

Abstract: The given work represents web-based monitoring framework for measuring performance of functioning of higher education establishment. High-level key performance measures for higher education are suggested. The models forming the basis for web-based monitoring framework and their goals are presented. Multi-agent architecture is suggested for implementation of the framework.

Keywords: Performance measurement, higher education, web-based monitoring, information retrieval

I. Introduction

Usually the goals and tasks of organization are identified in its business plan or strategic program of development. Their fulfillment is directly connected with the prosperity of organization. The management cycle in any company presupposes the analysis of the results of its functioning. Generally the success of company's policy can be expressed in terms of its performance. We consider performance as working effectiveness with respect to goals of organization. It characterizes the results of any business – whether they correspond to organization's long-term strategy.

Management in the field of higher education has its own peculiarities. First of all higher education is service business and in most cases it doesn't presupposes the production of any goods. Secondly, the aims of higher education establishments (HEEs) cannot be expressed only through financial indicators. Since one of the main goals of HEE is preparation of highly qualified professionals for needs of society and economy, HEE's activities can be partially considered as non-profitable.

To remain competitive among prospective university students and to extend its influence in academic community HEE has to analyze results of its work and make some changes in its development program based on this analysis. Performance measurement in higher education combines the evaluation of all types of activities in HEE. This includes teaching, research, and finances. Performance measurement must provide a comprehensive assessment of all aspects of outcomes of HEE's functioning.

Management of HEE needs to obtain continuously the feedback from implementation of working plans. To do this HEE has to organize appropriate monitoring and evaluation activities. The result of monitoring is the collection of data on performance indicators. Evaluation activities are directed on finding the difference between the current situation and the expected results according business plan as well as obtaining the performance estimates.

We can state that all business processes that take place in HEE influence its performance. This can be

expressed in terms of key performance indicators (KPI). The data necessary for finding of values of KPI should be obtained during monitoring activities. The data sources for such monitoring include internal and external ones. For example, if we consider return on investment value as KPI, then the data necessary for its calculation may be taken from the account department of HEE. However not all performance indicators are present in official records which are the internal sources of data.

In this work we suggest to consider the web as an external data source for performance measurement. Business processes of HEE find their reflection on the Internet. Different web-sites contain information related to the outcomes of universities work. Our goal is designing of monitoring system which can retrieve the web pages that are defined as data sources and to extract the data on KPI from them.

The rest of this paper is organized in the following way. Section 2 describes existing approaches of performance measurement and introduces performance measures in higher education. Section 3 outlines the main problems of web-based monitoring and their possible solutions. The agent realization of suggested framework is presented in section 4. The possible domain of experiments and prospects are given in section 5.

II. Performance measurement in higher education establishment

Performance measurement is an object of study in economy, project management and technical sciences. The formalization of this notion depends on particular model of enterprise management.

The most widespread approach for today in performance measurement is balanced scorecard (BSC) usage [7]. BSC allow translation of organization strategy into definite goals, tasks and indicators. It provides a powerful tool for performance measurement and control.

Another approach is Six Sigma strategy for business processes management [5]. It provides tools for performance reporting. According to Six Sigma KPI must reflect business processes behavior and must be measured, controlled, analyzed and improved.

Often when the object of research is service business, performance is associated with customers' satisfaction. In this case such approaches as GAP analysis [10], CSI (Customer Satisfaction Index) [17], benchmarking [16], and SERVQUAL technique [9] can be applied. All of them introduce some indicators for assessment of satisfaction of services' consumers.

These approaches have found their application in performance measurement in higher education [3]. Summing up we can say that all the approaches don't emphasize on the sources of data necessary to form the values of KPI. Therefore they don't consider methods for such data retrieval.

Possible performance measures suitable for universities include financial performance, research performance, teaching performance, community service performance [4]. We suggest to use the following performance measures for higher education domain taken from the general model of performance measurement in service business [14]. University business performance can be considered as from the four aspects: customers, finances and market, business processes and human resources.

