

Public Policy Formulation through Non Moderated Crowdsourcing in Social Media

Yannis Charalabidis, Anna Triantafillou, Vangelis Karkaletsis, Euripidis
Loukis

► **To cite this version:**

Yannis Charalabidis, Anna Triantafillou, Vangelis Karkaletsis, Euripidis Loukis. Public Policy Formulation through Non Moderated Crowdsourcing in Social Media. Eftthimios Tambouris; Ann Macintosh; Øystein Sæbø. 4th International Conference on Electronic Participation (ePart), Sep 2012, Kristiansand, Norway. Springer, Lecture Notes in Computer Science, LNCS-7444, pp.156-169, 2012, Electronic Participation. <10.1007/978-3-642-33250-0_14>. <hal-01541473>

HAL Id: hal-01541473

<https://hal.inria.fr/hal-01541473>

Submitted on 19 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Public Policy Formulation Through Non Moderated Crowdsourcing in Social Media

Yannis Charalabidis¹, Anna Triantafillou², Vangelis Karkaletsis³, Euripidis Loukis¹

¹University of the Aegean, Information and Communication Systems Engineering Dept.,
Gorgyras and Palama Str., 83200 Karlovassi, Samos, Greece
{yannisx, eloukis}@aegean.gr

²Athens Technology Center, Rizariou Str. 10, 15233 Halandri, Athens, Greece
a.triantafillou@atc.gr

³National Center for Scientific Research "Demokritos", Institute of Informatics and
Telecommunications, P. Grigoriou and Neapoleos, 15310 Ag. Paraskevi, Athens, Greece
vangelis@iit.demokritos.gr

Abstract. The emergence of web 2.0 social media enables the gradual emergence of a second generation of e-participation characterized by more citizens' control, in which government agencies post content (e.g. short or longer text, images, video) to various social media and then analyze citizens' interactions with it (e.g. views, likes/dislikes, comments, etc.). In this paper we propose an even more citizens controlled third generation of e-participation exploiting web 2.0 social media as well, but in a different manner. It is based on the search by government agencies for content on a public policy under formulation, which has been created in a large set of web 2.0 sources (e.g. blogs and microblogs, news sharing sites, online forums) by citizens freely, without any initiation, stimulation or moderation through government postings. This content undergoes advanced processing in order to extract from it arguments, opinions, issues and proposals on the particular policy, identify their sentiments (positive or negative), and finally summarize and visualize them. This approach allows the exploitation of the vast amount of user-generated content created in numerous web 2.0 social media for supporting governments to understand better the needs, wishes and beliefs of citizens, and create better and more socially rooted policies.

Keywords: e-participation, public policy, crowdsourcing, social media, sentiment analysis, visual analytics.

1 Introduction

For long time it has been widely recognised that representative democracy, in which citizens' representatives identify needs and problems of society and design public policies and services for addressing them, should be combined with continuous citizens' involvement; this is based on an understanding of the growing complexity and multi-dimensionality of social problems and needs, which necessitates the exploitation of citizens' extensive knowledge on them. This gave rise to a new model of democracy, which is termed "participatory democracy" [1] – [4], and combines

decision making by citizens' representatives with citizens' participation, with the latter not replacing but supporting and enhancing the former. Row and Frewer [5] define public participation as 'the practice of consulting and involving members of the public in the agenda-setting, decision-making and policy forming activities of organizations or institutions responsible for policy development'. Many governments all over the world have shown consistent interest in and made considerable efforts for enabling the participation of citizens in the formulation of public policies through various mechanisms and channels [6] – [8]. We can distinguish three main motivations of government for this: to produce better quality policies (proportionality and fairness), to build trust and gain acceptance of policy-making processes (awareness and consensus-building), and to share responsibility for policy-making (inclusion and participation).

The development and increasing penetration of information and communication technologies (ICT), and the Internet in particular, provides an efficient and inclusive channel for the extensive application of the above principles. This gave rise to the development of electronic participation (e-participation) [9] – [13]. According to the OECD [10] e-participation is defined as the use of ICTs for supporting the provision of information to the citizens concerning government activities and public policies, the consultation with the citizens and also their active participation. However, the first generation of e-participation has been highly government-controlled, based on official websites of government agencies, which defined and controlled the topics and rules of all electronic discussions taking place there. The use of these official e-participation spaces by the citizens has been in general limited and below the initial expectations [14]; most of these official e-participation spaces were largely unknown to the general public due to the high costs of promotion and the slow pace of dissemination, while the topics dealt with were sometimes distant from people's daily problems and priorities, so that content contributions by non-experts was inhibited.

