



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Recovering the Past. Eastern European Web Mining Platforms for Reconstructing Political Attitudes

Camelia Florela Voinea

Department of Political Science, International
Relations and Security Studies
Faculty of Political Science
University of Bucharest, Romania
camelia.voinea@fspub.unibuc.ro

Markus Schatten

Artificial Intelligence Laboratory
Faculty of Organizational Informatics
University of Zagreb
Croatia
markus.schatten@foi.hr

Date of submission: November 25th, 2014

Date of acceptance: January 23rd, 2015

Abstract

During the past half century, the political attitude of the Eastern European people toward the state, government and society changed dramatically. So did their value systems. Inglehart's materialist vs. post-materialist comparative analysis gives a measure of this value change, but not enough as to fully characterize the phenomena underlining the differences in political culture before and after the Fall of Berlin Wall. Little has left from the communist regimes to prove how this change actually occurred and where we are as compared to the stable democratic regimes. With rare exceptions, no public survey has been developed in the Eastern European countries between 1950-1990 able to mirror people's true beliefs and values. In order to understand the current value systems and political attitudes of the people in the Eastern Europe, we have to recover the past. One way to do that is to identify key concepts in the texts, discourses, audio and video recordings of the past times. The present paper provides the rationale of this approach and describes a system which works on dynamically collecting content-based items from library and web references and resources. The system currently works on concepts described by single words or compound

Corresponding Author: Dr. Camelia Florela Voinea, Associate Professor

Affiliation: Department of Political Science, International Relations and Security Studies, Faculty of Political Science, University of Bucharest

Address: #8, Spiru Haret Street, 010175 Bucharest, Romania

e-mail: camelia.voinea@fspub.unibuc.ro

Copyright © 2015, Camelia Florela Voinea and Markus Schatten

European Quarterly of Political Attitudes and Mentalities - EQPAM, Volume 4, No.1, January 2015, pp. 22-39.

ISSN 2285 – 4916

ISSN-L 2285 – 4916

expressions, and could be extended so as to work on multimedia items, like words, images, and sounds (voices, music, audio signals, etc.). Our approach aims at constructing a dynamic system and an open access repository of content-based collections of the past and offers a research instrument to the students of political attitudes toward democracy and freedom of the people in Eastern Europe. We approach the problem of recovering the historical process of political change in the Eastern European societies known as the Fall of Berlin Wall in terms of political attitude change modeling and simulation. Modeling makes intensive use of web and data mining technologies for identifying political attitude structural configurations in patterns of value and belief change. Based on web-extracted political attitude configurations, simulation provides a clue on how political attitude structure looks like, and how political attitude change emerges in macro level political change phenomena.

Keywords: political attitudes, shadow configuration, culture modeling, web mining

1. Introduction

There are few computational and simulation modeling approaches of the complex political change phenomena known as the Fall of Berlin Wall. Though intensively investigated since the 1990s, this research has uncovered a complex milieu of phenomena. These phenomena have been studied rather separately so that, as the time passes, the overall picture becomes more and more complicated.

In order to understand the traces left in the political cultures of the Eastern European societies by more than half century of authoritarian communist regimes, political culture research has been massively oriented toward surveys on issues of major interest, like the political attitudes toward democracy and governance, EU membership and euroscepticism, voting behavior, public opinion dynamics, or political socialization.

The issue of political change has been approached from various perspectives (Russell and Welzel, 2014). Some authors explain the current political attitudes toward democracy by the political culture inherited from the communist regimes (Kubik, 2012; Stamenova, 2000a,b, 2002, 2003; Karklings, 2005). Others look for explanation in the economic situations of the Eastern European countries, many of these countries currently undergoing difficult processes of transition to democracy and to market economy (Rohrschneider, 1999; Rohrschneider and Schmitt-Beck, 2003). Ronald Inglehart (1997; 1977; 1971) and Christian Welzel (2005) explain the value change as a natural consequence of the changes in the economic conditions and in the institutional structure, which have been subjected to deep transformations.

Political culture theory (Almond and Verba, 1963; Verba and Pye, 1965; Converse, 1964) suggests that political change phenomena could be explained in terms of changes in values, beliefs, and attitudes at the individual level. Other authors suggest that revolutions or political change from authoritarian regimes to democratic ones could be explained in terms of recurrent (macro-level) phenomena based on path-dependent processes of change (Laitin, 1995; Tilly, 2001). William Mishler and Detlef Pollack (2003) describe the political culture with two concepts: "thick", for culture aspects which are resistant to change, and "thin" for culture aspects with a high degree of volatility, instability or superficiality. Taken together, *thick & thin* culture aspects should describe a culture continuum which would help identify the role played by each aspect in inducing a political change. Mishler and Pollack (2003) describe this culture continuum as a one-dimensional structure, in which thick aspects are at one extremity of the axis, acculturation at the other extremity, and the thin aspects – in between. The main advantage of modeling the major political change phenomena in the political culture continuum paradigm is that the political change itself is operationalized in terms of change in value, beliefs, attitudes and norms. The difficulty comes from

operationalizing the culture continuum concept in the case when empirical data is missing, which is the case of Eastern European communist regimes before the Fall of Berlin Wall in 1989.

In this paper we approach the problem of modeling and simulation of the historical process of political change in the Eastern European societies known as the Fall of Berlin Wall in terms of political attitude reconstruction. Modeling makes intensive use of web and data mining technologies for identifying political attitude structural configurations in patterns of value and belief change. Based on web-extracted political attitude configurations, simulation provides a clue on how political attitude structure looks like, and how political attitude change emerges in macro level political change phenomena.

2. Eastern Europe Communist Era: How to Understand Value Change without Empirical Survey Data?

Modeling and simulation approaches on 1989-political change in Eastern Europe have however to meet a major challenge: modeling needs to combine both the economic and culture mechanisms whose workings are supposed to converge to the dynamics of political change phenomena. The challenge concerns the means to obtain the necessary data to run and validate such a model. Empirical data do not provide sufficient support for replicating the dynamics of the real system by the model system. Game-theoretical approaches are often too much restrictive in achieving a rational output so that the dynamics of the model system might be sometimes at odds with the real system. Bottom-up approaches use generative data mechanisms in order to evaluate the explanative power of the model system (Voinea, 2012). A bottom-up approach cannot however avoid comparing the model system's output against real data and evaluating how believable are its assumptions and results.