Customers as stakeholders of HEE are represented by student (or graduates), government and society [2]. From students' point of view performance of HEE is expressed through their level of satisfaction. The attributes of their relation to obtained education show themselves on personal web-sites, blogs, social networks. Government is interested in keeping the balance of man power in economy, i.e. there should be neither shortage nor abundance of specialists of particular profession on the labor market. All KPI from the point of view of public authorities can be found easier in official statistics, than on the web. Society joins simply all people and enterprises. Corporate web-sites or sites of some business-projects may contain information about their employees and participants, which reflects the performance of HEE. For instance, information about some common projects between university and some enterprise indicates the achievements of HEE management.

Financial aspect of performance measurement takes into account financial characteristics (like return on equity, return on investment) and marketing indicators (such as market share, percent of new services on the market). Some of this KPI may be present on the web pages with the news, statistical reviews or marketing researches.

Among key business processes in HEE we can distinguish educational, research and methodical processes. Educational process is the basic one and its characteristics can be found on the Internet. Research business-processes are connected with organization of conferences, approbation of research results at conferences, publication of papers and mon-

ographs, membership in scientific communities, participation in research projects. All these activities are highlighted on the web on corresponding web-sites. Methodical business process deals with production of courseware and materials that can be useful for wide readership. The existence of such literature and information about its authors has an image on the web on review web-sites.

When considering human resources as performance measure we take into account turnover rates, suggestion rates and employees satisfaction. Data on corresponding KPI can be found in social networks, blogs, news and so on.

So it can be noticed that business processes of university can be evaluated through KPI and are reflected on the web. In this work we suggest to measure performance of HEE within web-based monitoring framework which will link KPI and data stored on the Internet.

III. Problems of web-based monitoring

The sources of data for web-based monitoring are distributed over web pages of different types. As it was discussed previously necessary data on KPI is concealed on the personal web-sites, web pages of social networks, blogs, conferences and journals, abstract databases, news, universities, business projects and corporative web-sites.

The goal of suggested web-based monitoring framework is finding of possible sources of data on KPI, retrieving these data and its processing for defining values of KPI (fig. 1).

The input for such system is determined by HEE business processes in the form of KPI. Depending on the management goals and management problems to be solved different KPI can be considered in different situations. Therefore the suggested framework represents a dynamic model of monitoring process.

Measurement Model provides statistical methods for calculation of values of KPI. Model of Sources specifies web pages that can be treated as sources of data on KPI, since initially they are not known. Further these sources are used as seed URLs for crawling process. A Crawler is tool for Internet browsing and collection of all web pages for their further indexing. The task of Clustering Model is diversification of collected web pages and grouping those of them that contain necessary data on KPI. Information Retrieval (IR) Model provides techniques for extraction of these data and their preliminary processing to be the input of measurement model. Evaluation component provides background data on the efficiency of crawling and IR process, accuracy of clustering.

So the monitoring process is composed of four stages: defining of sources for monitoring, finding

the necessary data starting with those sources, retrieval of these data and its processing. All of these problems have to be formally modeled and solved based on information retrieval and statistical techniques.

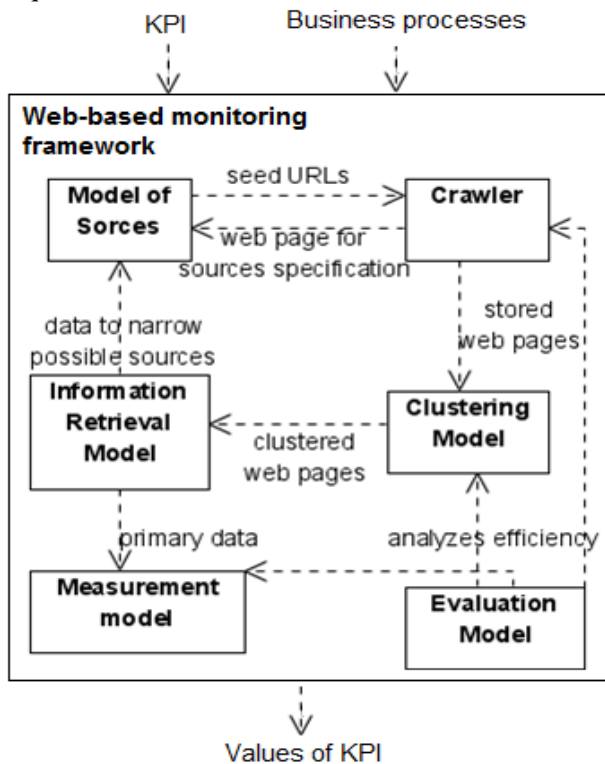


Fig. 1. Web-based monitoring framework

Since Model of Sources must provide a limited set of web pages that should be further used as seed URLs for web crawling, we assume that the process of such set construction requires experts' participation. An expert is a representative of HEE management who sets PMI and understands business processes. An expert must help to determine possible sources for monitoring of KPI. In this case we deal with topical information retrieval problem, which means that initially expert cannot formulate the correct and full request. That's why in the process of IR expert clarifies his request and obtains more and more improved responses from IR Model. Such type of problem is known as topical information retrieval. The possible methods to solve this problem include HITS, SALSA algorithms and others [6].

A Crawler must provide continuous and distributed searching of new web pages starting with the given initial pages. Web crawling algorithms are oriented on optimization of crawling by means of cutting off the paths of web graph which are found to be non-relevant. Among them we can name Fish search, Shark search, Context Graph search, PageRank and others [8].

Clustering Model is based on flat and hierarchical clustering methods [6]. Its goal is to throw away all noise web pages and to form the set of tar-

geted pages with the data on KPI. Clustering methods include algorithms based on TF-IDF measure, latent semantic indexing, topic models, Bayesian and support vector machines techniques [6, 1].

IR Model is based on XML retrieval and web scrapping methods [6, 15]. It must implement the procedure of extraction of necessary data from the definite web page for further processing. The reliability and validity of such data must be checked.

Evaluation Model combines measures of effectiveness of web crawling, clustering, XML retrieval and generally IR process. For example, precision, recall, NDCG (normalized discounted cumulative gain) estimates reflect the efficiency [6]. This information is needed for internal self-assessment of web-based monitoring framework. It can be used for system adjustment.

Measurement Model provides processing of statistical data on KPI. In fact we are talking about the latent variable models and estimation of their parameters. Such models include factor analysis model and its extensions [13], Item Response theory models [12], Bayesian network models [11] and others.

Generally the models that form the basis of monitoring framework incorporate extremely intelligent methods and algorithms. Their execution in continuously changing environment imposes constraints on realization of corresponding models.

IV. Multi-agent implementation of web-based monitoring framework

We suggest to use the concept of agent as an intelligent software entity to implement models of web-based monitoring framework. The multi-agent architecture of web-based monitoring system relies on the work of the set of agents that communicate via agent communication protocol and use the common knowledge base and ontology (fig. 2).

Topical search agent implements the Model of Sources and provides interactive procedure for topical IR. The work of crawlers is organized in parallel to make the web crawling process more distributed and efficient. Clustering and web scrapping agents are responsible for Clustering and IR Models implementation correspondingly and work with the common database. Measurement agent realizes statistical algorithms of KPI values estimation. Evaluation agent is used by agent management system to assess its work.

Such architecture is targeted at continuous collection and processing of data from the web. Such intelligent features as autonomy, interactivity, proactivity, ability to cooperate and to learn are supposed to help in realization of intelligent system of web-based monitoring.