The emergence of the new Web 2.0 social media enables the gradual emergence of a second generation of e-participation, taking place in popular social media where citizens choose to discuss and generate content, which is characterized by less government control and more citizens' control. In this new paradigm of e-participation government agencies post content (e.g. short or longer text, images, video) to various social media on their policies under formulation or implementation, and then collect and analyze citizens' interactions with it (e.g. views, comments, likes/dislikes, retransmissions, etc.). Therefore, in this second generation of e-participation governments continue defining and controlling the discussion topics to some extent (through making postings and initiating discussions on the topics they choose), but do not control the rules of the discussion (which are the ones defined by the social media). It has been widely recognised that governments have to exploit the numerous users-driven Web 2.0 virtual spaces, which have been launched through citizens' initiatives with dramatic success in terms of adoption and usage, for widening and enhancing e-participation. In many of these social media there is already significant 'bottom-up' political activity initiated by the citizens and not by government agencies [15] – [18]. Many political discussions are taking place there, political information and news are exchanged and propagated, and also off-line political events and initiatives (e.g. movements, demonstrations) are organized and promoted. For these reasons some governments have already started adopting this

new e-participation paradigm and exploiting social media (e.g. Facebook, Youtube, Twitter, Blogger, etc.) in many different domains of government, including public participation [19] – [22].

In this paper we propose an even more citizens driven and controlled third generation of e-participation, which again exploits web 2.0 social media, but in a different way than the second generation. It is based initially on the search by government agencies for content on a public policy under formulation, which has been created in numerous web 2.0 sources (e.g. blogs and microblogs, news sharing sites, online forums, etc.) by citizens freely, without any initiation, stimulation or moderation through government postings. This content then undergoes linguistic processing and semantic analysis in order to extract from it opinions, arguments, issues and proposals on the particular policy, identify their polarity (positive or negative), and finally summarize and visualize them. This approach allows the exploitation of the vast amount of user-generated content that is created in numerous web 2.0 social media, in order to supporting governments in understanding better the needs, wishes and beliefs of citizens, and creating better and more socially rooted policies. In this third generation of e-participation governments control neither the discussion topic (they simply 'listen' what topics and issues are mentioned by citizens in a broad policy area, which correspond to needs and problems that citizens, and not governments, identify), nor the discussion rules (they accept the ones defined by the various social media they exploit). The research presented in this paper has been conducted as part of the research project NOMAD ('Policy Formulation and Validation through nonmoderated crowdsourcing', which is partially funded by the 'ICT for Governance and Policy Modelling' research initiative of the European Commission.

The paper is structured in five sections. In section 2 the theoretical foundations of the proposed methodology are outlined. Then in section 3 the methodology is described, while in section 4 is presented the technological architecture of the ICT platform that will support this methodology. Finally, section 5 summarizes the conclusions and proposes future research directions.

2 Theoretical Foundations

In this section the theoretical foundations of the proposed methodology are outlined: a) bottom-up public policy formulation (in 2.1), and b) crowdsourcing (in 2.2).

2.1 Bottom-up Public Policy Formulation

The formulation of public policies for addressing problems and needs of society has been traditionally a top-down process. Governments and citizens' representatives have had the main role in identifying needs and problems of society and designing public policies and services for addressing them. The participatory democracy and public participation ideas gave some role to the citizens, who are provided some information on government activities and policies, and are asked to participate in

consultations and contribute through opinions and responses on issues and questions defined by government. According to OECD [6] – [8] the development of public participation includes three stages. The first two of them are:

- Information Provision, which establishes an ‘one-way relation’ between government and citizens, in which government produces and delivers information to be used by citizens.

- Consultation, which establishes an asymmetric ‘two-way relation’ between the government and the citizens, in which citizens provide views and feedback to government on issues and questions that government has previously defined.

These first two stages of public participation have a strong top-down orientation, and are fully controlled and initiated by government agencies. However, according to the above OECD reports (which are based on the experience of its numerous member countries), they can be followed (depending on the whole political context) by a third stage termed as ‘Active participation’, in which a more symmetric ‘two-way relation’ between government and citizens is established. In this stage citizens have a wider role in proposing new policy options and discussion topic, in addition to the ones proposed by government, and in shaping the policy dialogue in general, though the government still has the responsibility for the final decisions. This form of public participation is less government controlled, and can result in bottom-up public policies formulation providing significant benefits: it can provide new fresh ideas concerning needs and problems of society, ways of addressing them, advantages and disadvantages of the latter, and lead to the identification of cultural, social and economic changes that require government action, and to overcoming obsolete government agencies stereotypes in understanding social problems and needs and designing solutions.

According to the corresponding OECD reports on e-participation [9] – [10], the development of e-participation follows a similar trajectory. Its first two stages have a strong top-down orientation, aiming to support through ICT, and especially exploiting the Internet, the provision of information to the citizens on government activities and policies, and the consultation with the citizens on issues and questions defined by government. However, its third stage has a bottom-up orientation and focuses on supporting and facilitating a more active participation of citizens, through online tools and discussion formats which enable them to set the agenda for discussion (e.g. raise new issues that have to be discussed, in addition to the ones raised by the government), submit their own proposals and policy options and in general shape the final outcomes. Similarly, United Nations [23] – [24] suggest a three-step plan for enhancing eParticipation, which includes: (a) increasing e-information provision to citizens for decision making, (b) enhancing e-consultation for deliberative and participatory processes (both these steps having a top-down orientation) and (c) supporting e-decision making by increasing the input of citizens in decision making (this step having a bottom-up orientation).