One solution might be a different way of obtaining the data: if neither empirical observation, nor the data generative engines can provide it, then dynamic linking of web items might be a solution to obtaining the necessary data, and also to identifying associative patterns within the data which might convey more meaning in huge web-linking generated data collections. This method is not new: it is intensively used in content analysis or in sentiment analysis. What is really new in using it concerns its intrinsic associative and relational meaning-construction capability. On this idea we base the present research paper.

The paper approaches the issue of achieving a model of culture continuum which can explain the political regime change and, in general, polity change phenomena. It introduces a perspective over the culture continuum modeling which employs web and data mining technologies in harvesting the necessary data.

Harvesting the data on the 1989-Eastern European political change from the web resources might seem a trivial task. However, it is not. During the communist regimes, the individuals in the Eastern European societies have not been surveyed on their true beliefs, values and attitudes toward the political authority, regime, governance, etc. Public opinion was an area of massive ideological propaganda and manipulation. After a period of time of almost fifty years, during which the intellectual elites were practically eliminated from these societies, the value systems have considerably changed, but no one has actually measured this change until it practically "exploded" as a major political regime change. Moreover, in the absence of survey data, knowledge about value and political attitude change has remained ever since mostly qualitative or simply speculative, which makes the modeling quite difficult.

In order to understand the current trends in the Eastern European political attitudes toward democracy and governance, their past expressions and evolution should be recovered and studied. There

are two major problems concerning the recovering of past political attitudinal expressions: both of them concern the constructivist idea of building up simulation (virtual) models.

The first one will be called in this paper the "no initial reference" problem, and regards the lack of survey data on the individual values, beliefs and attitudes from the period before 1989 in the Eastern European communist countries. As empirical data is missing, there is no reference for identifying the items pertaining to the recovering and/or reconstruction of past attitudinal expressions and objects. Hence, the "no initial reference" problem appears as a problem of identifying an "issue" without knowledge about how this "issue" actually looks like. There is a large consensus on considering that, during the communist times in the Eastern European regimes, political attitudes are characterized by duality of social expression. The duality of social expression is, on the one hand, induced by propaganda and maintained through the coercive (authoritarian) political regime. On the other hand, duality of social expression is also the expression of deep social cleavages produced by the elimination of traditional elites, and the introduction of new elites usually lacking the argument of competence or political legitimacy (Precupețu, 2008).

The second one will be called the "no end reference" problem, and regards the lack of control over the process of data harvesting, where by "control" we denote the knowledge and/or ability to influence the items identification process so as a constructivist solution can be validated.

The solutions to the two problems define the goal of our approach: we aim at developing a theoretical approach on political attitude reconstruction in the "no initial reference", and "no end reference" scenario. The complexity of this approach resides in generating versions of political attitudes which are consistent with the final outcome of a change process: the major political regimes change in 1989 in all Eastern European countries.

In this paper we approach the problem of recovering the historical process of political change in the Eastern European societies known as the Fall of Berlin Wall in terms of political attitude change modeling and simulation. There are two ways of recovering the past: dynamically content-based (re-)collection, and generative linking. The former is based on data mining, web mining and content-analysis techniques for searching the multimedia library sources, like document texts, video and audio recordings, pictures and symbols. The latter is based on generative techniques of linking instant data from individual sources, like interviews, commentaries, personal stories and experiences, values and beliefs.

3. Repository of Eastern European Political Culture Items

There is an increased interest in the public opinion repositories which harvest data from online social networking platforms. Such repositories include forms of spontaneous expressions of human values, beliefs, sentiments which allow for the identification of dynamic patterns of preferences (Schatten, 2012). Some advanced approaches to the modeling of democracy, participation and emergence of social organization which intensively use online data harvesting from social networking platforms, like for example the ChartLeaders Media¹, employ theories of complexity, societal emergence and self-organization of the items pertaining to the configurations of macro patterns of organization and leadership in companies and nations. When concerning political issues like participation in democratic societies, such macro patterns could be used in modeling the voting behavior and predicting the electoral campaigns' outcomes:

"... rather than waiting for electoral campaign organizations to interpret the state of the nation on behalf of the voters, the electorate should have its own means of identifying and articulating the evolving priorities of the state and of the

¹ Platform online at: www.chartleaders.com

public and private institutions of which the state is made up. In other words, for voters to make the right choices at election times, they should ideally have a broad model that helps them consistently interpret, monitor, and share their needs from one era to the next.”

(Marc van der Erve and Erik Klein Nagelvoort)²

Our interest in building up and employing such repositories in the reconstruction of past political attitudes has been stimulated by the idea that these repositories can provide clues on how attitudinal expressions “rise” from the dynamically evolving relationships between values, beliefs and feelings, behavior, norms, and knowledge.

Our approach on modeling political attitudes formation and change is based on the generative paradigm of methodological individualism with some relevant modifications which regard the introduction of a relational extension into the classic paradigm. Such modifications do not represent a new issue: they have been theoretically defined in Social Simulation theory by Peter Hedström (2005) and in Political Methodology by Charles Tilly and his collaborators (McAdam, Tarrow and Tilly, 2001).

