TOWARDS THE SEMANTIC E-GOVERNMENT*

Robert BUCHMANN
Radu MEZA

Robert BUCHMANN
Lecturer, Business Information Systems Department,
Faculty of Economics and Business Administration,
"Babeş-Bolyai" University, Cluj Napoca, Romania
Tel.: 0040-724-801.812
E-mail: robert.buchmann@econ.ubbcluj.ro

Radu MEZA
Assistant Professor, Journalism Department,
Faculty of Political, Administrative and Communication Sciences, "Babeş-Bolyai" University, Cluj Napoca, Romania
Tel.: 0040-264-431.505
E-mail: radu.meza@polito.ubbcluj.ro

Abstract

E-government is quite a late achievement of the information society, although, in theory, it could have been the first application of the Internet, since Internet itself started as a governmental project. Many factors have been delaying the implementation of e-government, the most important ones being the digital divide, the privacy and security concerns and the availability of common data models on a national and international scale. Since its implementation comes much later than other on-line organizational models (e-business, social networks etc.), e-government also lags technologically. In Romania, but not only, e-government and public administration applications are data-driven rather than knowledge-driven, since databases are quite popular and robust as a standard model for storing the information behind traditional Web applications. On the other hand, the World Wide Web authorities invest a great deal of effort in raising the awareness of the so-called "knowledge society", defining formalizations for knowledge-aware applications. The paper discusses the threshold between data and knowledge, the implications of the knowledge society on e-government and public administration, and proposes a knowledge generation methodology based on existing data-driven systems.

Keywords: e-government, semantic web, knowledge, rules, databases.

* This paper presents the results of the research project 2443/2009 funded by the UEFISCSU-CNCSIS Program IDEI, managed by Lect. Robert Buchmann Ph.D.

1. Introduction

During the last decade, notions such as “information society” and “knowledge society” evolved in abstract notions widely used as rhetoric elements for stating the relevance and actuality of some organizational strategy, mostly from a management viewpoint, where the term “knowledge” gained in the early 1990s some weak formalization (Ikujiro, 1991), quite different from the more formal view provided earlier in epistemology and artificial intelligence. Such disparate perspectives fail to make a distinction that is valuable under today’s requirements for an e-society. That is not to say that such terms are ill-defined or indistinct, but rather that their definitions rely heavily on interpretation as long as they are not understood by way of technical formalization.

The data-knowledge distinction becomes clearer in the context of a recent paradigm called “the Semantic Web” (or “Web 3.0”), which aims to convert the World Wide Web from a network of documents to a network of knowledge, with search engines acting as expert systems interfaces, answering questions based on a cloud of interoperable formal knowledge and reasoning mechanisms, rather than simply searching for word combinations in interlinked text documents. Our paper provides intuitive insight regarding this paradigm and the data-knowledge distinction, with the aims of discussing their impact on the future of e-government, and of promoting a technical solution to support this migration.

The next section will cover the problem statement, debating what e-government should become in order to synchronize with the technological trends. Section three discusses the data-knowledge dichotomy, with insight into what sets these notions apart from a functional viewpoint, and emphasizes the formal knowledge representation as an alternative to data-driven systems. Section four describes the proposed methodology for obtaining knowledge from existing data sources, while the last section discusses the implications on public administration seen as a test bed for implementation. A demonstrative tool was also developed in order to support a collaborative SWOT analysis.

2. Problem statement and background

First, we need to stress the motivation of our research – the study started from an endeavor for finding out what branches of public administration in Romania have integrated semantic technologies and formal knowledge in their business processes, reporting and information models, but we could not identify any. All existing systems are rooted in Web 2.0 and data-driven systems. Even when migrations occur, the results are still data-driven systems, maybe some that are better performing or more service-oriented, but failing to involve the interoperability provided by formalized knowledge. Even projects that are under development (such as the unique student enrollment registry (National Students Enrollment Registry)) are aimed to be implemented in traditional, service-oriented architectures. One reason is that such systems are improvements of existing models rather than being redesigned visions that might affect the process landscape rather than simply upgrading infrastructure. Apparently, there is a lack of awareness regarding the potential of the Semantic Web/Linked Data paradigm (Linked Data Design Issues).

Lately, the European Commission stresses the need to raise such awareness. An important initiative in this regard is the FInES Cluster Task Force, grouping more than 30 FP6 and FP7 projects, aimed towards the interoperability of small and medium enterprises, but promoting principles that are equally applicable to public administration (FInES).

The key term used by the European Commission is "the Future Internet", with aliases such as "the Internet of Things" (as opposed to the current state of an "Internet of Documents"). Several projects, such as LOD2 (Linked Open Data 2), aim to demonstrate the benefits of interlinked formal knowledge over traditional Web documents and databases. However, tool development is only one side of the problem. User awareness and training is also of essence in order to support this shift, because, ultimately, formal knowledge must be generated by the users. Public administration is the ideal application area, due to the networked nature of its organizational models: formal knowledge must be stored where is generated, but exposed (published) for querying to any partners that might need it, in a fashion not unlike the way that web pages and forms are published nowadays.