This bottom-up dimension of e-participation and policy formulation, despite the abovementioned benefits it provides, is the least exploited and explored. Only some types of ICT tools having been investigated for this purpose, such as e-petition spaces (in which citizens can enter petitions concerning new government actions (e.g. services, programs, regulations) to the government or parliament, and solicit support and signatures) [25] – [26], electronic discussion groups supporting the development

of new policy options and the deliberation on them and online referenda. Also, the analysis of e-participation initiatives of many countries has shown that they focus mainly on e-information provision and e-consultation (i.e. on top-down oriented e-participation and policy formulation), and much less on the electronic support of citizens public policy related initiatives (i.e. bottom-up oriented e-participation and policy formulation) [27] – [28]. The methodology proposed in this paper contributes to filling this gap, as it allows the exploitation of user-generated content created freely by citizens in numerous web 2.0 social media on a public policy under formulation, without any government initiation or stimulation, in order to identify relevant needs, problems, proposed government actions, and also advantages and disadvantages of the latter.

2.2 Crowdsourcing

While previously the design and problem solving activities of firms were regarded as tasks performed exclusively by highly knowledgeable professionals, recent literature argues that these critical tasks can be performed even better by large, diverse and pluralistic teams of less knowledgeable individuals, giving rise to new distributed group-based multi-disciplinary design and problem solving practices [29]. The capability of a large network of people, termed as ‘crowd’, networked through web technologies, to perform difficult design and problem solving activities is gradually recognized by management researchers and practitioners. This is referred to as ‘collective intelligence’, which is defined as a ‘form of universally distributed intelligence, constantly enhanced, coordinated in real time, and resulting in the effective mobilization of skills’ [30]. Since ‘no one knows everything, everyone knows something, [and] all knowledge resides in humanity’ [30], digitization and communication technologies must become central in this coordination of far-flung genius. In [31] and [32] are described and analyzed several cases of crowd wisdom at work, which results in successful solutions emerging from a large body of solvers. Based on the analysis of such cases [32] finds that ‘under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them’. This is due to the diversity of opinion, independence, decentralization and aggregation that characterize such a crowd.

These ideas lead to the emergence of crowdsourcing, which is defined by J. Howe, one of the pioneers of this domain, as ‘the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals’ [33]. Another practical definition of crowdsourcing by D. Brabham is ‘a new web-based business model that harnesses the creative solutions of a distributed network of individuals’, in order to exploit ‘collective wisdom’ and mine fresh ideas from large numbers of individuals [32].

Crowdsourcing has started being applied initially in the creative and design industries, and then it expanded into other industries, for solving both mundane and highly complex tasks. It gradually becomes a useful method for attracting an interested, motivated group of individuals, which can provide solutions superior in

quality and quantity to those produced by highly knowledgeable professionals. As reported by several case studies [31] – [34] such a crowd can solve scientific problems that big corporate R&D groups cannot solve, outperform in-house experienced geophysicists of mining companies, design original t-shirts resulting in very high sales, and produce highly successful commercials and fresh stock photography against a strong competition from professional firms. These success stories indicate that the crowd, at least in some design and problem solving activities and under appropriate circumstances, can outperform the traditional industry faster and cheaper. This can result in a paradigm shift and new design and problem solving practices in many industries.

Limited research has been conducted on crowdsourcing, which is mainly descriptive, presenting mainly success stories of using this approach for performing design and problem solving tasks. However, there are some studies attempting to generalize and identify trends and patterns in this area. A typical example is the study presented in [34], which, based on the analysis of several case studies, identifies four dominant crowdsourcing types/approaches: i) the knowledge discovery and management approach (= an organization tasks crowd with finding and reporting information and knowledge on a particular topic), ii) the broadcast search approach (= an organization tries to find somebody who has experience with solving a rather narrow and rare empirical problem), iii) the peer-vetted creative production approach (= an organization tasks crowd with creating and selecting creative ideas), and iv) distributed human intelligence tasking (= an organization tasks crowd with analyzing large amounts of information). Also, there are some studies that aim to develop methodologies for the practical application of crowdsourcing in various types of organizations, and for the motivation of individuals to participate [35] – [37].

Though crowdsourcing ideas have been applied initially in the private sector, it is interesting and useful to investigate their applicability in the public sector. There are some first studies concluding that government organizations can apply these ideas, and use ‘citizen-sourcing’ for collecting information on citizens’ needs and for the solution of difficult problems [38] – [43]. According to [38] traditionally government agencies provide services to citizens, who consume them without questioning about them or taking part in decisions that led to their design and provision; social media can drive and facilitate new paradigms of government services design based on citizens-sourcing, in which citizens’ roles change, so that government become a consumer to whom citizens provide information, knowledge or even useful professional services. Citizen-sourcing can lead to the application of open innovation ideas in the public sector [34], and gradually result in ‘co-production’ of public services by government and citizens in cooperation [30].

