EKW: eGov – THE SYSTEM TO SUPPORT PROCESS OF GRANTING VISAS

Witold Abramowicz, Andrzej Bassara, Agata Filipowska, Marek Wisniewski
The Poznan University of Economics
Poland

ABSTRACT

The purpose of the paper is to present the system that can support national authorities in the process of granting people with visas by enhancing the information about the people who apply for them. The process of gathering such information can be enhanced with external information filtered out from external sources (dedicated sources in the Internet) and derived from databases that are not accessible to the public. Moreover, the article touches the subject of e-government as a compilation of e-administration and e-democracy and the Polish state-of-the-art in the e-Government solutions.

E-GOVERNMENT

Modern e-administration should use three major concepts (see Socitm and Idea Report, p.5):
. E-services: it is the area which all the above mentioned programs take care of. It is aimed at the introduction of the administrative services.
. E-governance: it focuses on the inclusion of society for the purposes of administration.
. E-knowledge: the cornerstone for the information society. It is designed to stimulate the information cycle among the parties involved.

Depicted view can be seen from many aspects. Gartner’s definition of e-administration (see Socitm and IdeA Report, p.4) covers three areas of interest: customer services, internal efficiency, and citizens engagement.

Gartner’s vision is even more focused on e-democracy as the last aspect is nothing less than e-democracy itself. Therefore, derived from e-administration, e-democracy will be defined as a willingness of the community to participate in the e-government structures and work. E-democracy is the crucial aspect of the whole e-government policy success.

E-democracy is not a levelled tendency. It can take place on very different levels. Four basic ways of providing e-democracy can be differentiated:
. common involvement in the realisation of conceptualism
. focusing on the most important issues for citizens
. collaboration and project refinement in order to achieve better results
. seek for innovation

Several approaches have been developed in order to introduce the vision of e-government in Poland. Most of them, notably the most comprehensive ones, are the government’s documents. Three major programs presented in the following initiatives can be listed: e-Polska (e-Poland) - based on e-Europe document, Narodowy Plan Rozwoju (National Development Plan), and Wrota Polski (Gates of Poland). It is not the purpose of this paper to present the idea of each one of them, however, some conclusions can be drawn. Although all of the above mentioned initiatives are sound and comprehensive, they influence only the fundamentals of e-government, they create structures of e-administration. One thing that ought to be stressed is that none of the ways of providing e-democracy have been really developed. All initiatives covered earlier are general documents. Although they touch indirectly some of the issues, first policy and later particular actions are needed.

The current Polish actions concentrate on integration of existing databases and other sources of information about citizens, real estates, cars - so called national registries. The
record of all Polish citizens and the people with long-term permission for living in Poland, which is called PESEL has existed in Poland since 1975. Now the registry of all real estates for the purpose of cadastral tax is being implemented. We believe that national registries are the background for other e-government projects.

**NATIONAL REGISTRIES**

A national registry is a special kind of inventory that is maintained by the authorities for their own needs. Definition of such registry can be divided into three separate parts: why and what for the registries are created, how they are organized and what the features of single registry are.

The purpose why the registries are created is connected with the state policy – governmental plans ( economical, social, etc.) need a complex information about the situation in the country which should be accurate and up-to-date. Such data stored in registries is the most important source of information about current trends. These facts can be observed and thanks to them, government can take some actions that counteract.

National registries store the data about citizens, things possessed by them, companies and other things in the country, that are important for authorities in the process of administering the country. The architecture of such registry, the way of maintaining and using the data is regulated by law. Because of the importance of such data, it is not fully accessible for the public, although most of the data stored in registries is shared with people on their demand.

The important feature of the national registry is that it is updated on the basis of the official document when something changes, but the previous data is not overwritten, it’s also stored and the previous owner of, for example real estate, can be identified. Moreover, such way of managing the data enables multiple analyses whose outcomes are usefull for the authorities. Other important thing is that the data stored in the registry is supposed to be accurate from the legal point of view. It means that data in the registry reflects the reality.