According to Dawes (2008), in the early 1990s the "reinventing government" movement strove for changes that they hoped would eventually lead to a shift from bureaucratic government to an "entrepreneurial government that is enterprising, catalytic, mission and customer driven, and results oriented". The term "digital government" was adopted in 1999 by the National Science Foundation (U.S.); a few years later the Organization for Economic Co-operation and Development (OECD) emphasized the use of information and communication technologies (ICTs) and specifically the Internet in order to achieve better government; the World Bank definition stresses citizen empowerment as part of what had come to be referred to as e-government by that point (Dawes, 2008).

The expression "information society" was employed not as a metaphor, but as a statement of fact, capturing the essence of a culture: a shift from material products to informational products, "an explosion of information and information systems" (Anderson, 2008).

In the 1990s, the emergence of the phrase "knowledge society", which "primarily referred to economic systems where ideas and knowledge functioned as commodities" (Anderson, 2008), brought some confusion, as many people could not draw a line between the terms "information" and "knowledge". David and Foray (2003) provided a very useful basic distinction: while knowledge is "a matter of cognitive capability" in the sense that it offers the capacity for intellectual or physical action, information is usually "structured and formatted data that remain passive and inert until used by those with the knowledge needed to interpret and process them". The true distinction is actually made when discussing reproduction costs of knowledge and information, due to the need of cognitive capabilities (inference rules) in the first case.

Perhaps one of the most significant changes that make the shift from the information society paradigm to the knowledge society paradigm visible is the emphasis on interconnectivity and technologies that support the development of communities as vessels for knowledge production and sharing. If the advent of the information society hailed the database as the preeminent type of text that would replace the narrative in the postmodern age, the knowledge society, faced with information overflow, both structured and unstructured, stresses the importance of developing automated sense-making systems that use formal knowledge in order to support decision-making.

In terms of managing the flow of knowledge between public administration workers, citizens and businesses, very little is done. Even provided with Internet access and a plethora of information systems delivering services at all four levels (information, one-way interaction, two-way interaction and full electronic case handling) it is highly debatable up to what point citizens and businesses will engage with the public authorities in these information transactions as long as there is little or no knowledge sharing to support decision-making on both sides (government on the one side and citizens and businesses on the other side).

The shift towards the knowledge society paradigm also brings a shift in approaches to issues relating to e-government and e-democracy. We will stress the importance of knowledge sharing as a basis for decision-making and the need for e-government systems that also provide environments for knowledge acquisition, not only information transactions. In what follows, we also refer to the advent of a new concept in the scientific community: e-cognocracy.

E-cognocracy is a new framework for decision-making in e-democracies that exploits the power of Internet based public knowledge (Moreno-Jiménez and Polasek, 2003). This "cognitive democracy is oriented towards the extraction and democratization of the knowledge related with the scientific resolution of public decision making problems associated with the governance of society" (Moreno-Jimenez, Cardenosa and Gallardo, 2009). Papers outlining the concept and describing some possible applications that combine the representative and participative models by aggregating preferences, priorities and positions of political parties' representatives and citizens in conjoint discussions on forums or social network sites (Moreno-Jimenez, Cardenosa and Gallardo, 2009; Moreno-Jimenez *et al.*, 2008) describe some still unperfected methods of generating knowledge in the e-democracy framework. We must notice that tools such as social network sites, wikis (Decman, 2009), are more and more referred to as working models that can be used either as platforms or as knowledge sources in order to improve existing governance or e-governance models. We stress the importance of ICT in enabling the users' knowledge transfer processes rather than striving to supply enormous amounts of information through new channels and disregarding the general public's capacity to process it. A significant number of authors try to draw from established Web 2.0 models and propose solution for e-government (Decman, 2009; Decman, Stare and Klun, 2010; Neris *et al.*, 2009).

Our approach proposes a methodology for generating formalized knowledge from existing information storages, such as the databases that drive most of existing e-government implementations. It is based on the RDF knowledge formalization model (Resource Description Framework) defined as part of the recent Semantic Web paradigm (Berners-Lee, Hendler and Lassila, 2001) and the Formal Concept Analysis methodology for a taxonomy generation, employed as a segmentation technique. A demonstrative tool was designed in order to support a collaborative SWOT analysis with 28 participants, whose results are also described.

The literature shows quite a lot of efforts invested in Web knowledge representation. A couple of years ago, Wikipedia content has been translated to a formal knowledge repository known as DBPedia (DBpedia). The World Wide Web Consortium is making efforts to standardize a data-to-knowledge conversion methodology (The RDB Direct Mapping W3C Working Draft) and mapping language (The RDB Direct Mapping W3C Working Draft; The R2RML W3C Working Draft). Our approach emphasizes the need to break down data tables in taxonomy of entities that can be reused as formal knowledge, and connected with statements from other sources. The FCA methodology has been formalized both mathematically (Ganter, Stumme and Wille, 2005) and algorithmically (Ganter, 2010; Kuznetsov and Obiedkov, 2001). One of our previous papers (Buchmann, Meza and Pulcher, 2011) takes a more technical approach and applies the methodology in the development of enterprise resource planning systems.

