hierarchy is also very useful for different socioeconomic reasons. At the same time, this identification of key collective actors (represented here through URLs) allows a better comprehension of leadership, influence process, and power-related structures. For social practitioners, it is a good way to identify key informants in a community through whom disseminating useful and important information occurs.

Indegree in Table 2 also shows a very unequal distribution of power of the URLs cited by users in the topic of globalization of agriculture, represented by an important accumulation of inlinks. Only 10 URLs represent an important, accumulated indegree).

Regarding other actors in the networks, the users, for the identification of key actors that disseminate and share URLs, as the previously cited *Mritiunjoy*, it is important to determine from where key elements that structure the network emerge. Is it possible to explain why ‘that’ greatly important actor is in the network of globalization of agriculture? Key actors in this type of network could configure and reconfigure the evolution of the network, structure, and even manipulate the type of interchange of resources in Delicious or in similar bookmarking sites.

Their prominence has something to say to social researchers, practitioners, etc. Is it by chance? Are most prominent actors in the type of website like Delicious corresponding to a profile of very active and participative people? Do they usually work (or have as a hobby) in this area, which could explain the accumulation and tagging of so many URLs in Delicious? These and other questions could be answered in further steps of the research, depending on the concrete goals at each moment.

5.2. Central Tags: Users producing Tags

In the process of linking URLs in Delicious, the majority of users selected tags suggested by the website or added new tags in a creative way for describing or qualifying the URLs that they were recommended. A ‘tag cloud’ was built in order to have a visual approach to the language that was employed by users in their descriptions. As we have focused on the retrieval of information regarding the topic of globalization of agriculture, the question now is to wonder what we could know about this topic through the extracted tags. From a total of 1,700 tags, two words were the main ones, as most cited when users labelled URLs. It is important to note that each user could label a URL with an unlimited number of tags (average 12 tags per user, max 433 and min 2). The most frequently used tags were the words: ‘economics’ (350 citations out of 1,700 tags, 20.6%) and ‘environment’ (273, 16%).

Other, very frequent tags were: sustainability (153), politics (152), economy (144), trade (131), business (99), poverty (97), culture (84), farming (84), africa (83), health (78), and development (76); in relative terms, these 13 tags represent one out of four labelled tags surrounding the topic (25.9%).

Discovering the importance of these words make us wonder not only of the reasons for the prominence of the first two tags regarding the globalization of agriculture but also for the rest. In addition, as 1,700 tags were also found qualifying and describing webpages regarding the topic of globalization of agriculture, some analyses are possible to know if some tags are used on an interchangeable basis, considered as synonyms, as was reported as one of the problems of collaborative tagging or the suggested vague-meaning problem (Kroski, 2005; Golder *et al.*, 2006; Hope *et al.*, 2007; Marchetti *et al.*, 2007).

The same thing could be done regarding 'term variations'. For instance, economy and economics are two important words in the topic of globalization of agriculture. Are these tags used in a similar or equivalent way at tagging? Why is the word economics sometimes used, and why, at other times, is economy used? Are they used in the same way at classifying the URLs?

By limiting the analysis to a particular period of time, tags could be associated to the use of language at a particular moment. They could even be a good representation of the ideological and terminological approach to the topic in the international arena, at that moment, and be useful for the study of the evolution and usage of language in a topic over time. On the other hand, the use of some tags at classifying URLs connected to the globalization of agriculture, and the distinction among users in the way they use some words as labels, could yield other types of results. Are scientific users utilizing the same tags as other professionals or general users? Perhaps different scientists or other users produce different labels around the same topic. Perhaps the first people tagging a topic are influencing the following tags that are incorporated in Delicious, etc. Nevertheless, some of this analysis can be limited by the information available from users. Other possible studies, going into more depth, may retrieve the pages that were labelled and undertake a content analysis to determine what kind of content is labelled through concrete tags. This is a cognitive way to see how users summarize and represent in short and definite words what could be broad and detailed content. It could be a way to see keywords that remind them about something. Through this, different applications could be suggested (for instance, in advertising, mobilizing, etc.).

Although this article has been more focused on the retrieval and illustration aspects, we have not shown networks of tags linked by users or by URLs tagged by users; a complementary, detailed analysis could help to identify users that have the same patterns at tagging or URLs that were similarly labelled. This opens a door to study structural equivalences and considering, for instance, applications for particular types of users.

Other questions emerge, as to why some labels are present but not others? Is it a question of language usage? Is it a question of traditions at tagging in Web 2.0? Is it a fashion? Supposedly, if people use Delicious for collaborative reasons, tags must at least be understandable for other users, unless the user prioritizes their own usage of Delicious.

6. CONCLUSIONS AND FUTURE RESEARCH

The main objective of this paper was presenting a methodology to use big data from Web 2.0 in social research. We had an interest in illustrating the extraction of data from a social bookmarking site (*Delicious*) and showing the type of results that this type of analysis could offer to social scientists. As it concerned big series of data crawled from a large social tagging

system, the analysis could have an important impact in the discovering of latent patterns, which form the basis in order to provide effective recommendations to different actors.

Our approach represents an important first step towards the development of empirical techniques capable of automatically differentiating groups of individuals with common interests, and individuals who occupy a more central position. This research is also of interest to make recommendations on the knowledge base of individual interests.

In addition, our analysis offers a previously unavailable understanding in the definition of recommendation services. To be able to identify a short list of the most centralized users and ties is extremely useful for researchers attempting to understand a community of more than a thousand links. This is particularly important for researchers interested in formulating strategies for intervention and mobilization, but practitioners and companies could also make use of this. The discovery of the central elements in a network (users and URLs) and the tags employed by users could be a key to the design of future strategies for the dissemination of messages, while also helping to achieve more success in communications, such as making use of important keywords to attract greater attention, etc. At the same time, if we know other interests of the users belonging to a network - through, for instance, other webpages that they link, and others tags that they label - we would be able to make recommendations, as done by other systems such as Amazon.

With regards to the process of retrieval of information, the method presented here was somewhat complex but easy to apply if there is some computer knowledge. Nevertheless, working in interdisciplinary teams could greatly help to develop this kind of knowledge, as it was in our case. Though the technical process described was successful, improvements are necessary in the future, at least regarding the retrieval methods and the implementation of IR and RS techniques in social commerce and social media contexts.

On the other hand, the relation between users and resources, which is largely employed by traditional Recommender Systems, changes into a ternary relation between users, resources, and tags, which is more complex to manage. This article has laid the first stone in the difficult process of understanding and discovering patterns in the process that characterizes users tagging URLs for collaborative reasons. The application was made under the topic 'globalization of agriculture'.

Some of the first contributions in the area of globalization of agriculture were that tags used to describe URLs in the Delicious's social bookmarking site were mostly concentrated around a few terms. The approximation to this topic in the future through other bookmarking sites (for instance, dominated by Spanish-speaking users) will allow the researchers to know if the recommended URLs are again media-based or are even the same webpages; or, for instance, if there is a semantic change concerning tags used for describing and classifying URLs.

Lastly, we do not want to close the article without clarifying that researching the topic of globalization of agriculture in a systematic and broad way may require the consideration of other starting points for the retrieval of information, at least to compare and contrast results. This is a limitation of this work. Nevertheless, the search yielded 1,700 different tags that have been used in the period of only a month to qualify and describe the phenomenon. It is a large number of tags, but we found a great concentration of them, as the centrality measure showed. The same argument could be made regarding URLs, in that they were extremely concentrated in mass media. Of course, other analysis in the future could be made with a longer period of time, along with other explorations. This is only a beginning.

ACKNOWLEDGEMENTS

Special thanks to José Carpio (University of Huelva) for his help in collecting the data used for this study by the Perl program. We are also deeply grateful to Antonio Regidor, who helped us with filtering data with his expertise in Haskell computing. We also have received some interesting comments to a draft of this work from Ainhoa de Federico (University of Toulouse 2) and Teresa González (University of Huelva). The methodology was presented as working paper at the CIEO, Centre for Spatial and Organizational Dynamics, at the University of the Algarve. At the Universidad of Huelva, a preliminary version of this text was discussed with other colleagues. We want to thank all of them for their suggestions, especially to Marielba Zacarias and Paula Ventura Martins (University of the Algarve), and Andrea Capilla and Mónica Carmona (University of Huelva).

REFERENCES

- Ackland, R., and O'Neil, M. (2011) Online Collective Identity: The Case of the Environmental Movement. *Social Networks*. **33**: 177-190.
- Ames, M. and Naaman, M. (2007). Why we tag: motivations for annotation in mobile and online media. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. San Jose, California, USA.
- Bao, S., Xue, G., Wu, X., Yu, Y., Fei, B., and Su, Z. (2007). Optimizing web search using social annotations. In: *Proceedings of the 16th International Conference on World Wide Web, WWW 2007*. New York: ACM, pp. 501-510
- Bar-Ilan, J., Zhitomirsky-Geffet, M., Miller, Y., and Shoham, S. (2010). Tag, cloud and ontology based retrieval of images. In: *Proceeding of the Third Symposium on Information Interaction in Context, IiX 2010*. New York: ACM, pp. 85-94.
- Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*. **286**: 509-512.
- Barabási, A.L., Jeong, H., Ravasz, E., Neda, Z., Vicsek, T., and Schubert, A. (2002). Evolution of the Social Network of Scientific Collaborations. *Physica A*. **311**: 590-614.
- Bateman, S., Gutwin, C., and Nacenta, M. (2008). Seeing things in the clouds: The effect of visual features on tag cloud selections. In: *Proceedings of the ACM conference on hypertext and hypermedia*. New York: ACM Press, pp. 193-202.
- Bender, M., Crecelius, T., Kacimi, M., Michel, S., Neumann, T., Parreira, J.X., Schenkel, R., and Weikum, G. (2008). Exploiting social relations for query expansion and result ranking. In: *Data Engineering for Blogs, Social Media, and Web 2.0, ICDE 2008 Workshops*. Cancun, Mexico, pp. 501-506.
- Biancalana, C., and Micarelli, A. (2009). Social tagging in query expansion: A new way for personalized web search. In: *Proceedings IEEE CSE 2009, 12th IEEE International Conference on Computational Science and Engineering*. Vancouver: IEEE Computer Society, pp. 1060-1065.
- Blattner M., Zhang Y.-C., and Maslov S. (2007). Exploring an opinion network for taste prediction: An empirical study. *Physica A*. **373**: 753-758.
- Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*. **92**: 1170-1182.
- Borgatti, S.P. (2005). Centrality and Network Flow. *Social Networks*. **27**(1): 55-71.
- Borgatti, S.P., Carley, K., and Krackhardt, D. (2006). On the Robustness of Centrality Measures Under Conditions of Imperfect Data. *Social Networks*. **28**(2): 124-136.

- Carmagnola, F., Cena, F., Cortassa, O., Gena, C., and Torre, I. (2007). Towards a tag-based user model: How can user model benefit from tags? In: C. Conati, K. McCoy and G. Paliouras (eds.), *UM 2007. LNCS (LNAI)*. **4511**: 445–449. Springer, Heidelberg.
- Carmel, D., Zwerdling, N., Guy, I., Ofek-Koifman, S., Har'el, N., and Ronen, I. (2009). Personalized social search based on the user's social network. In: *Proceeding of the 18th ACM conference on information and knowledge management*. Hong Kong, China, pp. 1227–1236. 1646109: ACM. doi: 10.1145/1645953.1646109.
- Cattuto C., Loreto, V., and Pietronero, L. (2007a). Semiotic dynamics and collaborative tagging. In: *Proceedings of the National Academy of Sciences*. **104**: 1461–1464.
- Cattuto, C., Schmitz, C., Baldassarri, A., Servedio, V.D.P., Loreto, V., Hotho, A., Grahl, M., and Stumme, G. (2007b). Network properties of folksonomies. *AI Communications*. **20**(4): 245–262.
- Cress, U., Held, C., and Kimmerle, J. (2012). The Collective Knowledge of Social Tags: Direct and Indirect Influences on Navigation, Learning, and Information Processing. *Computers and Education*. <http://dx.doi.org/10.1016/j.compedu.2012.06.015>
- Dattolo, A., Ferrara, F., and Tasso, C. (2012). On social semantic relations for recommending tags and resources using folksonomies. In: Z.S. Hippe, J.L. Kulikowski and T. Mroczek (eds.), *Human-Computer Systems Interaction. Backgrounds and Applications*. **98**: 311–326.
- Diani, M. (2003). 'Leaders' or Brokers? Positions and Influence in Social Movement Networks. In: M. Diani and D. McAdam (eds.), *Social Movements and Networks. Relational Approaches to Collective Action*. Oxford University Press. New York.
- Domínguez J. A, Merino, B., and Aledo, A. (2010). The Identity of Sociology or what to do when the Universe is Unknown: Qualitative solutions against the quantitative obsession, *Spacial and Organizational Dynamics*, **5**: 7-22.
- Duan W., Gu, B, and Whinston, A.B. (2008). Do online reviews matter?—an empirical investigation of panel data. *Decision Support Systems*. **45**: 1007–1016.
- Durao, F., and Dolog, P. (2009). A personalized tag-based recommendation in social web systems. In: *Proceedings of International Workshop on Adaptation and Personalization for Web 2.0 (AP-WEB 2.0 2009)*. Trento, Italy.
- Freeman, L.C. (1979). Centrality in Social Networks: Conceptual Clarification. *Social Networks*. **2**: 215-239.
- Fu, W. (2008). The microstructures of social tagging: A rational model. In: *Proceedings of the ACM conference on computer supported cooperative work*. San Diego: ACM, pp. 229–238.
- Fu, W., Kannampallil, T. G., and Kang, R. (2009). A semantic imitation model of social tag choices. In: *Proceedings of the IEEE international conference on computational science and engineering*. Los Alamitos: IEEE Computer Society, pp. 66–72.
- Fu, W., Kannampallil, T. G., and Kang, R. (2010). Facilitating exploratory search by model-based navigational cues. In: *Proceedings of the international conference on intelligent user interfaces*. Hong Kong: ACM, pp. 199–208.
- Garrido, M., and Halavais, A. (2003). Mapping Networks of Support for the Zapatista Movement: Applying Social Network Analysis to Study Contemporary Social Movements. In: M. McCaughey and M. Ayers (eds.), *Cyberactivism: Online Activism in Theory and Practice*, London: Routledge, pp. 165-184.
- Golder, S., and Huberman, B. A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*. **32**(2): 198–208.
- Hanneman, R.A, and Riddle, M. (2005). *Introduction to social network methods*. University of California, Riverside, CA (published in digital form at <http://faculty.ucr.edu/~hanneman/>)

- Hassan-Montero, Y., and Herrero-Solana, V. (2006). Improving tag-clouds as visual information retrieval interfaces. In: *InScit 2006: International Conference on Multidisciplinary Information Sciences and Technologies*. Mérida, Spain.
- Herring, S., Kouper, I., Scheidt, L., and Wright, E. (2004). Women and Children Last: The Discursive Construction of Weblogs. In: L. Gurak, S. Antonijevic, L. Johnson, C. Ratliff and J. Reyman (eds.), *Into the blogosphere: Rhetoric, community, and culture of weblogs* (available at <http://blog.lib.umn.edu/blogosphere/>).
- Hotho, A., Jaschke, R., Schmitz, C., and Stumme, G. (2006). Information retrieval in folksonomies: Search and ranking. In: Y. Sure and J. Domingue (eds.), *ESWC 2006. LNCS. 4011*: 411–426. Springer, Heidelberg.
- Janetsko, D. (2009). Nonreactive Data Collection on the Internet. In: N. Fielding, R. M. Lee and G. Blank (eds.), *SAGE Handbook of Online Research Methods*. London: Sage.
- Jin, S., Lin, H., and Su, S. (2009). Query expansion based on folksonomy tag co-occurrence analysis. In: *2009 IEEE International Conference on Granular Computing*. IEEE, Los Alamitos, pp. 300–305.
- Jones, S. Millermaier, S., Goya-Martínez, M., and Schuler, J. (2008). Whose Space is MySpace? A Content Analysis of MySpace Profiles. *First Monday. 13*(9).
- Kammerer, Y., Nairn, R., Pirolli, P., and Chi, E.H. (2009). Signpost from the masses: Learning effects in an exploratory social tag search browser. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. New York: ACM, pp. 625–634.
- Kang, R., and Fu, W. (2010). Exploratory information search by domain experts and novices. In: *Proceedings of the international conference on intelligent user interfaces*. Hong Kong: ACM, pp. 329–332.
- Kang, R., Kannampallil, T., He, J., and Fu, W. (2009). Conformity out of diversity: Dynamics of information needs and social influence of tags in exploratory information search. In: D. Schmorow, I. Estabrooke, and M. Grootjen (eds.), *Foundations of augmented cognition. Neuroergonomics and operational neuroscience*. Berlin/Heidelberg: Springer, pp. 155–164.
- Kannampallil, T., and Fu, W. (2009). Trail patterns in social tagging systems: Role of tags as digital pheromones. In D. Schmorow, I. Estabrooke, and M. Grootjen (eds.), *Foundations of augmented cognition. Neuroergonomics and operational neuroscience*. Berlin/Heidelberg: Springer, pp. 165–174.
- Knautz, K., Soubusta, S., and Stock, W.G. (2010). Tag clusters as information retrieval interfaces. In: *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences, HICSS 2010*. IEEE Computer Society, Washington, DC, pp. 1–10.
- Kumar, A., and Thambidurai, P. (2010). Collaborative web recommendation systems a survey approach. *Global Journal of Computer Science and Technology. 9*(5): 30–36.
- Kwai Fun, I.P.R., and Wagner, C. (2008). Weblogging: a study of social computing and its impact on organizations. *Decision Support Systems. 45*: 242–250.
- Liu, D., Hua, X.S., Wang, M., and Zhang, H. (2009). Boost search relevance for tag-based social image retrieval. In: *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo, ICME 2009*. IEEE Press, Piscataway, pp. 1636–1639.
- Marlow, C., Naaman, M., Boyd, D., and Davis, M. (2006). HT06, tagging paper, taxonomy, Flickr, academic article, to read. In: *Proceedings of the seventeenth conference on Hypertext and hypermedia*. Odense, Denmark.
- Maslov, S., and Zhang, Y.-C. (2001). Extracting Hidden Information from Knowledge Networks. *Physical Review Letter. 87*: 248701-248705.