This paper makes a contribution in this direction, presenting a methodology for the application of crowdsourcing by government agencies, which follows the first of the above mentioned crowdsourcing approaches proposed by [34] (discovery and reporting of information and knowledge on a public policy under formulation). An innovative feature of the proposed methodology is that it supports an ‘extended crowdsourcing’, which extends even before the ‘open call’ for information, knowledge and solutions for a particular topic or issue that characterizes crowdsourcing (see the above mentioned definition of it in [33]), having both ‘passive’ (non-moderated) and ‘active’ (moderated) components. In particular, it includes initially a search for

relevant content on a public policy under formulation, which has been created in numerous web 2.0 sources (e.g. blogs and microblogs, news sharing sites, online forums, etc.) by citizens freely, without any initiation, stimulation or moderation through government postings (= passive crowdsourcing). Based on the advanced processing and analysis of this content some first conclusions on the needs, issues, opinions, proposals and arguments of citizens on this public policy under formulation, which will provide a basis for an 'open call' (= active crowdsourcing), having the form of postings to various social media, that solicits citizens' interactions (e.g. likes/dislikes) or comments (e.g. remarks, opinions, positions, proposals) on more specific topics and questions.

3 Methodology Description

The proposed methodology includes four stages (Figure 1), which enable the policy-makers users of it (e.g. government organizations, members of parliament, politicians) to effectively LISTEN and monitor what citizens say in social media, ANALYZE those conversations and get the main stakeholders positions and opinions, RECEIVE all responses and data properly processed and displayed for an effective use and exploitation, and finally ACT on this information by proceeding to a more active crowdsourcing through postings to various social media. The results of the policy-maker's actions will also be monitored and undergo the same processing in order to lead to refinements until the finalized policy is formulated.

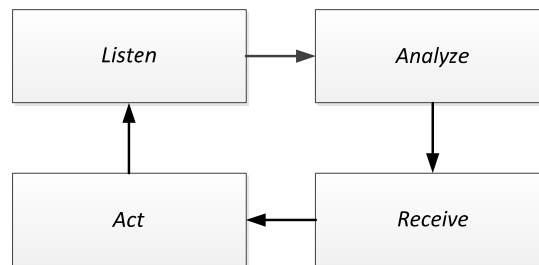


Fig. 1. The four stages of the proposed methodology

In particular, the first stage, called LISTEN, includes listening and monitoring what people say, what their needs, their opinions, positions and proposals are on a topic (e.g. a public policy under formulation) defined by the policy-maker. For this purpose a focused crawler is necessary, that is, a program which browses the Web in an automated, organised fashion, and attempts to visit solely content that is relevant to a previously known topic (or set of topics). For the purposes of the discussed platform, the crawler should reach relevant textual information from a variety of sources, with emphasis on environments that provide sizable user-created content, like blogs, micro-blogging sites, social networking platforms, discussion forums, comment-enabled news, corporate or video sharing sites etc.

The second stage, called ANALYZE, includes advanced processing and analysis of

the content, from which are identified relevant citizens' needs, opinions, concerns, proposals, sentiments and other information hidden within the citizens' conversations. As the majority of the available information is in textual form, the processes related to this stage will mainly make use of linguistic techniques. In particular, each web page found by the crawler in the previous stage will go through a series of automated analysis processes:

- *Language Detection*, which will recognize the language used in the page.
- *Opinion and Argument Extraction*, with the use of the appropriate semantic similarity measures and inference mechanisms that will allow the identification of analysed content that is pertinent to the particular policy.
- *Sentiment Analysis*, using smart sentiment classifiers analyse the mentions, which recognize the polarity (positive, neutral, negative) of the elements identified above.
- *Argument Summarisation*, with the use of the appropriate algorithms for generating qualitative information about opposing arguments, in the form of anonymity-preserving and automatically-generated summaries.

The RECEIVE stage aims to present to the end-user (policy maker) the knowledge acquired from the previous phases in a complete, coherent and usable manner. The platform will provide an aggregated view of the discovered opinions, their polarity, their association with policy concepts and statements, and statistical indications of their significance and impact. With the use of visual analytics all related data will be presented into a visible form that highlights important features, including commonalities and/or discrepancies. In this context, all the data that comes from sources as diverse as blogs, online opinion polls and government reports are properly displayed for an effective use by the policy-maker.