The benefits can be mapped to the points 1, 2 and 5 from the requirements formulated by Netchaeva (2002) for e-government:

1. setting up different governmental websites with respect to both centralized and decentralized state institutions offering basic information to their users;
2. allowing for interaction with governmental authorities;
3. providing the users access to communication tools such as forums and opinion polls and taking their feedback into account;
4. offering online services such as payment of fines, taxes, renewal of licenses etc.; and
5. bundling all these in a single, complex e-government portal.

3. Formal knowledge on the Web

Currently, e-government applications are built around data repositories following the relational model. This model is the technical foundation of data-driven systems. When a user fills a form or requests a report, data is generally stored in tables and managed through a “closed world approach”.

Table 1: Example of a tabular data storage

ID	Name	BirthDate	Email	FatherOf
XB077789	John Smith	12.07.60	js@email.com	Mary Smith
KX076669	Mary Smith	13.05.80	ms@email.com	

Every time a user performs a search, he or she browses a list of results or filters them based on some key values, under the wraps the Web site translates his actions to formal queries for selecting data sets based on specified row/column filters.

While Web 2.0 still employed the same technology, shifting the emphasis on the user as content generator, Web 3.0 requires a migration towards the networked knowledge model called RDF (Resource Description Framework) that is more fit to the networked nature of the Web. This is the structural foundation of the so-called knowledge society, knowledge-driven systems and the knowledge-based governance.

On the Web, formal knowledge comes in two shapes – statements and rules – which are managed through an “open world approach”. The statements follow the basic linguistic structure with three sentence parts – a subject, a predicate and an object[1]:

:JohnSmith :isA :Citizen.

:JohnSmith :isFatherOf :MarySmith.

:MarySmith :hasMaritalStatus :Divorced.

When a part occurs in multiple statements, they can be connected in a graph, with subjects and objects as nodes.

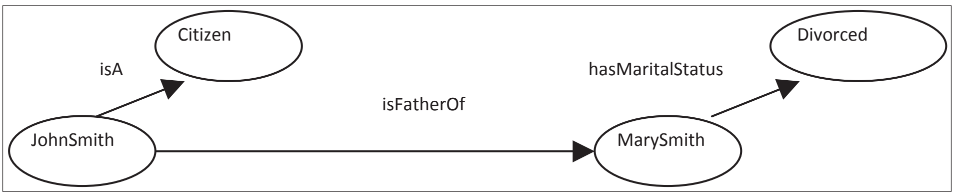

Figure 1: Interconnected formal statements

Just as with the table, any parts of the graph can be extracted with formal queries (or a public Web form programmed to translate user selections to such queries). One of the most important advantages that this model brings is the fact that different parts of the graph may belong to different organizations. The aggregated graph forms a cloud of knowledge that can be exposed to all partner organizations and target users. Browsers can navigate it in a fashion similar to navigating web pages. In businesses, certain organizations might withhold certain information from exposing it in the knowledge cloud, but public administration provides ideal scenarios for this, due to its networked organizational model – territorial agencies, various branches of the public administration would benefit greatly from sharing among them knowledge that was acquired locally. In order to specify which agency “made” each statement, the terms must be qualified with the URL address of their “owner” organization. If :JohnSmith and :MarySmith are terms defined by different agencies, instead of …

:JohnSmith :isFatherOf :MarySmith.

…the knowledge cloud would store the statement as:

<http://population.gov/JohnSmith> <http://population.gov/isFatherOf> <http://otheragency.gov/MarySmith>.

By defining such identifiers for all things (people, relationships, assets etc.) we also provide semantic links between different organizations: besides the explicit semantics

1 Several shared vocabularies (RDFS, OWL, FOAF etc.) provide standard terms that can be used in such statements. Discussing such vocabularies is out of our scope; therefore the examples will not fully reflect the syntactical formalisms, for better expressivity.

of this example, it also indicates that further statements about MarySmith may be acquired by following the link http://otheragency.gov/MarySmith. The link, instead of providing a traditional Web page, would provide all the statements that otheragency.gov has about MarySmith, and so on. Declarations of equivalence are also possible, when different organizations use different identifiers (terms) for the same things.