- Mathes, A. (2004). Folksonomies - cooperative classification and communication through shared metadata (Acceded in 13th of August 2011, on the Web site of: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>)
- Milicevic, A., Nanopoulos, A., and Ivanovic, M. (2010). Social tagging in recommender systems: a survey of the state-of-the-art and possible extensions. *Artificial Intelligence Review*. **33**: 187–209. 10.1007/s10462-009-9153-2
- Millen, D., Yang, M., Whittaker, S., and Feinberg, J. (2007). Social bookmarking and exploratory search. In: L. Bannon, I. Wagner, C. Gutwin, R. Harper and K. Schmidt (eds.), *Proceedings of the European conference on computer-supported cooperative work*. London: Springer, pp. 21–40.
- Nardi, B., Schiano, D., and Gumbrecht, M. (2004). Blogging as Social Activity, or, Would You Let 900 Million People Read Your Diary?. In: *Proceeding CSCW '04 Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM: New York, pp. 222–231.
- Niwa, S., Doi, T., and Honiden, S. (2006). Web page recommender system based on folksonomy mining for ITNG 2006 submissions. In: *Proceedings of the Third International Conference on Information Technology: New Generations*. IEEE Computer Society, Washington, DC, pp. 388–393.
- O'Reilly, T. (2007). *What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software*. *International Journal of Digital Economics*. **65**: 17–37.
- Parameswaran, M., and Whinston, A. (2007a). Research issues in social computing. *Journal of the Association for Information Systems*. **8**: 336–350.
- Parameswaran, M., and Whinston A. (2007b). Social computing: an overview. *Communications of the Association for Information Systems*. **19**: 762–780.
- Rattenbury T., Good, N., and Naaman, M. (2007). Towards extracting Flickr tag semantics. In: *WWW '07: Proceedings of the 16th international conference on World Wide Web*. ACM Press. New York, USA, pp. 1287–1288.
- Rendle, S., and Lars, S.T. (2010). Pairwise interaction tensor factorization for personalized tag recommendation. In: *Proceedings of the Third ACM International Conference on Web Search and Data Mining, WSDM 2010*. ACM. New York, USA, pp. 81–90.
- Rivadeneira, A.W., Gruen, D.M., Muller, M.J., and Millen, D.R. (2007). Getting our head in the clouds: Toward evaluation studies of tagclouds. In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM Press. New York, USA, pp. 995–998
- Rogers, R., and Zelman, A. (2002): Surfing for knowledge in the information society. In: G. Elmer and L. Rowman (eds.), *Critical Perspectives on the Internet*. Lanham, MD.
- Schueler, B., Sizov, S., and Staab, S. (2007). Management of Meta Knowledge for RDF Repositories. In: *International Conference on Semantic Computing*, pp.543-550. <http://doi.ieeecomputersociety.org/10.1109/ICSC.2007.79>
- Schuler D. (1994). Social computing. *Communications of the ACM*. **37**: 28–29.
- Scott J. (2000). *Social Network Analysis*. Sage Publication. London.
- Shneiderman B., Bederson, BB, and Drucker, S.M. (2006). Find that photo!: interface strategies to annotate, browse, and share. *Communications of the ACM*. **49**(4): 69–71.
- Shumate, M., and Dewitt, L. (2008). The North/South Divide in NGO Hyperlink Networks. *Journal of Computer Mediated Communication*. **13**(2): 405–428 (available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2008.00402.x/pdf>).

- Siersdorfer, S., and Sizov, S. (2009). Social recommender systems for web 2.0 folksonomies. In: *Proceedings of the 20th ACM Conference on Hypertext and Hypermedia, HT 2009*. ACM. New York, USA, pp. 261–270.
- Simpson, E., and Butler, MH. (2009). Analyzing Communal Tag Relationships for Enhanced Navigation and User Modeling. *Collaborative and Social Information Retrieval and Access: Techniques for Improved User Modeling*. IGI Global, pp. 43-64. doi:10.4018/978-1-60566-306-7.ch003
- Smith, G. (2008). *Tagging: People-powered metadata for the social web*. New Riders Press. Berkeley, CA.
- Stiglitz, J.E. (2006). *Making Globalization Work*. WW Norton and Company. New York, NY.
- Trant, J. (2009). Studying social tagging and folksonomy: A review and framework. *Journal of Digital Information* **10**(1). <<http://journals.tdl.org/jodi/article/view/269/278>> Retrieved on 01.01.12.
- Van Velsen, L., and Melenhorst, M. (2009). Incorporating user motivations to design for video tagging. *Interacting with Computers*. **21**: 221–232.
- Vander Wal, T. (2004). *You down with folksonomy?* Accessed in 17th of July 2011, on the web site of: <http://www.vanderwal.net/random/entrysel.php?blog=1529>
- Vander Wal, T. (2007). Folksonomy coinage and definition. Accessed in 17th of July 2011, on the web site of: <http://vanderwal.net/folksonomy.html>
- Vickery G., Wunsch-Vincent, S. (2007). *Participative web and user-created content: Web 2.0, wikis and social networking*. OECD Publishing.
- Viégas, F.B., Wattenberg, M., and Feinberg, J. 2009. Participatory Visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics*. **15**(6): 1137-1144.
- Wang, J., Davison, B.D. (2008). Explorations in tag suggestion and query expansion. In: *Proceeding of the 2008 ACM Workshop on Search in Social Media, SSM 2008*. ACM. New York, USA, pp. 43–50.
- Wang, Q., and Jin, H. (2010). Exploring online social activities for adaptive search personalization. In *Proceedings of the 19th ACM international conference on information and knowledge management ACM*. Toronto, ON, Canada, pp. 999–1008. doi:10.1145/1871437.1871564.
- Wasserman, S., and Faust, K. (1994). *Social Network Analysis*. Cambridge University Press. Cambridge.
- Wetzker, R., Zimmermann, C., Bauckhage, C., and Albayrak, S. (2010). I tag, you tag: translating tags for advanced user models. In: *Proceeding WSDM '10 Proceedings of the third ACM international conference on Web search and data mining*. ACM. New York, NY, USA, pp. 71–80.
- Xia, X., Zhang, S., and Li, X. (2010). A personalized recommendation model based on social tags. In: *Database Technology and Applications (DBTA), 2010 2nd International Workshop on*. IEEE Conference Publications, pp. 1–5
- Xu, S., Bao, S., Fei, B., Su, Z., and Yu, Y. (2008). Exploring folksonomy for personalized search. In *Proceedings of the 31st annual international ACM SIGIR conference on research and development in information retrieval*. ACM. Singapore, Singapore, pp. 155–162. ACM. doi:10.1145/1390334.1390363.
- Yao, A. (2009). Enriching the Migrant Experience: Blogging Motivations, Privacy and Offline Lives of Filipino Women in Britain. *First Monday*. **14**(3).
- Zacarias, M., and Ventura, P. (2011). Collaborative Methods for Business Process Discovery, *Spacial and Organizational Dynamics*, **7**: 45-55.

Appendix 1
Seed sites

	Wikipedia	External links
1	http://es.wikipedia.org/wiki/Walden_Bello	http://waldenbello.org/
2	http://es.wikipedia.org/wiki/Jos%C3%A9_Bov%C3%A9	http://www.jose-bove.eu/
3	http://es.wikipedia.org/wiki/Noam_Chomsky	http://www.chomsky.info/
4	http://es.wikipedia.org/wiki/Michel_Chossudovsky	http://www.globalresearch.ca/
5	http://es.wikipedia.org/wiki/Enrique_Javier_D%C3%ADez_Guti%C3%A9rrez	http://diezgutierrez.wordpress.com/
6	http://es.wikipedia.org/wiki/Christian_Felber	http://www.christian-felber.at/
7	http://es.wikipedia.org/wiki/Ram%C3%B3n_Fern%C3%A1ndez_Dur%C3%A1n	http://laexplosiondeldesorden.wordpress.com/
8	http://es.wikipedia.org/wiki/Eduardo_Galeano	http://eduardogaleano.org/
9	http://es.wikipedia.org/wiki/Susan_George	
10	http://es.wikipedia.org/wiki/Ha-Joon_Chang	http://hajoonchang.net/
11	http://es.wikipedia.org/wiki/Naomi_Klein	http://www.naomiklein.org/
12	http://es.wikipedia.org/wiki/Michael_Moore	http://www.michaelmoore.com/
13	http://es.wikipedia.org/wiki/Raj_Patel	http://rajpatel.org/
14	http://es.wikipedia.org/wiki/Nicanor_Perlas	http://www.nicanor-perlas.com/
15	http://es.wikipedia.org/wiki/James_Petras	http://petras.lahaine.org/
16	http://es.wikipedia.org/wiki/Ignacio_Ramonet	
17	http://es.wikipedia.org/wiki/Dani_Rodrik	http://rodrik.typepad.com/
18	http://es.wikipedia.org/wiki/Arundhati_Roy	http://www.weroy.org/arundhati.shtml
19	http://es.wikipedia.org/wiki/Jos%C3%A9_Luis_Sampedro	http://www.clubcultura.com/clubliteratura/clubescritores/sampedro/home.htm
20	http://es.wikipedia.org/wiki/Fred_Scholz	
21	http://es.wikipedia.org/wiki/Joseph_Stiglitz	http://www.josephstiglitz.com/
22	http://es.wikipedia.org/wiki/Vandana_Shiva	http://www.navdanya.org/
23	http://es.wikipedia.org/wiki/Carlos_Taibo	http://www.carlostaiibo.com/
24	http://es.wikipedia.org/wiki/Esther_Vivas	http://esthervivas.com/
25	http://es.wikipedia.org/wiki/John_Zerzan	http://www.johnzerzan.net/
26	http://es.wikipedia.org/wiki/Jean_Ziegler	

Source: Authors

A MATURITY MODEL FOR HIGHER EDUCATION INSTITUTIONS

Duarte Duarte
Paula Ventura Martins

ABSTRACT

The adoption of business process improvement strategies is currently a concern of most organizations. The quest for the benefits of this improvement on resource optimization and the responsiveness of the organizations has raised several proposals for process improvement methodologies. These approaches differ both in the principles that support them, and in the specific area to which they are intended. However, proposals and results of scientific research on process improvement in higher education institutions, extremely complex and unique organizations, are still scarce. This research project intends to propose the extension of a process improvement model for this particular type of organization. We propose to undertake a review of process areas, goals and practices used in reference maturity models, such as Capability Maturity Model Integration or Business Process Maturity Model, to determine which ones apply to academic organizations and which should be included, adapted or deleted. The resulting maturity model will be further validated in a Portuguese higher education institution. This study is being developed under the University of Algarve Informatics Engineering Doctorate program.

Keywords: Maturity; Education; Process; Improvement

JEL Classification: I23

1. INTRODUCTION

Higher Education Institutions are complex organizations. Although being autonomous, they have to execute a number of functions and develop a variety of procedures, so as to ensure the fulfilment of its duties, which inevitably raise constant challenges. The number of functions they perform and the variety of procedures developed under its autonomy to ensure the accomplishment of all its duties, raise constant management and administration challenges. Difficulties on procedure systematization and on workflow analysis, evaluation, and optimization carry problems not only to management itself, but also to information systems design (Zacarias & Martins, 2011).

Although many of its processes may be documented, these organizations, as well as many others, do not have systematic ways to verify if the graphical descriptions or workflows models are actually pleased in the actual implementation of activities. This problem can jeopardize not only the result of these procedures but also the validity of the requirements of any information system that meets the need of automation.

It is common to find newspaper articles on Information Technology reporting information systems failures that have huge execution times and costs (Bill Curtis & Alden, 2007). While some of these failures can be attributed to technology issues, the causes of many of these failures come from the organization that requested the service. Many, or perhaps most of

these problems, stem from weaknesses in organizational processes that are the automation system target (Bill Curtis & Alden, 2007).

The need for process optimization is not an exclusive problem of academic institutions. Organizations in many different areas, aware of workflow optimization benefits on time and cost control, have adopted strategies and methodologies for process improvement.

Today we can find a broad range of process improvement approaches, distinct from each other, either on its principles and techniques, or on the target area on which the improvements are focused. The most common approaches were initially developed and applied to software development organizations (e.g., Capability Maturity Model (CMM) (SEI, 1995) and Capability Maturity Model Integration (CMMI) (SEI, 2010)). However, inspired by these, other methodologies have been created for wider fields of application, allowing other institutions to reap the benefits of these initial approaches (e.g., Business Process Maturity Model (BPMM) (OMG, 2008)). Others were created as extensions of the most recognized models (e.g., Safety and Security Extensions to CMMI (Bofinger, Robinson, & Lindsay, 2002; Ibrahim et al., 2004)) in order to meet the specificities of a particular business area.

This research project aims, through the analysis of different process maturity models and through the evaluation of specific, higher education features of institutions, to suggest an extension of one of these models to this type of organization.

The next section presents concepts and specific characteristics of academic organizations from the point of view of its process areas. Moreover, we define the concept of business process, describe the concept of organization, approach the specificity of academic organizations and its process areas, identify the concept of business process and explain the main approaches to process improvement. The following section analyses the main process maturity models and those proposed for the education field. Afterwards, we identify the research problem, outline objectives and present the future work.

2. ACADEMIC INSTITUTIONS AND PROCESS IMPROVEMENT NEEDS

Laudon and Laudon (Laudon & Laudon, 2007) consider two main approaches to the concept of organization: the behavioural and the technical. In the first perspective, the organization is a balanced collection of rights, privileges, obligations and responsibilities. Individuals in these organizations develop their own ways to do their work, create ways of social networking and informally agree with their superiors and subordinates regarding ways, deadlines and conditions for the development of their tasks. Most of these agreements are obviously not documented, because they are informal. From the technical perspective, an organization is a social structure that receives formal and stable environment resources and processes them to produce outputs, which may be products or services. Capital and labour are the primary factors of production provided by the environment. The organization transforms these inputs into products and services in a production function. These products and services are produced by the environment in exchange for inputs. The set of procedures to organize the sequence to transform inputs into outputs are called business processes (Laudon & Laudon, 2007).

Higher Education Institutions are complex organizations with multiple power decision centres that bring together a wide range of heterogeneous interests. Mintzberg (Mintzberg, 1999), on his classification of the organization's structural configurations, places universities in the Professional Bureaucracies group, i.e., in the not centralized bureaucratic organizations group. In these organizations, the work developed by professionals is complex and standardized, predictable or predetermined. However, "in the Professional Bureaucracy,

often coexist two parallel hierarchies: one for professionals, directed upside, the democratic, and another to the functions of logistics support, directed downside, with the characteristics of a Machine Bureaucracy” (Mintzberg, 1999). This configuration is characterized as very dependent on the hierarchical structure defined in institutional organization, in which information flows are regulated and always run through the organization in a formal way. This model is based on work processes standardization in which all individuals have their roles rigidly defined.

The amalgam of functions and the required proximity to the students make the professionals of these organizations (the teachers) benefit from considerable autonomy. The professionals’ autonomy, coupled with the diversity of services and departments that integrate these institutions, make these structures very decentralized which hinders the formulation and adoption of comprehensive strategies and may increase the inertia of this type of organization.

In a professional bureaucracy, the flow of information between professionals is scattered and less formal. Various subsystems coexist in this type of organization, and there are a number of contact points with the central system. As a consequence, information flows are complex and not very systematic and structured. In the administrative structure, organized in a centralized bureaucracy with formalized procedures, the flow of information is more defined.

Within a single organization we can thus identify two structures with totally different attitudes regarding its information management: the first is a frame of teachers, which presents a decentralized structure with poorly defined information flows; the second is a centralized and formalized administration support. The definition of strategies for the management and optimization of processes must be different in each scenario.

Portuguese higher education institutions develop a diversity of functions, some of them connected but quite distinct. These functions include degrees and the completion of other courses, educational environment creation, research and scientific support, knowledge economic valuation and transfer, vocational training, community service, cultural exchange and cooperation with other national and international institutions, contribution to international cooperation, rapprochement between peoples, production and dissemination of knowledge and culture.

To comply with this multiplicity of functions, these organizations have a statutory, educational, scientific, cultural, administrative, financial and disciplinary regime which results in a variety of institutional organizations, also enshrined in the statutory scheme that fits these organizations. Universities and polytechnics can integrate autonomous instruction and research units, research facilities, libraries, museums and others. Any approach aimed at assisting these institutions in improving or optimizing their workflows must take into account the special characteristics of such organizations and the specific areas in which they are involved.

2.1 Business areas of Portuguese higher education institutions

An approach that intends to enhance process improvement in academic institutions has first to ascertain on which specific areas these organizations focus their activity. A first analysis of the academic institutions’ business areas can be made from a survey of the features of information systems used by Portuguese and foreign universities. The modules that integrate these systems may allow us to assess areas or large groups of business processes in which these organizations engage. In this section we analyse the FenixEdu system, developed and used by Instituto Superior Técnico (IST), and SAP Student Life Cycle, recently acquired by the University of Algarve and PeopleSoft Enterprise Campus Solutions, developed and marketed by Oracle. Various universities around the world use the two latter systems.

The three information systems presented, although showing its functionalities are organized and distributed in different forms, meet many of the same core business areas that are common to all higher education institutions:

- Student admission which manages student applications, admission and enrolment;
- Pedagogical management, that deals with specific aspects of teaching and learning course content;
- Assessment management, related to the processes of launching and monitoring student's classifications;
- Monitoring registration of student course progression;
- Fee payment;
- Course change and transfer management, which may include equivalence processes;
- Scientific activity management;
- Scholarships management;
- Human resource management;
- Physical resources management;
- Financial management;
- Internal Assessment which, besides the internal evaluation process, may include alumni monitoring;
- Community Relationship.

