Expert Systems With Applications 72 (2017) 18–29

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Research stakeholders identification using an mobile agent's framework

Mihai Horia Zaharia*, Florin Alexandru Hodorogea

"Gheorghe Asachi" Technical University, Romania

ARTICLE INFO

Article history:
Received 9 May 2015
Revised 11 October 2016
Accepted 4 December 2016
Available online 6 December 2016

Keywords:
Distributed computing
Artificial intelligence
Business process re-engineering

ABSTRACT

In the business process of reengineering, one of the main problems consists in the organizational reintegration of the employee. New instruments for the automatic identification of stakeholders among existing staff evaluation are needed. It is difficult to identify a stakeholder, especially if we refer to science stakeholders, and the proper evaluation of a researcher's value and position at international level is needed. There are large contradictions concerning the validity of various types of indexing services and other techniques used to classify the researchers. This paper proposes the use of corporate Human Resource Management (HRM) specific instruments when it comes to the researcher as informational stakeholder identification. The main idea behind the method is to use some form of information retrieval techniques over a set of research papers which contain, mostly, the same key words and to extract the hierarchy of the researchers involved/referenced over that set. A distributed approach was used due to its inherent scalability. The use of an inference engine was also deemed necessary due to the complexity of the problem. The system may be used in the automatic assessment of research performance in public or private institutions. It can be easily modified to support complex sets of rules depending on each user's needs in personnel evaluation. The main research field of this papers concerns distributed artificial intelligence.

1. Introduction

The concept of business process re-engineering (BPR) was introduced more than twenty years ago (Davenport & Short, 1990). The main idea was to convert classic ways of doing business into a more flexible and a more efficient form using information technology (IT) (Hammer & Champy, 1993). BPR was related to the organizational changes that were needed when the first wave of IT adoption began. Since then, IT involvement in the organizational workflows at various levels has continuously increased. As a result, the BPR concept has begun to be reanalyzed (Zamudio, Zheng, & Callaghan, 2012).

Nowadays the concept applies to the changes that any organization must undergo in order to merge with another one using an advanced software platform like the one based on service oriented architecture (SOA). The new SOA based approach gives a new dimension to BPR because the concepts must be adapted to the new architectural characteristics. The complexity of the changes required to integrate a new organization into a larger one needs the involvement of artificial intelligence (AI) (Yu, Mylopoulos, & Lespérance, 1996). Because nowadays most of the international organizations have geographically distributed branches, an approach based on intelligent agents is recommended, especially due to the advancement of multi-agent systems (Gorodetskii, 2012).

Yet, there are some problems concerning BPR, among which personnel efficiency still remains an issue (Terziovski, Fitzpatrick, & O' Neill, 2003). When the BPR is applied within an organization, some Human Resource Management (HRM) changes must also take place. This raises the classic problem of identifying the organization informational stakeholders, especially the latent or hidden ones (Mitchell, Agle, & Wood, 1997). This is critical because the advantage provided by BPR can be quickly lost if a stakeholder is moved from his area of expertise or, in the worst case, fired.

Friedman and Miles (2001; 2006) define the stakeholder as an individual who has the knowledge to influence a business process. Roberts and Mahoney (2004), after conducting an inquiry among managers, found that only about 65% of the participants used the term stakeholder correctly.

Mitchel classified the stakeholders by their influence but also by the urgency of their claims on the organization. The danger of neglecting the claims of some stakeholder categories has been

* Corresponding author at: Faculty of Automatic Control and Computer Engineering, Dept. Computer Engineering, Str. Prof. dr. doc. Dimitrie Mangeron, nr. 27, Iaşi 700050, Romania. Fax: +40 232 231343.
E-mail addresses: mike@cs.tuiasi.ro (M.H. Zaharia), florin.hodorogea@siliconservice.ro (F.A. Hodorogea).

http://dx.doi.org/10.1016/j.eswa.2016.12.006
0957-4174/

clearly presented e.g. Mitchell et al. (1997). Turner proposed some methods which can be used to properly manage stakeholder satisfaction (Turner, Grude, & Thurloway, 1999). Probably the most important type of stakeholder is the *hidden* or *dormant* one. Due to external or internal factors, subjective or objective, dormant or inactive stakeholders may suddenly change their state and become active from an organizational point of view. There is a 50% chance that this activation can be beneficial for the organization. For instance, they may begin to increase their performance, involvement and help by offering solutions to various local problems. However, in other cases, the newly activated stakeholders may use their knowledge to disturb the activity of the local branch. This kind of involvement can be difficult to prove, even when this may even destroy a project or a newly initiated organizational change.

Consequently, new ways of identifying stakeholders must be developed. Unfortunately, there are no accurate ways of measuring the stakeholders' real influence. The mere identification of part of the hidden or dormant stakeholders would be a great asset for the organization. In the classic approach, based on face-to-face interviews and performance evaluation, stakeholder identification is a slow process, with many opportunities to err (Bryson, 2004; Zhao & Hou, 2010). The employees associate this type of evaluation with further personnel decisions (dismissal, transfers and promotions) and engage in simulated behavior strategies, in order to keep their jobs. The large scale integration of workflows using IT would make this kind of evaluation partially or totally automatic.