In our modeling approach, we develop and introduce modifications concerning the issue of generating relationships between items which are structural components of political attitudes, like values, beliefs, sentiments, norms, behavior, and knowledge. The generative paradigm allows for the simulation of individual interactions under political influence, as it is viewed as ubiquitous at the micro level (amongst the interacting individuals). Moreover, the generative paradigm allows for the simulation of the interaction between macro level agents (institutions, groups, society at large) and micro level (individual) agents. While the former type of simulation supports the study of emergence of political attitudes (upward causation), the latter type supports the study of the complex effects generated by macro level phenomena and/or constructs onto the micro level agents and behaviors (downward causation). In both cases, the relations between the political attitudes and their (potential) structural components are essential for the modeling of the dynamics of attitudes emergence and change. For studying this dynamics, two alternatives can be tested: one concerns the use of empirical survey data, and the other one concerns the use of generated data. The former is the classic empirical paradigm of modeling. The latter is a generative modeling paradigm employing web and data mining technologies. It has two alternatives which are worth investigating: generating data from (a) online (“live”) interviews or other methods of opinion harvesting from social networks, blogs, forum platforms, etc. (Marc van der Erve, ChartLeaders platform, 2014) or (b) from web and data mining systems (Schatten, 2013; 2012; 2011; Schatten, Rasonja, Halusek and Jakelić 2011).

4. Reconstructing the Past with Web Mining Techniques

One possibility to recover the past is to collect the data dynamically from multimedia sources: texts, videos, audio recordings and pictures (see **Figure 1**).

² Marc van der Erve and Erik Klein Nagelvoort, “*Democracy 2.0 – Time to look into a clean mirror*” (Submitted for publication), available online on the [Marc van der Erve's LinkedIn Profile](https://www.linkedin.com/in/marcvandererve/), accessed on November 2nd, 2014.

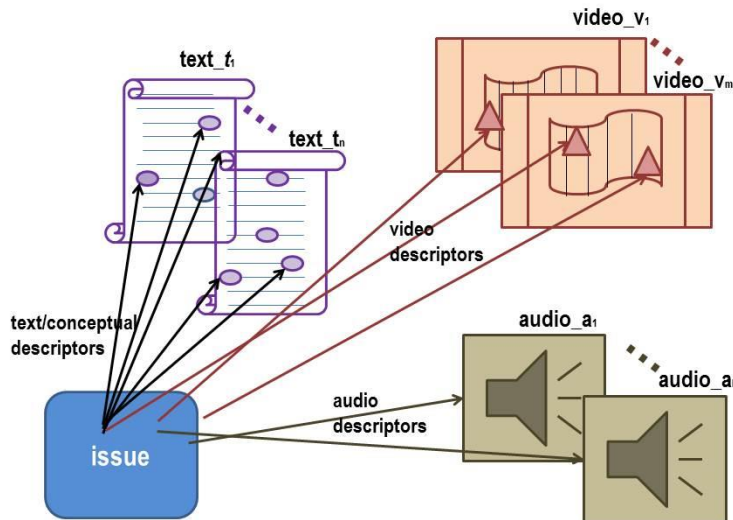


Figure 1.

Collections of multimedia descriptors are used to (re-)construct a historical context and (class of) phenomena.

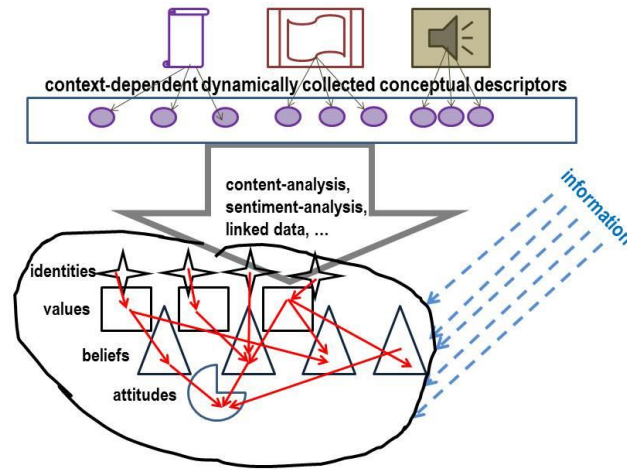
While their interpretation might appear as a complex task of combining linguistics, semiotics, and conceptual modeling, this task might be supported and, moreover, simulated by using a web mining system.

Reconstructing the past expression of attitudes is a task of re-constructing the relational complexity of a "milieu" of values, beliefs, normative constraints and contextual knowledge. Without proper knowledge of the original milieu itself, it either cannot be reconstructed at all or, if reconstructed, there is no means to prefer one reconstruction solution to any other unless a comparative basis could be employed. If comparative basis is missing, it can be replaced by perceptions of the past social context and individual preferences.

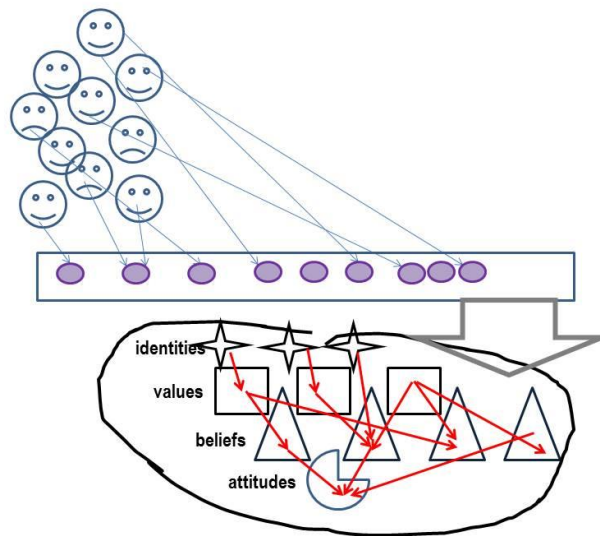
Working with political perceptions over the expression of past political attitudes involves however a difficult task of political perception construction. As constructs, both political attitudes and political perceptions of (past) political attitudes require web mining and special web harvesting techniques, the latter including online opinion and preferences expressions.

5. Modeling the processes of political attitude reconstruction

Since empirical data from public surveys on values and beliefs associated to democratic political participation in the Eastern European countries before 1989 cannot be found or extracted by web mining techniques, the theoretical and experimental settings described and reported in the present paper are aimed at political attitude reconstruction from (a) data extracted by web mining systems and (b) data provided from perceptions about past political attitudes (see Figure 2 a, b).