The three stages above perform a 'passive crowdsourcing', where the policy maker collects and analyses content which has been created without his/her intervention or moderation. In the fourth stage, ACT, the policy maker uses the collected cluster of problems, needs and proposals to formulate a draft-policy agenda, which can be tested out against social opinion. This is going to be achieved by 'active crowdsourcing' where policy text is posted to various social media (e.g. blogs, Twitter, Facebook, YouTube, etc.), soliciting citizens' remarks, opinions, positions and proposals on it.

4 Platform Technological Architecture

The methodology proposed in the previous section will be implemented and supported using an advanced ICT platform, which is going to be developed as part of the NOMAD project mentioned in the introduction. Its architecture is shown in Figure 2. Its main components are:

I. Policy Modeling Tool: Using advanced semantic representation technologies, it will model the particular policy, its domain and also topics and arguments relevant to the policy. It will be a user-friendly visual tool for semantic authoring, to be used by a domain expert with limiting expertise in knowledge representation technologies. It will be based on the ELEON Ontology Authoring Environment [44], developed by the National Center for Scientific Research 'Demokritos', which participates as a partner in the NOMAD project consortium. ELEON supports editing ontologies and

relating such ontologies with linguistic resources that can be used to extract structured ontological information from text. It supports the author with a number of innovative methods for ontology checking [45] and auto-completion [46].

II. Data Acquisition Module: The envisioned system will include a set of topical crawling modules in order to access web information from a variety of Web 2.0 applications. The topics will be defined via the appropriate keywords or phrases. In accordance with the proposed methodology, we will focus on sources that ensure the acquisition of user-created content, in particular Twitter, Facebook, Google+, YouTube, blogs constructed with Blogger and WordPress, forums and community networks, etc. The different sources have distinct content characteristics, in terms of the median size of each semantically meaningful text segment (article, comment, status, tweet etc.) and the general characteristics of the participating public. Thus, the various crawling modules will use a differentiated method for determining the relevance of a segment to the policy model concepts. Furthermore, the crawlers will provide source-level demographic information for the acquired content (by analyzing the available metadata and incorporating source information by platforms like Alexa).

III. Linguistic Processing Pipeline: The acquired textual content will be processed by a linguistic pipeline, comprising all the necessary processes for extracting information from the examined sources. In general, standard linguistic analysis will be applied (sentence splitting, tokenization, POS-tagging) in order to produce a structured form of the available data for further processing. An important step for the association with the policy model concepts is the identification of named entities and in particular entities relevant to the domain under examination. Thus, a Named Entity Recognition and a Co-reference Resolution process are required in order to achieve the semantic association of the available content and the constructed policy model. For the implementation of the linguistic pipeline, there will be parameterized and used the relevant components and tools available in the Ellogon platform developed by NCSR 'Demokritos' [47] – [48], an open source, general purpose natural language engineering platform, which provides a large set of linguistic analysis components. The system will take into account content in three languages (English, Greek and German) for the components that are language-dependent.

IV. Opinion Mining & Argument Extraction: The process is three-fold; the first step is to identify the presence of an opinion (i.e. subjective statements) related to a concept pertinent to a given policy model; the second step is to extract arguments that relate to and support the identified opinion; the final step is to determine the polarity (positive, negative, neutral) of the statement towards the aforementioned concept. The identification of text segments that bear an opinion and –possibly- related arguments will rely on machine learning techniques, where the machine learning system is trained with annotated documents in selected topics and with domain-agnostic content (a subset of the initial acquired web content). The system will also retain demographic information derived from the characteristics of the content's source and linguistic characteristics of the actual text. The sentiment analysis sub-component will be based on existing tools developed by NCSR 'Demokritos' [49, 50] as well as on other open source tools and resources. The implemented algorithms take into account various intricacies of the language forms commonly used in the context of user-generated web content, such as metaphors, nuances, irony etc.

V. Argument Summarization:The collected arguments and opinions and their relations to the various aspects and topics of the policy under formulation will be used in order to compile multi-dimensional quantitative data. The aggregation of individual arguments and opinions will be done in a manner which guarantees that the served data cannot be used to identify individual opinions and arguments. Furthermore, topical and geographical trends in each source (e.g., the tendency to discussions that pertain to particular places or topics) will be used to annotate sources, and the data crawled from them, with broad demographic and geographical properties. For example, a source can be annotated as a forum where mostly young people post, or one where mostly engineers post, or mostly Greeks, or mostly students of literature, etc. Such annotations will be exploited by the visualization component in order to offer a rich a multi-faceted presentation of the aggregated data.

VI. Social Reaction Visualization: This component will provide the policy maker with multi-faceted aggregates based on the results produced by the previous components, intuitively presented using information visualization and visual analytics techniques. We are going to explore the exploitation of visual analytics ideas and methods in the public policy formulation domain. The term visual analytics was introduced in the middle of the previous decade (2004-2005) in the research and development agenda [54, 55]. It describes a new multidisciplinary analytics field that combines various research areas including visualization, human computer interaction, data analysis, data management, geo-spatial and temporal data processing, spatial decision support and statistics. Application areas can be found wherever large amounts of data have to be analyzed. These include physics and astronomy, climate simulation, business intelligence and many more as stated in the European research roadmap [56]. However, currently there is no research concerning the exploitation of visual analytic tools and methods for supporting the policy making life cycle. Our research aims at filling this research gap.

