



Passive crowdsourcing in government using social media

Yannis Charalabidis, Euripidis N. Loukis and
Aggeliki Androutsopoulou

*Department of Information and Communication Systems Engineering,
University of the Aegean, Samos, Greece*

Vangelis Karkaletsis
*National Center for Scientific Research "Demokritos", Athens, Greece, and
Innovation Lab, Athens Technology Center, Athens, Greece*

Passive
crowdsourcing

283

Received 17 September 2013

Revised 31 December 2013

Accepted 21 February 2014

Abstract

Purpose – The purpose of this study is to develop a novel approach to e-participation, which is based on “passive crowdsourcing” by government agencies, exploiting the extensive political content continuously created in numerous Web 2.0 social media (e.g. political blogs and microblogs, news sharing sites and online forums) by citizens without government stimulation, to understand better their needs, issues, opinions, proposals and arguments concerning a particular domain of government activity or public policy.

Design/methodology/approach – This approach is developed and elaborated through cooperation with potential users experienced in the design of public policies from three countries (Austria, Greece and the UK), using a combination of quantitative and qualitative techniques: co-operative development of application scenarios, questionnaire surveys, focus groups and workshops and, finally, in-depth interviews.

Findings – A process model for the application of the proposed passive crowdsourcing approach has been developed, which is quite different from the one of the usual active crowdsourcing. Based on it, the functional architecture of the required supporting information and communication technologies (ICT) infrastructure has been formulated, and then its technological architecture has been designed, addressing the conflicting requirements: low response time and, at the same time, provision of sufficiently “fresh” content for policymakers.

Practical implications – Taking into account that traditionally government agencies monitor what the press writes about them, our research provides a basis for extending efficiently these activities in the new electronic media world (e.g. newspapers websites, blogs and microblogs, online forums, etc.) to understand better the needs, issues, opinions, arguments and proposals raised by the society with respect to important domains of government activity and public policies.

Social implications – The proposed approach provides a new channel for the “voice” of the society to be directly communicated to the government so that the latter can design its policies and activities based on the social needs and realities and not on oversimplified models and stereotypes.

Originality/value – Our paper proposes a novel approach to e-participation, which exploits the Web 2.0 social media – but in a quite different way from previous approaches – for conducting “passive crowdsourcing”, and elaborates it: it develops an application process model for it and also an ICT infrastructure for supporting it, which are quite different from the ones of the existing “active crowdsourcing” approaches.

Keywords Web 2.0, Public policy, Social media, E-participation, Crowdsourcing

Paper type Research paper



1. Introduction

Government agencies are increasingly interested in exploiting the capabilities provided by the information and communication technologies (ICT), and especially the Internet, for engaging the citizens in their public policymaking processes, and this has lead to the gradual development of the e-participation domain (OECD, 2004a; Saebo *et al.*, 2008; Loukis *et al.*, 2011). This offers to governments highly important capabilities to collect valuable information, knowledge, ideas and opinions from large number of citizens on the complex problems and needs of modern societies, which can be quite useful for designing better quality policies, and, at the same time, to build trust and gain more acceptance of their policies.

The first generation of e-participation was based on the development and operation of official government e-participation spaces, in which electronic discussions of various levels of structure were conducted (Loukis and Wimmer, 2012). It was highly government-controlled, as the owner government agencies defined and controlled the topics and rules of these electronic discussions taking place there. The use of these official e-participation spaces by the citizens has been, in general, limited (Chadwick, 2009; Ferro and Molinari, 2010). For this reason, the emergence of the Web 2.0 social media and their high popularity lead to the emergence of a second generation of e-participation, in which government agencies are posting policy-related content (e.g. text and video) to their accounts in several social media, and then analyze various types of citizens' interactions with it (e.g. views, likes and comments) (Margo, 2012; Bertot *et al.*, 2012; Snead, 2013). Government agencies initially adopted simpler forms of social media exploitation, which consisted of operating manually accounts in some social media, posting relevant content to them (e.g. concerning current and future policies and activities) manually and then reading citizens' interactions with it to draw conclusions from them. However, gradually they start experimenting with more sophisticated forms of social media exploitation, which are characterized by higher both technological and political complexity (Charalabidis and Loukis, 2012; Kokkinakos *et al.*, 2012; Wandhöfer *et al.*, 2012). In particular, from a technological perspective, they tend to use the application programming interfaces (API) of the social media for posting policy-related content to them, and then for the automated retrieval of citizens' interactions with this content (e.g. views, likes, comments), and/or other relevant content (instead of manual operations); finally they make various types of advanced automated processing of these interactions/content to support drawing conclusions from them. From a political perspective, they tend to use numerous social media as communication channels with a wider range of citizens' groups, and especially with the ones not traditionally involved in politics, to explore a wider range of opinions, needs and concerns with respect to complex and critical societal problems and public policies. This second generation of e-participation is less government-controlled: governments continue defining and controlling the discussion topics to some extent (through making postings and initiating discussions on the topics they choose), but they do not control the rules of the discussion (which are the ones defined by the social media).

This study develops a novel approach to e-participation, which can lead to a third generation of it (that can be however combined with the first two), based also on the exploitation of the social media by government agencies, but in a totally different manner, for conducting "passive crowdsourcing". It can be defined as the search by government agencies for content on a public policy under formulation, which has been

created in a large number of predefined Web 2.0 social media sources (e.g. blogs and microblogs, news sharing sites, online forums, etc.) by citizens freely, without any initiation, stimulation or moderation through government postings, and the extraction from it of relevant knowledge, ideas and opinions. Therefore, in this approach, the government is not active in conducting crowdsourcing (because it is in the previous generations of e-participation, posing to citizens particular questions or discussion topics), but it remains passive (not posing particular questions or discussion topics, but just “listening” to what citizens discuss, and analyzing the content they freely produce). In particular, this content is retrieved using the API of the predefined sources, and then undergoes advanced linguistic processing to extract from it needs, issues, opinions, arguments and proposals raised by citizens on a particular domain of government activity or 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 to support governments in understanding better the needs and opinions of citizens, and extract from them more knowledge and innovative ideas on the problems of modern societies to create better and more socially rooted policies. It is even less government-controlled than the abovementioned second generation of e-participation, and more citizens driven and controlled, as in this approach 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). It allows governments to move beyond their own official e-participation spaces (first generation of e-participation) and social media accounts (first generation of e-participation), and explore a wider range of Web 2.0 sources. However, it has some notable technological and political complexities and difficulties. Technologically, it uses the API of numerous social media to retrieve content from them (so it relies on platforms controlled by several other players, which are “fluid”, changing quite often – see [Charalabidis and Loukis \(2012\)](#)), and also it uses advanced linguistic processing techniques (that require extensive linguistic resources) for extracting opinions, arguments, issues and proposals on a particular domain or policy. Politically, it extracts content from a bigger variety of Web 2.0 sources (e.g. political blogs and micro-blogs of quite different orientations and ideologies), corresponding to much wider range of citizens’ groups and political approaches, so it explores (and therefore has to understand and manage) a wider and more heterogeneous range of opinions, needs and concerns with respect to important societal problems and public policies (with all the inherent problems that this creates).

In particular, this paper makes a threefold contribution: it outlines the basic idea of the aforementioned novel approach to e-participation, based on passive crowdsourcing, and also describes its theoretical foundations; it develops and elaborates this idea, initially by developing a process model for its application (meant as a model of process to be followed by government agencies for the practical application of it, consisting of a sequence of specific activities to be executed); and then by developing the functional architecture of an ICT platform supporting the application of this approach, and its technological architecture addressing the existing conflicting requirements (low response time and, at the same time, provision of sufficiently “fresh” content for policymakers).

All these have been developed through cooperation with potential users experienced in the design of public policies, using a combination of quantitative and qualitative techniques. This paper is – to the best of our knowledge – the first paper proposing this concept of passive crowdsourcing in government, and, at the same time, a process and an ICT infrastructure for its practical application.

The research presented in this paper has been conducted as part of the research project NOMAD (“policy formulation and validation through non-moderated crowdsourcing” – for more details see www.nomad-project.eu/), which is partially funded by the “ICT for governance and policy modeling” research initiative of the European Commission. We expect that the results of this research will be useful for both researchers in the domains of public administration, e-government/e-participation, ICT and political and social sciences, and also for government practitioners and consulting firms interested in the exploitation of Web 2.0 social media in the public sector.

The paper is structured in five sections. In Section 2 the theoretical foundations of the proposed approach are outlined, while in Section 3 the research methodology is described. Then in Section 4 a description of this approach is provided, followed by a process model for its application in Section 5. Next, the functional and technological architecture of an ICT platform supporting its application are described in Sections 6 and 7, respectively. Finally, Section 8 summarizes the conclusions and proposes future research directions.

2. Theoretical background

In this section, the theoretical foundations of the proposed approach are outlined: a) bottom-up public policy formulation; and b) crowdsourcing.

2.1 Bottom-up public policy formulation

The formulation of public policies for addressing problems and needs of societies has been traditionally a top-down process. Governments and citizens' representatives traditionally had the main role in identifying problems and needs of society, and in designing public policies and services for addressing them. The participatory democracy and public participation ideas gave some role in the public policies formulation processes to the citizens, who are initially provided some information on government activities and policies, and are asked to participate in relevant consultations and contribute through comments and opinions on issues and questions defined by government.

According to OECD, 2001 and 2004a the development of public participation includes three stages. The first two of them are “Information Provision” (establishment an “one-way relation” between government and citizens, in which government produces and delivers information to be used by citizens), and “Consultation” (establishment of an asymmetric “two-way relation” between the government and the citizens, in which citizens provide opinions and views 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 OECD reports (which are based on the experience of the numerous OECD member countries), they might be followed 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, proposing new policy options and discussion topics, in addition to the ones proposed by

government, and 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, offering significant benefits: it can provide new fresh ideas concerning needs and problems of society, ways of addressing them, advantages and disadvantages of the latter or even can lead to the identification of changes and new trends at the cultural, social and economic level that require government action.

The development of e-participation, according to the corresponding OECD reports on e-participation ([OECD, 2003](#) and [2004b](#)), follows a similar trajectory. Its first two stages have a strong top-down orientation, aiming to support through ICT, and especially 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 that 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), and submit their own proposals and policy. Similarly, [United Nations \(2008\)](#) suggest a three-staged plan for the development of e-participation, which includes:

- (1) increasing e-information provision to citizens for decision-making;
- (2) enhancing e-consultation for deliberative and participatory processes (with both these steps having mainly top-down orientation); and
- (3) supporting e-decision-making by increasing the input of citizens in decision-making (with this step having a bottom-up orientation).

However, this bottom-up dimension of e-participation 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 existing or proposed new government actions (e.g. services, programs and regulations) to the government or parliament, and solicit support and signatures) ([Santucci, 2007](#); [Cruickshank et al., 2010](#)), and electronic discussion groups supporting the development of new policy options and the deliberation on them. Also, the analysis of e-participation initiatives and practices of many countries has shown that they focus mainly on e-information provision and e-consultation (i.e. on the 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) ([Medaglia, 2008](#); [Panopoulou et al., 2012](#)). The approach proposed in this paper can support and promote bottom-up public policy formulation, so it contributes to filling the research and practice gaps.

2.2 Crowdsourcing

The capability of a large network of people, termed as “crowd”, networked through Web technologies, to perform difficult problem solving and design activities, which were previously performed exclusively by professionals, 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" ([Levy, 1997](#)). These ideas lead to the emergence of crowdsourcing, which is defined 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" ([Howe, 2006](#)) or as "a new Web-based business model that harnesses the creative solutions of a distributed network of individuals" to exploit "collective wisdom" and mine fresh ideas from large numbers of individuals ([Brabham, 2008](#)). While the use of the collective intelligence of a large group of people to help solve problems is an approach that has been used for long time ([Surowiecki, 2004](#), [Howe 2008](#)), it is only recently that businesses and other entities have started increasingly turning to crowdsourcing as a means of obtaining external expertise, accessing the collective wisdom and creativity resident in the virtual crowd. [Brabham \(2008\)](#) analyzes several cases of crowd wisdom at work, which result in successful solutions emerging from a large body of solvers and concludes that "under the right circumstances, groups are remarkably intelligent, and are often smarter than the smartest people in them", due to the diversity of opinion, independence, decentralization and aggregation that characterize such a crowd.

Crowdsourcing 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. 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 ([Surowiecki, 2004](#), [Howe, 2006, 2008](#), [Brabham, 2008, 2012](#)). 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.

There is a growing information systems (IS)- and business-related research on crowdsourcing (it is reviewed by [Pedersen et al. \(2013\)](#) and [Tarrell et al. \(2013\)](#)). Initially, it focused on analyzing successful case studies, while later it started generalizing, based on the experience of multiple cases, to identify patterns and trends in this area and also to develop effective crowdsourcing practices. A typical example in this direction is the study by [Brabham \(2012\)](#), which, based on the analysis of several case studies, identifies four dominant crowdsourcing approaches:

- (1) the knowledge discovery and management approach (i.e. an organization tasks crowd with finding and reporting information and knowledge on a particular topic);
- (2) the broadcast search approach (i.e. an organization tries to find somebody who has experience with solving a rather narrow and rare empirical problem);
- (3) the peer-vetted creative production approach (i.e. an organization tasks crowd with creating and selecting creative ideas); and
- (4) the distributed human intelligence tasking (i.e. an organization tasks crowd with analyzing large amounts of information).

Hetmank (2013), based on a review of crowdsourcing literature, identifies a basic process model of it, which consists of ten activities: define task, set time period, state reward, recruit participants, assign tasks, accept crowd contributions, combine submissions, select solution, evaluate submissions and grant rewards. Also, he identifies a basic pattern with respect to crowdsourcing IS structure, which includes four main components performing user management (providing capabilities for user registration, user evaluation, user group formation and coordination), task management (providing capabilities for task design and assignment), contribution management (providing capabilities for contributions evaluation and selection) and workflow management (providing capabilities for defining and managing workflows). Furthermore, there are some studies that attempt to generalize the experience gained from successful applications of crowdsourcing ideas to develop effective practices for crowdsourcing and motivating individuals to participate (Brabham, 2009; Stewart *et al.*, 2009).

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 perform “citizen-sourcing” for collecting information on citizens’ needs and for the solution of difficult problems (Torres, 2007; Chun *et al.*, 2010; Hilgers and Ihl, 2010; Nam, 2012). According to Nam (2012), traditionally government agencies provide services to citizens, who consume them without questioning about them or taking part in decisions that lead to their design and provision; social media can drive and facilitate new paradigms of government services design based on citizen-sourcing, in which citizens’ roles change and become more active, so that government becomes a consumer to whom citizens provide information and knowledge (concerning their needs and the required public services’ content, form and provision).

This paper makes a contribution in this direction, developing a novel e-participation approach based on the application of crowdsourcing ideas by government agencies in a “passive” manner. It is based on search for content on a domain of government activity or a public policy under formulation, which has been created in numerous predefined 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, so it constitutes a “passive crowdsourcing” (in contrast to the “active crowdsourcing” that characterizes the previous generations of e-participation, in which government agencies are more active, posing to the citizens particular questions or discussion topics). Based on the advanced processing and analysis of this content, some conclusions on the needs, issues, opinions, proposals and arguments of citizens on this domain of government activity or public policy under formulation can be drawn. This passive crowdsourcing can be combined with active crowdsourcing, as the former can provide a basis for the better and more efficient application of the latter; it can provide a basis for a subsequent “open call” (i.e. active crowdsourcing), having the form of postings to various social media that solicit citizens’ interactions (e.g. likes/dislikes) or comments (e.g. remarks, opinions, positions and proposals) on more specific topics and questions, which are based on the findings of the passive crowdsourcing.

3. Research methodology

To develop and elaborate this government passive crowdsourcing approach, through the development of an application process model and supporting ICT infrastructure, we

cooperated with potential users experienced in the design of public policies in the countries of the three user partners of the NOMAD project (Austria, Greece and UK), using a combination of different quantitative and qualitative techniques. In particular, the research methodology we adopted for this purpose consisted of the following steps:

- (1) Initially four application scenarios of this approach were developed by the user partners of the NOMAD project (Greek Parliament, Austrian Parliament, European Academy of Allergy and Clinical Immunology). Each application scenario constitutes a detailed realistic example of how this passive crowdsourcing concept could be applied for supporting the formulation of a particular public policy, and describes how various types of users involved in this might use the supporting ICT platform. The topics of these application scenarios were selected so that, on one hand, they reflect current debates and interests of the user partners, and, on the other hand, they cover quite different and diverse policy domains. In particular, the four application scenarios were:
 - “Effective changes in Greek Labor Law, and their direct impact on social security and employment” (Greek Parliament);
 - “European policies on the treatment of allergy and declaration of immunotherapy as an effective treatment for the disease” (European Academy of Allergy and Clinical Immunology);
 - “The Greek strategy for energy planning” (Greek Parliament); and
 - “Information management and involvement of civil society within the legislative process” (Austrian Parliament).

These application scenarios will be used in the next phases of the project as basis for the design of corresponding pilot applications of the proposed passive crowdsourcing approach, which will be implemented by the corresponding user partners (using the process model and the ICT platform described in the following sections) and then evaluated.

- (2) A questionnaire was distributed electronically to a sample population of potential users. The questionnaire was structured in four sections of questions referring, respectively, to:
 - respondent's personal information;
 - general citizens' participation information (in his/her organization);
 - current use of social media in policymaking processes; and
 - specific requirements and NOMAD general assessment.

The responses were analyzed statistically (the results can be found in the Deliverable D2.2 “Report on User Requirements” of the NOMAD project).
- (3) Organization of focus groups and workshops with the participation of potential users. This allowed in-depth discussion among people experienced in the design of public policies, with different backgrounds and mentalities, about this new concept, ways and processes of its practical application, required ICT functionalities and at the same time possible problems and barriers.
- (4) Organization of in-depth interviews based of a series of fixed questions concerning attitudes towards this new concept, and ideas concerning its particular application.

-
- (5) A review of similar systems, which offer at least a part of the functionalities that our ICT platform is going to offer.

In the above activities participated a mix of different types of actors involved in public policymaking: government decision-makers, policy advisors, governmental agencies employees of various hierarchical levels, representatives of non-governmental organizations and trade unions, communication as well as technology experts. This cooperative approach allowed us to develop and elaborate this idea of a passive crowdsourcing for supporting public policymaking, specify realistically how it might practically work and also collect attitudes, needs and requirements from a wide range of potential users and understand their motives and expectations.

Based on the above application scenarios developed by the user partners of the project (Step 1), the remarks on them expressed by the participants in the focus group discussions, workshops (Step 3), in-depth interviews (Step 4) and also the results from processing the responses to the Section 4 of the questionnaire (Step 2), the application process model presented in Section 5 was developed. Then based on it, and taking also into account the functionality requirements obtained from Steps 2, 3 and 4, and also the functionalities offered by similar systems (Step 5), the functional architecture of the supporting ICT platform presented in Section 6 was developed.

4. Approach description

The proposed approach enables government policymakers to exploit the extensive political content created in multiple Web 2.0 sources (e.g. blogs and microblogs, news sharing sites, online forums) by citizens on many policy-related topics, for gaining a first understanding of the needs, issues, opinions, proposals and arguments of them with respect to a particular domain of government activity (e.g. education) or a particular government policy (e.g. a proposed new education policy); this can be used then as a guide for proceeding to more active communicative actions, such as soliciting citizens' feedback and contributions on more specific questions, positions or proposals (e.g. defined through postings in government accounts in various social media, or in official government e-participation websites). This can be done through the four stages cycle shown in [Figure 1](#), which enables government policymakers to effectively LISTEN and monitor what citizens say in social media concerning a domain of government activity or a government policy, ANALYZE those conversations and get the main stakeholders' positions and opinions, RECEIVE the results of this analysis displayed properly for an effective use and exploitation, and finally ACT on this information by proceeding to relevant postings to various social media or official government e-participation websites. This cycle can be performed several times, until a sufficiently complete and detailed picture is formulated about citizens' needs, issues,

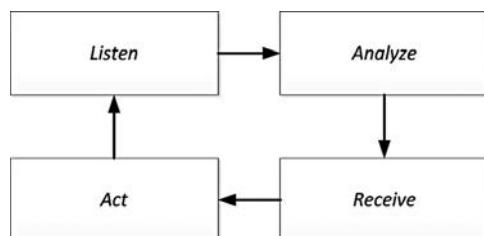


Figure 1.
The four stages of the proposed approach

opinions, proposals and arguments, in any of the stages of the policymaking cycle (agenda setting, policy analysis, policy formulation, policy implementation, monitoring and evaluation, according to [OECD, 2003, 2004a, 2004b](#)).

In particular, the first stage, called LISTEN, includes listening and monitoring what people say, what their needs, opinions, concerns, positions and proposals are on a topic (e.g. a domain of government activity or a public policy under formulation) defined by the policymaker. For this purpose a “focused crawler” is used, that is, a program which browses the Web in an automated, organized fashion, and retrieves solely content that is relevant to this topic from a variety of predefined sources (that often host discussions on this topic, selected by the user (government policymaker) based on his/her previous knowledge and experience), such as blogs, micro-blogging sites, social networking platforms, discussion forums, comment-enabled news websites, corporate or video sharing sites etc.

The second stage, called ANALYZE, includes advanced processing and analysis of the retrieved content, from which are identified relevant citizens’ needs, opinions, concerns and proposals, and other information hidden within the citizens’ conversations. As the majority of this content is in textual form, the processes of this stage will mainly make use of linguistic processing techniques (for a review of them see [Maragoudakis et al. \(2011\)](#)). In particular, each content unit retrieved by the crawler will go through a series of automated processing steps:

- *Language detection*, which will recognize the language used in it.
- *Opinion and argument extraction*, using appropriate semantic similarity measures and inference mechanisms that will allow the identification of elements of the analyzed content that are pertinent to the particular domain or policy.
- *Sentiment analysis*, using smart sentiment classifiers that recognize the polarity (positive, neutral or negative) of the elements identified above.
- *Argument summarization*, using appropriate algorithms for generating qualitative information about opposing arguments, in the form of anonymity-preserving and automatically-generated summaries.

The third stage, receive, aims to present to the end user (policymaker) the knowledge acquired from the previous stages in a complete, coherent and usable manner. The platform will provide an aggregated view of the results of the processing, their polarity, their association with policy concepts and statements, and also statistical indications of their significance and impact. With the use of visual analytics ([Wong and Thomas, 2004; Thomas and Cook, 2005; Keim et al., 2010](#)) all these results will be visualized, so that policymakers can easily understand them with minimal cognitive effort (e.g. in a familiar word cloud form), which is quite important due to the high information overload the policymakers usually experience.

The three stages perform a “passive crowdsourcing”, in which the policymaker collects and analyses content that has been created without his/her stimulation, intervention or moderation. In the fourth stage, ACT, the policymaker uses the needs, issues, opinions, proposals and arguments of citizens identified in the previous stage to formulate more specific questions, positions or proposals about the particular policy and then solicit citizens’ feedback and contributions on them. This is going to be achieved through “active crowdsourcing”, i.e. by making relevant stimulating postings to various social media (e.g. blogs, Twitter, Facebook, YouTube, etc.), and also to official government e-participation

websites; this can be further improved if there are links between the former and the latter postings, so that citizens participating in these social media discussions can easily move and participate in the official e-participation websites' discussions and vice versa.

Therefore, the proposed “passive crowdsourcing” approach does not replace the previous “active crowdsourcing” approaches, but can be combined with them to support them and increase their efficiency and effectiveness. In general, this third generation of e-participation does not aim to replace the first one (based on the official government e-participation websites) and the second one (based on social media), which have been outlined in the Introduction, but can be perfectly combined with them, as shown schematically in [Figure 2](#), leading to a powerful combination of e-participation actions.

5. Application process model

In close cooperation with potential users experienced in the design of public policies, as explained in more detail in Section 3 describing our research methodology, we examined how the passive crowdsourcing concept can be practically applied, and what is the most appropriate process model for its application. There was wide agreement that since the domains of government activity and the public policies for them are quite complex and multi-dimensional entities, it is not possible to search for content on them using just a small number of keywords. So it was concluded that the best solution for addressing this complexity is to develop a model of the specific domain, for which a policy is intended, which will consist of the main terms of it and the relations among them (a kind of “structured thesaurus” of this domain). An example of such a domain model for the energy domain, which has been developed based on the documents of the “Greek Strategy for Energy Planning”, is shown below in [Figure 3](#).

Based on such a domain model, we can then build a policy model, by adding to the nodes of the former: the “policy statements” (i.e. the specific policy objectives and actions/interventions that a policy includes) and also positive and negative arguments in favour or against them, respectively. An example of such a policy model for the energy domain is shown in [Figure 4](#) (including three policy objectives, one concerning the whole national energy planning, and two concerning the renewable energy sources, six positive arguments and nine negative ones).

These two models (domain and policy) will be used for searching for and retrieving relevant content concerning the main terms of a domain, or the policy statements and the arguments of a policy. This search has to be performed at regular time intervals to keep the retrieved content updated, and the results will be stored in a content database, and then undergo the advanced processing mentioned in the previous section (ANALYZE stage).

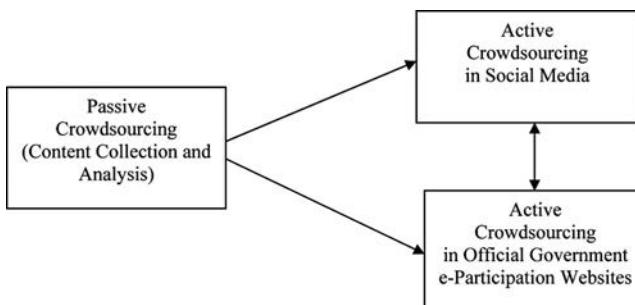
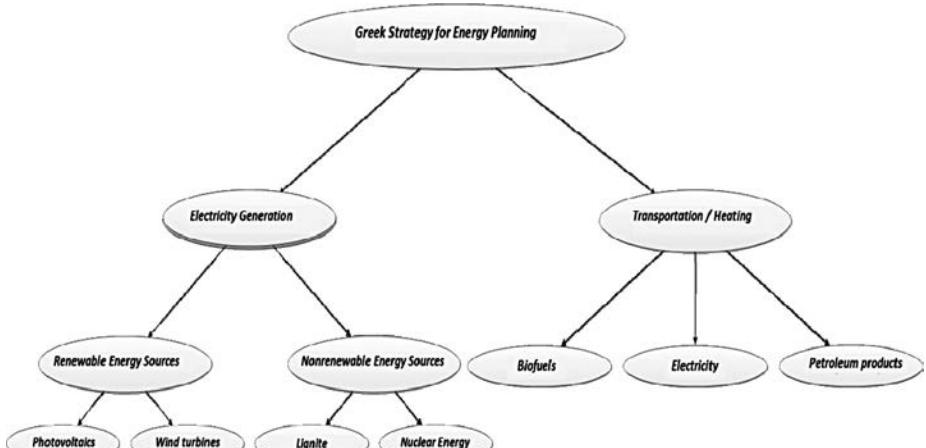


Figure 2.
Combination of the proposed approach with previous e-participation approaches

Figure 3.

Energy domain model based on the “Greek Strategy for Energy Planning” documents



The authorized policymakers will have the capability at any time to explore the results of this advanced processing and view a visualization of the most frequently mentioned terms – topics with respect to a particular domain or policy model (e.g. in a tag cloud form). Also, most of the potential users we interviewed mentioned that it is important to view citizens' sentiment with respect to these frequently mentioned terms – topics (whether they regard each of them as positive, negative or neutral) or even with respect to the individual policy statements and arguments of a policy model. However, our interviewees noted that all the above (i.e. frequently mentioned terms – topics and sentiments) may differ significantly between different citizens groups (e.g. between age, gender, education and region groups), so policymakers should have the capability to view them for particular citizens' groups or to view comparisons between different citizens' groups. Furthermore, because public stance changes rapidly, it was mentioned that policymakers should have the capability to view all the information for particular user-defined time periods, or to compare between different time periods, while future forecasts of them would be quite useful.

Therefore, a model of the process to be followed by government agencies for the practical application of this passive crowdsourcing approach consists of the following nine activities:

- (1) Development of a domain model.
- (2) Development of a policy model.
- (3) Definition of Web 2.0 content sources.
- (4) Search of these content sources at regular time intervals.
- (5) Process retrieved content.
- (6) Policymaker views polarized tag clouds with the most frequently mentioned terms – topics with respect to a particular domain or policy model and the corresponding sentiments for a predefined time period.
- (7) Policymaker views the sentiments with respect to the individual policy statements and arguments of a policy model.

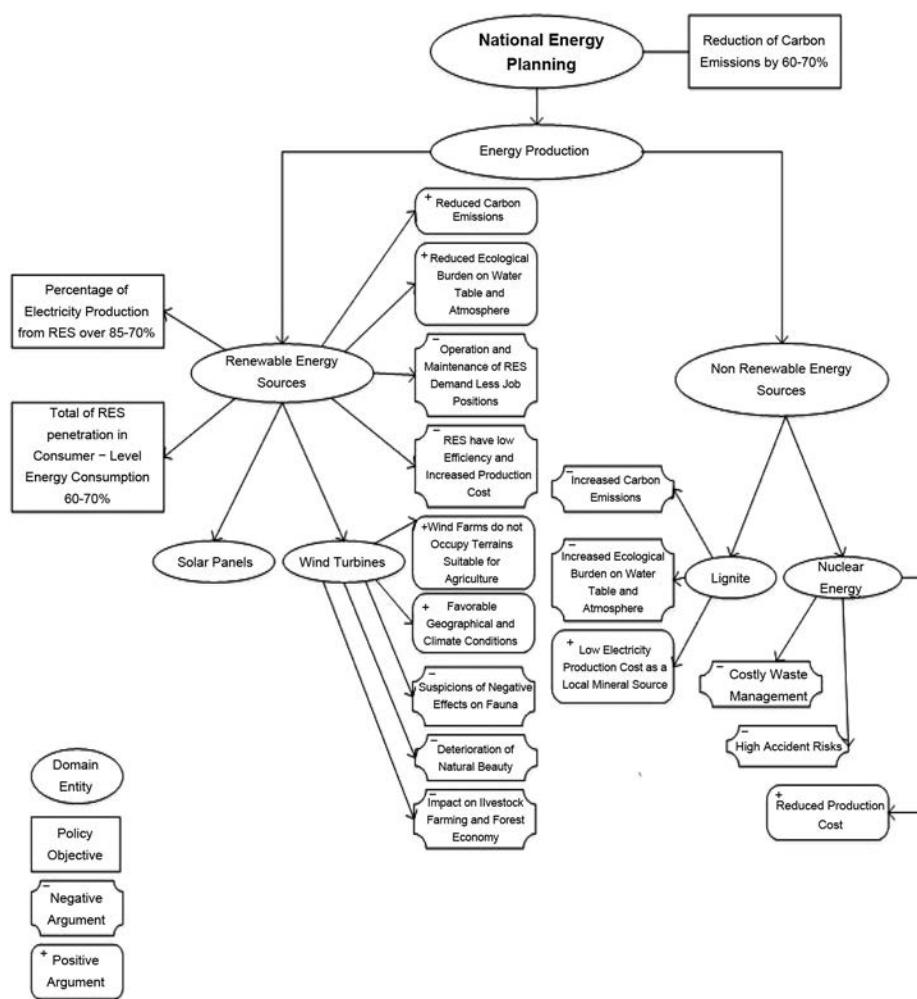


Figure 4.
Energy policy model
based on the energy
domain model (including
policy statements and
arguments)

- (8) Policymaker views the above for particular citizens' groups, and then makes comparisons between different citizens' groups or with other time periods.
- (9) Policymaker views short-term future forecasts of the above.

We remark that this application process model of the passive crowdsourcing is quite different from the one of the active crowdsourcing identified by Hetmank (2013), which has been outlined in Section 2.2.

Finally, we identified four roles which are required for this application process model:

- (1) *Domain models author*: This role will create domain models and also modify existing ones.

- (2) *Policy models author*: This role will create policy models based on existing domain models (i.e. add to their nodes policy statements and argumentations) and also modify existing ones.
- (3) *End user/policymaker*: This role will view the results of processing the content retrieved from the Web 2.0 sources in all the abovementioned forms.
- (4) *Platform administrator*: This role will have full access to all platform functionalities, monitor platform operation, manage the set of users accessing the platform and their access rights to the offered services and functionalities.

6. Platform functional architecture

Based on the application process model, we proceeded to the design of the functional architecture of an ICT platform that supports the application of this approach. In particular, we defined, in more detail, the functionality to be provided to each of the above four roles:

- (1) Domain models author:
- Creation of new domain models (i.e. definition of main terms of the domain and of the relations among them).
 - Modification of existing domain models.
 - Import of external domain models (e.g. having the form of ontology files in OWL).
 - Export of domain models (e.g. in the form of ontology files in OWL).
- (2) Policy models author:
- Access to domain models.
 - Creation of new policy models (using existing domain models by adding to their nodes policy statements and arguments).
 - Modification of existing policy models.
 - Import of external policy models (e.g. having the form of ontology files in OWL).
 - Export of policy models (e.g. in the form of ontology files in OWL).
- (3) End user/Policymaker:
- View the most frequently mentioned terms – topics with respect to a particular domain or policy model for a predefined time period, citizens' group and sources subset (see [Figure 5](#) for a first design of the corresponding screen).
 - View sentiment for these terms – topics.
 - View sentiment for each policy statement and argument of a particular model.
 - View differentiations of the above over time.
 - View differentiations of the above across citizens' groups.
 - View differentiations of the above across sources subsets.
 - View short-term future projections of the above.
- (4) Platform administrator:
- Users and roles management.

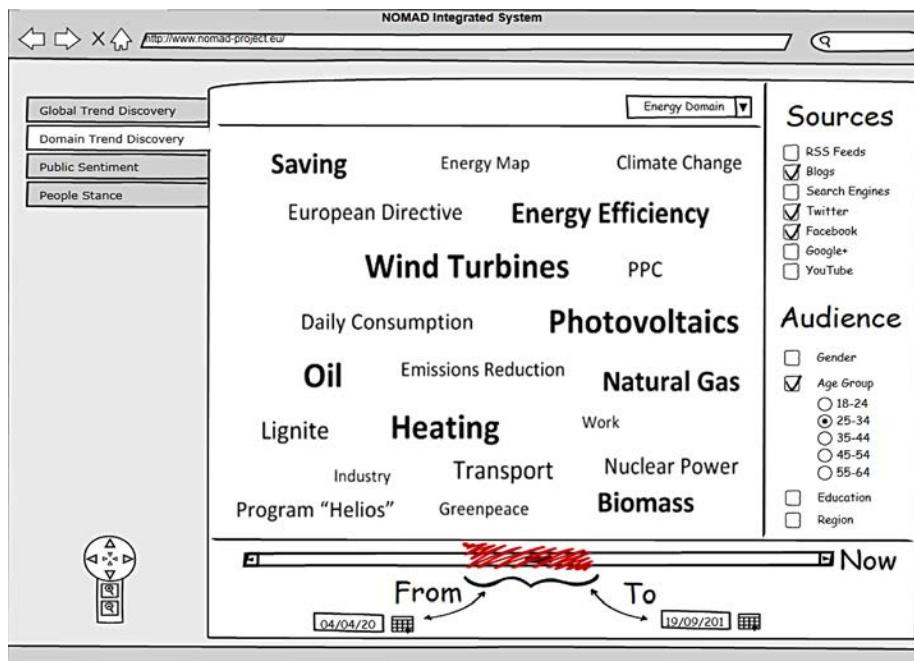


Figure 5.
View of the most frequently mentioned terms – topics with respect to a particular domain for a predefined time period, citizens' group and sources subset

- Domain and policy roles management.
- Monitoring and administration of all platform services.

7. Platform technological architecture

Based on the functional architecture of the platform, its technological architecture was designed. The objective of this design was to provide this functionality with an acceptable response time. Because this could not be achieved through online retrieval of content from the multiple and heterogeneous predefined sources (e.g. numerous blogs, news websites, Facebook, You Tube and Twitter accounts) and processing of it at the time a user initiates a search, the only solution was to perform a scan of the predefined sources at some regular time intervals (e.g. every six hours) to retrieve new content to store it in a database and then process it. Whenever the user performs a search, the results will be produced in a very short time, using this database. This separation between sources scanning and content processing, on one hand, and users searches processing, on the other, allows a low response time, and at the same time, sufficiently “fresh” content for policymakers (i.e. allows addressing these two conflicting requirements).

The aforementioned design leads to a three-layered technological architecture of the platform, which consists of a storage layer, a processing layer and a presentation layer, as shown in [Figure 6](#). Each of them includes a number of components, performing different tasks, which act as services coordinated by an orchestration component.

In particular, the *data storage layer* includes the repositories where the raw and processed content is stored:

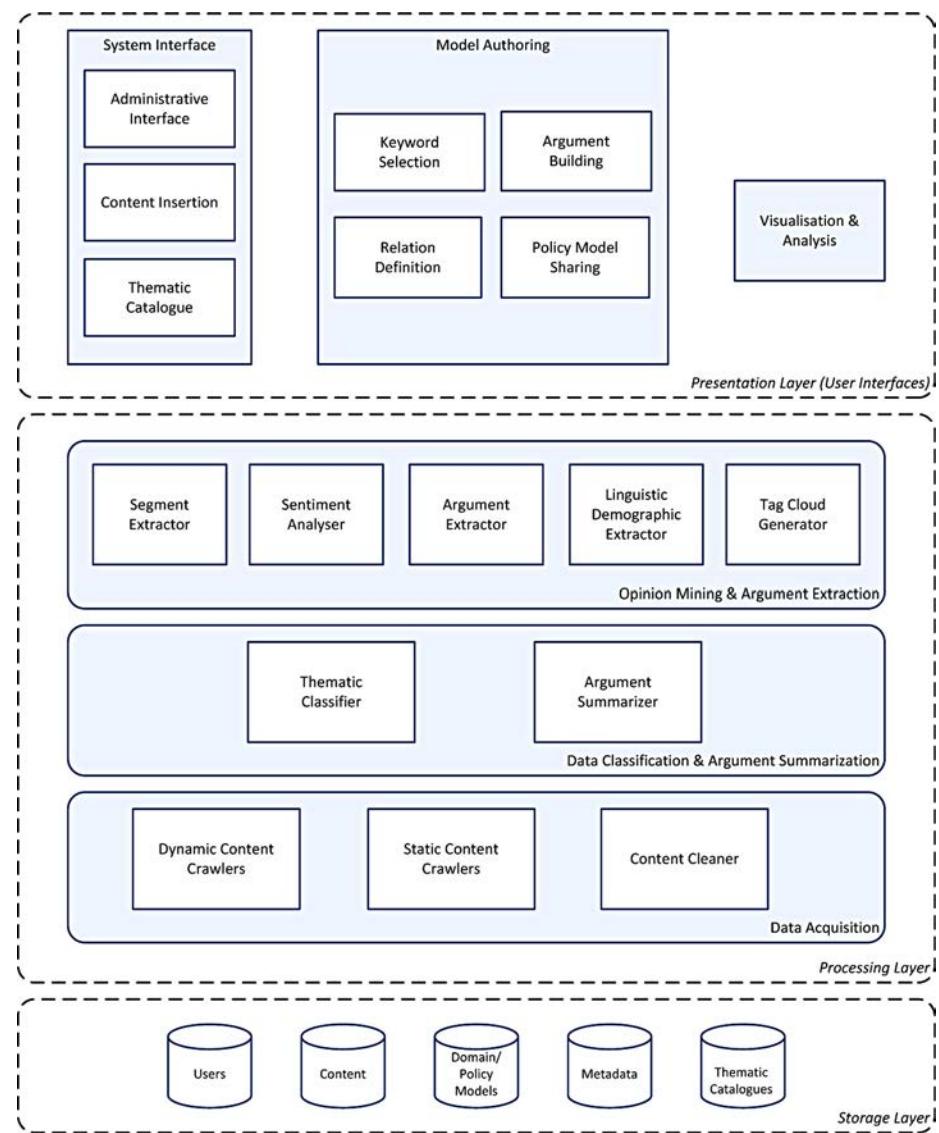


Figure 6.
Technological architecture
of ICT platform

- The *content repository*: It stores the raw content retrieved from the Web 2.0 sources, the cleaned content derived from the raw data, the content uploaded by users and the results of the linguistic analysis associated with each content unit.
- The *model repository*: It stores in a structured form the domain and policy models entered by users with domain expert and policy advisor roles.
- The *metadata repository*: It stores the metadata retrieved or calculated for our sources.

-
- The *thematic catalogues*: It stores a representation of the thematic categories used by the platform to characterize each content unit.
 - The *users repository*: It contains information about the roles and the users of the platform.

The *processing layer* includes all the components that retrieve and process the content from the predefined sources, which are organized in three sub-layers:

- (1) The *data acquisition layer*, which includes the crawling components for fetching content from the sources, using their APIs, as well as the modules responsible for cleaning the fetched content and obtaining the actual textual information from it (static content crawlers, dynamic content crawlers and content cleaner).
- (2) The *data classification and argument summarization layer*, which includes the thematic classifier, which processes the available content and associates it with one or more of the defined thematic categories in the thematic catalogues and the result summarizer, which processes the available results and provides a summarization that allows their presentation in a condensed manner.
- (3) The *argument extraction and opinion mining layer*, which includes all the components that process the available content and extract segments, arguments and sentiments (segment extractor, argument extractor, sentiment analyzer, linguistic demographic extractor and tag cloud generator).

The *presentation layer* includes all the components that either require input from the user or present to him/her the results:

- The *thematic catalogue* interface, for entering or updating the available thematic categories and also terms associated with each category.
- The *keyword selection* interface, which allows entering keywords/terms for creating domain models.
- The *relation definition* interface, which allows the user to introduce relations between the above keywords/terms for the definition of domain models.
- The *argument building* interface, which allows the user to insert in natural language statements and arguments supporting or objecting to policy statements of policy models.
- The *policy model-sharing* interface, which provides a catalogue of the policy models created by the user and allows defining them as visible to others.
- The *admin interface*, which provides the means to an administrator to manage the configurable aspects of the system.
- The *visualization and analysis module*, which utilizes the results of the processing layer to provide the user with a view of domain and policy models, and also various visualizations of the results of users searches, enabling also the selection of sources, demographic characteristics and time periods.

We remark that the structure and the components of this passive crowdsourcing ICT platform are quite different from the ones of the active crowdsourcing IS identified by Hetmank (2013), which have been outlined in Section 2.2.

The domain and policy-modeling components of the presentation layer (thematic catalogue, keyword selection, relation definition, argument building and policy model-sharing interfaces) will be based on the [ELEON Ontology Authoring and Enrichment Environment](http://www.iit.demokritos.gr/~eleon) (www.iit.demokritos.gr/~eleon), developed by the National Center for Scientific Research “Demokritos”, which participates as a partner in the NOMAD project. It supports editing ontologies and relating such ontologies with linguistic resources that can be used to extract structured ontological information from text, and also supports the author with a number of innovative methods for ontology checking ([Bilidas et al., 2007](#)) and auto-completion ([Konstantopoulos et al., 2011](#)). The sentiment analyzer will be based on existing tools developed by “Demokritos” as well, which use algorithms that 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. (for more details see [Rentoumi et al. \(2009\)](#) and [Rentoumi et al. \(2010\)](#)). For the summarization task, the “n-gram graph framework” ([Giannakopoulos et al., 2008](#); [Giannakopoulos and Karkaletsis, 2009](#)) will be used, which is a statistical, domain agnostic and language-independent framework that allows the analysis of texts as character n-gram graphs.

8. Conclusions

The previous sections have developed a novel approach to e-participation, which is based on the exploitation of Web 2.0 social media by government agencies, in a quite different way from the previous e-participation approaches, for conducting “passive crowdsourcing”. It is based on the search by government agencies for content on a domain of government activity or on a public policy under formulation, which has been created in numerous predefined 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 and stored, and then undergoes advanced processing using various linguistic processing techniques to extract from it needs, issues, opinions, proposals and arguments raised by citizens, identify their sentiments (positive or negative) and finally summarize and visualize them, in a way that is easily understandable by government policymakers with minimal cognitive effort. The proposed approach allows the exploitation of the vast amount of user-generated political content created continuously in numerous Web 2.0 social media to enable governments to understand better the needs and opinions of citizens, and extract from them more information, knowledge and ideas on the problems of modern societies to create better and more socially rooted policies. Therefore, it follows the first of the main crowdsourcing approaches identified by [Brabham \(2012\)](#), mentioned in Section 2.2, focusing on the knowledge discovery and management. Its application can lead to a new third generation of e-participation, which can be, however, perfectly combined with and enhance the first two generations (both oriented towards “active crowdsourcing”).

The proposed approach is less government-controlled than the first two generations of e-participation, and more citizens-driven and controlled, as in this approach, governments control neither the discussion topic (they simply “listen” to what topics and issues are mentioned by citizens in a broad policy area and what needs and problems the citizens identify), nor the discussion rules (they accept the ones defined by the various social media they exploit). It allows governments to

move beyond their own official e-participation spaces (first generation of e-participation) and social media accounts (second generation of e-participation) and explore a wider range of Web 2.0 sources. In this sense, it can promote bottom-up public policy formulation. However, at the same time, it is characterized by higher levels of technological and political complexity and challenges than the second generation of e-participation.

In particular, in cooperation with potential users experienced in the design of public policies, this government passive crowdsourcing concept has been developed and elaborated. We developed initially a process model for its application, which describes how it can practically work (including a sequence of specific activities to be executed), and then based on it, we designed the functional architecture of an ICT platform that supports its application. An important element of both the process model and functional architectures is the creation and use of domain and policy models. The inherent complexity and multidimensionality of public policies makes it difficult to search for relevant content in the predefined Web 2.0 sources using a small number of keywords, and this difficulty can be addressed through domain and policy models, which are used for searching relevant content. It has also been concluded that the practical application of this new concept will require four main roles: domain models author, policy models author, end user/policymaker and technical administrator; for each of them, the functionality required from the above supporting ICT platform were determined.

Finally, based on the above requirements the technological architecture of an ICT platform that fulfils these functional requirements, and in general supports the application of this government passive crowdsourcing concept, was designed. To achieve a low response time and at the same time provide sufficiently "fresh" content for policymakers (being two basic conflicting requirements), its design should be based on the separation between sources scanning and retrieved content processing, on one hand, and users' searches processing, on the other. This leads to a three-layered technological architecture of the platform, which consists of a storage layer, a processing layer and a presentation layer. Each layer includes a number of components performing different tasks, which function as services coordinated by an orchestration component.

Our analysis showed that the objectives and the application process of this passive crowdsourcing approach and also the structure and components of the required supporting ICT platform differ significantly from the ones of the existing active crowdsourcing approaches, which have been identified by [Hetmank \(2013\)](#) and outlined in Section 2.2. In particular, this approach does not involve the main tasks of the active crowdsourcing (problem definition, open call for contributions, search for and motivation of contributors, evaluation of contributions and finally reward of the most successful of them), but has a quite different task structure (including domain and policy modeling, definition of the Web 2.0 sources to be used and sophisticated processing of the retrieved content). It requires more extensive and complex ICT support than the existing active crowdsourcing approaches, which is based on the use of API of numerous Web 2.0 sources (which are controlled by other players and are heterogeneous and change frequently), in combination with advanced linguistic processing techniques.

The research presented in this paper has interesting implications for research, practice and society. It opens up a new direction of research on this novel passive crowdsourcing approach, from both e-participation and crowdsourcing

perspectives. This research should aim at developing ways and processes of applying it in different types of organizations of the public sector, and also ICT infrastructures for supporting them and at investigating their effectiveness, benefits and business value, their combination with active crowdsourcing practices and at the same time their application barriers and risks (especially from cultural and legal viewpoint). The present study provides a basis for this future research. With respect to practice, taking into account that traditionally government agencies monitor what the press writes about them, our research provides a basis for extending efficiently these activities in the new electronic media world (e.g. newspapers, websites, blogs and microblogs, online forums, etc.) to understand better the needs, issues, opinions, arguments and proposals raised by the society with respect to important domains of government activity and public policies and, in general, with respect to the big problems of modern societies. It allows government agencies to identify trends and changes in the above, and also differences between various citizens' groups (e.g. age, gender, education and region groups), which provides a sound basis for designing appropriate public policies for addressing them. Finally, the proposed approach has significant social implications, as it provides a new channel for the "voice" of the society to be directly communicated to the government, so that the latter can design its policies and activities based on the social realities, and not on oversimplified models and stereotypes.

The main limitation of this study is that its main developments (i.e. the main elements of the proposed approach, the application process model and the functional and technological architecture of supporting ICT platform) have been based exclusively on our cooperation with potential users highly experienced in the design of public policies. Therefore, they have to be evaluated through "real life" pilot applications of them, which might lead to improvements (at the level of the main elements of the proposed approach, or/and the application process model or/and the functional and technological architecture of the supporting ICT platform). This is already in progress as part of the aforementioned project NOMAD. The application scenarios mentioned in Section 3 will be used in the next phases of the project as basis for the design of corresponding pilot applications of the proposed passive crowdsourcing approach, which will be implemented by the corresponding user partners (using the process model presented in Section 5 and the ICT platform presented in Sections 6 and 7) and then evaluated. This evaluation will focus on the main technological and political complexities and challenges it faces, which have been mentioned in the Introduction, using the methodology proposed by Ferro *et al.* (2013a). In particular, the technological evaluation will be based on the software platforms and ecosystems theory (Gawer, 2010; Tiwana *et al.*, 2010), and will assess to what extent the API of the used Web 2.0 social media sources provide all the required capabilities for retrieving from them the required content with all necessary descriptive fields, and, in general, to what extent the main preconditions of the platform-based software development paradigm are fulfilled. Also, it will assess to what extent the advanced linguistic processing techniques we use have a good performance for the various types of texts found in these heterogeneous Web 2.0 sources. The political evaluation will be based on the wicked problems theory (Rittel and Weber, 1973; Conklin and Begeman, 1989; Conklin, 2003), and will assess to what extent it is useful for reaching wider citizens' groups and identifying particular problems/issues they perceive concerning the particular policy under discussion, possible

solutions to them and relevant advantages – positive arguments – and disadvantages – negative arguments. Furthermore, as the proposed approach is an important innovation for government, we will evaluate it from an innovation diffusion perspective using the methodology proposed by Ferro *et al.* (2013b) and assess to what extent it has the fundamental preconditions for a wide diffusion proposed by the diffusion of innovation theory of Rogers (2003) (relative advantage, compatibility with existing values and processes, reasonable complexity, trialability and observability). In general, extensive research is required on passive crowdsourcing, examining for it all the multiple aspects that have been researched for the existing active crowdsourcing paradigm (see reviews of this research by Pedersen *et al.*, 2013; Tarrell *et al.*, 2013).

References

- Bertot, J.C., Jaeger, P.T. and Grimes, J.M. (2012), "Promoting transparency and accountability through ICTs, social media, and collaborative e-government", *Transforming Government: People, Process and Policy*, Vol. 6 No. 1, pp. 78-91.
- Bilidas, D., Theologou, M. and Karkaletsis, V. (2007), "Enriching OWL ontologies with linguistic and user-related annotations: the ELEON system", *Proceedings of IEEE International Conference on Tools with Artificial Intelligence (ICTAI-07)*, Patras.
- Brabham, D.C. (2008), "Crowdsourcing as a model for problem solving an introduction and cases", *Convergence: The International Journal of Research into New Media Technologies*, Vol. 14 No. 1, pp. 75-90.
- Brabham, D.C. (2009), "Moving the crowd at threadless – motivations for participation in a crowdsourcing application", *Information, Communication and Society*, Vol. 13 No. 8, pp. 1122-1145.
- Brabham, D.C. (2012), "Crowdsourcing: a model for leveraging online communities", in Delwiche, A. and Henderson, J. (Eds), *The Routledge Handbook of Participative Cultures*, Routledge.
- Chadwick, A. (2009), "Web 2.0: new challenges for the study of E-democracy in an era of informational exuberance", *I/S: A Journal of Law and Policy for the Information Society*, Vol. 5 No. 1, pp. 9-41.
- Charalabidis, Y. and Loukis, E. (2012), "Participative public policy making through multiple social media platforms utilization", *International Journal of Electronic Government Research*, Vol. 8 No. 3, pp. 78-97.
- Chun, S.A., Shulman, S., Sandoval, R. and Hovy, E. (2010): Government 2.0). "Making connections between citizens, data and government", *Information Polity*, Vol. 15 No. 1/2, pp. 1-9.
- Conklin, J. (2003), "Dialog mapping: reflections on an industrial strength case study", in Kirschner, P., Buckingham Shum, S. and Carr, C. (Eds), *Visualizing argumentation: software tools for collaborative and educational sense-making*, Springer Verlag, London.
- Conklin, J. and Begeman, M. (1989), "gIBIS: a tool for all reasons", *Journal of the American Society for Information Science*, Vol. 40 No. 3, pp. 200-213.
- Cruickshank, P., Edelmann, N. and Smith, C. (2010), "Signing an e-petition as a transition from lurking to participation", in Tambouris, E., Macintosh A. and Glassey O. (Eds), *Second IFIP International Conference on Electronic Participation – ePart 2010, Lausanne, Switzerland*.
- ELEON Ontology Authoring and Enrichment Tool, available at: www.iit.demokritos.gr/~eleon
- Ferro, E. and Molinari, F. (2010), "Making sense of gov 2.0 strategies: no citizens, no party", *Journal of eDemocracy and Open Government*, Vol. 2 No. 1, pp. 56-68.

- Ferro, E., Loukis, E., Charalabidis, Y. and Osella, M. (2013a), "Evaluating complex forms of social media use in government", *Proceedings of the 19th Americas Conference on Information Systems-AMCIS 2013*, 15-17 August, 2013, Chicago, IL.
- Ferro, E., Loukis, E., Charalabidis, Y. and Osella, M. (2013b), "Analyzing the centralised use of multiple social media by government from innovations diffusion theory perspective", *Proceedings of the IFIP Fifth International Conference on e-Participation-ePart 2013*, 17-19 September, 2013, Koblenz.
- Gawer, A. (2010), "The organization of technological platforms", *Research in the Sociology of Organizations*, Vol. 29, pp. 287-296.
- Giannakopoulos, G. and Karkaletsis, V. (2009), "N-gram graphs: representing documents and document sets in summary system evaluation", *Proceedings of Text Analysis Conference TAC 2009*, 16-17 November, 2009, Gaithersburg, MD.
- Giannakopoulos, G., Karkaletsis, V., Vouros, G. and Stamatopoulos, P. (2008), "Summarization system evaluation revisited: n-gram graphs", *ACM Transactions on Speech and Language Processing*, Vol. 5 No. 3, pp. 1-39.
- Hetmank, L. (2013), "Components and functions of crowdsourcing systems – a systematic literature review", *Wirtschaftsinformatik Proceedings 2013*, Leipzig.
- Hilgers, D. and Ihl, C. (2010), "Citizensourcing: applying the concept of open innovation to the public sector", *The International Journal of Public Participation*, Vol. 4 No. 1, pp. 67-88.
- Howe, J. (2006), "The rise of crowdsourcing", *Wired*, Vol. 14 No. 6, available at: www.wired.com/wired/archive/14.06/crowds.html
- Howe, J. (2008), *Crowdsourcing, Why the Power of the Crowd is Driving the Future of Business*, Crown Business, New York, NY.
- Keim, D.A., Kohlhammer, J., Ellis, G.P. and Mansmann, F. (2010), *Mastering the Information Age – Solving Problems with Visual Analytics*, Eurographics, Goslar.
- Kokkinakos, P., Koussouris, S., Panopoulos, D., Askounis, D., Ramfos, A., Georgousopoulos, C. and Wittern, E. (2012), "Citizens collaboration and co-creation in public service delivery: the COCKPIT project", *International Journal of Electronic Government Research*, Vol. 8 No. 3, pp. 44-62.
- Konstantopoulos, S., Karkaletsis, V., Vogiatzis, V. and Bilidas, D. (2011), "Authoring semantic and linguistic knowledge for the dynamic generation of personalized descriptions", in Sporleder, C., Bosch, A.V and Zervanou, K. (Eds), *Language Technology for Cultural Heritage*, Springer Verlag, Berlin.
- Levy, P. (1997), *Collective Intelligence: Mankind's Emerging World in Cyberspace*, Plenum, New York, NY.
- Loukis, E. and Wimmer, M. (2012), "A multi-method evaluation of different models of structured electronic consultation on government policies", *Information Systems Management*, Vol. 29 No. 4, pp. 284-294.
- Loukis, E., Macintosh, A. and Charalabidis, Y. (2011), "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*, Vol. 13 No. 1, pp. 1-12.
- Maragoudakis, M., Loukis, E. and Charalabidis, Y. (2011), "A review of opinion mining methods for analyzing citizens' contributions in public policy debate", *Proceedings of the IFIP Third International Conference on e-Participation-ePart 2011*, August 29 - September 1, 2011, Delft.
- Margo, M.J. (2012), "A review of social media use in e-government. administrative sciences", *Administrative Sciences*, Vol. 2 No. 2, pp. 148-161.

- Medaglia, R. (2008), "Measuring the diffusion of eParticipation: a survey on Italian local government", *Information Polity - An International Journal of Government and Democracy in the Information Age*, Vol. 12 No. 4, pp. 265-280.
- Nam, T. (2012), "Suggesting frameworks of citizen-sourcing via Government 2.0", *Government Information Quarterly*, Vol. 29, pp. 12-20.
- Organization for Economic Co-operation & