

# 2011 International Conference on Internet Technology and Applications

## iTAP 2011

*Proceedings*



<http://www.itapconf.org/2011>

Aug. 16~18, 2011, Wuhan, China



# Methods of the content mining as the feedback mechanism in e-government

Rasim M. Alguliev

Department of Information Society problems  
Institute of Information Technology of ANAS,  
IIT of ANAS  
Baku, Azerbaijan  
[secretary@iit.ab.az](mailto:secretary@iit.ab.az), [depart17@iit.ab.az](mailto:depart17@iit.ab.az)

Farhad F. Yusifov

Department of Information Society problems  
Institute of Information Technology of ANAS,  
IIT of ANAS  
Baku, Azerbaijan  
[farhadysifov@gmail.com](mailto:farhadysifov@gmail.com)

**Abstract**— The paper describes the brief overview of the basic principles of formation and management of e-government. A content mining method of improving the efficiency of governance indicators of e-government has been proposed.

**Keywords**— *Information society, Electronic Government, Electronic Governance, Transformation, Public Management, content mining*

## I. INTRODUCTION

Structure of the government can be examined as a developed information system, as main principals and process of public management are based on reception, processing, analysis and information dissemination. At present, we can note that, without use of information-communication technologies (ICT), activity of the government in information society will not meet requirements and principals presented by the society. Introduction of ICT in government organizations, in such manner, is an answer to requirements presented for operating effectiveness of governance and is connected to concept of new public management.

Electronic readiness enhancement and acceleration of development of electronic government are the most noticeable trends of global informatization and formation of world information society. Regular monitoring of these trends is the main task of many international and national organizations [1-6].

Electronic government (E-government) is a conception of state management implementation that characterizes information society [7-8]. It is based on capacity of information-telecommunication technologies and values of open civil society. The e-government service consumers can be citizens, non-government organizations or government authorities.

One of the important tasks at the e-government creation level is realization of effective governance of e-government. Application of the method of content mining may improve the efficiency of governance indicators of e-government.

## II. INFORMATION ENVIRONMENT AND MONITORING OF E-GOVERNMENT

Ideally, e-government should provide all the citizens with fast and simple access to all services of government institutions and an opportunity to realize all their constitutional rights in participation in the public administration. The essence of e-government is to provide people with access to information.

At very early stages of e-government development program realization the future e-government work principles need to be followed. At this stage an important task is formation and development of information environment [4,7-9].

Generally an information environment can be presented by three basic component:

- Languages, displaying information, which are used for information interaction in the country;
- content volume, meaning content and information value;
- information-communication infrastructure that is material base of information interaction.

It may be said that contents and characteristics of information-communication infrastructure in the aggregate determine the information space of a country.

Conception of information environment covers both the society en masse and any scope of activity. Information environment of science or of its separate areas, information environment of economy, culture etc. can be referred to as information environment of the world, country, region, city.

In modern society information environment has mainly a network structure. Major communication form between the points of information environment is Internet that provides technological opportunity for each individual to have information interaction with another individual.

Development of information environment in different countries is going on irregularly in space and time subject to the level of their development [5-7,9]. Also, subject to the level of development of country and its information environment in whole the e-government realization program can be assessed.

E-government development process tracking and assessment of its effectiveness are the special tasks because

subject of research is dynamic and under explored. Majority of the researches are aimed at study of e-government potential i.e what can citizens, businesses and government authorities get from use of information-communication technologies upon providing and consumption of service. Suggestions on formation of e-government monitoring system [4, 8, 10-12]:

- indicators of e-government in a country is to be formed considering the level of interactivity of on-line government services;
- indicators should be formed depending on level of governance;
- upon developing e-government and change of external situation the indicators should be adjusted;
- indicators should be used related to the level of interactivity, quality etc.

At the stage of development and governance of e-government the list of indicators should include: quantity of services provided online, the quality of these services and demand for the services.