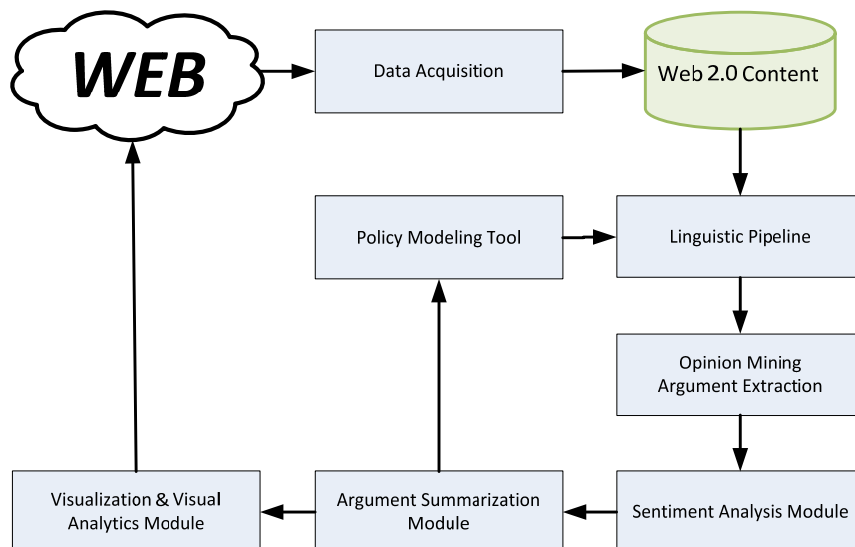


Fig. 2. An abstract depiction of platform's technological architecture

5 Conclusions

The first generation of e-participation has been based on the construction of official websites of government agencies, which provided to the citizens information on their activities and policies, and also capabilities for participation in consultations, whose topics and rules were defined fully by the government. Therefore it was highly government-controlled and had a top-down orientation. The limited use of these official websites and the emergence of web 2.0 social media leads to the gradual emergence of a second generation of e-participation, which is based on content (e.g. short or longer text, images, video) posted by government agencies to various social media, and then on the analysis of citizens' interactions with it (e.g. views, likes/dislikes, comments, etc.). This second generation of e-participation is less government-controlled and more citizens-controlled (as governments continue controlling the discussion topics to some extent - through making postings and initiating discussions on the topics they choose - but do not control the rules of the discussion - which are the ones defined by the social media). In the previous sections, a methodology has been proposed that leads to a third generation of e-participation, exploiting also the web 2.0 social media, and characterised by even more citizens' control. It includes initially the search by government agencies for content on a public policy under formulation, which has been created in numerous web 2.0 sources (e.g. blogs and micro-blogs, news sharing sites, online forums, etc.) by citizens freely, without any initiation, stimulation or moderation through government postings. This content is retrieved, centrally stored and processed using various advanced techniques in order to extract from it citizens' opinions, issues, arguments and proposals on the particular policy, identify their sentiments (positive or negative), and finally summarize and visualize them.

The proposed methodology allows the exploitation of the vast amount of user-generated content created in numerous Web 2.0 social media for supporting governments in understanding better the needs, wishes and beliefs of citizens, and creating better and more socially rooted policies. It can enable a more bottom-up e-participation and public policy formulation, which can offer significant benefits: it can provide new fresh ideas concerning needs and problems of society, ways of addressing them, advantages and disadvantages of the latter; also, it can lead to the identification of cultural, social and economic changes that require government action, and to overcoming obsolete government agencies stereotypes in understanding social problems and needs and designing solutions. Also, it provides guidance for the application of crowdsourcing by government agencies, in a novel 'extended' form, which includes both 'passive crowdsourcing' (non-moderated) and 'active crowdsourcing' (moderated) stages. In this way more valuable 'tacit knowledge' on important social problems and needs, and policy options for addressing them, which is possessed by large numbers of citizens, can be transformed into 'explicit (codified) knowledge' that can be used by government for designing better and more socially rooted policies.

The proposed methodology will be further validated in the NOMAD project. Initially, based on user requirements analysis, the ICT platform that will support the implementation of the methodology (see Figure 2) will be developed. Then, the

proposed methodology and the above platform will be validated through a number of pilots in real life conditions, so that their added value in the policy making process can be assessed. In particular, in the above pilots it will be examined to what extent the platform is user-friendly and also functions correctly and effectively, to what extent this methodology can provide the abovementioned benefits and value, and what improvements of them should be made. In general, further research is required concerning methodologies and tools for the application of crowdsourcing by government agencies for increasing the bottom-up dimensions of public policy formulation.