(a)



(b)

Figure 2 (a) and (b).
 Collection of instantly generated data (perceptions of the past).

In order to understand how political attitudes form and change in the authoritarian political regimes, we have dynamically generated relationships (links) between values, beliefs, norms and political attitudes. These dynamic relationships help in constructing individual perceptions over the past political attitudes and in generating the so-called “shadow configurations”, that is configurations of structural relationships which can support the identification and extraction of the items which play essential roles in the emergence or change of political attitudes toward government and governance, state and institutions, and society at large.

5.1 "Shadow Configurations" of Political Culture

The present approach started from the hypothesis that in authoritarian (strongly coercive) political regimes, individuals do not discuss openly about their true beliefs and feelings toward the regime itself. Fear and responsibility for their own lives and those of their families, on the one hand, or low levels of trust in the government and state, on the other hand, might have been factors which have influenced many people to not discuss openly about their true beliefs, to not overtly manifest their political attitudes when these were opposing the regime itself. A huge amount of research literature on these issues convey the idea that much of what people really thought about the political regime (dictatorship, institutions, legal environment, etc.) is still unknown or known from self-reporting data provided long time after properly experiencing opposition or distrust toward the political regime. Scarcity of empirical data resources makes very difficult the study of political attitudes formation and change during strongly coercive political regimes. Investigating day-to-day political attitudes toward such political regimes long time after they have been removed requires either access to survey data or the possibility to qualitatively reconstruct the data from the past documents. This task is usually performed by identifying and searching library and/or multimedia resources and results in the content-based analysis of retrieved data. When such documents are missing, this task may be replaced by qualitative modeling. One modeling solution is to replace content-based analysis of real document/data by simulation modeling of the studied phenomena, generating data and performing the content-based analysis on the experimental data. Such experimental data should consist in dynamic configurations (patterns) of relationships between political culture items like values, beliefs, political attitudes, norms, and knowledge. We have called this experimental data the political culture "shadow" configurations.

5.2 Generating "Shadow Configurations" of Political Culture Items

The shadow configurations have been obtained in two kinds of experimental settings:

The first one is a generative simulation modeling (Voinea, 2013) of the interactions between individual agents which hold personal beliefs, are attached to specific value systems, and express political attitudes toward the state/government/society at large. The individual agents are embedded in social environments which have an explicit economic dimension modeled as a set of public resources, rules for accessing the resources, and levels of social/individual welfare. The individual interactions are thus subjected to a double influence: the political influence exercised by economic means, and political influence exercised by political culture means. The former consists in the political influence exercised by the individual agents amongst themselves (micro level), and the political influence exercised by the macro-level political institutions onto the individual behaviors developed at the micro level. The latter consists in the political influence exercised by the macro-level institutions on the individual agents by means of public resource management and the welfare distribution mechanisms. This simulation experimental settings use generated data for the political culture items, namely the values, beliefs, and attitudes held by any individual agent. Modeling the political attitudes formation and change consists in simulations of self-organizing processes of political influence in which values, beliefs, affect and norms dynamically change, and their changes finally result in the political attitude emergence or in the political attitudes change (Voinea, 2013a; 2013b).

The second one is a generative simulation modeling based on harvesting data from online web multimedia resources. Using data mining techniques for identifying and harvesting appropriate data, collections of political culture items (values, beliefs, and political attitudes) are generated. A collection consists in political attitude structural items which are interlinked through a basic set of relationships. As the relationships are dynamically generated, that is they are generated if the web resources are accessible

(i.e., path-dependent processes), and the generated items are only contextually connected (i.e., context-dependent processes), the resulting collections may prove a high variability of configurations: they vary from one to the other in content, size, and complexity of connections. Their variability is therefore the reason for which they have been called “shadow configurations”, as they represent context-dependent instances of real political attitudes.

A shadow configuration is described by means of descriptor types. There are three types of descriptors used in the experimental trials: conceptual (text), video (motion picture), audio (sound), and photography (static picture). In the simplest version employed in the experiments, a shadow configuration might include four items: values, beliefs and/or sentiments, and political preferences (see **Figure 3**). Other items, like norms, are specified at the beginning of each experiment, and they do not change during single trials. Information flow or knowledge are both considered as dynamic input for each experiment and stimulate the change processes.

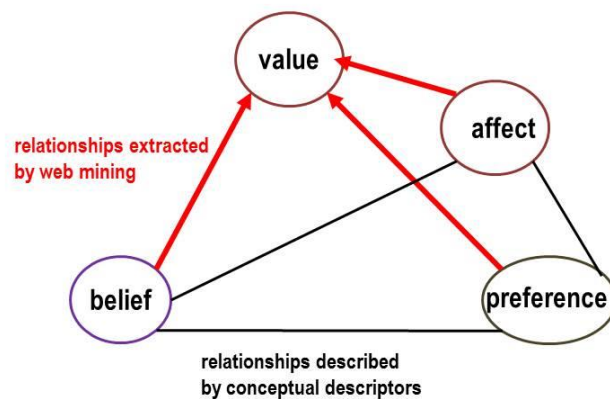


Figure 3.

Shadow configurations are collections of political attitude structural components and their relationships. In the initial experimental system, the structural components are “values”, “beliefs”, “sentiments”, and „preferences“. The web mining system has been used for extracting (1) the instances of values, beliefs, sentiments, and preferences, and (2) the instances of relationships between them. The extracted data is compared with the conceptual data and a synthesis is performed such that a final set of data is provided as outcome.

Since generating such data collections depends (heavily!) on the existing resources, the accessibility of resources and the web searching engines, the shadow configurations are highly unstable, requesting means for evaluating their validity and reliability. However weak they might appear, they have the advantage of providing clues about the real political attitudes and their potential structural configurations. This makes possible the study of political attitudes in the absence of empirical data (survey and field studies).