Analyzing the scientific literature in this field, we can conclude that “electronic government” consists of give main modules [4, 10-12]: G2G (government to government), G2B (government to business), G2C (government to citizens), G2N (government to non-government organizations) and G2K (government to knowledge). Among these components, we can distinguish two modules: G2N and G2K, which are very important for completeness of interrelations in social life. Main goal of G2N processes – is to forge interactions of the government with citizens during decision-making, a G2K – is interaction of government and scientific sector, which is directed to solving problems of application of innovative technologies and creation of new mechanisms of public administration by development and use of knowledge in the society.

Particularly, international organizations such as UN, ITU, OECD, World Bank, CID etc play a major role in research of e-government. Also, researches of companies such as Gartner, Microsoft, CISO, HP etc in the field of implementation of electronic government is a good sign, as given organizations and companies possess a sufficient material base for conduction of qualitative and important scientific-practical researches. In its turn, it allows to attract significant financial resources to the field of research of electronic government. It also must be noted that, given organizations and consulting companies are interested in certain coverage of electronic government problem by social-political goals.

In researches of UN (Benchmarking E-government: A Global Perspective), conception of “development index” of an e-government (e-Government Index) was applied and a certain value of that index for 169 countries of UN was introduced [5]. In accordance with the results of the year of 2009, e-Government development Index for 192 countries of UN was determined [6].

### III. SOURCES AND METHODS OF INFORMATION COLLECTION

At the stage of governance of e-government another important point is information collection and methods for efficient analysis. Generally the sources could be:

- Websites of government institutions;
- IT-division of government institutions;
- Users are citizens, organizations and civil servants.

As a matter of principle, study of the sites provides more reliable information. Two methods can be used for this:

- 1) Study of the sites by administrators (experts), who monitoring the site and give their opinion;
- 2) Users identification - who work with the sites and study the area of interest of the users.

Upon scale studies, when the number of sites and the users is very high, it would be reasonable to apply intellectual methods of analyzes such as methods of content mining [13-16]. Content is any informative signify loading of a server, information complex – texts, diagrams, multimedia. Considerable parameters of content are its volume, urgency and relevance.

In addition, rapid development of modern information technologies, the growth of Web resources available on WWW increases the necessity of the use of the method of content mining for administrators (experts), in order to extract and assess necessary information and make an optimal decision.

Registration files (log-files) on server have always been the main source of information on traffic through the Web-site and on users’ conduct. The overwhelming majority of control means and Web-site operation analyses are based on the use of these files. Such parameters as the volume of traffic received by the Web-site, visitors address, their access ways to the site and response to the site content are assessed based on information on access to this site. However there is more and more demand for in-depth and more detailed comparison with the basic information on the Web-sites.

At present, special attention is given to creation of methods and software for the analysis of log files data, to the monitoring of network activity, and also to tracking of viruses and hacker programs. The basic task of the Web-traffic analysis software tools is discovering useful information from the log files of the server. Detailed analysis of the Web-servers log files professional Log-analyzers are used. It should be noted that at the moment the state of Web-analyzers does not completely meet the listed requirements.

On the basis of application of a content mining method it is possible to use elementary general administration of local resources, content update monitoring, the statistics of attendance and forums for users and specialists that are characteristic of e-government as a control mechanism. The application of a content mining method helps to improve the efficiency of the governance indicators of e-government.

Use of content mining methods gives rise to the need of creation of server and client intellectual systems which can effectively acquire knowledge from the Internet. The application of content mining methods can help to solve

number of such problems as clustering and analysis of users, study of Web- resources content, activity of visitors with regard to use contents and others [15-17].

Also it should be noted that in order to increase transparency, effectiveness of the utilization of state funds and the effectiveness of services provision to citizens the public administration must use the tools of information communication technology (ICT). The services of government sector must ensure the transparent and attainable solutions, be interactive, and be on-line when necessary. ICT tools must be used in order to make policy more accountable and to make the best monitoring, evaluation and control possible.

#### IV. CONTENT MINING METHODS

With the advent of the automation tools, texts in the electronic form, effective 60s of past century an initial development of the content mining of the information of large volumes was obtained - the data bases and interactive media-sources. The traditional political use of content mining contemporary technologies was augmented by the unlimited list of headings and subjects that cover production and social spheres, business and finances, culture and science. This process was, in turn, accompanied by the large number of different program systems [13-18].