Besides statements, a knowledge base may also store rules. These are responsible for generating new statements through automatic reasoning patterns. Rules consist of a premise and a conclusion – every time the premise is found in the knowledge base, the conclusion is automatically generated. The concluded statement can have various uses for public administration scenarios:

- It can be just another statement added to the knowledge graph; one of the most relevant scenarios for public administration would be to detect when the same entity has different identifiers in different contexts. For example, if two different persons have the same mobile number, it's plausible that they are actually the same person:

If X :hasNumber Z and Y :hasNumber Z Then X :isSameAs Y

- It can be a statement signaling the need for action items:

If X :hasFilled :FormZ Then X :mustReceive :R

- It can be a contradiction, signaling that different agencies hold contradictory information about the same person/asset:

If X :isFatherOf Y and Y :isFatherOf X Then X :InContradictionWith Y

The main advantages that the graph model brings to form-based applications are reflected by the differences between the closed world and open world approaches:

- The database structure (the table headers) is defined early and stays rigid (costly to redesign). It defines a restrictive model by imposing what kind of data can be stored, what value ranges are allowed and so on. In the open world approach, knowledge models are flexible. Since there are no predefined fields or cells to be filled, new statements can be added at any time without costly structural changes and without wasted storage space (such as the bottom-right cell from Table 1, due to the lack of a certain piece of data).
- The database is "introverted": its main purpose is to serve the owner organization, while exposure to external usage is strictly controlled and costly to implement. The knowledge graph is extroverted: its purpose is to interconnect statements from various organizations and merge them implicitly. Explicit efforts must be made to limit the knowledge exposure.
- Databases create a rigid reflection of reality: anything that does not fit the model is rejected and anything that is missing does not exist; knowledge bases create a more realistic image: anything that is missing is possible to emerge at any moment, rules can be applied to generate new knowledge or to check coherency against knowledge from other sources.

In the following example, we consider several public administration agencies: lifevents.ro tracks life events and personal data registration for citizens; financeagency.ro tracks their fiscal activities; studentregister.ro tracks their educational activities.

Each agency assigns identifiers using its address and an internal identifier in order to form statements. All the statements about the same subject are available at an address similar to that subject identifier.

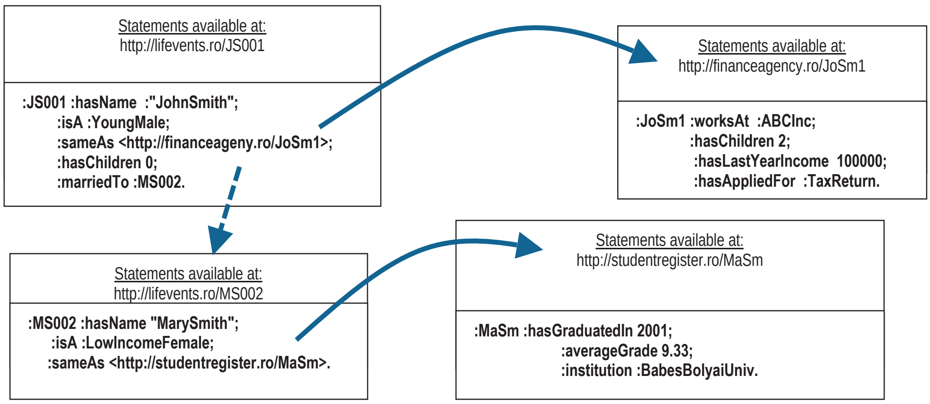

Figure 2: Linked Data – an example of interconnected formal knowledge sets

Links between agencies are defined by those statements that use subjects and objects from different sources: the "sameAs" associations declare the equivalence between two entities identified differently by two organizations (for example MS002 from lifevents.ro and MaSm recorded at studentregister.ro). Based on such links, a common web interface could retrieve and merge information about the same entities, coming from different sources. This model is also called Linked Data, seen as a stepping stone from data to knowledge.

Having the statements merged in a cross-organizational graph, a rule could detect that John Smith declared 0 children to one source, and 2 children to the financial agency (contradiction detection). Another rule could trigger action items on behalf of the financial agency, based on the premise statement that he applied for tax return. Another rule could state that the marriage is a reciprocal association, thus MarySmith is also married to John Smith (which could further trigger automatic transfer of some statements from husband to wife).

Labels such as :LowIncomeFemale, :YoungMale are actually concepts that can also act as subjects, and be further described by specific attributes (combinations of age range, gender, income range etc.). These can be defined by an automatically generated taxonomy obtained by clustering techniques such as FCA, as will be shown in the next section.

4. The proposed methodology

Even if the foundation of Semantic Web has been formalized and implemented during the last decade, the application of the paradigm is still in its inception. The first reason is that it does not provide an obvious advantage, easily pinpointed during the usage experience. Its benefits need to be validated on a larger scale, in the quality of search results and the cost of sharing within collaborative environments. The second problem is that it takes quite an effort to translate existing information into knowledge bases. While people and organizations were easily convinced to write Web pages in order to have a Web presence, it is much more difficult to convince them to provide formal statements and rules. This would be an activity just as important as social networking is nowadays – the organizations that would not expose formal knowledge will not be "knowledgeably connected". The exposure does not have to be 100% public, but needs to make sure that information will be reached effortlessly by everyone who needs it (and is entitled to it).