This analysis, though simple and lacking validation and further development at a later stage, shows that the areas in which higher education organizations engage are unique. This is the set of areas that will have to be addressed in order to improve operating processes in academic organizations, i.e., the activities in each of these areas that should be targeted for optimization in such organizations.

3. PROCESS MATURITY MODELS

How can we describe what a process is? There are various process definitions. Different disciplines characterize this concept in different ways, depending on the type of approach. In the context of information systems, the business process is the set of procedures or ways to organize the sequence for transforming inputs into outputs. This concept can be defined as how an organization coordinates and organizes a range of work activities, information and knowledge in order to produce a particular product or service (Laudon & Laudon, 2007) or simply as a set of tasks or activities performed to achieve a specific purpose or a particular result (OMG, 2008).

Process capability is the process measurable achievement outcome. An organization's process capability helps to predict the possibility of a project to achieve proposed goals. Projects undertaken by organizations with low process capacity cause large variations in costs, time, features and quality (Ibrahim et al., 2001).

Process improvement is a systemic approach that helps organizations optimize the sequence of activities so that they may improve their results. There are several approaches to process improvement. Kulpa and Johnson (Kulpa & A., 2008) summarize the existing approaches into five categories: Business Process Reengineering, Benchmarking, Process engineering/workflow management, Reverse Engineering and Model Based Process Improvement. The approach of this research is the Model Based, similar to other studies already developed in an educational context (e.g., Computing Education Maturity Model (Lutteroth, Luxton-Reilly, Dobbie, & Hamer, 2007), E-Learning Maturity Model (S. Marshall & Mitchell, 2002, 2003; Stephen Marshall & Mitchell, 2004; S. Marshall & Mitchell, 2006a, 2006b, 2008, 2009; S. J. Marshall & Mitchell, 2005, 2007)).

From a software development perspective, the capability maturity of an organization is defined as the power to “meet the demands of its customers in a reliable and repeatedly way” (Poppendieck, 2004) or as the degree to which an organization has established its procedures in order to repeatedly offer their clients high quality software within a given budget and timeframe (Chrissis, Kourad, & Shrum, 2003).

Maturity models are evolutionary roadmaps to the implementation of certain practices that are vital for one or more areas of an organization’s processes. Maturity levels guide the evolution of an organization from a state in which practices are poorly defined and incoherent to a level of innovation and continuous optimization (OMG, 2008). Capability maturity models are focused on improving processes in an organization. These models contain the essential components that effective processes must include for one or more disciplines and describe an evolutionary improvement path from immature or ad hoc processes, to mature and disciplined processes, with improved quality and efficiency (Chrissis, et al., 2003; SEI, 2010). These models allow us to evaluate the maturity level of an organization and, from there, develop a route for improving the capability of their processes.

Capability maturity models are repositories of practices that have proven effective through extensive application in industrial and government organizations. Organizational performance can be measured and improved, by comparing institutions’ practices with the essential practices contained in the models. Capability maturity models describe the key elements of an effective process for a particular discipline and the phases through which the processes can be established, implemented, evaluated and improved (Ibrahim, et al., 2001).

Capability maturity models are associated with one or more evaluation methods that help determine the ability of current processes and define the most critical issues for improving process quality and effectiveness (Ibrahim, et al., 2001).

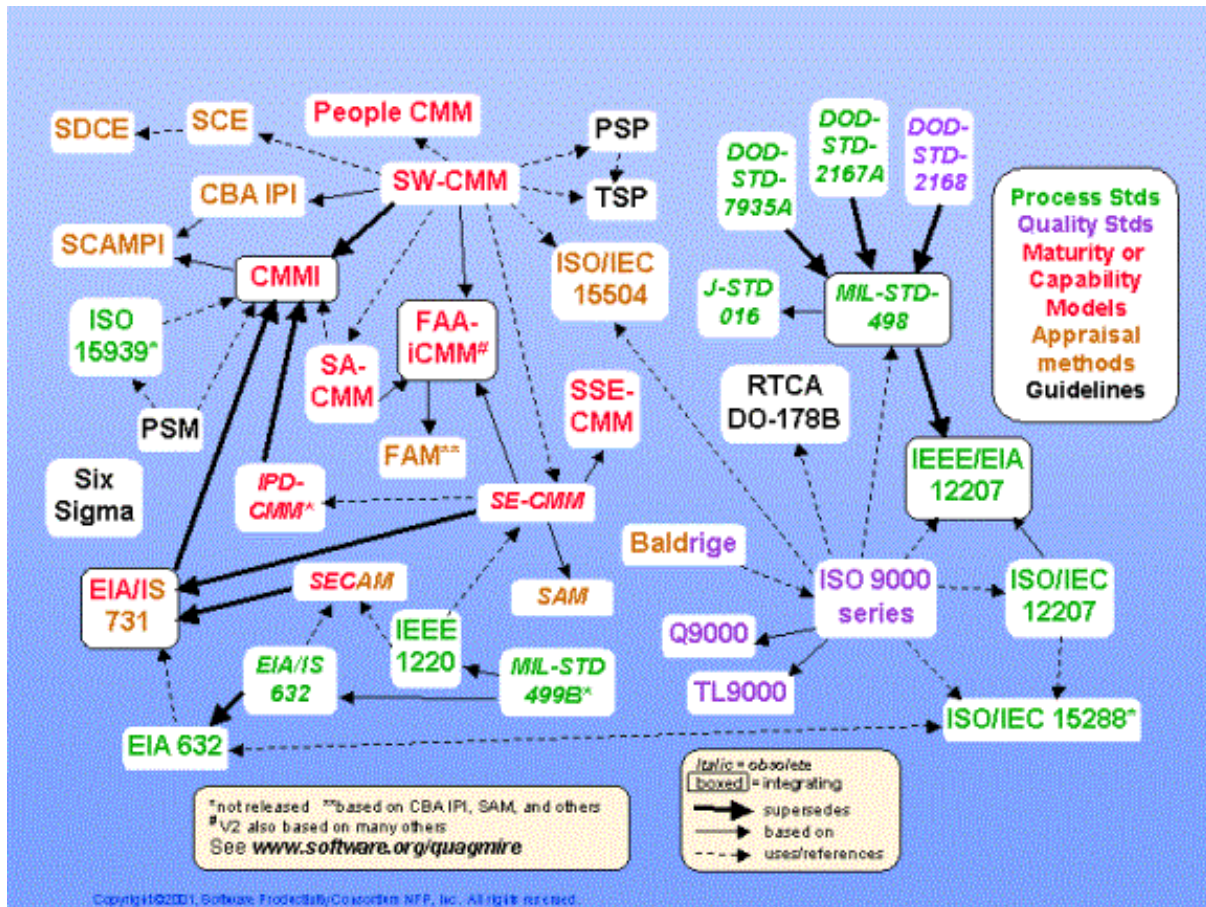
Capability Maturity Model Integration (CMMI) takes advantage of the guidance models geared to a single discipline (such as software and systems engineering), relates these disciplines, removes redundancy and integrates best practices into a common reference model, with common structure and terminology. An integrated capability maturity model can be used across different subjects (Ibrahim, et al., 2001). Maturity level refers to the performance state or degree that can be expected in an organization, in terms of its process capability (SEI, 2010).

3.1 Process Maturity Reference Models

The number of standards, recommendations, maturity models and other frameworks for process improvement that have been developed and then promulgated by governmental and trade organizations has hindered the selection of the best approach for an organization to improve their processes. In 1997, the Software Productivity Consortium created a Web page to help organizations understand which were the most important and how they related to each other. In 2001, Sheard (Sheard, 2001) updated this information and divided the approaches in categories: Maturity Models and Guidelines, Software Standards, Integrated Maturity Models, Systems Engineering Capability Models, Systems Engineering Standards, Measurement Standards and Quality Standards.

Figure 1 is an excerpt from Sheard’s scheme (Sheard, 2001) that shows the relationships between models, standards and assessment systems. Software standards were withdrawn from the original figure since they are not the main focus of this investigation.

Figure 1 - Excerpt from Sheard's frameworks quagmire (Sheard, 2001)



Sheard, S. A. (2001). Evolution of the frameworks quagmire

Sheard considered that the most important maturity models were the Capability Maturity Model (CMM), for organizations, and the Personal Software Process (PSP) and Team Software Process (TSP), for project development. All these models were developed by the Software Engineering Institute (SEI) in Carnegie Mellon University. The integrated models mentioned as the most important were Capability Maturity Model Integration (CMMI) also developed by SEI, which evolved from CMM, and FAA-iCMM, an integrated version of CMM developed by United States Federal Aviation Administration.

3.1.1 Capability Maturity Model Integration

CMMI is considered integrated, since it contains the necessary practices to maturity in various disciplines: Systems Engineering, Software Engineering, Integrated Product Development and Supplier Sourcing and Process. Currently, the CMMI has three versions: CMMI for development, focused on product and service development; CMMI for Services, directed to the processes of service organizations; and CMMI for acquisition, centred on acquisitions and supply of goods and services from others.

Process areas concerning organizations' generic processes are common to all versions. Process areas that differ from version to version are related to the business area to which each version is dedicated.

Each specific goal describes the characteristics that must be present for proper implementation of a process area. They are necessary to determine whether a process area is really implemented. A generic goal describes the characteristics necessary to institutionalize the processes that implement a specific process area. Specific practices and generic practices

are descriptions of activities considered important to meet specific and associated generic goals. Specific and generic practices describe the activities expected in order to meet the specific and generic goals of a process area.

These models have two representations: the continuous and the staged. The first allows the focus on specific processes that are considered important for the organization's immediate goals. The second allows the application of a standardized sequence of improvements that may serve as a basis for comparison regarding the maturity of different projects and organizations.

Table 1 - Comparison between CMMI representations

Level	Continuous representation Capability levels	Staged representation Maturity levels
0	Incomplete	
1	Executed	Executed
2	Managed	Managed
3	Defined	Defined
4		Quantitatively managed
5		In Optimization

Continuous representation establishes four capacity levels for each defined process area: Incomplete or Ad-hoc, Executed, Managed and Defined. In this representation, each process is analysed individually. The same organization can have coexisting processes with very different capability levels, since the analysis is performed independently for each process. This representation is used when an organization wants to achieve capacity levels in specific processes.

A process in level 0, Incomplete, is a process that does not run or runs partially. One or more specific objectives of the process area are not satisfied, and there are no general objectives, since incomplete processes are not institutionalized.

At level 1, the process is Implemented or executed, i.e., fulfils the tasks necessary to produce the result, and the specific goals of the process area are satisfied. However, the improvement achieved in this level may not be repeated in time if not institutionalized (if level 2 and 3 generic objectives are not reached).

At capacity level 2, the process is managed, i.e., is planned and executed in accordance with a policy. It employs staff with expertise to implement it and has the resources to produce controlled results. It involves relevant stakeholders (interested parties) and is monitored, controlled, reviewed and evaluated to ensure compliance with the process description.

A process in level 3 is a defined process, i.e. a managed process adapted from standard procedures according to the organization adjustment lines and its description and is maintained. At this level, the standards, process descriptions and procedures relative to a project are tailored from the set of organization standard processes to fit a particular project or organizational unit, in particular; therefore, they are more consistent, except for differences permitted by the adaptation guidelines. At this level, processes are usually more rigorously described than in capacity level 2. A defined process clearly states the purpose, inputs, entry criteria, activities, roles, measures, verification steps, outputs and exit criteria. At capability level 3, processes are more proactively managed, based on understanding of process related activities and process detailed measures, its work products and services.

The scaled representation, or maturity centred, establishes five maturity levels for classifying organizations: Initial, Managed or Repeatable, Defined, Quantitatively Managed and Optimization. Each maturity level sets objectives and practices that the institution should develop in each process area to be considered at that level. An organization is considered at a given level when the objectives of that level and all previous levels have been reached. For an organization to obtain a certain level of certification, it is necessary that all processes reach this maturity level.

At maturity level 1, processes are usually ad hoc. Organizations positioned at this level do not provide a stable environment to support processes. Success depends on the competence and the heroic efforts of the people and not on the use of proven processes. Despite this chaos, organizations in maturity level one can often produce products and services that work, though they often exceed budgets and fail to meet deadlines. Organizations at this level of maturity are characterized by a tendency to commit beyond their capacity, abandon processes in times of crisis, and for being unable to repeat their own successes.

At maturity level 2, the organization has to ensure that all processes are planned and executed in accordance with policy. They employ experienced and adequate resources to produce controlled outputs. All processes involve relevant stakeholders and are monitored, controlled, reviewed and evaluated to determine conformity with the process description. The procedural discipline reflected by maturity level 2 helps to ensure that existing practices are retained during times of stress. When these practices are in place, projects are performed and managed according to their documented plans.

At this maturity level, the status of the work products and service delivery are visible to management at defined points (for example, major milestones and completion of major tasks). Commitments with relevant stakeholders are established and revised as required. The work products are appropriately controlled. The work products and services satisfy processes descriptions, standards and specified procedures.

At maturity level 3, processes are well characterized and understood, and they have defined standards, procedures, tools and methods. The organization standard process set, which is the basis of maturity level 3, is established and improved over time. These standard procedures are used to establish uniformity within the organization. The processes of a project are defined by adaptation to the set of standard processes according to the organization's adaptation guidelines. At this maturity level, processes are typically more rigorously described than at maturity level 2. A defined process clearly states the purpose, inputs, entry criteria, activities, roles, measures, verification steps, outputs and exit criteria. At maturity level 3, processes are managed more proactively, based on understanding process-related activities and process-detailed measures, its work products, and services.

At maturity level 4, organization and projects establish quantitative, quality objectives and process performance, using them as a managing criterion. Quantitative objectives are based on the needs of customers, end users, organization and the ones accountable for processes' implementation. Quality and process performance are understood in statistical terms and managed throughout process life. Some sub-processes are selected for detailed performance measurements, statistically analysed and stored in an organization measurements database to support decision-making. Special causes of variations in the processes are identified and the sources of these causes are corrected to prevent recurrence. At this level, process performance becomes quantitatively and qualitatively predictable.

Maturity level 5 focuses on continuously improving process performance through incremental process improvements and innovations and technology. Quantitative process improvements are established for the organization and are continuously reviewed to reflect changes in strategic goals and then used as criteria in managing process improvement. Effects of deployed process improvements are measured and evaluated against the quantitative

process improvement. Both defined processes and standard processes are submitted, measurable, improvement activities. An important distinction between the maturity levels 4 and 5 relates to the type of process variation. At maturity level 4, organization is concerned with treating special causes of process variation and achieving statistical predictability of the results. Though processes may produce predictable results, the results may be insufficient to meet the targets set. At maturity level 5, the organization is concerned with treating common causes of process variation and promoting changes in the process (moving average process performance or reducing the observed process variation) in order to improve process performance and meet quantitative process improvement targets.

Organizational process maturity certification is not directly made through CMMI. The organization is evaluated using a specific method and receives a score of one to five. The score corresponds to the level of maturity where the organization is positioned (in staged representation); alternatively, it may refer only to isolated process areas, instead of maturity levels. In this case, processes are classified according to capability levels (continuous representation).

SCAMPI method (Standard CMMI Appraisal Method for Process Improvement) (SEI, 2001) is the assessment method for SEI's official classification of organizations. SCAMPI allows the processes' strengths and weaknesses to be identified and reveals the risks of developing or acquiring, while enabling the determination of capability and maturity levels.

3.1.2 FAA-iCMM

United States Federal Aviation Administration (FAA) Capacity and Maturity Model resulted from the need for strict procedure control to ensure the safety of 2 million passengers in about 30 000 daily commercial flights and 35 000 daily private flights (Ibrahim & Pyster, 2004). Given the inability of other models to meet all the needs of a system with so many disparate needs, FAA developed the ICMM (Ibrahim, et al., 2001) which integrates the principles and practices of 10 other models and standards: ISO 9001, ISO / IEC 12207, ISO/IEC 15288, ISO/IEC 15504, Malcolm Baldrige National Quality Award Criteria, EIA 731, CMMI, Software Acquisition CMM, CMM for Software and Systems Engineering CMM (Ibrahim, et al., 2001).

The maturity assessment method associated with this model is FAM (FAA-ICMM Appraisal Method). The method consists of 20 procedures organized in three phases: evaluation planning and preparation, results assessment and report. It involves interviews and documentary analysis, among other data collection techniques. Assessments are always conducted by external entities and organizations can focus only on a unit within a project, program or line of business or can involve all business processes of the company. This model will not be considered in this study, since there are no records of its use outside FAA companies.

3.1.3 Business Process Maturity Model

In 2008, Object Management Group (OMG) published process maturity model, based on CMMI principles called Business Process Maturity Model (BPMM) (OMG, 2008).

Although this model can be mapped to CMMI, it presents substantial differences. BPMM is geared towards the improvement of larger transactional business processes, for instance those constituting workflows that stretch beyond the boundaries of the organization, while opposed to project-oriented CMMI, which is more circumscribed (Bill Curtis & Alden, 2007).

BPMM includes five maturity levels that represent degrees of organization transformations to improve its processes. The sequence of steps is ordered so that each level provides the basis from which improvements can be triggered to operate at the next level. It allows the

identification of deficiencies in the processes of an organization and guides the improvements through logical and incremental steps (B. Curtis & Alden, 2006).

At level 1, Initial, an institution's practices and business processes are conducted inconsistently and sometimes in an ad-hoc mode. The results of processes are difficult to predict. Employees are often overworked, because management does not assign tasks evenly or does not provide the necessary resources for the implementation of activities. Management focuses on immediate case resolution, often hastily, and does not provide a stable work environment that allows the execution of tasks in a professional and disciplined way. Organization achievements are more due to heroic individual efforts than to an official set of sustainable processes.

At level 2, Managed, an organization uses forms of activity stabilization within each work unit to ensure that tasks can be performed repeatedly and so fulfil the essential obligations of each working group. The main concern in this level is management control over work units' environment to ensure that people can carry out their activities in a repetitive way and that are based on procedures that they are capable of performing. However, different work units can perform similar tasks through different procedures. The activity stabilization is still local; that is, the patterns are not yet established at the organization level. At this level, processes are defined concerning the current situation; that is, according to the way that each group conducts each process. This setting allows the identification of the best practices conducted by each working group, the analysis of the differences in the methods and procedures, the evaluation of the results and the assessment of the degree of commitments' fulfilment.