The interesting feature of content mining is the fact that until recently this methodology was attributed to a certain sphere of human activity (policy and sociology). Nevertheless, today content mining is increasingly used in many fields of political and economic life and this contributes to the larger applied relevance to use sociology and policy in the methodology of the content mining. Based on the received information conclusions are made and recommendations are formed at the stage of the data mining regarding making that very managerial decision which everything started with. After the preparation of report about the carried out estimation and "feedback" according to the results of estimation the final stage comes - making a decision considering the received information.

The distribution of the methodology of content mining to two branches is considered as universally recognized: qualitative and quantitative. The basis of quantitative content mining is frequency of occurrence of certain content characteristics in the documents. The method of qualitative content mining is based on the very fact of presence or absence of one or several characteristics of the content in the text.

The method of qualitative content mining is based on the fact that at any phase of quantitative content mining an expert can be involved for the evaluations of results. Thus, this method is intended to provide the expert with the necessary tools for conclusions and additional results. The expert with the aid of such tools can reveal specific properties of the part of information and verify them against the general text flow and extend general properties of the text flow to its specific subject part.

It should be noted that content mining is first of all a quantitative method that assumes the numerical estimation of some components of the text that can be supplemented also

with different qualitative classifications and revelation of various structural regularities. Therefore the most successful definition of content mining can be considered the one provided in the relatively recent literature – it is a systematic numerical treatment, estimation and interpretation of form and content of the information source.

From the point of view of ICT specialists the content mining is a typical example of the applied information text analysis that contributes to the extraction of some special components that are interesting to a researcher from the entire variety of existing information in it and their presentation in a form convenient for the perception and subsequent analysis. The numerous specific versions of content mining are distinguished depending on what these components are and what is precisely understood under the text [19, 20].

Thus, the idea of content-monitoring can be formulated in the simplest form as a constant implementation of content mining of continuous information traffics that is narrowly outlined by its problems. Let us emphasize that it is continuous reproduction of input data processing in time is the very characteristic of content-monitoring. Actually, content mining comes out here as composite, and content-monitoring has got its own problems and methods to solve applied problems.

Discovering of knowledge can be defined as finding and analysis of useful information. Use of methods of content mining gives rise to the need of creation of server and client intellectual systems which can effectively analyze the content of Web-resources.

#### V. CONCLUSION

In conclusion it can be said that the development of Information Society has been and remains the priority trend of different countries. Propagation contributes to economic growth, creation of new jobs, it connects the distant localities with the urban centers and improves standard of life.

Application and implementation of new methods in the field of management of e-government means not only the creation of effective governance of workflow, creation of internal information networks, and databases in the agencies of public administration and offices that makes it possible to improve interaction between different agencies of public administration, but also opportunity for population to access the services of governmental institutions and public sector through the information networks.

#### REFERENCES

- [1] O. Andronova and A. Nikolaev, "Electronic government in Europe and world", 2006, [http://www.ci.ru/inform22\\_01/p\\_0600.htm](http://www.ci.ru/inform22_01/p_0600.htm) (In Russian)
- [2] N.A. Antanovich, "Electronic governance» as the passage technology from public administration to effective public management", 2005, [www.law.bsu.by/pub/11/ant3.pdf](http://www.law.bsu.by/pub/11/ant3.pdf) (In Russian)
- [3] M. Finger and P. Rossel, "From e-government to e-governance: implications for technology management", 2006, <http://infoscience.epfl.ch/record/89522/files>
- [4] M. Kovalev, "Development of electronic government considering the international experience", Transactions Banking, Cherven, 2006, pp. 16–25 (In Russian)