Such registries are used on different levels of administering the country, they cover different subjects, but they store detailed data. The more detailed, the better. For that reason enhancing this data gathered only by clerks with some external data could improve the decisions and actions taken by the authorities.

**ENHANCING THE REGISTRY WITH ADDITIONAL INFORMATION**

The registry stores mainly the data that is changed in the process of administering the country or the region by the authorities. But that is not the whole information that can be found about real estate, citizen, etc. For example, information on real estate can be found in such registry, in the newspaper, in the real estate’s agent database, on the web page of the owner, etc. There are many different sources of information that can be used, but they should be identified at first. The idea of using the information that comes from heterogenous sources in order to enhance the information stored in the data warehouse was described in (Abramowicz, Kalczynski, Wecel, 2002). eKW (enhanced Knowledge Warehouse) concept is based on the idea of building the profiles of the data warehouse that depict the data stored and using the profiles in the process of filtering the data from the web.

The idea presented in (Abramowicz, Kalczynski, Wecel, 2002) is developed for the purpose of cadastral tax and the evaluation of the real estates value in the eVEREst (enhanced value estimation of real estates) project (see Bassara, Filipowska, Wisniewski, ITEE 2003). The purpose of the project is to enhance existing information on real estates (from different registries, land and mortgage books) with information filtered from the web to estimate the value of single real estate with greater accuracy. This concept can be used and modified based on different registries’ profiles. As the result, the process of granting people with long-term visa will be presented further in the article.

**THE SYSTEM PROPOSAL**

The Poznan University of Economics with collaboration of the Wielkopolska Voivodeship Office (the biggest unit of territorial administration in Poland) intends to build a system to support civil servants with
external data and information (i.e. data that do not exist in any other system that works within the office) on specific group of office’s applicants. The need for creation of such system is justified by dual nature of granting rights in Poland. Most permissions are granted after satisfying all prerequisites by applicant, the second group consists of permissions issued in creditable character. This group consists of rights that imply long-range after effects, but revoking of such rights is troublesome or even impossible (for instance nationalising). In this case, not only all formal prerequisites are to be satisfied, but also expert knowledge of clerk is being taken into consideration. It is the clerk, who regardless of fulfillment of prerequisites, can refuse granting allowance basing his decision on unclear, uncertain factors. In order to improve the quality of clerks’ decisions, in our system it has been decided to equip them in extra information on people soliciting for certain permissions (in case of this systems i.e. permission for permanent residence or naturalisation). This information is to originate from sources other then submitted application, conducted from interviews or offices registers.

Authors of the proposed system believe in multiplicity of information existing in external sources, such as databases of international organizations, Interpol or press notes that can influence both positively and negatively offices decision. The construction of such a system is going to have a positive impact on inner country security and accelerated satisfaction of office’s customers resulting from improved effectiveness of clerks’ work and transparency of undertaken decisions.

THE ARCHITECTURE OF THE PROPOSED SYSTEM

Processes

Defining such system and its architecture end-user needs should be defined at first. This definition construction cycle is based on the processes approach. Therefore the processes that the system will be based upon will be presented.

Process #1 Profiling

When a person applies for the visa, he needs to fill in a form. Based on the data from the form the clerk is to find information on the specific person. He checks the system to see if there is any information on the person, if there is – the profile of the specific person exits and it is possible that the additional information has been filtered out, the only action that happens is rebuilding the profile using the new data from the form. If the person hasn’t applied for the visa earlier, the clerk must bring in all the data to the system, and then the profile will be automatically created by Profiling Service that is connected with Information Repository. Based on the profiles stored in the Information Repository Web Crawler’s agents instantly search in the predefined sources for information matching the profiles.