At level 3, Standardized, standard processes are developed and formalized from best practices identified in the various working groups. Best practices that were identified at the local level, i.e., those that achieved the best results, are adopted in all units and are also incorporated into other practices, from external sources, to address identified problems. The implementation of common processes allows the organization to control and monitor these processes, evaluate performance, learn and share knowledge and develop common skills. Standard processes' adaptation guidelines are provided when it is necessary to apply them in different situations and business needs. Experiments using standard procedures serve as a support for learning and apply to other processes.

At level 4, Predictable, the main objective is to develop the ability to predict the results throughout the workflow, in order to understand and control the variations, so that the results of the process can be predicted from intermediate states. To make this possible, the organization endeavours to establish a statistically stable process, whose results can provide good indicators for predicting the final results. The use of common processes allows the reuse of knowledge, experience and artefacts produced in various business processes. Thus, conditions are set for effective knowledge management, i.e., the capacity to reuse what was learned and developed in different situations. The foundations of knowledge management have been built over the previous levels of maturity. The standardization that took place at level 3 allows the inclusion, in the fourth level, of a set of functional processes in unique and integrative workflows. This reengineering often powers extraordinary opportunities to make processes more efficient.

At level 5, Innovative, proactive and opportunistic improvement actions are developed in order to seek innovations that can bridge the gap between the current capacity of the organization and the capacity needed to achieve business goals. Predictability obtained at the fourth level may still not achieve the desired results of business processes. In this level, proactive activities are triggered to allow raising the capacities of different processes in order to achieve the capacity that will accomplish the desired objectives. Continuous improvement is institutionalized and change management becomes a habitual and encouraged process in

the organization. Emerging changes are evaluated continuously to avoid a negative impact on performance or resource consumption.

BPMM features 30 business areas. Maturity levels 2 through 5 are composed of process areas, varying in number from level to level which, collectively, allow the achievement of its level capability. Each process area aims to achieve development goals, support or maintenance of the desired features for this level. Each process area consists of an integrated set of best practices that indicate what should be done but not how it should be done. As such, organizations are free to set their own methods and approaches to meet the goals and objectives of each process area.

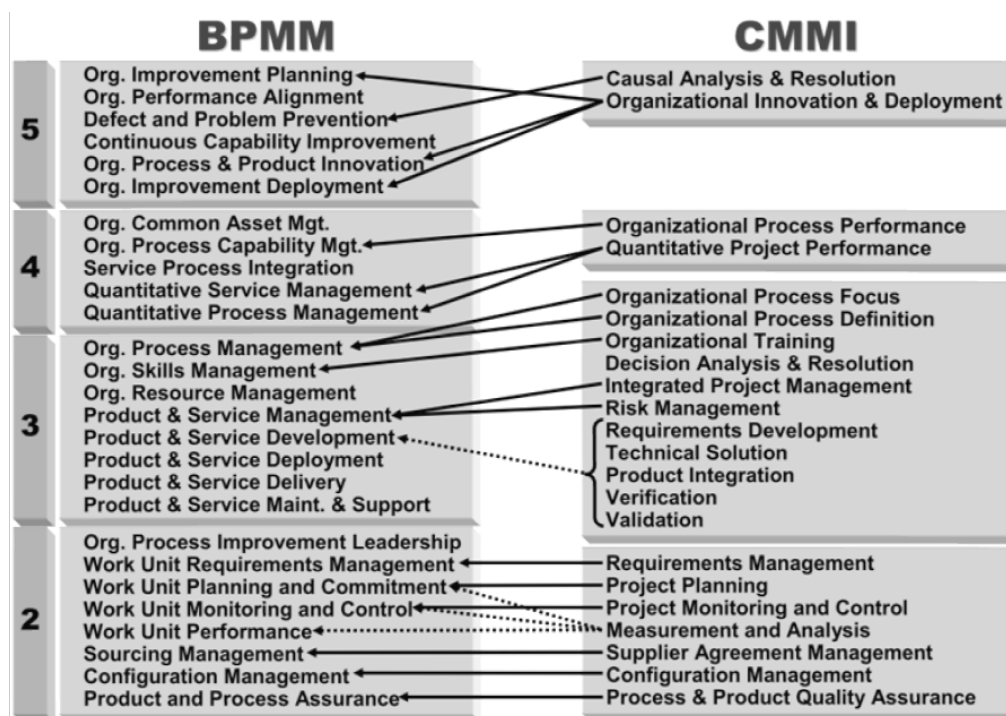
BPMM provides four ways for assessing an organization's BPMM compliance: Initial Assessment, a short procedure with reduced costs that only allows a general compliance review; Progress Assessment, which investigates all process areas and practices in order to determine maturity achievements within a maturity level or to anticipate the results of a confirmatory evaluation; Supplier Evaluation, which takes place during supplier selection and is identical to a progress assessment but does not include organization members' evaluation; and, finally, Confirmatory Assessment, which is a lengthy and thorough investigation of all process areas and practices within the evaluated maturity level.

The presented capability and maturity models have very similar characteristics. Since the three models were developed from the same principles, the maturity levels proposed are identical.

Although different, the three models base their approach on the adoption of best practices to achieve certain goals, grouped into process areas that vary in number, from maturity level to maturity level, and have similar ways to evolve from one maturity level to another. All featured models include one or more associated assessment methods.

The differences that stand out relate mainly to the content of the process areas proposed in each model, i.e., to the goals and practices defined for each maturity level, and to the way these goals and practices are grouped into process areas.

Figure 2 - Correlation between BPMM and CMMI, (B. Curtis, 2004)



Curtis, B. (2004). *Overview of the business process maturity model*

Still, it is possible to match the process areas of different models. BPMM process areas, for example, can be mapped to the ones suggested by CMMI. According to Curtis (B. Curtis, 2004), the contents of all SEI's model process areas are considered on one or more of OMG's model areas. Figure 2 presents a correlation between BPMM and CMMI process areas presented by the author.

Apart from process areas, the main differences between the models concern the structure of the paths that organizations have to travel to achieve higher levels of capacity or maturity. CMMI has two alternative representations: continuous and staged. The first allows capacity improvement in specific processes, and the second allows the improvement of all processes of an organization through maturity levels. FAA-ICMM and BPMM feature a single track, equivalent to the SEI's model continuous alternative. The concepts of goal and associated practices are also different between the FAA-ICMM and CMMI.

For CMMI and BPMM, specific goals describe particular characteristics that processes should have in order to meet the requirements of a process area. Generic goals describe the characteristics considered necessary to institutionalize the processes that implement a process area, i.e., for specific best practices to become routine and part of the institution's organizational culture. FAA-ICMM, rather than specific practices, proposes based targets that are considered fundamental to process execution. Generic practices, as well as describing the routines for process institutionalization, provide attributes that processes must possess so that their capacity can be improved.

Another important distinction has to do with the type of processes that are targeted to improve. BPMM is directed towards improving more transactional, business processes, involving workflows that go beyond the barriers of organization, unlike CMMI, which is oriented to defined projects (Bill Curtis & Alden, 2007). Curtis (B. Curtis, 2004) describes the BPMM not only as an evolution of the CMMI, in the sense that this was based on its principles, but also as a more comprehensive approach. While CMMI applies mainly to the development processes of software applications, BPMM applies to a broader spectrum of business processes and is therefore applicable to any business sector.

Although CMMI has been evolving in recent years, and each new update has presented innovations and different versions for different areas (development, acquisition or services), it is still a very generic model. Similarly, both FAA-ICMM and BPMM present general and abstract structures, ignoring certain business areas specificities. FAA-ICMM presents less detail on the engineering process areas of systems and its goals and practices. If, on one hand, the general and abstract approaches suggested by these models enable the use and adaptation to extended contexts and business areas, its application is limited due to the lack of suitability to some specific business sectors. Because of their unique characteristics and private purposes, these sectors cannot be addressed the same way as any other service organization or product manufacturer.

A maturity model adaptation or extension to enable process improvement in academic institutions requires, at first, the selection of one of these reference models. No records were found of FAA-ICMM use in companies outside the American Aviation Federation, and the model has not been updated since version 2, which dates from 2001.

Business Process Maturity Model seems to be more geared towards improvements associated with more organizational, transverse processes, making it more convenient for the alignment of management areas. In addition, the organizational perspective provided seems to be more interesting for the holistic improvement of an academic institution's business processes.

A staged representation of the Capability Maturity Model Integration for services also seems to be suitable for process maturity development in organizations that provide academic

institution services. Most maturity models' studies and proposals involving teaching and learning have been inspired by this model's principles.

3.2 Educational Maturity Models

The need to adopt process improvement strategies is also a global concern in education institutions. Over the past 10 years, some investigations have been conducted so as to focus on the search for maturity models in education.

White et al. (White, Longenecker, Leidig, Reynolds, & Yarbrough, 2003) launched the discussion about the applicability of CMMI to Information Systems Curriculum in the United States. The authors presented a proposal containing the features that educational institutions should develop, as well as a set of key process areas for each of the five levels of the CMMI maturity model, applied to the curriculum model: Level 1 - Initial; Level 2 - Repeated; Level 3 - Defined; Level 4 - Managed and Level 5 - Optimization. The model was supported by the major United States and international professional organizations like the Association for Information Systems (AIS), the Association for Computing Machinery (ACM) and the Association for Information Technology Professionals (AITP). It integrates a sequence of didactic units composed of objectives and specific content area. Each unit is defined by a set of skills that students must possess at the end of the courses, which are subject to measurement through mechanisms of evaluation and certification in the SI area.

Neuhauser (Neuhauser, 2004) presented a maturity model for online course design aiming to provide a tool to plan and evaluate these courses, based on a set of best practices. The proposed maturity model, Online Course Design Maturity Model (OCDMM), introduces in a phased manner a set of good practices at the institution and provides an integrated system for these practices to maturity since it guides the planner through best practices, learning principles, technologies, objectives and performance standards.

Just like CMMI, OCDMM provides a progressive procedure to transform face-to-face courses in online courses through 5 maturity levels or stages from Level 1, the initial level, where only e-mail and, occasionally, other online resources are used, to level 5, which integrates the best practices for implementing online courses. Each level includes 5 key process areas common to all levels: components (coverage) and appearance; individualization and customization; technology use; socialization and interactivity and assessment. Each of these areas contains a number of general guidelines and specific practices that distance-learning literature describes as being successful.

Thompson (Thompson, 2004, 2006) proposed a Learning Process Maturity Model (LPMM), based on CMM, to help students identify strengths and weaknesses in their learning activities and select the most appropriate strategies for learning. The author draws the parallel between software development and learning processes and presents a definition of maturity in the latter context, describing what may be a mature learner, based on education science literature. Levels presented in this model are similar to CMM. However, the model does not present key areas to consider, but just a set of skills that students should possess at each of these levels.

The Learning Process Maturity Model is still under development. The next step is the development of tools that allows students to transit between maturity levels.

Wang and Zhang (Wang & Zhang, 2007) proposed an IT service management model for Chinese universities based on ITIL (Information Technology Infrastructure Library). The authors present two reasons to justify the need to adapt the ITIL framework to higher education institutions: firstly, the existent model only provides a theoretical platform based on good practices and does not indicate ways for service development, since it has to be adapted to each organization's complexity; secondly, these models are oriented for commercial purposed organizations, the same way as the ones developed by Microsoft (Microsoft, 2011)

or HP (Hewlett-Packard, 2000), and not for higher education institutions, different in organization, culture and technology while having different recipients.

The research identified the differences between these IT systems and the discrepancies among commercial organization departments and university departments. From the comparison of these differences, the authors propose adapted models for organization, process management and technology in order to adapt the ITIL platform to the IT service management reality in Chinese universities.

Lutteroth et al. (Lutteroth, et al., 2007) proposed a maturity model for computer science teaching, also inspired by CMM, called Computing Education Maturity Model (CEMM), that helps computer science teachers by providing a set of best practices and strategies to improve teaching. The authors used the approach and the main concepts of CMM but did not follow it in detail, since they considered that it would not be possible to create a maturity model for education by pure analogy with the CMM. While CMM points out five stages for software project development, CEMM presents five stages for computer course development. The authors believe that, as the project in CMM, the course is a well-defined entity, usually with restricted costs and suffering few variations in time.

Dounos and Bohoris (Dounos & G., 2009) suggested the combined use of Total Quality Management (TQM) principles and the key concepts of CMMI for process improvement in higher education institutions. The authors suggest that TQM quality management principles, techniques used in industry and achieved benefits can be obtained in higher education institutions through the implementation of CMMI. At all five levels suggested in this model, the authors propose the use of TQM benchmarking techniques.

Marshall and Mitchell (S. Marshall & Mitchell, 2002, 2003; Stephen Marshall & Mitchell, 2004; S. Marshall & Mitchell, 2006a, 2006b, 2008, 2009; S. J. Marshall & Mitchell, 2005, 2007) proposed an E-Learning Maturity Model (EMM), also adapted from CMM. EMM divides an institution's capacity to support and provide e-learning into thirty-five processes, grouped into five broad categories or process areas: Learning, Development, Support, Evaluation and Organization. Processes are, however, inter-related through shared practices and different perspectives from five dimensions. Each process in the EMM is divided within each of these dimensions in practices that define how process results can be achieved by institutions. These practices are essential for processes to be successfully achieved, or they are simply useful for obtaining the particular process results. Practices are routines that define the fundamental concepts of different process dimensions. They are derived from empirical data resulting from scientific research in the e-learning field and from experts' opinions. These practices can be evaluated in the institutional context.

In this latest version, the authors proposed a method for the evaluation of processes in which the performance of each practice is rated on a 5-level scale from not suitable to fully suitable. Ratings are based on evidence collected by the institution to verify that the practices are being followed or not followed, to ascertain the quality as they are being performed, and to analyse the importance they have in the process.

After evaluating all process practices, a results average is determined, and a rating is assigned to each dimension. The proposed evaluation procedure identifies not only the degree of suitability of each dimension but also the practices that need to be improved. In his last job, Marshall and Mitchell mapped this assessment method to other guides and other sectorial and institutional standards for educational quality.

Petri, Garcia and Giraldo (Petrie, García, & Giraldo, 2009) proposed a model for higher education institutions certification, based on CMMI, to improve the capacity of processes in engineering and technology institutions, faculties and students. The authors tried to create a method to improve the levels included in engineering curriculum and in certification models to allow institutions in LACCEI (Latin American and Caribbean Engineering Consortium

of Institutions) to obtain accreditations and facilitate other accreditation processes. In the model proposal, Petri, Garcia and Giraldo include in each CMMI maturity level the documentation required for ABET (Accreditation Board for Engineering and Technology) accreditation. ABET is a U.S. organization that certifies higher education courses in applied science, computing, engineering and technology areas.

The model called Engineering Education Capacity Maturity Model (EECMM) uses the same levels of CMMI and identifies the capabilities and processes that are focused at each level of maturity.

Bass (Bass, 2010) developed a Maturity Model for Information and Communication Technologies in Educational Institutions in Developing Countries. The study aims to provide guidance for ICT infrastructure planning and to create a reference model to the necessary development phases for the efficient use of these resources.

This maturity model is based on documentary sources and on the analysis of a number of (Quintas & Gonçalves, 2010), colleges and universities in Ethiopia. Main documentary sources used were international benchmarks for ICT skills, such as International Computer Driving Licence (ICDL), European Computer Driving Licence (ECDL), scripts produced by Joint IEEE / ACM Computing Curriculum Task Force for Superior Education and Skills Framework for Information Age (SFIA), which provide a taxonomy that includes 86 ICT skill areas and a set of 290 associated tasks.

The proposed model consists of eight maturity levels, each identifying ICT infrastructure necessary to achieve institutional objectives and skills that students are expected to be able to develop at that level. To achieve a certain level of maturity, the institution should be able to achieve the objectives at lower levels. The model, which features eight levels of maturity, is not based on any reference to software development.

Table 2 presents the comparison of the educational maturity models described. Most models found are based on CMM or on the staged representation of CMMI. Although the various proposals intend to facilitate process maturity in different business areas, most of the presented models have the same five levels of maturity. They all suggest attributes that the organization should possess to be positioned at each stage. However, unlike the model in which they were based, most teaching maturity models do not explicitly identify any key process areas. Only the models developed by Dounos and Bohoris and by Marshal and Mitchel provide these areas as well as the methodologies and evaluation techniques to assess the fulfilment of requirements, to effectively place an organization in a certain level of maturity.

Moreover, the studied models give an insight of the processes of isolated business areas, i.e., they relate only to an informational entity such as the student, the course, the online course or the IT resource. None of these models present maturity practices that encompass the various entities or units, nor do they approach the processes that are crosswise to higher education institutions. On the other hand, most of the models present 'what to do', but none of them, perhaps with the exception of the model proposed by Duonos and Bohoris, presents 'how' an organization can effectively improve their processes allowing it to climb through the maturity ladder proposed. Table 3 synthesizes strengths and weaknesses of all the educational maturity models presented.