6. Theoretical and Experimental Settings

Re-collecting items from the web by means of data mining techniques for investigating library resources (usually these are text documents, books, reviews, etc.), and other public multimedia resources (video and audio recordings manipulated by broadcasting actors, like television, radio, and internet providers) is characterized by an associative dynamics. This dynamics characterizes the discovering of different connected items while searching on the web. This kind of dynamics addresses the size of a web collection (number of items), like for example the web collection for the issue of "kings' discourse" (see **Figure 4**). In this case, the dynamics of associating different web items, which are directly or indirectly concerned with the issue of "king's discourse", is the dynamics of the size of the collection of multimedia descriptors for the specific (fix) issue. Data mining techniques allow further analysis of the items/descriptors in the collection, which uses measurements of items' frequency, etc. Content-based analysis of a specific issue (in our case, "king's discourse") is further used in the analysis of the meaning of the investigated issue.

We could therefore generalize this procedure to a type of content-based analysis procedure which takes as inputs the dynamically (re-)collected web items, and provides as output a dynamic description of (a class of) phenomena called in this paper the "shadow configuration".

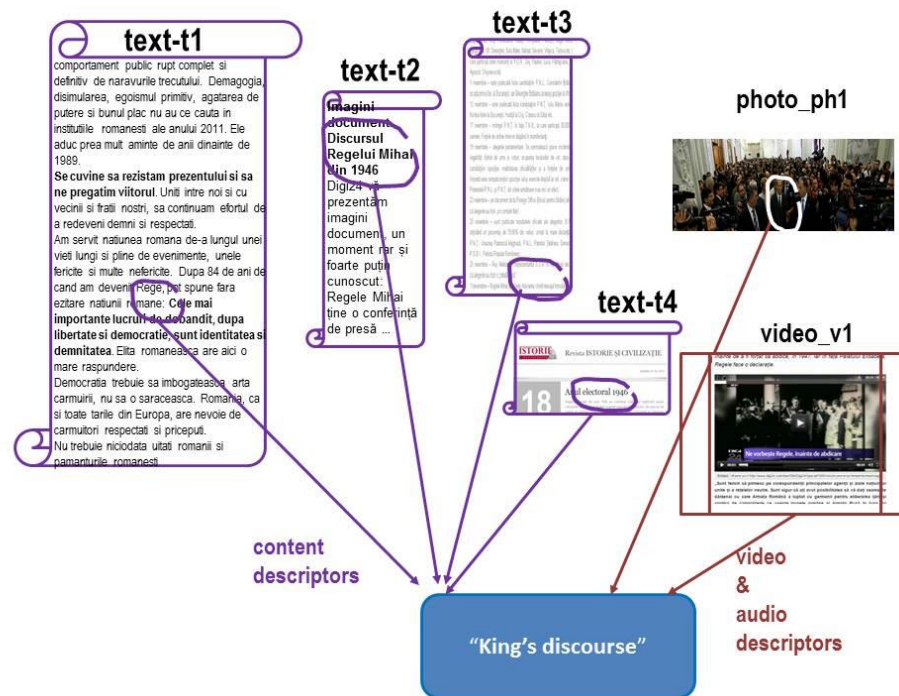


Figure4.

Web and data mining techniques provide the means for constructing multimedia item collections which include different categories of descriptors for some particular issue. Example: data mining on the issue of "King's discourse" might provide multimedia descriptors such as:

"**text_t1**" is a text fragment of King Michael's discourse in the Romanian Parliament in 2011, and "**photo_ph1**" is King Michael's photography in the Romanian Parliament in 2011 (*source: HotNews.ro; broadcasting date: October 25th, 2011; accessed on November 1st, 2014*);

“**text_t2**” is a text label for the “**video_v1**” video recording of King Michael’s discourse in the Romanian Parliament in 1946 (*source*: Romanian private television DIGI24; broadcasting date: May 10th, 2014; accessed on November 1st, 2014);

“**text_t3**” and “**text_t4**” are two text fragments of a political history article concerning the King Michael’s discourse in 1946 (*source*: the National History Museum of Romania as owner of the site: <http://www.comunismulinromania.ro> ; article published on August 10th, 2010; accessed on November 1st, 2014).

In our example (see **Figure 4**), multimedia relevant items can be generally described as:

Text/Conceptual Descriptors:	text_t1, text_t2, ..., text_tn
Video Descriptors:	video_v1, video_v2, ..., video_vm
Photo/Image Descriptors:	photo_ph1, photo_ph2, ..., photo_php
Audio Descriptors:	audio_a1, audio_a2, ..., audio_ak
Issue (Re-)Construction:	attribute_1, ..., attribute_q

where the issue re-construction is achieved by means of content-based analysis, and semiotics analysis.

In the “*king’s discourse*” experiment, various web resources have been searched with the explicit goal of extracting the structural components of the political attitudes toward the eventual come-back of the constitutional monarchy as a political regime in Romania. The goal of the experiments was to extract a configuration of the structural relationships between the components of a political attitude of accepting the constitutional monarchy as a political regime. The outcomes of the web extraction processes have been collected as dynamic structural configurations (shadow configurations).

The structural components of the political attitude have been described in terms of values (V), beliefs (B), sentiments (S), and preferences (P). On a content-based analysis of two discourses of the King Michael (one from 1946 and another one from 2011), the political attitude structure has been extracted as a collection of political culture items and the relationships between them. The content of this collection varies from one experiment to another depending on context knowledge and the accessible resources (see **Diagram 1**).

The output of this web mining experiment is called a “shadow configuration” (see **Figure 5**), as it provides only potential context-based constructs of the structure of political attitude toward constitutional monarchy in Romania.