Process #2 Information and Data Harvesting

Source Controller containing references queries periodically static web sites looking for any changes within structure of source and changes of presented information. If a new profile is created or one is rebuilt, Source Controller queries dynamic sites and external databases for specific data and information. Every document acquired is stored in Repository and indexed - all its metadata fields are filled in. Periodically and also on demand Reasoner processes acquired collection of documents in order to assign a person or a group of people to each document.

Process #3 Decision Support

The clerk wants to receive information on a specific person. He or she provides the system with a person identifier (name and surname, identification number), which triggers a query that results in updating the data and information from external sources and provides the user with a list of documents related to the person.

Architecture

The proposed system is to consist of the following components: Information Repository, Reasoner, Gateway, Source Controller, Web Crawler, Interfaces for Specific Sources.
Information Repository
The central unit of the system is the Information Repository. It contains both information and data retrieved from external sources as well as profiles. The basic unit of data and information storage is document. It consists of data or information body (i.e. plain text document retrieved from web site or xml document retrieved from external database) and set of predefined and loose (not included in system meta-data set collection) metadata fields. Besides fields such as title, retrieval date, type of source whose values are relatively easy to determine, there are also two fields - one for people described in the body of the document and the second for reliability of association of a given document with the specific person. Values of these fields are estimated by the Reasoner.

Reasoner
The most challenging mechanism to be developed is the Reasoner. This service is responsible for utilisation of huge amount of data stored within Information Repository in form of documents and profiles. Techniques such as text mining and shallow text processing with massive computing will be used in order to identify with great reliability persons involved in each information will be used (see Abramowicz and Piskorski, 2003, p.12).

Source Controller
Source Controller is the service containing pointers as well as all necessary data to access external sources. Pointers are mainly URI for external services which are: static web sites, dynamic services or external databases. Apart from the pointer to such a service, the source controller is equipped with all interfaces necessary to interact with the source (like: proper access rights or implementation of specific metaprotocols). Depending on the type of source, the source controller may check the source periodically for changes within the presented information (mainly concerning static web sites) or it can trigger the query on events like users interaction with the system or creation of a new profile (mainly concerning dynamic services). The Source Controller is equipped with two tools to harvest the information which are Interfaces (native protocol of source service) and the Web Crawler. All harvested information will be stored within the system. The proper infrastructure will be provided for the purposes of confidentiality. Conceptually, the encapsulation of system's components provides all but necessary mechanisms for the proper safety of stored data.
Web Crawler

Web Crawler is a robot capable of traversing through the web from specific point in form of URL within exact range in order to look for certain information. Web Crawler is equipped with Information Repository’s information needs expressed in form of subject profiles and source description coming from Source Controller. Every document retrieved by Web Crawler is submitted to Information Repository for further processing and storage.

Gateway

Polish law does not allow maintaining certain data only in electronic manner. This restriction conforms to personal data gathered by voivodeships’ offices. That approach imposes certain limitations (at least until law changes). At this time currently utilised systems are not allowed to work within the same networking system with other software. For this reason data input for building profiles is going to be manual. For the same reason Information Retrieval procedures are also going to utilise human resources. Gateway is going to be designed and implemented to make usage of common protocols, and data description, and manipulation languages (like Web Services) in order to be easily replaced with interface to other operational systems.

CONCLUSION & SUMMARY

In the paper some prerequisites for an e-gov system and the short overview of such a system that can support administration in the process of granting people with long-term visa was presented. Creating such solution that would meet the needs of its users depends on providing the system with comprehensive sources of information. If the information is available in any source, the system can use it matching it to specific profile. That is why, the more information sources system will use, the better its outcome will be. In the information society the most important is the knowledge derived from the information. Enhancing any process with additional information means actually eliminating accidental decisions that can be taken. And that is the future of e-government.

REFERENCES

Steven Clift, The E-Democracy E-Book: Democracy is Online 2.0
Socitm and IdeA (2002), Local e-government now: a worldwide view, Executive Summary