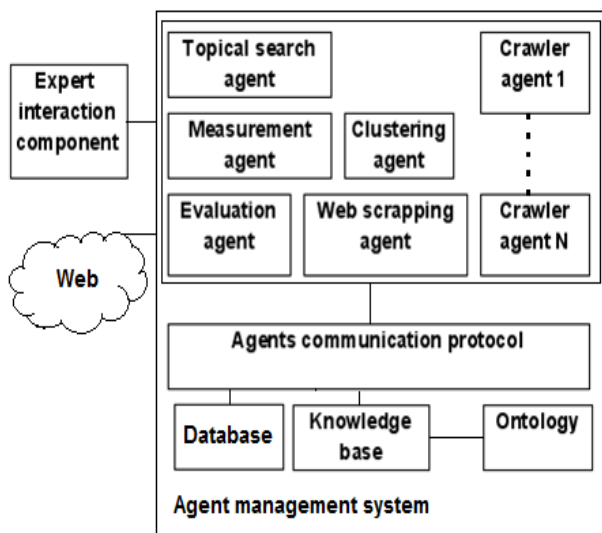


Fig. 2. Multi-agent architecture of web-based monitoring system

V. Discussion and future work

In our work we suggest measure the performance of research business process in HEE. The general tendencies in the development and organization of scientific work are actuated by university management. Therefore to build the right policy in this area management needs to estimate the results of research work, i.e. its performance. The scientific work of the whole HEE is expressed through the achievements of separate scientists working in it. KPI of research work may include the activities of HEE’s employees in participation in conferences and their organization, efforts to publish their works in journals, membership in scientific communities, and participation in collaborative projects with business and government.

The Model of Sources must help a representative of HEE management (expert) to define possible sources of data for further crawling. Topical IR presupposes that initially the request of an expert may

be “research performance indicators” and after iterative procedure of bounding of the set of possible sources an expert may choose for instance the web pages of conferences, abstract databases, and scientific societies. The Crawler uses those pages for web crawling and their clustering based on Clustering Model. After this only web pages containing useful information are processed by IR Model. For example, if we would like to estimate the activities of HEE scientists in conferences organization, we take as a seed URL a conference web page. A crawler retrieves all pages that it links to. And after application of algorithms of Clustering Model we get pages with program committees. IR Model extracts from program committee web page number of scientists with specified affiliation and Measurement Model provides the calculation of the corresponding KPI.

The challenges in constructing the described web-based monitoring system are caused by issues in implementation of each model. Topical IR underlying the Model of Sources requires efficient methods of relevance estimation of response web pages on each iteration of interaction with an expert. The choice of web crawling algorithm must be substantiated on the estimates of its performance compared to others. The clustering accuracy influences the probability to miss the useful web page. Web scrapping algorithms make challenge to unstructured data on the web. Finally, processing of data by Measurement Model requires estimation of reliability of obtained values of KPI.

The given work represents the idea of measurement of performance of HEE functioning based on data sources from the web. The suggested web-based monitoring framework for performance measurement introduces particular models to be realized. Implementation issues of each model mentioned above form the future work of current research.

References

1. Blei D.M., Lafferty J.D. Topic models. Retrieved May 03, 2013 from <http://www.seas.harvard.edu/courses/cs281/papers/blei-lafferty-2009.pdf>
2. Cherednichenko O., Riabko O. Information technology of higher education monitoring: a quality-based approach // Eastern European journal of enterprise technologies. – 2011. – Vol. 1/7 (49) . – P. 4-6.
3. Development of performance measurement instruments in higher education. Australian Government. Retrieved December 13, 2012 from http://www.innovation.gov.au/HigherEducation/Policy/Documents/DevelopPerformanceMeasuresHigerEd_Finaldiscussionpaper.pdf.
4. Lee A., Yuningsih Y. Multiple performance measures of higher education: instrument development and validation. Retrieved January 21, 2013 from [http://www.afaanz.org/openconf/2010/modules/request.php?moule=oc_proceedings&action=view.php&a=Accept](http://www.afaanz.org/openconf/2010/modules/request.php?moule=oc_proceedings&action=view.php&a=Accept+as+Paper&id=338)
5. Lowenthal J.N. Six Sigma project management: a pocket guided. Milwaukee, Wisconsin, ASQ Quality Press, 2002.
6. Manning C, Raghavan P., Schutze H. An introduction to information retrieval. Cambridge UP, 2008.
7. Nair M. Essentials of balanced scorecard. John Wiley & Sons, 2004.
8. Novak B. A survey of focused web crawling algorithms // Proceedings of SIKDD. – 2004. – P. 55-58.
9. Parasurman A., Zeithaml V., Berry L. A conceptual model of service quality and its implications for future research // Journal of Marketing. – 1985. – Vol. 49. – P. 41-50.
10. Predvoditeleva M. D., Balayeva O. N. Approaches for Service Quality Management: a Focus on the Con-

- sumer // Management in Russia and Abroad. – 2005. – Vol. 2. – P. 90-100.
11. Rebai F. Bayesian Network. InTech, 2010.
 12. Reeve B. An Introduction to Modern Measurement Theory. Retrieved November 1, 2012 from [http:// appliedresearch.cancer.gov/ areas/cognitive/immt.pdf](http://appliedresearch.cancer.gov/areas/cognitive/immt.pdf).
 13. Ross S.M. Introduction to probability and statistics for engineers and scientists. London, Elsevier Academic Press, 2004.
 14. Sapri M., Kaka A., Alias B. Performance measurement in the service business: the facilities management function. Retrieved April 02, 2013 from [http://eprints.utm.my/1139/1/ Maimunah.pdf](http://eprints.utm.my/1139/1/Maimunah.pdf) [Real Estate Educators and Researchers Association (REER) Seminar 2005].
 15. Schrenk M. Webbots, spiders, and screen scrapers: A guide to developing Internet agents with PHP/CURL. No Starch Press, 2012.
 16. Stapenhurst T. The Benchmarking Book: A how-to-guide to best practice for managers and practitioners. Elsevier, 2009.
 17. Zhang L., Han Z., Gao Q. Empirical study on the student satisfaction index in higher education // International Journal of Business and Management. – 2008. – Vol. 3. – N 9. – P. 46-51.

Чердниченко О.Ю., Янголенко О.В.

Модели веб-ориентированного мониторинга для оценивания производительности в сфере высшего образования

Аннотация: Данная работа представляет модель веб-ориентированного мониторинга для измерения ключевых индикаторов эффективности функционирования высшего учебного заведения. Предложено рассматривать веб-пространство в качестве внешнего источника данных для нахождения значений высокоуровневых индикаторов производительности работы университета. В работе представлены модели, составляющие основу веб-ориентированного мониторинга, и описаны их цели и назначение. С помощью модели источников данных определяются веб-страницы, которые используются веб-краулером в качестве стартовых страниц для поиска. Модель кластеризации позволяет отсортировать собранные страницы и получить множество страниц, содержащих необходимые данные. Модель извлечения информации предоставляет алгоритмы для извлечения этих данных и их первичную обработку. Модель измерений предоставляет статистические методы для расчета значений индикаторов на основе извлеченных данных. Оценочный компонент используется для обработки вспомогательных данных, касающихся эффективности работы веб-краулера и информационного поиска в целом, точности кластеризации, а также надежности и валидности полученных результатов. Для реализации предложенной модели веб-ориентированного мониторинга предложена мульти-агентная архитектура информационной системы. Согласно ней каждой модели соответствует программный агент. Агенты взаимодействуют с помощью коммуникационного протокола и функционируют в соответствии со знаниями, представленными онтологией и базой знаний.

Ключевые слова: измерение производительности, высшее образование, веб-ориентированный мониторинг, поиск информации