Table 2 - Comparison between educational maturity models

Model	Base	Business area	Levels	Process areas	Description	Improv. method	Assess. method	Focus
CMMI-ISC (White et al.)	CMMI	IS Curriculum	5	Variable number per level	Practices/ Features	-	-	IS Curriculum
OCDMM (Neuhauser)	CMM	On-line course planning and assessment	5	5 common to 5 levels	Practices/ Features	-	-	On-line course
LPMM (Thompson)	CMM	Learning	5	-	Skills	-	-	Student
ITIL-ITSMM (Wang e Zhang)	ITIL	Universities IT Services	-	-	-	-	-	IT services
CEMM (Lutteroth et al.)	CMM	Computer science teaching	5	-	Practices/ Features	-	-	Computer science course
CMMI-TQM (Dounos e Bohoris)	CMMI	Higher education institutions	5	-	Practices/ Features	TQM (Bench marking)		Higher education institution management
eMM (Marshall e Mitchel)	CMM/ CMMI	E-learning	5 dimensions	5 common to 5 dimens.	Practices/ Features	-	eMM Capability Assessment	On-line course
MRAIES (Petri, Garcia e Giraldo)	CMMI	Higher education institutions	5	-	Practices/ Features	-	-	Higher education institution management
ICTMMEI-DV (Bass)	-	ICT in education institutions	8	-	Practices/ Features	-	-	ICT Equip. and use

Table 3 - Strengths and Weaknesses of the educational maturity models

Model	Strengths	Weaknesses
eMM (Marshall e Mitchel)	Suggests attributes for an entity to be positioned in a specific maturity level; Methodology is based on a known and established reference model; Defines processes and groups them in defined process areas; Defines a set of good practices for each process to obtain success; Suggests an associated assessment method; Model has been continuously revised.	Focus on a specific entity: on-line course;
MRAIES (Petri, Garcia e Giraldo)	Process improvement is obtained in a staged, progressive way; Methodology is based on a known and established reference model; Suggests attributes for an entity to be positioned in a specific maturity level.	Doesn't present process areas; Levels are based on ABET specific requirements; Doesn't present maturity practices that encompass the various entities.
ICTMMEI-DV (Bass)	Improvement is obtained in a staged progressive way; Suggests attributes for an entity to be positioned in a specific maturity level.	Focus on a specific goal: provide guidance for ICT infrastructure planning in Educational Institutions on Developing Countries; Doesn't present process areas; Doesn't present maturity practices that encompass the various entities.

Model	Strengths	Weaknesses
CMMI-ISC (White et al.)	Process improvement is obtained in a staged progressive way; Methodology is based on known and established reference model; Suggests attributes for an entity to be positioned in a specific maturity level.	Focus on a specific entity: IS curriculum; Doesn't present maturity practices that encompass the various entities; Doesn't suggest ways to evolve from a level to another; The model has not yet been validated.
OCDMM (Neuhauser)	Process improvement is obtained in a staged, progressive way; Methodology is based on a known and established reference model; Suggests a set of good practices for an entity to be positioned in a specific maturity level.	Focus on a specific entity: on-line course; Doesn't present maturity practices that encompass the various entities; Doesn't suggest ways to evolve from a level to another; Model validation was obtained through the application of questionnaires to faculty representatives.
LPMM (Thompson)	Process improvement is obtained in a staged progressive way; Methodology is based on a known and established reference model; Suggests attributes for an entity to be positioned in a specific maturity level.	Focus on a specific entity: student learning skill; Doesn't present process areas; Doesn't present maturity practices that encompass the various entities; Doesn't suggest ways to evolve from a level to another; The model has not yet been validated.
ITIL-ITSMM (Wang e Zhang)	Suggests attributes for an entity to be positioned in a specific maturity level. Methodology is based on a known and established reference model;	Focus on a specific entity: university IT services; Doesn't present process areas; Doesn't present maturity levels or dimensions; Doesn't present maturity practices that encompass the various entities; The model has not yet been validated.
CEMM (Lutteroth et al.)	Process improvement is obtained in a staged, progressive way; Methodology is based on a known and established reference model; Suggests attributes for an entity to be positioned in a specific maturity level.	Focus on a specific entity: computer science course; Doesn't present process areas. Doesn't present maturity practices that encompass the various entities; Doesn't suggest ways to evolve from a level to another; The model has not yet been validated.
CMMI-TQM (Dounos e Bohoris)	Process improvement is obtained in a staged, progressive way; Methodology is based on a known and established reference model; Suggests attributes for an entity to be positioned in a specific maturity level; Suggests that the ways to evolve from a level to another should be based on benchmarking and Total Quality Management techniques.	Doesn't present process areas; Doesn't present maturity practices that encompass the various entities; The model has not yet been validated.

4. INVESTIGATION PROBLEMS

In order to limit the scope of this study and explain the major issues that this problem raises, this section identifies the research problems. The main problems identified were the following:

- The existing reference models have too broad of a spectrum, i.e., they are too general and ignore important, specific characteristics of higher education organizations, since they do not meet their business areas;
- The developed or adapted models in the educational field are focused on an isolated entity, sector or very specific business area, ignoring other academic institution areas;
- Most educational maturity models proposed do not suggest any process areas and its related goals and practices or the proposed process areas; goals and practices do not fit the reality of higher education institutions;

- The maturity models that focus on teaching processes do not give indications of ways of aligning the processes of management and administration of the institution with the processes related to the area of education that is the target of their attention;
- Existing reference models indicate which attributes an organization must show at each level of maturity and what best practices should be followed in each of them, but they do not indicate how they can or should be implemented so as to accomplish the improvement of processes within the organization.

This research aims to develop, adapt or extend one of the reference models presented in order to meet the following objectives:

- Review and compare the different, existing, business process maturity models and to highlight its inadequacy to the business areas of higher education institutions;
- Propose a business process maturity model which aligns management and teaching practices that coexist in academic institutions;
- Develop a set of new or adapted methodologies to provide ‘how’ an academic organization can improve its business processes, thus moving from one level to another in the proposed maturity model;
- Enhance the provision of educational maturity models;
- Validate the proposed maturity model in a typical, higher educational organization.

5. PRESENT AND FUTURE WORK

This work began with the study of the organizational structure of Algarve University, followed by the modelling of all business processes of one of its services - the Academic Services. Around fifty of the processes were analysed and modelled using Business Process Model and Notation (BPMN) (OMG, 2009) graphical representation. Process analysis enabled the identification of some improvement possibilities and demonstrated the variety of connections between this service’s sections and the links connecting them to the other units of the university and to outside entities.

The work also included the determination of the business areas within academic institutions, based on the review of the functionalities of three information systems used by Portuguese universities. The modules integrated in these systems, and the procedures they aim to systematize or automate, allowed the determination of a set of business process areas associated with these types of organizations. The analysis of FenixEdu, developed and operated by Instituto Superior Técnico (IST), Student Life Cycle (SAP), recently acquired by the University of Algarve, and PeopleSoft Enterprise Campus Solutions, developed and commercialized by Oracle, permitted the identification of 13 business areas: student admission; teaching/learning; student assessment; student progression; tuition management; student course change and transfer; scientific activity; scholarship; Human resources; physical resources; Finance; Internal Assessment and Community Relationship.

The methodology that will be used to create a maturity model suitable to these institutions is still being defined. However, it will necessarily include a more detailed analysis of the existing maturity models both for education and wider areas.

Finally, we propose to undertake a review of key process areas, goals and practices used in each model in order to determine which ones apply to academic organizations and which ones should be excluded, added or extended.

One possible approach is to select a set of well-defined informational entities that encompasses all the academic activities and to limit the analysis to each of them separately, as Lutteroth et al. (Lutteroth, et al., 2007) did with the *course* entity in CEMM.

The model validation will be developed at a later stage, through its application to two or more units of the same university, in order to verify the feasibility of its application, either individually or on the relationship between them. Alternatively, we can choose an approach based on informational entities, monitoring the use of the model on two defined entities to perform the validation.

This research is being carried out within the doctoral program in Informatics Engineering of the Electrical Engineering and Computer Science Department of the Science and Technology Faculty - University of the Algarve and is supervised by PhD Paula Ventura Martins.

REFERENCES

- Bass, J. M. (2010). *A New ICT Maturity Model for Education Institutions in Developing Countries*.
- Bofinger, M., Robinson, N., & Lindsay, P. (2002). *Experience with Extending CMMI for Safety Related Applications*. Paper presented at the 12th International Symposium of the International Council on Systems Engineering (INCOSE'02) Las Vegas, Nevada.
- Chrissis, M., Kourad, M., & Shrum, S. (2003). *CMMI - Guidelines for process integration and product improvement*. Boston: Addison Wesley.
- Curtis, B. (2004). *Overview of the business process maturity model*. Paper presented at the SPIN, San Antonio.
- Curtis, B., & Alden, J. (2006). Business Process Improvement Guided by the BPMM. *BPTrends*.
- Curtis, B., & Alden, J. (2007). The Business Process Maturity Model (BPMM): What, Why & How. *BPTrends*.
- Dounos, P., & G., B. (2009, 18-20 June, 2007). *Exploring the interconnection of known TQM process improvement initiatives in Higher education with key CMMI concepts*. Paper presented at the 10th QMOD Conference. Quality Management and Organizational Development. Our Dreams of Excellence., Helsingborg, Sweden.
- Hewlett-Packard. (2000). The HP IT Service Management Reference Model: Hewlett-Packard.
- Ibrahim, L., Bradford, B., Cole, D., LaBruyere, L., Leineweber, H., Piszczek, D., et al. (2001). The Federal Aviation Administration Integrated Capability Maturity Model (FAA-iCMM), Version 2.0 Federal Aviation Administration.
- Ibrahim, L., Jarzombek, J., M., A., Bate, R., Croll, P., Horn, M., et al. (2004). Safety and Security Extensions for Integrated Capability Maturity Models U. S. F. A. Administration (Ed.)
- Ibrahim, L., & Pyster, A. (2004, May/June 2004). A Single Model for Process Improvement - Lessons Learned at the US Federal Aviation Administration. *IT Pro*.
- Kulpa, M., & A., K. (2008). *Interpreting the CMMI: a process improvement approach* (Second ed.): Auerbach Publications.
- Laudon, K., & Laudon, J. (2007). *Management Information Systems: Managing the Digital Firm* (10 ed.). New Jersey: Person Prentice Hall.
- Lutteroth, C., Luxton-Reilly, A., Dobbie, G., & Hamer, J. (2007). *A maturity model for computing education*. Paper presented at the Proceedings of the ninth Australasian conference on Computing education - Volume 66, Ballarat, Victoria, Australia.

- Marshall, S., & Mitchell, G. (2002). *An E-Learning Maturity Model?* Paper presented at the 19th Annual Conference of the Australian Society for Computers in Learning in Tertiary Education.
- Marshall, S., & Mitchell, G. (2003, May 6-9). *Potential Indicators of e-Learning Process Capability*. Paper presented at the EDUCAUSE in Australasia 2003 Conference, Adelaide, Australia.
- Marshall, S., & Mitchell, G. (2004). *Applying SPICE to e-learning: an e-learning maturity model?* Paper presented at the Proceedings of the sixth conference on Australasian computing education - Volume 30, Dunedin, New Zealand.
- Marshall, S., & Mitchell, G. (2006a, 5-7th September). *Assessing Sector E-Learning Capability With an E-Learning Maturity Model*. In D. Whitelock and S. Wheeler, Eds. Paper presented at the 13th International Conference of the Association for Learning Technology, Edinburgh, Scotland.
- Marshall, S., & Mitchell, G. (2006b). *E-Learning Maturity Model Version Two: New Zealand Tertiary Institution E-Learning Capability: Informing and Guiding E-Learning Architectural Change and Development Project Report*. Report to the New Zealand Ministry of Education.
- Marshall, S., & Mitchell, G. (2008). *What are the key factors that lead to effective adoption and support of e-learning by institutions?* Paper presented at the HERDSA, Rotorua, New Zealand.
- Marshall, S., & Mitchell, G. (2009, 3rd-6th May). *Using the e-Learning Maturity Model to Benchmark Institutional Learning and Teaching Plans and Capabilities*. Paper presented at the Educause in Australasia Conference, Perth, Australia.
- Marshall, S. J., & Mitchell, G. (2005). *E-Learning Process Maturity in the New Zealand Tertiary Sector*. Paper presented at the EDUCAUSE in Australasia 2005 Conference, Auckland, NZ.
- Marshall, S. J., & Mitchell, G. (2007, 9 April - 2 May 2007). *Benchmarking International E-learning Capability with the E-Learning Maturity Model*. Paper presented at the EDUCAUSE in Australasia 2007, Melbourne, Australia.
- Microsoft. (2011). *MOF Process Model for Operations* Retrieved 20-1-2011, from <http://technet.microsoft.com/en-us/library/cc539260.aspx>
- Mintzberg, H. (1999). *Estrutura e Dinâmica das Organizações* (Terceira Edição ed.).
- Neuhauser, C. (2004). *A maturity model: Does it provide a path for online course design. The Journal of Interactive Online Learning*.
- OMG. (2008). *Business Process Maturity Model, Version 1.0*. In O. M. Group (Ed.).
- OMG. (2009). *Business Process Model and Notation Beta 1 for Version 2.0*.
- Petrie, M., García, V., & Giraldo, G. (2009, June 2-5). *Modelo de Registro y Acreditación de Instituciones de Educación Superior basado en el Modelo CMMI*. Paper presented at the Seventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI'2009), San Cristóbal, Venezuela.
- Poppendieck, M. (2004, February 2004). *Measuring Maturity*. *Software Development Magazine*.
- Quintas, H., & Gonçalves, J. (2010). *The Leadership of Schools in Three Regions in Portugal Based on the Findings of External Evaluation*. *Spatial and Organizational Dynamics*, 5, 43-60.
- SEI. (1995). *The Capability Maturity Model: Guidelines for Improving the Software Process*. Reading, MA: Addison-Wesley.

- SEI. (2001). Standard CMMI Appraisal Method for Process Improvement (SCAMPI), Version 1.1: Method Definition Document.
- SEI. (2010). CMMI for Development, Version 1.3. In S. E. I.-C. M. University (Ed.).
- Sheard, S. A. (2001). Evolution of the frameworks quagmire. *Computer*, 34(7), 96-98.
- Thompson, E. (2004). Towards a learning process maturity model.
- Thompson, E. (2006). *Using a subject area model as a learning improvement model*. Paper presented at the Proceedings of the 8th Australian conference on Computing education - Volume 52, Hobart, Australia.
- Zacarias, M., & Martins, P.V. (2011). Collaborative Methods for Business Process Discovery. *Spatial and Organizational Dynamics*, 7, 45-55.
- Wang, Z., & Zhang, X.-y. (2007, 20-22 Aug. 2007). *An ITIL-based IT Service Management Model for Chinese Universities*. Paper presented at the Software Engineering Research, Management & Applications, 2007. SERA 2007. 5th ACIS International Conference on.
- White, B., Longenecker, H., Leidig, P., Reynolds, J., & Yarbrough, D. (2003, 11-2003). *Applicability of CMMI to the IS Curriculum: A Panel Discussion*. Paper presented at the Information Systems Education Conference (ISECON 2003), San Diego, CA.

HOW TO MODEL PEOPLE WORK PRACTICES FROM ONTOLOGICAL TRANSACTIONS

*António Gonçalves
Marielba Zacarias
Pedro Sousa*

ABSTRACT

This paper proposes a model to represent people's work from ontological business transactions according to the concepts of Activity Theory and Enterprise Ontology. With this model we aim at providing a comprehensive and structured modelling approach that, at the same time, avoids excessively detailed descriptions of technology-supported activities. The proposed model is complemented with a set of rules that enables the interrelating transactions described with the Enterprise Ontology language and task descriptions based on Activity Theory. An important benefit of this model is enabling the analysis of manifested contradictions to identify the aspects needing changes within organizational activity systems. The proposed model and rules were applied in a study case.

Keywords: Activity Theory; Enterprise Ontology; Organizational Modelling

JEL Classification: D83, O31, O33

1. INTRODUCTION AND MOTIVATION

The evolution of the Information Systems (IS) field has been marked by the importance given to models and modeling activities as a means of facilitating the communication among systems' stakeholders. The high inter-dependence between IS and an organization's structure, culture and processes, as well as the need of aligning IS and organizations, has led to an expansion of the IS field that includes organizational modeling as part of the systems development process.

Modelling an organization is a complex task due to several aspects: (Anderson,1999; DeBoever,1997): (1) the dynamics of change that characterizes the respective organizations and ecosystems; (2) the different concerns or perspectives, as well as their inter-dependencies that need to be represented and (3) the emergence of economies based on information and knowledge, which are now acknowledged as essential organizational assets, leading to dually regarding people both as actors and knowledge resources.

In spite of its complexity, organizational modelling is a valuable effort, since it enables a representation of explicit, organizational knowledge. Having a model of the organization (hereby referred to as "The organizational model") can improve access to information needed for people to understand decision-making and operations of the organization. This knowledge about how the organization works allows the development of a "collective consciousness" by sharing different but inter-related concerns, such as business goals and processes, organizational structure and resources including technological artefacts, people and materials, among others.

Research in Information Systems has provided several organizational, modeling frameworks. Some focus on specific sectors such as government, defense or finance. Other

frameworks are generic; that is, they can be applied to a wide range of organizations of any size, vertical sector or industry, or degree maturity in the enterprise architecture discipline. Some well-known generic frameworks are the Integrated Framework Architecture (IAF) (Cap Gemini, 2007), the Enterprise Architecture Planning (EAP) (Spewak & Hill, 1992), the Open Group Architecture Framework (TOGAF) (Open Group, 2003), and the Enterprise Unified Process (EUP) (Ambler et al., 2005). The most commonly depicted enterprise perspectives are **process**, **organization**, **information**, **application** and **technology** perspectives. It is noteworthy that these approaches are mostly centered on business activities and their inputs and outputs. Within this perspective, there is research aiming at developing collaborative approaches to capture and model business processes (Ventura & Zacarias, 2009)

Other frameworks have been identified as Language-Action Perspective (LAP) approaches. LAP modeling approaches are communication-centered because they emphasize how people communicate with others and how language is used to create shared understandings, as well as to co-ordinate actions (Schoop, 2001). There is also research aiming at developing methodologies

A well-known LAP approach is the Dynamic Essential Modelling of Organizations (DEMO) (Dietz, 2006) for the (re)design of business processes. According to Dietz, it is possible to manage modelling complexity through the integration of three different concerns of the organization; (1) Ontological Model, (2) Realization, and (3) Implementation of the organization are defined as follows (Dietz, 2011; Jong & Dietz, 2010):

1. **Ontological Model:** The essential model of an organization's view of the business regarded in terms of the transactions among its human actors. The ontological model reveals the deep structure of the organization that is more stable over the life cycle of the organization business.
2. **Realization of the Organization:** The integration of Business, Information and Documents.
3. **Implementation of the organization:** The actual operation of organization by its people and information technologies.