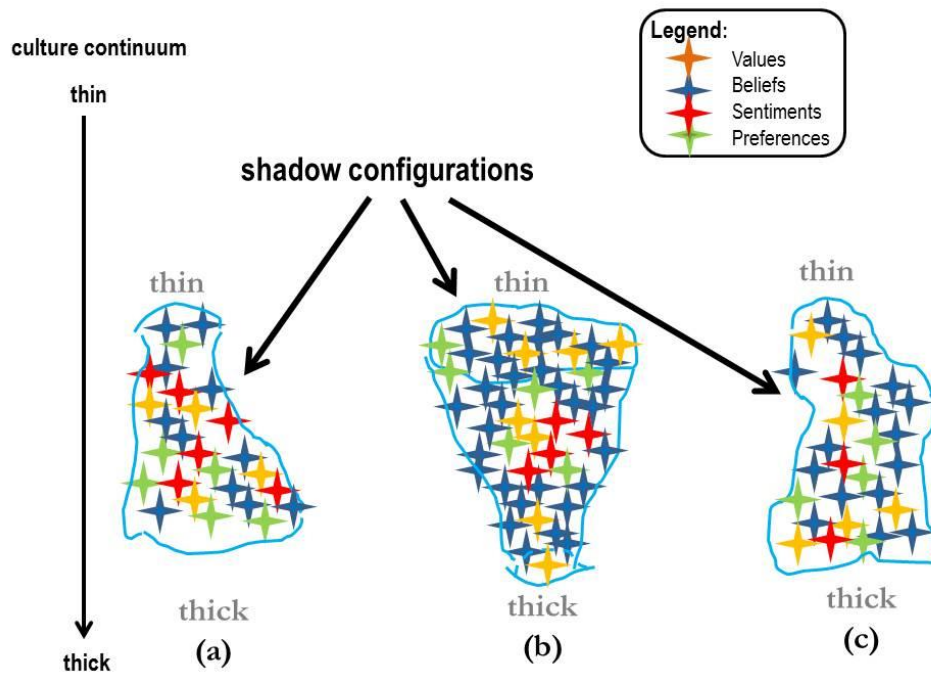


Figure 5.
Shadow configurations for the structure of the political attitude toward the constitutional monarchy as a political regime.

6.1 Romanian Experimental Settings and Preliminary Outcomes

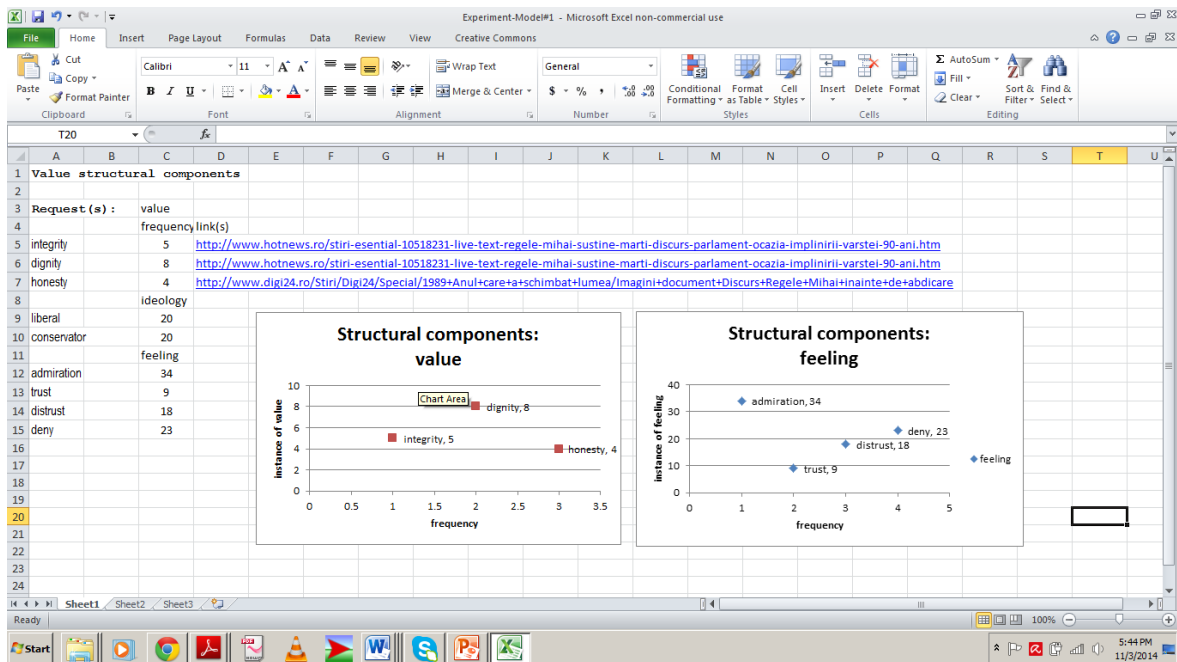
The shadow configurations provide the structural components of a political attitude toward the selected issue of study (constitutional monarchy):

Value structural components:

Request(s): value

What are the values pertaining to the emergence of acceptance/rejection political attitude?

Answer(s): instances of values obtained during the web mining experiments:



Value structural components

Request (s) :

value

frequency

link(s)

integrity

5

<http://www.hotnews.ro/stiri-esential-10518231-live-text-regele-mihai-sustine-marti-discurs-parlament-ocazia-implinirii-varstei-90-ani.htm>

dignity

8

<http://www.hotnews.ro/stiri-esential-10518231-live-text-regele-mihai-sustine-marti-discurs-parlament-ocazia-implinirii-varstei-90-ani.htm>

honesty

4

<http://www.digi24.ro/Stiri/Digi24/Special/1989+Anul+care+a+schimbat+lumea/Imagini+document+Discurs+Regele+Mihai+inainte+de+abdicare>

Diagram 1.

Web mining procedures return data analysis results, and links (sources of web extraction).

Processing of web mining results: structural components of the shadow configurations are listed and statistically analyzed.

The same type of request is submitted to the web mining system during the experiments for: beliefs, sentiments, behaviors, norms. System's output consists in collections of instances of beliefs, sentiments, behaviors, norms extracted from document texts. These outputs are further submitted for content-based analysis, sentiment-analysis, semiotic analysis, etc. (see *Section 6.1* and *Section 6.2* for more details). On the basis of the context data and shadow configurations, the outputs are used in simulating the political attitude formation and/or change.