Acknowledgements

The research leading to these results has received funding from the European Union 7th Framework Programme under grant agreement n° 288513 (NOMAD - Policy Formulation and Validation through non moderated crowdsourcing). For more details see <http://www.nomad-project.eu/>

References

1. Pateman, C.: Participation and Democratic Theory. University Press, Cambridge (1970)
2. Barber, B.: Strong Democracy. University of California Press, Berkeley (1984)
3. Held, D.: Models of Participation. Polity Press, Cambridge (1987)
4. Rowe, G., Frewer, L.J.: Public Participation Methods: A Framework for Evaluation. *Science, Technology & Human Values* 25(1), 3-29 (2000)
5. Rowe, G., Frewer, L.J.: Evaluating Public-Participation Exercises: A Research Agenda. *Science, Technology, & Human Values* 29(4), 512-57 (2004)
6. Organization for Economic Co-operation & Development: Citizens as Partners – Information, Consultation and Public Participation in Policy-Making. OECD Publication Service, Paris (2001a)
7. Organization for Economic Co-operation & Development: Engaging Citizens in Policy-making: Information, Consultation and Public Participation - PUMA Policy Brief, OECD Publication Service, Paris (2001b).
8. Organization for Economic Co-operation & Development: Evaluating Public Participation in Policy Making. OECD Publication Service, Paris (2004).
9. Organization for Economic Co-operation & Development (OECD): Engaging Citizens Online for Better Policy-making- Policy Brief. OECD Publication Service, Paris (2003)
10. Organization for Economic Co-operation & Development (OECD): Promise and Problems of e-Democracy: Challenges of Online Citizen Engagement. OECD Publication Service, Paris (2004)
11. Timmers, P.: Agenda for eDemocracy – an EU perspective. European Commission (2007)
12. Saebo, O., Rose, J., Flak, L.S.: The shape of eParticipation: Characterizing an emerging research area. *Government Information Quarterly* 25, 400-428 (2008).
13. Loukis, E., Macintosh, A., Charalabidis, Y.: Editorial of the Special Issue on E-Participation in Southern Europe and the Balkans: Issues of democracy and participation via electronic media. *Journal of Balkan and Near East Studies* 13(1), 1-12 (2011).
14. Ferro, E., Molinari, F.: Making Sense of Gov 2.0 Strategies: No Citizens, No Party. In: Prosser, A. and Parycek, P. (eds) *Proceedings of EDEM 2009* (2009).

15. Chadwick, A.: Guest Editor's Introduction: The Internet and Politics in Flux. *Journal of Information Technology and Politics* 6(3-4), 195-196 (2009)
16. Honeycutt, C., Herring, S. C.: Beyond Microblogging: Conversation and Collaboration via Twitter. In: *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS)*, Waikoloa, Big Island, Hawaii (2009)
17. Agarwal, N., Lim, M., Wigand, R.: Finding her Master's Voice: the Power of Collective Action among Female Muslim Bloggers. In: *Proceedings of the 19th European Conference on Information Systems (ECIS) 2011*, Helsinki, Finland (2011)
18. Larsson, A., Moe, H.: Who Tweets? Tracking Microblogging Use in the 2010 Swedish Election Campaign. In: *Proceedings of the 19th European Conference on Information Systems (ECIS) 2011*, Helsinki, Finland (2011)
19. Osimo, D.: *Web 2.0 in Government: Why and How?* JRC Scientific and Technical Reports. European Commission, Joint Research Centre, Institute for Prospective Technological Studies (2008), retrieved online at: <http://ftp.jrc.es/EURdoc/JRC45269.pdf>
20. Tapscott, D., Williams, A. D., Herman, D.: *Government 2.0: Transforming Government and Governance for the Twenty First Century*. nGenera Corporation (2008)
21. Meijer, A., Thaens, M.: Alignment 2.0: Strategic use of new internet technologies in government. *Government Information Quarterly* 27, 113–121 (2010)
22. Charalabidis, Y., Loukis, E.: Transforming Government Agencies' Approach to eParticipation through Efficient Exploitation of Social Media. In: *Proceedings of the European Conference on Information Systems (ECIS) 2011*, Helsinki, Finland (2011)
23. United Nations: *e-Government Readiness Knowledge Base*, 2007
24. United Nations: *e-Government Survey 2008: From eGovernment to Connected Governance*, 2008
25. Santucci, D.: Studying e-petitions: State of the art and challenges. In: *ESF-LIU Conference on Electronic Democracy*, Vadstena, Sweden (2007)
26. Cruickshank P., Edelman N., Smith C.: Signing an e-Petition as a Transition from Lurking to Participation. In Tambouris E., Macintosh A., Glassey O. (Eds.) *Second IFIP International Conference on Electronic Participation – ePart 2010*, Lausanne, Switzerland (2010).
27. Panopoulou, E., Tambouris, E., Sanchez-Nielsen, E., Zotou, M., Tarabanis, K.: Learning from eParticipation initiatives of regional and local level authorities in Greece and Spain. *Journal of Balkan and Near Eastern Studies* 13(1), 13-35 (2011)
28. Medaglia, R.: Measuring the diffusion of eParticipation: a survey on Italian local government. *Information Polity. An International Journal of Government and Democracy in the Information Age* 12(4), 265-280 (2008)
29. Mau, B., Leonard, J. and The Institute Without Boundaries: *Massive Change*. Phaidon, New York (2004)
30. Lévy, P.: *Collective Intelligence: Mankind's Emerging World in Cyberspace*. Plenum, New York (1997).
31. Surowiecki, J.: *The Wisdom of Crowds: Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. Doubleday, New York (2004)
32. Brabham, D.C.: Crowdsourcing as a Model for Problem Solving An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies* 14(1), 75–90 (2008)
33. Howe, J.: The Rise of Crowdsourcing. *Wired* 14(6), (2006), retrieved from <http://www.wired.com/wired/archive/14.06/crowds.html>.
34. Brabham, D.C.: Crowdsourcing: A Model for Leveraging Online Communities. In: A. Delwiche and J. Henderson (Eds.) *The Routledge Handbook of Participative Cultures*, Routledge (2012)