In this respect, our paper proposes an automated methodology of converting existing data tables to knowledge graphs. This is acquired in two steps:

- automated conversion of tabular data to formal statements; and
- automated segmentation of table records in classes, in order to generate entity taxonomies.

The result of the methodology can become the starting point of a knowledge base. The first step looks for natural associations occurring in any table and translates them to statements:

1. Every table has a record identifier column (primary key). The knowledge base also needs unique identifiers for things, but they must contain the URL address of the owner organization. The obvious solution is to juxtapose (concatenate) the domain address owned by the governmental agency with the table name and the key from the table (the social security number). This combination will generate unique values for all the "things" in any table. For Table 1, John Smith's identifier would be

http://www.govagency.com/citizens#XB077789

2. Every table is a set of its records. If each record represents an entity (a thing, a person), a table can be considered a class of objects/entities. Each of them belongs to the table, thus there is an IsA association between them:

:XB077789 :isA :Citizen.

3. A column name is an association between each entity and the corresponding value from that column (name is also a property, rather than an identifier, since its uniqueness is not guaranteed!):

:XB077789 :hasEmail "js@email.com".

:XB077789 :hasName "John Smith".

4. The columns with unique values would serve to provide rules for detecting the equivalence of things. For example, it is generally considered that e-mail addresses are not shared, and two persons with the same e-mail address are actually the same person. For this, a rule can be imposed to trigger this signal:

IF X :hasEmail Y and Z :hasEmail Y Then X :sameAs Y.

5. Another rule can be generated for every column label to designate classes for subjects and objects of statements:

If X :hasEmail Y Then X :isA Citizen and Y :isA EmailAccount.

This might look obvious, but in automated systems it is important if the system is able, via inference, to label something based on what statements it is involved in. By applying such rules, new information can be generated automatically, without explicit input from the user.

The second step of the proposed knowledge generation process is to apply Formal Concept Analysis on the content of each table. FCA can be seen as a hierarchical clustering technique, for classifying a population based on the attributes of individuals. Thus, its purpose is to further drill down a table and discover new statements expressing such classifications.

Formal Concept Analysis is computed on a formal context, a cross table indicating what attributes each entity has. The entities correspond to the original table rows, while the attributes can be chosen from table columns, by looking for relevant values (gender values) or value ranges (age ranges, birthdate ranges etc.).

Table 2: A formal context example

	A1	A2	A3	A4
O1	X	X	X	
O2		X		X
O3	X		X	X
O4		X	X	X

FCA algorithms are used to organize the entities in classes based on the attributes that they share. Such classes can be computed for all the attribute combinations.

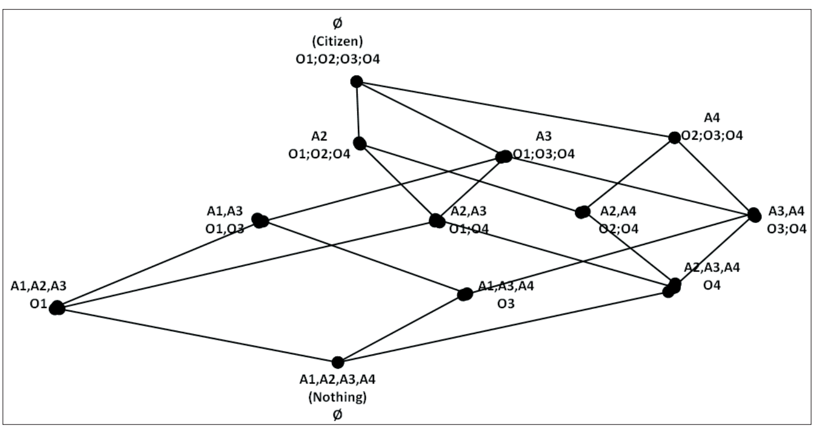

Figure 4: A Hasse diagram of formal concepts derived from Table 2

In the Hasse diagram, each node represents a class of entities sharing the attributes indicated. The highest node is the group of ALL entities in the table. The lowest node is the void set, which is obtained if at least two attributes are mutually exclusive (e.g.

Male and Female, disjoint age ranges). As the nodes are lower, they have more specific descriptions (more attributes, thus more statements about them). As the nodes are higher, they contain objects with less specific descriptions (less facts). Every arc of the diagram represents a subclass inclusion. Also, between any object and groups that it belongs to, an IsA association can be generated (the relationship of belonging to a class).

The main distinction between FCA classification and traditional segmentation applied in social sciences or marketing is that the latter's purpose is to detect relatively few groups which are stable when new data is added to the storage. Stable clusters mean that a reliable pattern has been detected in the population. In knowledge systems, new statements will allow to draw new conclusions, depending on where they fit in the class hierarchy. The conclusions that can be drawn from a taxonomy are usually new labels that can be applied to things.

5. Discussion and benefits for e-government

For demonstration purposes, we built an interface for formal knowledge browsing and filtering using SIMILE tools (The SIMILE Project).