Finally, the political attitude is reconstructed from the structural components (i.e., shadow configurations) which obtain high validity support from the simulation runs.

This is nevertheless a "happy" case because the items *do exist* in the web repositories. Moreover, these items can be searched, identified and finally harvested by complex data and web mining techniques. Construction of meaning in this case depends on the collection's size and on the items' relevance: the larger the collection of relevant items, the richer the meaning(s) of the studied issue.

6.2 Croatian Experimental System and Settings

The example briefly analyzed in Section 6.1 has been developed on the Romanian web mining platform.

A similar, however more advanced Croatian platform has been used for web mining as well as conceptual network analysis procedures of the Croatian Scientific Bibliography (CROSBI).³ The implemented system, which is still work in progress, used a number of technologies including web scraping (Scrapy),⁴ an object-relational database (PostgreSQL)⁵ an advanced scripting language (Python).⁶ The selection of technologies wasn't arbitrary, Scrapy allowed for easy implementation of a number of harvester agents that collected and extracted data from semi-structured documents that were stored in a specially designed PostgreSQL database. PostgreSQL was selected due to its unique text mining and natural language processing (NLP) capabilities that allowed for automated dictionary based text stemming. PostgreSQL uses dictionaries to eliminate words that should not be considered in a search (so called stop words), and to normalize words so that different derived forms of the same word will match (lexemes). Python was used to glue this technologies together and provide analysis related features.

CROSBI is a social application in which Croatian scientists provide bibliographic data about their publications (Schatten, 2013). A usual entry includes authors, title, type of publication, abstract, keywords, link to document (if available), language, databases the publication is abstracted in, scientific field, category, as well as a number of additional fields depending of document type like journal name or publisher.

In order to test a number of simple and advanced text search techniques all currently available⁷ bibliographic entries were harvested with a total of 385,236 distinct publications. Six searching techniques were tested and analyzed including:

simple (naive) keyword search – keywords were searched regardless of syntax and grammar;

NLP⁸ enhanced keyword search – keywords were normalized and then searched;

NLP enhanced title search – title was normalized and vectorized before searching;

NLP enhanced abstract search – abstract was normalized and vectorized before searching;

simple graph based keyword search – keyword graphs in form of folksonomy based conceptual networks (see Mika, 2007; Schatten et al., 2011; Schatten 2013 for more detailed descriptions) regardless of syntax and grammar were searched;

NLP enhanced graph based keyword search – as in the previous method, but this time keywords were normalized.

3 Available at <http://bib.irb.hr>

4 Available at <http://scrapy.org>

5 Available at <http://www.postgresql.org>

6 Available at <http://www.python.org>

7 Harvesting was conducted on October 10th, 2014

8 Due to a lack for support for the Croatian language by PostgreSQL, only English words were normalized in all cases.

configuration hypothesis. Future versions of this system might include sentiment analysis related features to analyze not only the semantics of a given corpus of text, but also the sentiments the authors had while writing them.

7. Conclusions and Future Work

The point in collecting the data from the web is neither their quantity, nor their variability, though all of them have a certain relevance, which cannot be dismissed. Collecting this data would be aimed basically in getting clues about the ways they generate patterns of political attitudes structural components which can be characterized by attributes like intensity, stability, resistance to change, etc.



Map Source: Map No. 3877 rev.7 United Nations, Dept. Of Field Support. Cartographic Section, September 2008, (accessed on November 5th, 2014).

Figure 7.

A Pan-Eastern European Repository Project. Member countries.

Such repository has the attribute of self-organization as an intrinsic attribute: the items within are dynamically linking to each other, changing these bindings as new perceptions, and knowledge are arriving into the system. The instant data self-organizes as new associations between identities, values, beliefs, norms and attitudes are included into the dynamically generated collections. The repository becomes therefore a living image of the past and present.

Current approach has several limitations. One concerns the necessity to extend the shadow configurations to including multimedia descriptors (video, audio, picture). Another one concerns the simulations experiments, which should develop individual interactions among agents on both culture and economic dimensions simultaneously.

Further developments address the question of identifying a value system description in terms of the minimal set of values which can describe the political change mechanism (Rokeach, 1973). The relationships between structural components should be defined and a theoretical approach is necessary in order to define the modeling paradigm of political attitude change in political culture terms. Previous approach (Voinea, 2013c) provides a modeling paradigm inspired by bottom-up approaches from social simulation. However, methodological individualism is not sufficient for explaining political macro-level change in the continuum culture model assumed in this paper.

The present paper presents the preliminary joint research work of Romanian and Croatian Team. The University of Belgrade and the Serbian Mihajlo Pupin Institute's Team (Sanja Vranes, Valentina Janev, Uroš Milošević) has recently joined this research project. The project is going to be extended to other Eastern European countries (see **Figure 7**) and create a Repository of the History of Eastern European Political Attitudes and Mentalities for the period 1945-2014.

References

- Almond, G.A., Verba S. (1963) *The Civic Culture: Political Attitudes and Democracy in Five Nations*, Princeton, NJ.
- Atanasov P., Simoska E. 2013. *Old/New Identities – The Case of Macedonia*, *European Quarterly on Political Attitudes and Mentalities* Vol.2 No.3, (July 2013), pp.40-52.
- Converse P. ([1964] 2006) *The Nature of Belief Systems in Mass Publics* (1964), *Critical Review: A Journal of Politics and Society*, Vol. 18, Nos.1-3, pp. 1-74. Available online, accessed on May 7th, 2013.
- Hedström, P. (2005) *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.
- Inglehart R. (1997) *Modernization and Postmodernization: Cultural, Economic and Political Change in 43 Societies*, Princeton University Press.
- Inglehart R., Welzel Ch. (2005) *Modernization, Cultural Change and Democracy*, NY, Cambridge University Press.
- Inglehart, R. (1971) *The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies*, *American Political Science Review* No. 65, p. 99-1017.
- Inglehart, R. (1977) *The Silent Revolution, Changing values and Political Styles among Western Publics*, Princeton: Princeton University Press.
- Karklins, R. (2005) *The System Made Me Do It - Corruption In Postcommunist Societies*, Rasma Karklins, M. E. Sharpe, New York.
- Kubik J. (2012) *Cultural approaches*, in: Raj Kollmorgen, Wolfgang Merkel and Hans-Jürgen Wagener (Eds.) *The Handbook of Transformation Research*, Publisher: VS-Verlag für Sozialwissenschaften.
- Laitin, D.D. (1995) *The Civic Culture at 30*, *American Political Science Review* Vol.89, No.1, pp. 168-173.
- Marc van der Erve and Erik Klein Nagelvoort, "Democracy 2.0 – Time to look into a clean mirror" (Submitted for publication), available online on the Marc van der Erve's LinkedIn Profile, accessed on November 2nd, 2014.
- McAdam D., Tarrow S., Tilly Ch. (Eds.) (2001) *Dynamics of Contention*, Cambridge University Press.
- Mika, P. (2007) *Ontologies are us: A unified model of social networks and semantics*. *Web Semantics: Science, Services and Agents on the World Wide Web*. 5(1), pp. 5–15.