35. Stewart, O., Huerta, J., Sader, M.: Designing crowdsourcing community for the enterprise. In: Proceedings of the ACM SIGKDD Workshop on Human Computation ACM New York, USA (2009).
36. Li, Z., Hongjuan, Z.: Research of crowdsourcing model based on case study. In Proceedings of 8th International Conference on Service Systems and Service Management (ICSSSM) (2011).
37. Brabham, D. C.: Moving the Crowd at Threadless – Motivations for Participation in a Crowdsourcing Application. *Information, Communication & Society* 13(8), 1122 – 1145 (2009)
38. Nam, T.: Suggesting frameworks of citizen-sourcing via Government 2.0. *Government Information Quarterly* 29, 12–20 (2012)
39. Hilgers, D., Ihl, C.: Citizensourcing: Applying the concept of open innovation to the public sector. *The International Journal of Public Participation* 4(1), 67–88 (2010)
40. Bovaird, T.: Beyond engagement and participation: User and community coproduction of public services. *Public Administration Review* 67(5), 846–860 (2007)
41. Torres, L. H.: Citizen sourcing in the public interest. *Knowledge Management for Development Journal* 3(1), 134–145 (2007)
42. Lukensmeyer, C. J., Torres, L. H.: Citizensourcing: Citizen participation in a networked nation. In: Yang, K., Bergrud, E. (Eds.) *Civic engagement in a network society*, pp. 207–233. Information Age Publishing, Charlotte, NC (2008).
43. Chun, S. A., Shulman, S., Sandoval, R., Hovy, E.: Government 2.0: Making connections between citizens, data and government. *Information Polity* 15(1/2), 1–9 (2010)
44. ELEON Ontology Authoring and Enrichment Tool, <http://www.iit.demokritos.gr/~eleon>
45. Bilidas, D., Theologou, M., Karkaletsis, V.: Enriching OWL Ontologies with Linguistic and User-related Annotations: the ELEON system. In: Proceedings of IEEE International Conference on Tools with Artificial Intelligence (ICTAI-07), Patras, Greece (2007)
46. Konstantopoulos, S., Karkaletsis, V., Vogiatzis, V., and Bilidas, D.: Authoring Semantic and Linguistic Knowledge for the Dynamic Generation of Personalized Descriptions. In: C. Sporleder et al. (eds), *Language Technology for Cultural Heritage*. Springer Verlag, Berlin/Heidelberg (2011)
47. ELLOGON Linguistic Analysis Platform, <http://www.ellogon.org>
48. Petasis, G., Karkaletsis, V., Paliouras, G., Androutsopoulos, I., Spyropoulos, C. D.: Ellogon: A New Text Engineering Platform. In: Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002), Las Palmas, Canary Islands, Spain, (2002)
49. Rentoumi, V., Petrakis, S., Karkaletsis, V., Vouros, G.: A Collaborative System for Sentiment Analysis. In: Proceedings of SETN 2010 (2010)
50. Rentoumi, V., Giannakopoulos, G., Karkaletsis, V., Vouros, G.: Sentiment analysis of figurative language using a word sense disambiguation approach. In: Proceedings of RANLP 2009 (2009)
54. Wong, P. C., Thomas J.: Visual analytics. *IEEE Computer Graphics and Applications*, 24(5), 20–21 (2004)
55. Thomas, J., Cook, K. (Editors). *Illuminating the Path: Research and Development Agenda for Visual Analytics*. IEEE Press, Los Alamitos, CA (2005)
56. Keim, D. A., Kohlhammer, J., Ellis, G. P., Mansmann, F.: *Mastering The Information Age – Solving Problems with Visual Analytics*. Eurographics, Goslar Germany (2010).