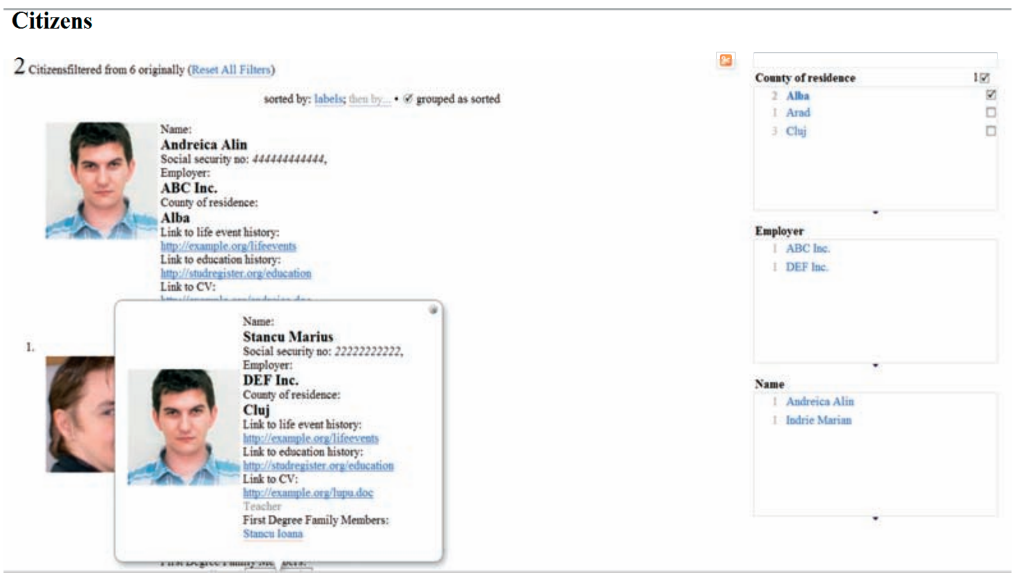

Figure 5: Interface of the demo tool

The tool aggregates statements from multiple local servers and turns their predicates into properties that can be displayed about each entity. Some predicates can be turned into filters, visible on the right side of Figure 5, or sorting criteria. For each relevant type of records (life event, employer etc.) similar pages can be built and linked to each entity.

Based on this tool, we developed a collaborative SWOT analysis via a survey over 28 subjects who were asked to suggest items for each SWOT chapter. The items were synthesized in a common list, then each subject was asked to grade each of them according to how convincing was the tool in reflecting it: the maximum (5) means that the subject saw higher benefits (for strengths/opportunities) or dangers (for weaknesses/threats) of the tool over traditional data-driven systems, the minimum (1) means that the subject finds the aspect less relevant in the proposed tool, the average (3) means that the subject does not perceive relevant differences from traditional systems. Table 3 shows the average grades for each SWOT aspect.

Table 3: Strengths/Weakness/Opportunities/Threats as identified by test users

	Aspects identified	Avg	Discussion
S	1. A common interface to multiple information sources 2. Contradiction detection rules 3. Low cost of interoperability 4. Action triggering/decision support rules 5. Fact discovery rules	4.07 4 1.61 3.68 2.43	The main perceived benefits are the common gateway, the ability of detecting contradiction in the knowledge graph and the ability to trigger action signals when certain information appears from other sources. Cost advantages have not been perceived since most subjects did not have the technical understanding of the implementation requirements or were able to identify low-cost technology for traditional system implementations as well.
W	1. Usability and responsiveness 2. Performance 3. Privacy and security (seen as internal feature) 4. Nonsynchronicity	3 3.61 3.5 2.07	Querying formal knowledge has 10-30 times lower performance than extracting data from tables. This is perceived by users as latency in interface responsiveness and general system performance. The demonstration implementation did not have any privacy and security features, users were worried regarding the full exposure of interconnected statements. The proposed methodology does not provide a synchronous image of the table data, but a redundant copy that must be regenerated periodically. Again, only subjects with technical knowledge could identify this weakness. Solutions such as (the D2R server) address this issue.
O	1. Collaboration and data integration opportunities 2. Information tracking for the citizen and public administration 3. Knowledge discovery	3.93 4.06 2.57	The population tends to get involved in social platforms and individuals make data about themselves available to the public, in manners rather unrestricted. Authorities can fathom such a wealth of information to gain insight regarding social trends and to classify people based on their self-descriptions. The Semantic Web promises interoperability without precedence, over different platforms and organizations. In theory, government agencies could get, through a simple search, aggregate results of public records, shopping history, preferences and so on, without having citizens fill any explicit declarations. Most e-activities rely on the basic use cases: searching, browsing and filtering and the results of these are significantly affected by the new model – the information is detected based on interorganizational associations over the entire "Future Internet".
T	1. Total governmental control and abuse 2. Privacy and security (seen as external threat potential) 3. Refusal of institutions to adopt experimental models	3.89 3.29 1.75	Users are worried about how public administration or external intervention could abuse such interoperability. Currently, the Internet supports anonymity and lack of responsibility for on-line statements and actions. Formal knowledge would allow connecting these to their originator, through simple searches (rather than costly traffic monitoring). Most users didn't perceive the refusal of adopting experimental models as a significant future deterrent.