- [5] Benchmarking. E-government: A Global Perspective. Assessing the progress of the UN member states. 2002, <http://unpan1.un.org/intradoc/groups/public/documents/-UN/-UNPAN021547.pdf>
- [6] United Nations E-Government Survey 2010, [http://www2.unpan.org/egovkb/global\\_reports/10report.htm](http://www2.unpan.org/egovkb/global_reports/10report.htm)
- [7] R. De, "Electronic governance theory", ICEGOV'2008, December 01–04, 2008, pp. 11–15
- [8] R.C. Joseph and D.P. Kitlan, "Key Issues in E-Government and Public Administration" [https://irma-international.org/downloads/excerpts/-reference/IGR5202\\_1YPRPC8baU.pdf](https://irma-international.org/downloads/excerpts/-reference/IGR5202_1YPRPC8baU.pdf)
- [9] V. Godse and A. Garg, "From E-Government to E-Governance", ICEG'2007, December 28–30, 2007, pp. 13–20
- [10] T.F. Gordon, "eGovernance and its Value for Public Administration", [www.tfgordon.de/publications/Gordon2004a.pdf](http://www.tfgordon.de/publications/Gordon2004a.pdf)
- [11] S.C. Jain Palvia and S.S. Sharma, "E-Government and E-Governance: Definitions/Domain Framework and Status around the World", ICEG'2007, December 28–30, 2007, pp. 1–12
- [12] R. Traunmüller and C. Leitner, "e-Government: State and Perspectives", ICEGOV'2008, December 01–04, 2008, pp. 4–7
- [13] G. Stumme, A. Hotho, and B. Berendt, "Usage mining for on the Semantic Web", <http://www.kde.cs.uni-kassel.de/stumme/papers/2004/-berendt04usage.pdf>, 2004.
- [14] R. Cooley, "The Use of Web Structure and Content to Identify Subjectively Interesting Web Usage Patterns", ACM Transactions on Internet Technology, Vol. 3, No. 2, pp. 93–116. 2003.
- [15] R. Meo, P. L. Lanzi and M. Matera, "Integrating Web Conceptual Modeling and Web Usage Mining", Proc. of WebKDD'04 (International Workshop on Web Mining and Web Usage Analysis), Seattle, USA, August, ACM Press. <http://maya.cs.depaul.edu/webkdd04/final/meo.pdf>, 2004.
- [16] M. Spiliopoulou, C. Pohle and M. Teltzrow, "Modelling and Mining Web Site Usage Strategies", In Proceedings of the Multi-Konferenz Wirtschaftsinformatik, Nurnberg, Germany, [http://www.iw.uni-karlsruhe.de/fgecommerce/archiv/download/tagung\\_erlangen/-Track5\\_2\\_Spiliopoulou.pdf](http://www.iw.uni-karlsruhe.de/fgecommerce/archiv/download/tagung_erlangen/-Track5_2_Spiliopoulou.pdf), 2002.
- [17] A. Abraham, "Business Intelligence from Web Usage Mining", Journal of Information & Knowledge Management, Vol. 2, No. 4, <http://arxiv.org/pdf/cs.AI/0405030>, 2003
- [18] K. Balog, P. I.Hofgesang, W. Kowalczyk, "Modeling Navigation Patterns of Visitors of Unstructured Websites", The Twenty-fifth SGAI International Conference on Innovative Techniques and Applications of Artificial Intelligence (Springer Verlag), [http://www.cs.vu.nl/ci/DataMine/DIANA/pages\\_eng/publications.html](http://www.cs.vu.nl/ci/DataMine/DIANA/pages_eng/publications.html), 2005
- [19] N. Shadbolt, T. Berners-Lee, and W. Hall, "The Semantic Web Revisited", IEEE Intelligent Systems. 21(3), pp. 96-101, [http://eprints.ecs.soton.ac.uk/12614/01/Semantic\\_Web\\_Revisted.pdf](http://eprints.ecs.soton.ac.uk/12614/01/Semantic_Web_Revisted.pdf), 2006.
- [20] J. P. Norguet, E. Zimanyi, and R. Steinberger, "Semantic Analysis of Web Site Audience", Proceedings of the 2006 ACM symposium on Applied computing. <http://code.ulb.ac.be/dbfiles/media260.pdf>, 2006.