- Mishler W., Pollack D. (2003) Thick-and-Thin: A Neo-Cultural Synthesis, in: Pollack D., Jacobs J., Muller O., Pickel G. (Eds.) Political Culture in Post-Communist Europe. Attitudes in New Democracies, Ch.13, pp.237-256. Aldershot: Ashgate.
- Precupețu, I. (2008) On the nature and causes of corruption in Romania. Evidence from a Grounded Theory approach. Discussion paper, University of Konstanz, 2008, available online, accessed on June 23rd 2012 at: <http://www.unikonstanz.de/crimeandculture/docs/Discussion_Paper_No_13_ICCV_April_2008.pdf>
- Pye, L., Verba, S. (1965) Political Culture and Political Development, Princeton.
- Rohrschneider, R. (1999) Learning Democracy. Democratic and Economic Values in Unified Germany. Oxford: Oxford University Press.
- Rohrschneider, R., and Schmitt-Beck, R. (2003) Trust in Democratic Institutions in Germany: Theory and Evidence Ten Years After Unification, in: German Politics 11 (3), pp.35-58.
- Rokeach, M. (1973) The nature of human values. New York: Free Press.
- Russell, D., Welzel, C. (Eds.) (2014) The Civic Culture Transformed: From Allegiant to Assertive Citizens, Cambridge: Cambridge University Press.
- Schatten, M., Rasonja, J., Halusek, P., Jakelić, F. (2011) An Analysis of the Social and Conceptual Networks of CECIS 2005 – 2010, in: Hunjak, T., Lovrenčić, S., Tomičić, I. (Eds.) Proceedings of the 22nd Central European Conference on Information and Intelligent Systems, Varaždin : Faculty of Organization and Informatics, pp.259-264.
- Schatten, M. (2011) Mining Social and Semantic Network Data on the Web, FOI, University of Zagreb, Croatia.
- Schatten, M. (2012) Opinion Mining on News Portal Comments. Towards a Croatian Sentiment Database, SlidePresentation, FOI, University of Zagreb, Croatia.
- Schatten, M. (2013) What do Croatian Scientist Write about? A Social and Conceptual Network Analysis of the Croatian Scientific Bibliography, *Interdisciplinary description of complex systems* Vol.11 No.2, pp. 190-208.
- Stamenova, Sv. (2000a) Determinants of Political Support in Bulgaria, in: Political and Economic Orientations of the Bulgarians Ten Years After the Beginning of the Systemic Change (Analyses of Survey Results), Georgi Ganev (Ed.), Sofia: Center for Liberal Strategies.
- Stamenova, Sv. (2000b) Work Values Shift Tendencies during Post-communist Transformation Period (The Polish and The Bulgarian examples), in: European Value Systems, J. Koralewicz (Ed.), ISP, PAN.
- Stamenova, Sv. (2002) Conceptions of Post-Communist Social Change: 1990-1999, in: Sociological Perspectives, Nikolai Tilkidjiev and Stefan Dontchev (Eds.), Sofia: Professor Marin Drinov Academic Publishing House, pp. 294- 304.
- Stamenova, Sv. (2003) Peculiarities of Political Culture in Post-Communist South East European Countries (The examples of Macedonia, Serbia and Montenegro, and Bulgaria), in: New Europe College Yearbook 2001-2002, Bucharest: Humanitas, pp. 211-250 <<http://library.nec.ro/papers/regional2001-2002/yearbook.htm>>
- Tilly, Ch. (2001) Mechanisms in Political Processes, *Annual Review of Political Science* No. 4, pp.21–41.
- Voinea C.F. (2012) Briberyscape: An Artificial Society-Based Model of Corruption Emergence and Growth, in: C.F. Voinea (Ed.) Political Attitudes and Mentalities. The Historical heritage of Europe: A Challenge for the Future of Political Analysis, ArsDocendi-Bucharest University Press, pp. 113-148.
- Voinea C.F. (2013a) Why Individual Interactions Are Not Enough? Agent-Based Simulation of Polity: An Operational Concept, in: C.F. Voinea, B. Todosijevic, G. Boella (Eds.) Political Attitudes and Mentalities. Eastern European Political Cultures: Modeling Studies, ArsDocendi-Bucharest University Press, pp.77-122.
- Voinea, C.F. (2013b) Eastern European Political Cultures and the Need for Explanation. The Look from Inside, in: C.F.Voinea, B. Todosijevic, G. Boella (Eds.), Political Attitudes and Mentalities. Eastern European Political Cultures: Modeling Studies, ArsDocendi-Bucharest University Press.
- Voinea, C.F. (2013c) Patterns of Paradigmatic Preferences and Tendencies in the Political Culture Studies of the Eastern European Political Phenomena, in: C.F.Voinea, B. Todosijevic, G. Boella (Eds.), Political Attitudes and Mentalities. Eastern European Political Cultures: Modeling Studies, ArsDocendi-Bucharest University Press.