The work of Dietz focuses mostly on the former concern (Ontological Model) and only briefly refers to aspects related to realization and implementation. Dietz & Hoogervorst (2008) argue that the development of technological artefacts are based on a notion of theological¹ systems that is disconnected from the business construction and operation. However, the development of business artefacts is based on an ontological notion, which means that it is at the core of the construction and operation of the business itself. Despite this separation, cohesion is needed in the different approaches given that, in an organization, people (supported by technology) are collectively responsible for the construction and operation of the organization. They are also jointly responsible for the company's evolution (adaptation to the changing needs). These responsibilities can only be supported if people have adequate organization knowledge.

Dumay (2005a; 2005b) argues that Dietz's model is not sufficiently detailed to capture human interactions and does not explain how the work done by humans, supported by technological artefacts, can aid the achievement of a new organization's business goals, or improve its performance (Paternò et al, 1997; Norman, 1986; Diaper & Stanton, 2004; Busbach, 1996). Reijswoud and Rijst (1994) show how Dietz's work can only be used for analysis of high-level businesses. Lyytinen (2004) states that whereas it can be used for analysis of business processes models, it cannot analyse the actual implementation of business process.

¹ It is concerned with the function and the external behaviour of a system.

Our research aims at defining a model that complements Dietz's Ontological Model in order to capture the actual work done by people in organizations and their interactions with technological artefacts, allowing a cohesive representation of the Ontological Model of the organization and its Implementation. The proposed complementary model aims at enabling more detailed analysis of the Implementation of the organization. Understanding such implementation is essential in managing human resources, an important aspect of organizational performance (Gomes et al, 2010). An important requirement of our work is to achieve a balance between developing a precise, but technically difficult, formal model, and an easier, but excessively vague, informal model.

Our approach involves the use of Activity Theory (Leont'ev,1978), which aims at analysis and redesign of human activities. Activity Theory incorporates notions of intentionality, history, mediation, motivation, understanding, communication, culture and context of the people and technology. This theory allows understanding the environment in which technological artefacts are used and support the work done by people in the organization. Several attempts have already being made to operationalize Activity Theory (Kaptelinin & Nardi,2006; Korpela et al.,2002; Martins & Daltrini,1999; Mwanza, 2001) in areas such as work analysis and human-computer interaction.

Aligning the ontological point of view with a view capturing work practices has been acknowledged in several works as a means of achieving a better knowledge of the people's work supporting business (Bødker,1989; Nardi,1996; Redmiles,2002; Kaptelinin & Nardi,2006). Our proposal aims at achieving such alignment by providing a model for the analysis of the organization that integrates the principles of Activity Theory and the Ψ^2 theory (the base of Dietz's approach).

This article is organized as follows: section 2 provides an overview of the DEMO³ methodology and theory Ψ . Section 3 introduces our proposal described in terms of a set of rules and supporting method, which can be applied to modelling the organization implementation. In section 4, the proposed solution is applied to a case study. Finally, in section 5, a discussion is made of the results obtained, along with the conclusions and future work.

2. THEORETICAL FRAMEWORK

2.1 DEMO and Ψ Theory

According to Dietz, there is an essential difference between organizations and other systems, such as technical systems. The technical systems belong to the category of rational systems. This implies that they do not make decisions, but only calculations, and in doing so support the decision only (Liu et al.,2003). Organizations belong to the category of social systems in which people in their category of social subjects have the ability to undertake commitments to each other through social interaction (Dietz,2006), concerning things they bring to reality.

The DEMO methodology (Dietz,2006), developed by Dietz, provides a means of dealing with the complexity of the representation of an organization and its dynamics, and it favours the Complexity Theory (Weinberg,1975), to the detriment of deterministic models of organizations⁴. DEMO provides an immaterial⁵ specification of organization through an ontological model of organizations. DEMO emphasizes the description of the core business of the organization and is based on firm Ψ theory.

² Acronym for *Performance in Social Interaction*

³ Acronym for *Design & Engineering Methodology for Organizations*

⁴ System that attaches to the providential action determinations of human.

⁵ This is because abstracts are the people who perform specific acts in the social organization and the technology they use to do so.

The Ψ theory finds its roots in the scientific fields of the philosophy of language, particularly in the Language Action Perspective (LAP) (Barjis et al., 2002), in Austin's acts of communication (Austin, 1978) and in systemic ontology of Bunge (1979). It recognizes the dynamics, the incompleteness and uncertainty of the reality of the organization, as well as the multiple connections between the components of this reality, and focuses on the use of language to achieve mutual agreement and understanding between people.

According to the Ψ theory, through their social interactions, people engage in obligations relating to actions to be taken and agree on the results of these actions (Dietz, 2003; 2006). This is done via acts of coordination, through language that can be understood as issuing a sentence seen as an action. In this case, the act is called a performative utterance of contractual act, and it creates new facts or actions or part of an action. By stating the act, the announcer does not describe or even state the performing of an action. He is really performing it. The performative utterances do not describe or verify something, are not true or false, are not only the saying or stating, but are part of the action. In general, this means: When we say something, through a locutionary⁶ act, with the intent or effect of changing the world (or act upon the world), we are somehow performing illocutionary⁷ and perlocutionary⁸ acts that cause the intended change.

The Ψ theory consists of several axioms. A summary of the axioms of Ψ theory is shown in Table 1. A complete overview of the theory and associated methodologies is available in Dietz's book (Dietz, 2006) and a number of articles (Dietz & Hoogervorst, 2008; Dietz & Albani, 2005; Dietz, 2006b; Dietz & Hoogervorst, 2007).

In Ψ theory, the transaction axiom indicates that the acts performed by agents occur always and only in universal standards and business transactions and call the result of the execution of a transaction a fact. (Dietz, 2006). The default transaction consists of the following acts: request, commitment, affirmation and acceptance of acts of coordination. It features two actors, each with a distinct role: The initiator, which initiates and completes the transaction; and the performer, who performs the act of production acts.

Table 1. Performance in Social Interaction Theory (Albani & Dietz, 2011)

AXIOM	DESCRIPTION
Operation	The activities of the actors constitute the operation of a business. In exercising, these subjects perform two kinds of acts: production or coordination acts. These acts produce definitive results: production and coordination facts.
Transaction	The transaction axiom defines the relation between acts, so the transaction can be defined as a universal standard in which there are executed acts of coordination and production, involving two actors always aiming to achieve a certain result.
Composition	Describes the interrelationships between transactions
Distinction	The axiom of distinction establishes the existence of three human capacity determinants in implementing the roles of the actors during operation; these capabilities known as Performa, Inform and Form: Form: formal aspects of communication and information; Inform: content aspects of communication and information; Performa: relates to the appearance of new things, directly or indirectly, through communication.

An example of standard business transactions looks like the negotiation of buying fruit at a grocery store:

⁶ Corresponds to the act of uttering something.

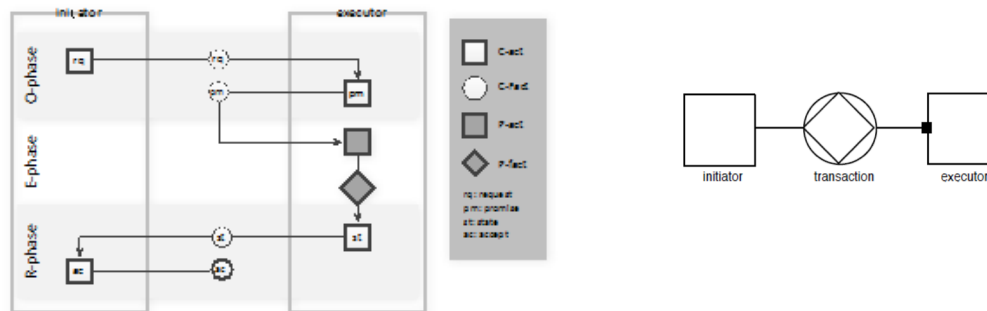
⁷ They represent the actions that take place within language, containing the value in the statements.

⁸ The effects obtained by the use of language (to convince, persuade), i.e., the result of the actions performed by the language

1. Person A asks (order), person B that wants to buy a certain amount of fruit;
2. Person B makes a commitment (promises) to person A to fill a bag of fruit with the amount requested;
3. **<act of production: effective delivery of the “bag of fruit”>;**
4. The person B states to person A that has manufactured the “bag with fruit”;
5. Person A accepts from person B the bag with the fruit in accordance with his expectations

A transaction is a journey through this whole pattern, and each business process in every organization is a collection of transactions linked together in different phases of the acts of coordination. Figure 1 is a graphical representation of the standard ontological business transaction.

Figure 1. Standard Business Transaction (right) and its constituent acts (left), adapted from (Dietz,2006)



2.2 Activity Theory

Activity Theory, developed from the work of Vygotsky, Leont'ev and Luria is understood as a systemic and collective theory, with a structure that aims at analysing human activities. Activity Theory takes into account the interaction between the participants in the sociocultural context in which they operate (Vygotskiï & Cole, 1978; Veer, R., Valsiner, 1993). Interest in applying Activity Theory in the field of Information Systems arose in the 90s, and it is reflected in various publications under the topics of Human-Machine interface and Cooperative Work (Bødker, 1989; Nardi, 1996; Redmiles, 2002; Kaptelinin & Nardi, 2006).

Leont'ev postulated that human activity is always social and cooperative, thus it is collective, and takes place within a division of labor (Leont'ev, 1978). The collective activity is linked to the object (purpose) of the activity and the subjects performing it, of which community members (individually) are not often aware. The concept of the object of the activity is subsumed in the concept of activity, in the sense that there is no activity without an object. One thing or a phenomenon becomes the object of an activity as a means to satisfy a given human need or motivation.

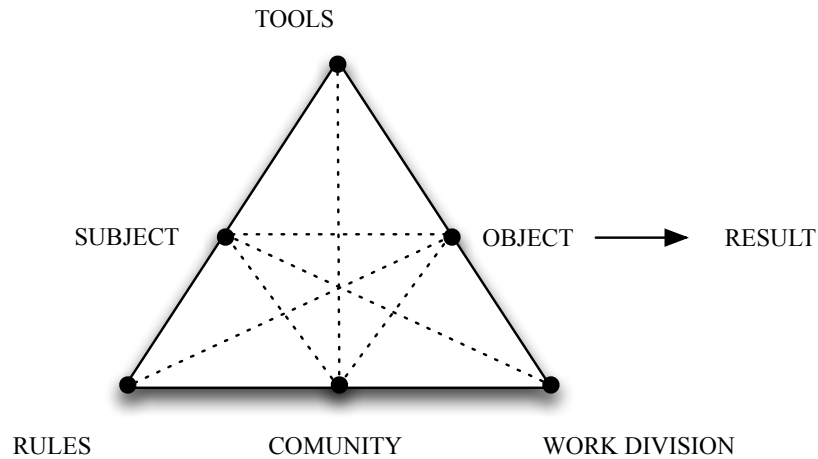
Table 2. Hierarchical levels of activity proposed by Leont'ev (1978)

LEVEL	ORIENTED	COMPOSITION	PERFORMED BY
Activity	Objects that satisfy a need or desire (motive)	Made by actions	Community
Action Can only be understood in the context of the activity to which it belongs	Directed towards a conscious goal	Composed of other actions or operations	Individual or group
Operation The means used to implement the actions	Instrumental conditions	Are initiated by specific environmental situations	Individual or machines

An activity produces outcomes and is performed through actions. Nevertheless, the activity can not be reduced to actions, which are temporary and have a clearly defined beginning and end. Rather, an activity is developed throughout time within a socio-historical process. Individual actions are linked to specific targets or goals that are more or less conscious (Leont'ev,1978). In Activity Theory, goals are not fixed. As subjects act, new goals can be formulated or revised from existing goals. Actions are performed through operations. Operations are performed in an automatic, unconscious fashion and are not clearly related to goals. Operations depend on the conditions in which actions are performed. In this context of three levels, Leont'ev proposed structure for an activity as shown in Table 2.

Engeström departed from the theoretical basis of Vygotsky and expanded his studies to emphasize the notion of mediation. For him, the evolution of the activity occurs through various forms of interaction between organisms and their environment (Engeström,1987; 2000; 2006). The author proposes a system of representation of human activity that encompasses the various components of an activity and their interdependences. According to this author, individuals participate in activities defined by conditions provoked by the division of labor, even without being fully aware of the object and reasons of such activities. Activities are performed by subects through actions to satisfy the object of the activity, which is in turn driven by given purposes or motifs (Engeström,1999). Activities are normally mediated by instruments or tools. Tools can be physical (e.g. a hammer) or psychological (e.g. signs). The model proposed by Engeström to represent the structure of the activity is illustrated in Figure 1, where the constituent elements of activities and their inter-relationships are graphically depicted.

Figure 2. Triangular Model of Activity Theory, adapted from Engeström (1999)



The constituent elements of activities are defined below:

- **Object.** This component reflects the nature of human activity, which enables control of behaviour in order to meet the results identified.
- **Subject:** Participants in the execution of activities. Element representing the individual and social nature of human activity. Includes discussion and collaboration to achieve a common goal. Subjects are individuals or a group of individuals involved in the activity that are guided by a purpose or reason (the object). Hence, subjects form a community share the same general object and constitute a community distinct from others. This element is part of the activity that allows investigating the socio-cultural context in which individuals operate.
- **Tools:** The relationship between subjects and the object is mediated by the use of tools. Tools can be psychological (influence behavior) and/or physical (influence the manipulation of objects). Tools are resources used to transform the object in order to get a result. They can be any resource used during the transformation process: hardware, software, models, methods, theories or even language. Tools themselves can change and, in turn, change the activity. Tools are an essential notion in capturing work practices within an organization.
- **Community:** Social and cultural context of the subjects in which the activity is developed. The community consists of all individuals sharing the same object and, hence, includes all activity stakeholders. This element is also important in the study of the socio-cultural context of subjects.
- **Rules, norms and sanctions:** Boundaries (rules and regulations) affecting the direction of development activities. Rules can be explicit and implicit (e.g. standards of social behaviour within a specific social community). Rules, norms and sanctions specify and regulate, explicitly and implicitly, providing the correct procedures and acceptable interactions between participants within the system activity. Rules mediate the relationship between subjects and their community.
- **Division of labour:** Refers to the allocation of responsibilities. Framing the role to be played by each subject in the development of an activity in the community. Both the horizontal division of tasks between the members of the community, and the vertical division of power and status, continuously mediate negotiated distribution of tasks,

possession of power and responsibilities between the community and the object of the activity system.

According to the Engeström's Triangular Model of Activity Theory, the focal point of the analysis of an activity system is the midpoint of the right side of the triangle (the production of something), which happens when the activity takes place. In the production of outcomes of any activity, participation includes: the subject, the object of the activity, the tools used and the actions and operations that act on the object and produce the result (Engeström, 1999).

The triangular model developed by Engeström's analysis suggests the possibility of multiple relationships within the triangular structure activity; however, the main task is always to understand the whole rather than their separate connections. For Engeström, Activity Theory is an important framework to understand the totality of human work and its praxis, since work cannot be understood or analysed outside the context in which it occurs. Thus, when analysing human work, we must take into consideration not only the actions of individuals, but also who is involved, what are their motives and goals, what rules and procedures exist and the community where the activity occurs.

3. PROPOSED APPROACH

Our approach allows analysing the work done by people from the ontological transactions of Dietz's Enterprise Ontology. The solution consists of a set of rules, which integrate representations of ontological transactions using Dietz's Enterprise Ontology (section 2.1) with notions drawn from Activity Theory (section 2.2). Section 3.1 describes the set of rules, and section 3.2 describes a model with a graphic representation of the concepts and relationships defined in such rules.

3.1 Proposed Rules

An integrated view of Activity Theory and Enterprise Ontology is achieved through the following set of rules:

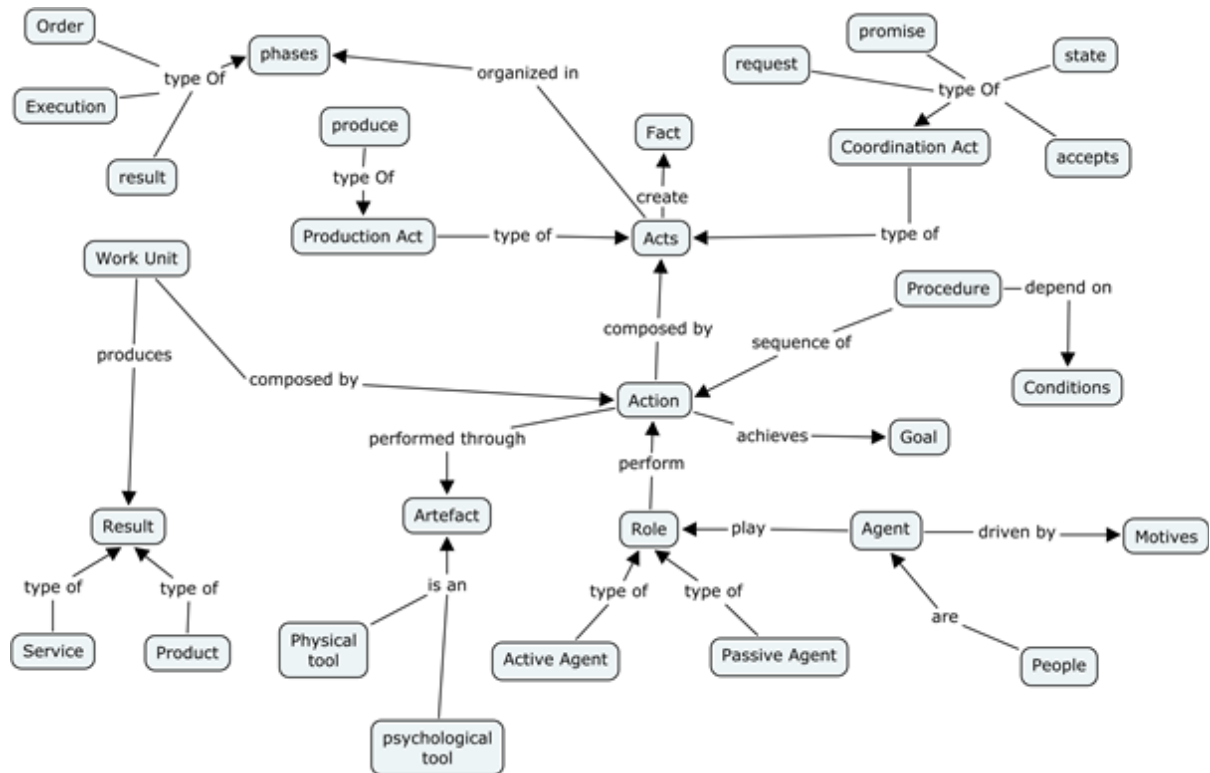
- **Rule R1 (Work Unit [WnT]):** Understanding work practices entails defining and analysing minimum units of work (which we designate simply as **Work Unit**) involving work (which we designate by acts). Acts are actions performed by members of the organization that are considered relevant for the organization. Acts can be performed consciously or unconsciously. It also entails identifying the following elements: **stakeholders**, who are the people that request a product or service, and the people who satisfy such request. This rule is drawn from Activity Theory, where work always results in a collective effort (e.g. always involves the participation of several people performing acts towards a common result). We associate the product (or service) of the work unit to the result (product or service) of the ontological transactions. It is noteworthy that people working in a **WnT** are not always aware of its result. **WnT** have a dynamic nature, constantly changing over time, because the performance of acts is understood as part of a socio-historical process.
- **Rule R2 (Operational Classification):** The acts performed by people in a **WnT** can be acts of **coordination** or **production**. In **production acts**, people contribute to the achievement of the result of each **WnT**. By performing **coordinating acts**, people meet the commitments related to **production acts**. Once done, the **act** results in the

creation of a **fact**, which will be a production or coordination fact according to its related act. This policy stems from the concepts present in the Ψ theory;

- **Rule R3 (Life Cycle of Acts):** In a **WnT**, coordination and production acts follow a sequence of **acts**: *request*, *promise*, *produce*, *state* and *accept*. These acts are organized into three distinct phases: the **Order Phase** (O-phase), the **Execution Phase** (E-phase) and the **Result phase** (R-phase). In O-phase, people try to reach an agreement on the desired outcome in the **WnT**, which is the intended **fact** by that work unit and encompasses *request* and *promise* acts. In the E-phase, production is performed, comprising the **acts** of *producing* a result of interest to business. In the R-phase, comprising the *state* and *accept* acts, people try to reach an agreement whether the result is that which was agreed in the O-phase. The first and last step are *performative* in the sense that they do not affirm or deny anything, but that they perform an act when they are announced, and they are used to reach an agreement, respectively, on the *request* and on its *acceptance*. The intermediate phase is the action associated with the request (and triggered by it). This policy stems from the concepts present in the Ψ theory.
- **Rule R4 (Decomposition of the act):** Every act (e.g. request, promise, producer, state and acceptance) consists of a sequence of actions and operations that are created and maintained in the organization. Actions are held in a conscious way. Operations are conducted in an unconscious manner, always subordinated to actions and dependent on execution conditions (environmental and sociocultural). An action will be performed by a set of operations. The operation of an action is not a fixed plan or prescription of the work to be performed. Rather, it is only a roadmap that can be modified according to the context in which it is executed. Due to the unconscious and contextual nature of operations, it is unlikely to describe a priori what operations will be part of the actions that people take. Still, from an afterthought and observation of real work, it becomes possible to conduct a survey of the most common operations used by the subject during the work. This rule follows the hierarchical structure of an activity proposed by Leont'ev in Activity Theory;
- **Rule R5 (Mediation):** Actions and operations on the environment are performed through artefacts that can be physical or mental tools such as rules, signs, contexts and work divisions used to control behaviour, improve communication and motivate people. An artefact evolves over time, thereby adding to the history of development within a **WnT**. Since actions and operations depend on mediating artefacts, they both constrain or enhance the way that people perform them. Outside the context of the **WnT** in which the artefacts are used, it becomes difficult to understand their usefulness. Ideally, people should not be aware of artefact usage, unless those who develop the artefact desire it. This rule results from the analysis of the systemic model proposed by Engeström to represent the structure of social activity;
- **Rule R6 (Agents):** People who participate in a work unit play the role of agents in an organization. **Agent roles** specify the responsibilities, competencies and authority necessary to participate in given **production** or **coordination** acts. Agents can be active or passive. Active agents have coordination roles. Passive agents are limited to production roles. The purpose of this rule is to model the work carried out by people, mediated by tools, rules and division of labour, taking into account the roles cast in ontological business transactions.

Figure 3 depicts a graphic representation of the concepts and their relationships described in the previous rules.

Figure 3. Model Proposed

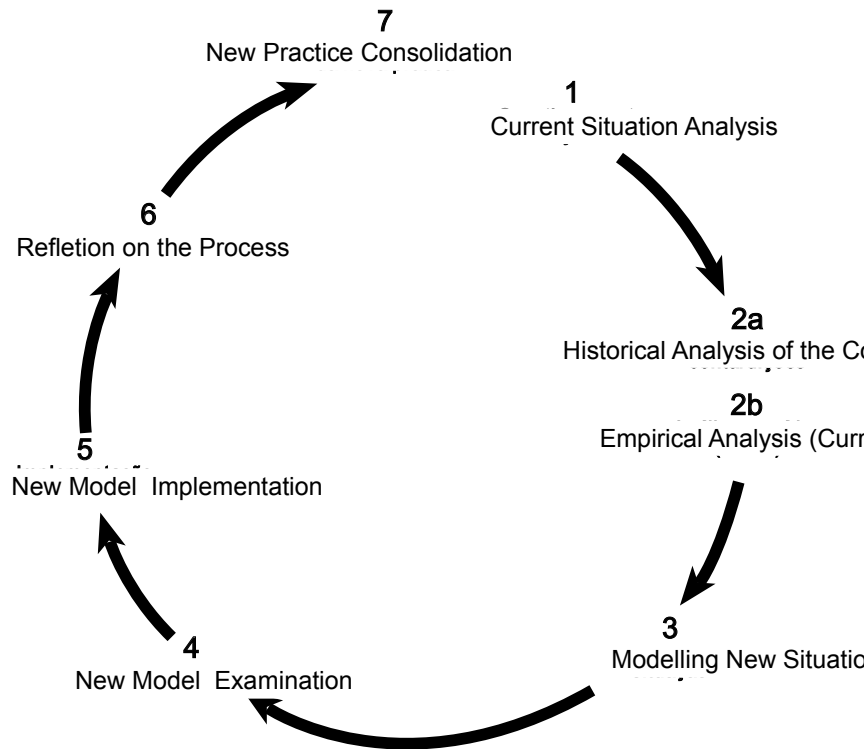


3.2 Supporting Method: Finding contradictions

For our proposed method, the usage of the notion of contradictions was chosen, which, according to Engeström, constitutes a key element in analysing human work. Contradictions are historically accumulated, structural tensions within and between work units (Engeström, 2000). The notion of contradiction was first stated by Hyenkov (Carelli, 2003) as being the one that could potentially destabilize people's work. Hyenkov states that contradictions are a sign of work mobility. On the other hand, Engeström (2000) described that contradictions should be conceptualized as tensions or unbalances manifested by failures, problems or errors, which can be detected by analysing the speech of the people in the organization (Engeström & Sannino, 2011). Contradictions cause the appearance of an expansive learning cycle (see figure 4) that leads to changes in work units. This change is cyclical and can produce other contradictions (Engeström, 1987).

The expansive cycle allows us to understand the evolution of work units through the identification of time periods marked by innovation, transformation and change, as well as others marked by cultural reproduction and/or the learning of the innovations produced. To make it possible to integrate the learning cycle in the context of the work units, it was decided to incorporate it in the Boyd Decision Cycle illustrated in figure 5 (Breton & Rousseau, 2005; Bunge, 1979).

Figure 4. Expansive Learning Cycle



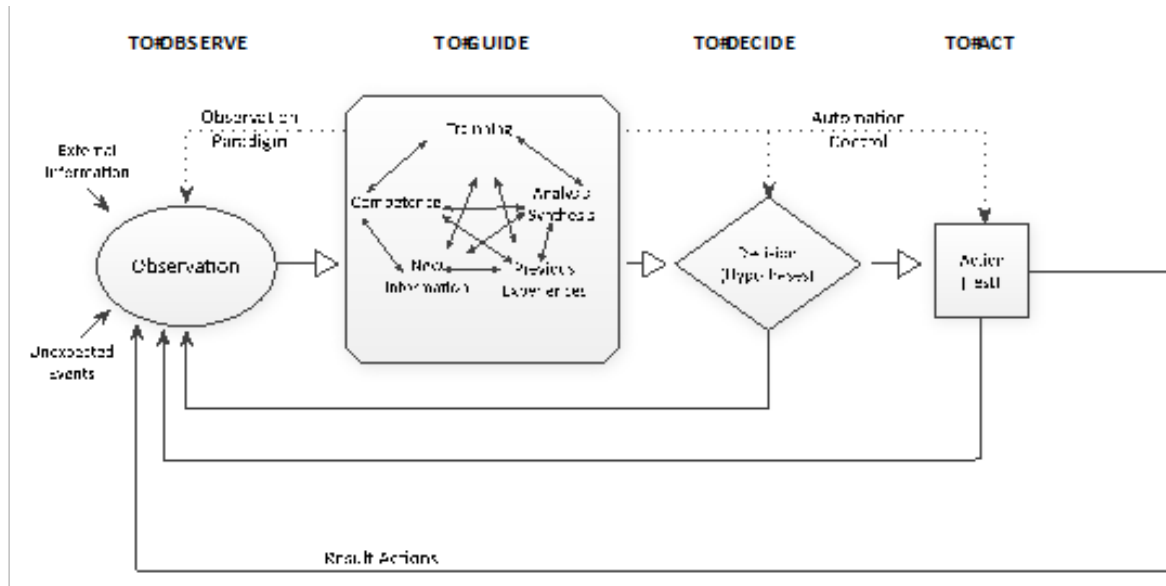
The Boyd Decision Cycle includes the perception that the process comprises a set of steps executed iteratively. These steps are: “observe - orient - decide - act.” Each phase of the cycle is explained below:

1. **Observation** includes the collection and compilation of data and information about the organization, particularly in ontological process diagrams, new requests made by organizational structures, external information about existing technologies to support business, historical knowledge of the solutions found over time to the issues raised, as well as the knowledge of their weaknesses and strengths;
2. **Orient** is to interpret scenarios based on observations, previous experience, an organizational view, an organizational culture, viewpoints, etc. The orientation results in the construction of diagrams that represent the reality in order to make sense of the actions to be performed. The orientation is highly dependent on the existing view, which in turn is dependent on the tacit knowledge that each element has of a team. Helping a team to observe and get a global sense of what is observable, respecting the particular vision of each element, is a key task. The orientation is a way of promoting the sharing of tacit knowledge that each individual has and that is usually difficult to be formalized or explained to others, because it is subjective and inherent to the abilities of a person. It leads directly to the decisions but also configures observation and action, allowing the definition of alternative solutions;
3. **Decision** is made from the image produced in the orientation step, and it defines the most appropriate response to the current situation. The outcome of the decision may

flow into both sides: Immediate action, or return to observation if there is not enough information for a decision;

4. **Action** is a decision that will then be implemented through the selected action.

Figure 5. Boyd Decision Cycle



This method yields better results through quick iterations. The speed of iteration overcomes the quality of iteration, which is achieved from a process of continuous improvement of the solutions encountered. Based on this vision, we define the steps (see table 3) that incorporate rules and the expansive learning cycle:

Table 3. Steps of the proposed method

STEPS	NAME	DESCRIPTION
1.1	Observation	Map each Ontological Transaction to a Work Unit (WnT), as defined in Rule D1.
1.2		Associate the object of the WnT to the product of the Ontological Transaction
1.3		Associate the result of the WnT acceptance product to the Ontological Transaction, specifically the fact of the transaction when it is run successfully.
2	Orientation	Through observation, decompose the stages of the cycle of acts (O-step, E-step and R-step) in the list of people who perform the actions and identify the list of possible procedures to perform the actions. Identify the tools that mediate iteration between people and actions according to the Rules D2, D3, D4 and D5.
3	Decision	Historical analysis of the tasks and procedures and used tools based on the operation of the organization to detect changes that may be proposed (Expansive Learning Cycle 1, 2 and 2b).
4	Action	Reflection about operation and consolidation of the new practice and impact assessment (Expansive Learning Cycle 3-7).

4. EXAMPLE

We illustrate our approach by applying the model and method to an example described by Dietz in his book (Dietz,2006), described below:

A person wants to buy a bouquet of flowers, and may use different ways to order it (eg by telephone, email, going to the store). To order it is necessary to mention information for the invoice (ie name, tax number, corporate headquarters) and information about the type of bouquet it proposes to buy.

When the bouquet is ordered in the internet it indicates the day and time that it will be picked up at the store. The person who manages the customers requests follow them up through the reception of the different means whereby requests are received and checks if the information provided is complete. If not, communicates with the client in order to complete the data. If an application is complete, the bouquet is produced.

Then, the person responsible for the invoices calculates the amount that has to be paid and prepares the issue of the invoices for payment, receives and checks money, and delivered presentially, if the customer is at the store. Payments must be made by bank transfer or cash. Once payment is received the bouquet is delivered to the customer in the store

4.1 Example Ontological Model

As aforementioned, the starting point is the Ontological Model of the organization, which was built using the DEMO methodology. Figure 6 provides a general view of transactions, which in this example are transactions T1 (Bouquet Order) and T2 (Bouquet Payment). Both transactions involve the actors A1 (Client) and A2 (Organization). Transaction T1 is initiated by actor A1 and executed by actor A2 (i.e. the Bouquet Order transaction is initiated by the Client and executed by the Organization). Conversely, transaction T2 is initiated by actor A2 and executed by actor A1 (i.e. the Bouquet Payment transaction is initiated by the Organization and executed by the Client).

Figure 6. Ontological transactions of the example

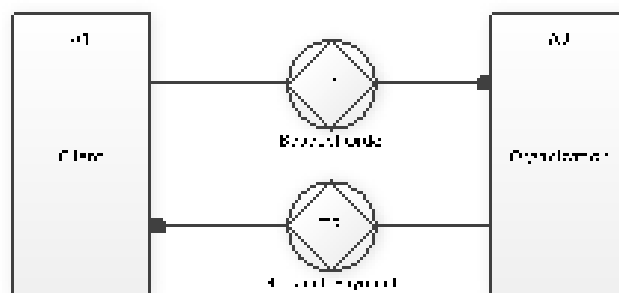


Figure 7. Internal structure of transactions T1 and T2

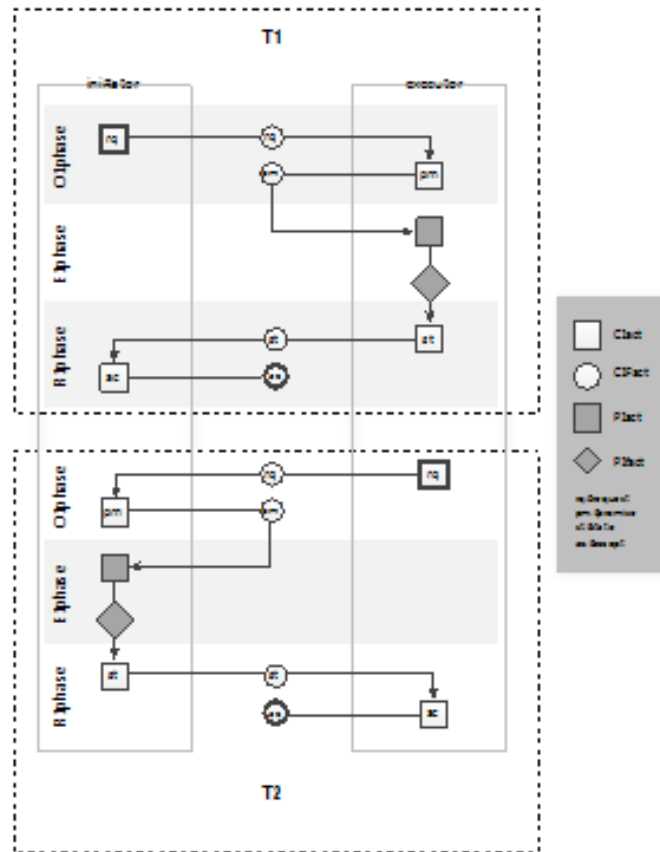


Figure 7 illustrates the internal structure of each transaction in terms of phases, as well as the associated acts and facts. The O-phase of Transaction T1 is initiated by the coordination act request *rq* (*client requests a bouquet*). The request creates a coordination fact *rq* (*bouquet ordered*) that coordinates a response from the organization expressed through a coordination act *pm* (*promiste bouquet for a given date*) and its corresponding coordination fact (*bouquet promised for...*). The latter coordination act triggers a production act (*ellaborate bouquet*) and its corresponding production fact (*bouquet ellaborated*), which correspond to the E-phase. The R-phase encompasses a coordination act *st* (*the organization states that the bouquet is ready*) corresponding to a coordination fact *st* (*information about the bouquet stated to the client*). Once the client knows that the bouquet is ready, it triggers a coordination act *ac* (*the client accepts the bouquet*) and a related coordination fact *ac* (*bouquet accepted*).

Based on the ontological model, we proceed to complement it by adding associated work units with information specific to work practices, using the proposed method described in section 3.2.

4.1.1 Observation

Observation describes the Work Units that will be subject to detailed analysis. As aforementioned, a mapping is established between work units (WnT) and corresponding ontological transactions work. This includes associating work units with ontological transactions, and identifying the work unit object, result, and participating agents (active and passive).

Table 4. Mapping ontological transactions and work units

WORK UNIT	OBJECT	RESULT	ACTIVE AGENTS	PASSIVE AGENT
Bouquet order	Bouquet order	Bouquet is ordered	Clients and organization members responsible for taking bouquet orders	People who elaborate the bouquet
Payment order	Payment	Bouquet is paid and the customer receives it	Organization members responsible for bouquet payments and the client, who is responsible for making the payments	

4.1.2 Orientation

In the orientation phase, each identified **WnT** is described in detail. This detailed description entails identifying the **WnT** goals, actions, procedures and tools for each phase of the transaction. As an example, we present a detailed description of the **WnT** “Order Bouquet”.

Table 5. Bouquet custom order: O-PHASE

WNT: ORDER BOUQUET: O-PHASE			
GOALS	ACTIONS	PROCEDURES	TOOLS
1) <u>MAKE REQUEST</u>	1.1 The customer makes a request for a bouquet	1.1.1 Access a bouquets catalogue 1.1.2 Make choice 1.1.3 Send choice to Company	Catalogue Telephone Email Log book
2) <u>ACCEPT REQUEST</u>	1.2 The organization accepts the request to make the bouquet	1.2.1 Access Registration Customer Choice and day and time of delivery 1.2.2 Check if the choice made by the customer is possible 1.2.3 Store in the organization indication of when the customer comes to get the order 1.2.4 Confirm to the customer that the bouquet will be done	Catalogue Telephone Email Log book

Table 6. Bouquet custom order: E-PHASE

WNT: ORDER BOUQUET: E-PHASE			
GOALS	ACTIONS	PROCEDURES	TOOLS
1) <u>MAKE THE BOUQUET</u>	1.1 Manufacturing	1.1.1 Make the bouquet according to the logbook	Flowers Mechanical tools Log book
		1.1.2 Store the bouquet	

Table 7. Bouquet custom order: R-PHASE

WNT: ORDER BOUQUET: R-PHASE			
GOALS	ACTIONS	PROCEDURES	TOOLS
1) <u>DELIVER THE BOUQUET</u>	1.1 The Organization delivers the bouquet	1.1.1 Access to the register of bouquets 1.1.2 Find bouquet 1.1.3 Register Bouquet Delivery	Log book
2) <u>ACCEPT THE BOUQUET</u>	2.1 Client accepts bouquet	2.1.1 Customer confirms the reception of bouquet 2.1.2 The organization registers the Delivery of Bouquet	Log book

4.1.3 Decision and Action

In the decision phase, manifestations of contradictions were analysed from the observation of the tensions in interactions between people. We sought to analyse and categorise the types of manifestations of people involved in the operation of the organization. As a way of analysis, we tried to find the manifestations of conflicts at each stage of Unit Work: Order Phase, Execution Phase and (Result) Delivery Phase.

In the Order phase, the manifestations found were related to the inability of the customer and organization reaching an agreement about the type and the desired Bouquet or because the catalogues were not updated within the range of existing bouquets or, otherwise, the prices were not the same because they have changed.

In the Execution phase, we found a difficulty in understanding the composition of special bouquets that were ordered: those who had the task of building Bouquets did not perceive the language used by those who registered the order. This difficulty was reflected in a bouquet composition that was different from the type that the customer had ordered.

In the Delivery phase, there were two types of stresses, especially in clients who made the order by phone: The orders were not ready on the day and time agreed, and the product delivered was different from what was requested by the client. Customers didn't appear to collect the product; and, because the payment was made only at the time of delivery, the result was a loss to the organization. There were also delays in the delivery of the bouquet that was ordered by phone, because the bouquets of customers who went to the store were made while the customer was in the store, which delayed the production of bouquets ordered by phone.

From the previous contradictions, it was decided to implement the following solutions:

1. Customers who made the request by telephone have gained access to a web catalogue with access to a set of standard bouquets.
2. There was no possibility to compose different types of bouquets by phone.
3. It became mandatory for customers to pay before the elaboration of the product when the order was made by phone.

This decision led to the adoption of new tools of mediation, including the availability of an electronic Catalogue on the Internet, together with a company website and an order management application, which included payment verification. However, new tensions emerged due to the amendments made that will require a second iteration of the proposed cycle.

5. CONCLUSION AND FUTURE WORK

This article addresses the description of work practices using Dietz's organizational ontological model of Ψ Theory and Activity Theory. The Ψ theory used to represent essential (ontological) transactions is composed of business processes and identifying the actors participating in such transactions. Actors of an organization perform two kinds of acts: production and coordination acts that produce corresponding production and coordination facts. Acts are organized in patterns defined as transactions that ultimately create a relevant business result (a production fact).

The Ontological Model provides an initial base of analysis but does not provide information regarding labour division, which technologies are used, or the specific actions and operations executed that result from specific socio-cultural rules. Activity Theory is used to complement information related to the particular implementation of a transaction. In particular, we use Engeström's diagrams to describe the key elements of activities that allow the analysing of human practices such as tools used, subjects involved, as well as actual actions and procedures resulting from specific socio-cultural rules. To this end, we adopted a set of rules and applied them to allow the capture and analysis of these elements.

The interconnection of these theories provides a basis to analyse the organizational activities both from the perspectives of its essential transactions and the particular way people perform their work through actions in the organization. We use the Activity Theory notion of contradictions to identify activity aspects requiring changes. In other words, the focus of the analysis is the contradictions present during activity execution. The identification of contradictions allows identifying problems and proposing solutions.

In the future, we plan to expand the information captured in work units by explicitly including socio-cultural rules governing actions and procedures. Furthermore, the cycle of expansive learning will be explored as means for a continuous analysis and resolution of tensions and contradictions in an organization according to a base ontological model. Our purpose is to define a systematic and socio-technical approach of organizational analysis through a redesign so that it is possible to inter-relate abstract models with concrete situations through critical discussions, rejections, reformulations, and the proposal of proper solutions. This approach aims not only at identifying areas of improvement but also learning from them, by first determining if they have been identified and resolved in the past. Since solving contradictions cannot take place at the individual level due to their social nature, we plan to include graphical models of activity contradictions in the DEMO methodology in order to support their communication and discussion.

REFERENCES

- Albani, A., Dietz, J. (2011). Enterprise ontology based development of information systems. *International Journal of Internet and Enterprise*.
- Ambler, S. W. Nalbone, J. and Vizdos, M. (2005). *Enterprise Unified Process: Extending the Rational Unified Process*. Prentice-Hall.
- Anderson, P.: Perspective (1999). Complexity theory and organization science. *organization Science*. 10, 3, 216–232.
- Austin, J.L., Urmson, J.O. (1978). *How to do things with words*. Harvard University Press.
- Barjis, J. et al. (Eds) (2002). *Seventh International Workshop on the Language-Action Perspective on Communication Modeling*. TU Delft.
- Brehmer, B. (2005). The dynamic OODA loop: Amalgamating Boyd's OODA loop and the cybernetic approach to command and control. *Proceedings of the 10th International Command and Control Research Technology Symposium*.
- Breton, R., Rousseau, R. (2005). The C-ODA: A cognitive version of the OODA loop to represent C2 activities. *Proceedings of the 10th International Command and Control Research Technology Symposium*.
- Bunge, M.: *Treatise on Basic Philosophy* (1979). Volume 4: *Ontology II: A World of Systems*. Springer-Verlag GmbH.
- Busbach, U. (1996). *Activity Coordination in Decentralized Working Environments*.
- Bødker, S. (1989). A Human Activity Approach to User Interfaces. *Human-Computer Interaction*. 4, 3, 171–195.
- CAP Gemini. Integrated architecture framework IAF, 2007. URL http://www.capgemin.com/resources/thought_leadership/architecture_and_the_integrated_architecture_framework/.
- Carelli, I.M. (2003). *Estudar on-line: análise de um curso para professores de inglês na perspectiva da teoria da atividade*. Pontifícia Universidade Católica, São Paulo.
- DeBoever, L. (1997). Concept of “Highly Adaptive” Enterprise Architecture, Enterprise Architecture keynote address, Dezembro 1997.
- Diaper, D., Stanton A., N. (2004). *The handbook of task analysis for human-computer interaction*.
- Dietz, J. (2003). Designing technical systems as social systems. *Proceedings of the 8th International Working Conference on the Language-Action Perspective on Communication Modelling (LAP 2003)*. pp. 187–207 Citeseer.
- Dietz, J.: Enterprise Engineering Manifesto, <http://ciaonetwork.org/publications/EEManifesto.pdf>, (2011).
- Dietz, J. (2006). *Enterprise Ontology: Theory and Methodology [Hardcover]*. Springer.
- Dietz, J. (2006b). The deep structure of business processes. *Communications of the ACM*. 49, 5, 58.
- Dietz, J., Albani, A. (2005). Basic notions regarding business processes and supporting information systems. *Requirements Engineering*. 10, 3, 175–183.
- Dietz, J., Hoogervorst, J. (2007). Enterprise Ontology and Enterprise Architecture, How to let them evolve into effective complementary notions. *GEAO Journal of Enterprise Architecture*. 2, 1.

- Dietz, J., Hoogervorst, J.A.P. (2008). Enterprise ontology in enterprise engineering. Proceedings of the 2008 ACM symposium on Applied computing - SAC '08. p. 572 ACM Press, New York, New York, USA.
- Dumay, M. (2005a). Demo or practice: critical analysis of the language/action perspective. 10th International Conference on the Language-Action Perspective.
- Dumay, M., Dietz, J. and Mulder, H. (2005b). Evaluation of DEMO and the Language/Action Perspective after 10 years of experience. The Language Action Perspective on Communication Modelling, Kiruna, Sweden, June 19-20.
- Engeström, Y. (2006). Activity Theory and Expansive Design. Theories and Practice in Interaction Design Lawrence Erlbaum Associates. 3–24.
- Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*. 43, 7, 960–974.
- Engeström, Y. (1987). Learning by expanding: an activity-theoretical approach to developmental research. Helsinki : Orienta-Konsultit Oy.
- Engeström, Y. et al. (1999). Perspectives on Activity Theory. Cambridge University Press.
- Engeström, Y., Sannino, A. (2011). Discursive manifestations of contradictions in organizational change efforts: A methodological framework. *Journal of Organizational Change Management*. 24, 3, 368–387.
- Jong, J., Dietz, J. (2010). Understanding the realization of organizations. *Advances in Enterprise Engineering IV*. 31–49.
- Gomes, Jorge F. S., Coelho, Joaquim P., Correia, Anabela, Cunha, Rita C. M, (2010). Development and validation of an instrument measuring the strength of the human resource management system. *Spatial and Organizational Dynamics*, 5.
- Kaptelinin, V., Nardi, B. (2006). Acting with technology: Activity theory and interaction design. MIT Press.
- Korpela, M. et al. (2002). Information systems development as an activity. *Computer Supported Cooperative Work (CSCW)*. 11, 1, 111–128.
- Leont'ev, A. (1978). Activity, consciousness, and personality. Prentice-Hall.
- Liu, K. et al. (2003). Modelling dynamic behaviour of business organisations—extension of DEMO from a semiotic perspective. *Knowledge-Based Systems*. 16, 2, 101–111.
- Lyytinen, K. (2004). The Struggle with the Language in the IT—Why is LAP not in the Mainstream. Proceedings of the 9th International Working Conference on the Language-Action Perspective on Communication Modelling (LAP). Rutgers University, New Brunswick, NJ, USA, June 2-3.
- Martins, L., Daltrini, B.M. (1999). An approach to software requirements elicitation using precepts from activity theory. 14th IEEE International Conference on Automated Software Engineering. pp. 15–23 IEEE Comput. Soc.
- Mwanza, D. (2001). Where theory meets practice: A case for an Activity Theory based methodology to guide computer system design. *Human-Computer Interaction*.
- Nardi, A. (1996). Context and consciousness: activity theory and human-computer interaction. The MIT Press.
- Norman A., D. (1986). In user centered system design: new perspectives on human computer interaction.
- The Open Group. (2003). The Open Group architectural framework (TOGAF) - version 8.1, 2003. URL <http://www.opengroup.org/architecture/togaf8-doc/arch/>.

- Paternò, F. et al. (1997). ConcurTaskTrees: A Diagrammatic Notation for Specifying Task Models.
- Redmiles, D. , (2002). Introduction to the Special Issue on Activity Theory and the Practice of Design, <http://www.springerlink.com/content/r9l92tag4w8m5y41/>.
- Reijswoud, V. Van, Rijst, N. van der. (2004). The Analysis of Planning Activities Using the DEMO Transaction Process Model. Maastricht research school of Economics of Technology and Organizations.
- Schechtman, G.M. (1996). Manipulating the OODA Loop: The Overlooked Role of Information Resource Management in Information Warfare.
- Schoop, M. (2001). An introduction to the language-action perspective. ACM SIGGROUP Bulletin, 22(2):3–8.
- Spewak, S. H. and Hill, S. C. (1992). Enterprise Architecture Planning: Developing a BluePrint for Data, Applications and Technology. John Wiley & Sons.
- Veer, R., Valsiner, J. (1993). Understanding Vygotsky: a quest for synthesis. Blackwell.
- Ventura, P. and Zacarias, M. Collaborative methods for Business Process Discovery. Spatial and Organizational Dynamics, 7.
- Vygotskiĭ, L.S., Cole, M. (1978). Mind in society: the development of higher psychological processes. Harvard University Press.
- Weinberg, G.M. (1975). An introduction to general systems thinking. Wiley-Interscience.

EDITORIAL NORMS
JOURNAL OF SPATIAL AND ORGANIZATIONAL DYNAMICS

In order to simplify the editors' task, authors are urged to **adopt the norms listed below** for the Journal publication. Please note that the article should be sent in its final version.

Being so, the final document should have the following editorial norms:

Page layout:

Paper size A4, 2,5cm left, right, bottom and top margins.

Document Format:

Manuscripts should be submitted in Word file using font Times New Roman and single line spacing. The document should not have more than 25 pages (including bibliography), and a PDF version of the document must be provided.

Biographical Authors' note:

The text must be in English, on a single page, with a text up to 100 words. The information given should include academic career, present professional position and research interests, if applicable. Should also mention affiliation and personal e-mail address. Use size 11, regular, justified.

Title:

Should be concise and informative with size 15, bold, left aligned and in capital letters. After title add author(s) name(s) in size 11, italic, left aligned.

Abstract:

The abstract should have between 150 to 200 words and should not contain any undefined abbreviations.

Use size 11, regular, justified.

Keywords:

Must be provided up to 4 keywords separated by coma (,) at the end of the abstract.

An appropriate **JEL code(s)** must be provided with minimum of 3 digits, for example Q21. This classification system is prepared and published by the *Journal of Economic Literature*. For more information, please visit the website at <http://www.aeaweb.org/econlit/jelCodes.php?view=jel>.

Plain Text body:

For plain text body use size 11, regular, justified.

Subtitles 1st level - size 11, bold, capital letters, left aligned;

Subtitles 2nd level - size 11, bold, low case, left aligned;

Subtitles 3rd level - size 11, italic, low case, left aligned.

Table of Contents:

A Table of Contents should be provided at the beginning of the manuscript. Use size 11, regular, left aligned.

Abbreviations:

Abbreviations should be defined at first mention and used consistently thereafter.

Footnotes:

Footnotes can be used to give additional information. They should not contain any figures or tables and should be in size 8, regular, left aligned.

Footnotes to the text are numbered consecutively.

Footnotes to the title of the article are given with the reference symbol (*).

Endnotes can not be used.

Acknowledgments:

Acknowledgments of people, grants, funds, and others, should be placed in a separate section before the reference list. The names of funding organizations should be written in full.

Citations:

After quoting a text extract, cite the reference giving only the author's name and publication year in parentheses. Ex: (Flores *et al.*, 1988; Winograd, 1986; Cunha and Cintra, 1996)

But if you are citing the author inside the text, add only the publication year between parentheses.

Ex: Winograd (1986) describes ...

References List:

The list of references should only include works that are cited in the text.

Do not use footnotes or endnotes as a substitute for a reference list.

Reference list entries should be alphabetized by the last name of the first author of each work.

To format reference list use size 11, regular, left aligned.

Ideally, the names of all authors should be provided, but the usage of *et al.* in long authors list will also be accepted.

Ex: Pierzynski, G. *et al.* (1994). *Soils and environmental quality*. Lewis Publishers. Florida.

Scientific Article:

Last name of the author, First initial. (Publication year). Article title. *Title of the Journal or Review*.

Volume(Issue): first page-last page.

Ex: Sadiq, M. and Alam, I. (1997). Lead contamination of groundwater in an industrial complex. *Water, Air and Soil Pollution*. **98(2)**: 167-177.

Book:

Last name of the author, First initial. (Publication year). *Book title*. Additional information. Edition number, Publishing house. Publishing place.

Ex: Costa, J. (1995). *Caracterização e constituição do Solo*. 5th edition, Foundation Calouste Gulbenkian. Lisbon. ISBN: 000 000 000 0000

Book Chapter:

Ex: Silko, L.M. (1991). The man to send rain clouds. In: W. Brown and A. Ling (eds.), *Imagining America: Stories from the promised land*. Persea. New York.

Online Document:

Last name of the author, First initial. (Publication year). *Document title*. Accessed in: day, month, year, in: URL.

Ex: Chou, L., McClintock, R., Moretti, F. e Nix, D.H. (1993). *Technology and education: New wine in new bottles – Choosing pasts and imagining educational futures*. Accessed in 24th of August 2000, on the Web site of: Columbia University, Institute for Learning Technologies: <http://www.ilt.columbia.edu/publications/papers/newwine1.html>.

Dissertation:

Ex: Tingle, C.C.D. (1985). *Biological control of the glasshouse mealybug using parasitic hymenoptera*. Ph.D. Thesis. Department of Biological Sciences, Wye College, University of London. 375 pp.

Tables, Figures, Graphics and Boards:

All tables, figures, graphics and boards are to be numbered using Arabic numerals and should have a title explaining its components above the body, using size 9, bold, centred.

The source and year of the information given in tables, figures, graphics and boards should be included beneath its body, centred, size 8, regular. For tables and boards contents use size 8.

Figures and graphics must be in JPEG format (image).



ISSN: 2183-1912

Authors

António Gonçalves

Duarte Duarte

Estrella Gualda

Juan D. Borrero

Marielba Zacarias

Paula Ventura Martins

Pedro Sousa