For each chapter, hierarchical clustering was applied, with the following results.

Table 4: Clustered attitudes for the SWOT analysis

	Aspects identified	C1	C2	C3	Cluster description and count (excluding outliers)
S	1. A common interface to multiple information sources 2. Contradiction detection rules 3. Low cost of interoperability 4. Action triggering/decision support rules 5. Fact discovery rules	4.9 4.6 1.4 5.0 3.9	3.9 3.9 1.5 3.2 2.0		C1: 7 subjects – the cluster of enthusiasts, 3 strengths graded near maximum C2: 19 subjects – more moderate grades, still significantly over the average for some strengths Both clusters were skeptical about the cost benefits, thus cost metrics and evaluation are required for future development of the study.
W	1. Usability and responsiveness 2. Performance 3. Privacy and security (seen as internal feature) 4. Nonsynchronicity	2.6 4.1 4.4 2.1	3.8 3.1 1.7 1.4		C1: 17 subjects – encountered performance problems and expressed worries regarding privacy and security features C2: 9 subjects – did not question privacy and security, emphasized bad usability Few subjects were interested in how the relationship between the original database and the knowledge graph is maintained on long term or did not see it as a specific weakness.
O	1. Collaboration and data integration opportunities 2. Information tracking for the citizen and public administration 3. Knowledge discovery	3.9 4.4 1.7	4.2 4.5 3.4	3.5 2.7 2.7	C1: 11 subjects – moderate enthusiasts, identifying collaboration/tracking scenarios for public administration but considered that knowledge discovery is already handled well by traditional systems C2: 11 subjects – the enthusiasts, identifying relevant opportunities for knowledge management C3: 6 subjects – did not see anything special compared to existing systems
T	1. Total governmental control and abuse 2. Privacy and security (seen as external threat potential) 3. Refusal of institutions to adopt experimental models	4.6 3.2 1.4	3.5 4.5 3.3	3.0 2.5 1.1	C1: 13 subjects – strong worry regarding authority abuse and extreme control C2: 6 subjects – strong worry regarding unauthorized "knowledge attacks" C3: 8 subjects – did not see anything significantly different from the current state Most subjects did not perceive the refusal of adoption of such technologies as a specific threat.

6. Conclusions

The proposed solution adds value to legacy systems and prepares the public administration infrastructure for the emergence of decision-making based on automated processing of formal semantics. The benefits of exposing existing data tables as formal statements emerge from the ability to connect statements from different sources (financial records, police records, workplace records etc.) and have computers reporting conclusions automatically, based on formal rules reasoning over the merged statements. Rules can detect contradictory information which can be interpreted as falsehood in declarations: people working/being in different places at the same time, different persons having the same ID card/phone number/email address (or any other personal attribute), the same person having different sex, residence (or other attributes that should be unique at any given time). Bureaucratic action can be generated automatically when certain statements appear in the system.

Not everything should be exposed as formal knowledge and access to exposed knowledge should be properly channeled because, as previously mentioned, Web 3.0 is implicitly extroverted. On the other hand, especially in Romania there are significant problems in assuring the interoperability of data that is already available and bringing it together in an efficient manner, with automated consistency checking. The semantic e-government would provide the basis for much more effective authority, with information of higher opportunity and relevance available to queries and reporting engines. The same effectiveness would also serve the citizen, who will be able to access his records on-line from a single interface, independent of the data owner, to detect contradictions triggered by bad governmental management or to easily detect associations between his public records and laws, news or knowledge of public interest.

A formal approach to the foundations of the knowledge society is of essence now, when computers are able to handle knowledge representations that are closer than ever to natural language, with the ability to do rule-based reasoning over facts emerging from various sources.

European projects are already investing efforts in rendering Internet content as formal knowledge, but the sheer quantity of content reclaims automated ways of converting the existing data as well.