A problem may appear when one believes that the experience of the HR expert cannot be replaced. This is only partially true, because the HR expert will remain useful to the organization, with the computer speeding up his/her work in order to increase their efficiency. A part of one's experience will be used to generate the rules needed by an inference engine involved in document flow analysis.

A possible approach to hidden stakeholder detection would be to extract information about the real structure of document flow. There are situations when a document does not follow the expected processing flow. For example, it is sent to a third party for a supplementary check. Due to the high amount of information to be processed this deviation in document flow structure can be detected only by automatic processing.

In the scientific world, we may witness various differences between the quality and output of private and academic research. As far as private research is concerned, the pressure on employees may be high. However, it may bear very good results and offer a high income to the researcher. Performance evaluation is done using corporate specific techniques. In the public funded research, some quality factors are observed, such as the Hirsch index, but they alone are not enough to evaluate the whole performance of an individual within the organization (Priem, Taraborelli, Groth, & Neylon, 2010). In most cases, the criteria used in evaluations vary a lot. As a result, only indirect methods of performance analysis can be used. Most of the common methods of performance evaluation take into account external funding and some visibility indexes, such as the Hirsch evaluation, or they use a weighted sum, based on the ISI impact factor values of each publication. If the evaluation criteria used in HR management are correct, the approach may be successful. Unfortunately, within each organization there may be deviations regarding its evaluation principles. An internal instrument is needed; one that could analyze and also retrieve other research papers from the Internet in order to properly rate the importance of a researcher.

The researcher himself/herself can be viewed as a stakeholder for their organization (Lapointe, Mignerat, & Vedel, 2010). In this case, the evaluation must be made at the level of his/her group. To accomplish that, the group needs to be identified in the first place and then the position of the researcher can be estimated. The Achilles' heel in the case of researcher performance evaluation is the individual's real influence in the domain, which may be smaller than expected in case of a large number of citations. The idea is to analyze the scientific papers selected from a domain, using a combination of keywords and to extract the influence tree of the author.

Until a few decades ago, this method was not applicable because there was not enough available information in electronic form. Nowadays, this is no longer a problem, and the proof that the needed information is available is represented by existing plagiarism detection systems. So, only a few changes in knowledge extraction are enough to obtain a better representation of the real influence of a researcher in a specific field. The same technique can be simultaneously used to analyze the same researcher's influence in the R&D department of an organization. Finally, both results can be merged to obtain a pertinent performance evaluation of the whole work of the researcher.

The research presented in this paper is related to the field of document (scientific papers) clustering based on citations. The term clustering refers in this research the process of finding a cluster of authors that cite one with each other. The approach is not new; it actually preceded current research in social networks (Morlacchi, Wilkinson, & Young, 2005). The first attempts were related to the identification of the researcher's network structure created by the citing process (Newman, 2001a) and to the application of graph theory to the obtained network (Newman, 2001b). After the development of data mining, the research was mostly focused on clustering in order to capture the critical topics in the research papers (Aljaber, Stokes, Bailey, & Pei, 2010) or on combining citation semantics (Tong, Dinakarpandian, & Lee, 2009).

The original problem of using citations to identify the most influential researchers can be solved nowadays by creating the citation tree and analyzing the differences in the frequencies of citations (Cribbin, 2011). In this paper we present the use of a document workflow analysis system for the extraction of the research relationships among the researchers. This will be done by analyzing the domain of their published papers and also the references used in the latter. This may help to properly identify the stakeholders from a domain in conjunction with other quantitative methods that are already used. The systems presented belong to the area of computer assisted decision support and can be used by HRM staff in the phase of employee evaluation.

As is previously mentioned there are some aspect related to personnel evaluation that are hard to quantify. Even when standard measurement techniques are used, the specific training of an HR expert, make it difficult to properly use them. The main aim of this research is to offer an instrument that can provide a supplementary and intuitive way of assessment for the HR experts in the area of scientific throughput efficiency. The presented instrument is in the area of distributed artificial intelligence. It was created on top of an existing general framework for distributed computing based on intelligent agents (developed using JADE). For this research supplementary agents were developed accordingly with the analyzed problem.

There are some tools that can estimate the throughput of a researcher. The most important are Google scholar and Microsoft academic. The research gate also offers some facilities for researchers. The analysis is more detailed if the 'Publish or Perish' tool is used. This tool uses previously mentioned solutions to compute not only the h-index or i10-index but more complex ones, such as the contemporary h-index and some supplementary analysis, the age-weighted citation rate and an analysis of the number of authors per paper (Harzing, 2016). Anyhow, these tools compute some statistical parameters for one individual but do not offer an all citation relationship graph for an author. Those tools are focused only on computing the universally accepted statistical