References:

1. Anderson, R.E., 'Implications of the Information and Knowledge Society for Education', 2008, in Voogt, J. and Knezek, G. (eds.), *International Handbook of Information Technology in Primary and Secondary Education*, New York: Springer Science+Business Media, LLC, 2008, pp. 5-22.
2. Berners-Lee, T., Hendler, J. and Lassila, O., 'The Semantic Web', 2001, *Scientific American*, vol. 284, no. 5, pp. 34-43.
3. Buchmann, R.A., Meza, R. and Pulcher, D., 2011. 'An Approach to the Semantization of ERP Systems', 2011, *Business Information Systems, Lecture Notes in Business Information Processing*, vol. 87, part 7, pp. 218-229.
4. David, P.A. and Foray, D., 'Economic Fundamentals of the Knowledge Society', 2003, *Policy Futures in Education*, vol. 1, no. 1, pp. 20-49.
5. Dawes, S.S., 'The Evolution and Continuing Challenges of E Governance', 2008, *Public Administration Review*, vol. 68, Issue Supplement, pp. S86-S102.
6. *DBpedia*, [Online] available at http://wiki.dbpedia.org/, accessed on October 18, 2011.
7. Decman, M., Stare, J. and Klun, M., 'E-Government and Cost-Effectiveness E-taxation in Slovenia', 2010, *Transylvanian Review of Administrative Sciences*, no. 31E, pp. 48-57.
8. Decman, M., *Web 2.0 in e-Government: The Challenges and Opportunities of Wiki in Legal Matters*, in Hahamis, P. and Remenyi, D., *9th European Conference on E-Government, Proceedings*, Academic Conferences Ltd, 2009, pp. 501-509.
9. *FInES Cluster*, [Online] available at http://www.fines-cluster.eu/fines/jm/, accessed on October 14, 2011.
10. Ganter, B., 'Two Basic Algorithms in Concept Analysis', 2010, *Formal Concept Analysis*, vol. 5987, pp. 312-340.

11. Ganter, B., Stumme, G. and Wille, R., *Formal Concept Analysis. Foundations and Applications*, Berlin: Springer-Verlag, 2005.
12. Ikujiro, N., 'The Knowledge-Creating Company', 1991, *Harvard Business Review*, November, pp. 96-104.
13. Kuznetsov, S. and Obiedkov, S., 'Algorithms for the Construction of Concept Lattices and Their Diagram Graphs', 2001, *Principles of Data Mining and Knowledge Discovery*, vol. 2168, pp. 289-300.
14. *Linked Data Design Issues*, [Online] available at http://www.w3.org/DesignIssues/LinkedData.html, accessed on September 25, 2011.
15. *Linked Open Data 2*, [Online] available at http://cordis.europa.eu/fetch?CALLER=PROJ_ICT&ACTION=D&CAT=PROJ&RCN=95562, accessed on September 25, 2011.
16. Moreno-Jiménez, J.M. and Polasek, W., 'E-democracy and Knowledge. A Multicriteria Framework for the New Democratic Era', 2003, *Journal of Multi Criteria Decision Analysis*, vol. 12, no. 2-3, pp. 163-176.
17. Moreno-Jimenez, J.M., Cardenosa, J. and Gallardo, C., 'Arguments That Support Decisions in e-Cognocracy: A Qualitative Approach Based on Text Mining Techniques', in Lytras, M.D., Carroll, J.M., Avison, D., Dale, A., Sipior, J., Damiani, E., Tennyson, R.D., Naeve, A., Lefrere, P., Tan, F. and Vossen, G. (eds.), *Visioning and Engineering the Knowledge Society: A Web Science Perspective, Proceedings*, Berlin: Springer-Verlag, 2009, pp. 427-436.
18. Moreno-Jimenez, J.M., Escobar, M.T., Toncovich, A. and Turon, A., 'Arguments That Support Decisions in E-Cognocracy: A Quantitative Approach Based on Priorities Intensity', in Lytras, M.D., Carroll, J.M., Damiani, E., Tennyson, R.D., Avison, D., Vossen, G. and Ordonez De Pablos, P. (eds.), *Open Knowledge Society: A Computer Science and Information Systems Manifesto*, Berlin: Springer-Verlag, 2008, pp. 649-658.
19. *National Student Enrollment Registry*, [Online] available at http://www.rmu.ro/home.aspx, accessed on October 18, 2011.
20. Neris, V.P.D., Almeida, L.D.A., de Miranda, L.C., Hayashi, E.C.S. and Baranauskas, M.C.C., *Towards a Socially-constructed Meaning for Inclusive Social Network Systems*, in Zang, H., Feng, X., Liu, K. and Jiang, G. (eds.), *Information Systems in the Changing Era: Theory and Practice*, Marrickville: Aussino Academic Publishing House, 2009.
21. Netchaeva, I., 'E-government and E-democracy', 2002, *International Communication Gazette*, vol. 64, no. 5, pp. 467-477.
22. *The D2R Server Official Site*, [Online] available at http://www4.wiwiss.fu-berlin.de/bizer/d2r-server/, accessed on October 18, 2011.
23. *The R2RML W3C Working Draft*, [Online] available at http://www.w3.org/TR/r2rml/, accessed on October 18, 2011.
24. *The RDB Direct Mapping W3C Working Draft*, [Online] available at http://www.w3.org/TR/rdb-direct-mapping/, accessed on October 18, 2011.
25. *The Resource Description Framework*, [Online] at http://www.w3.org/2001/sw/wiki/RDF, accessed on October 18, 2011.
26. *The SIMILE Project*, [Online] available at http://simile.mit.edu/, accessed on October 18, 2011.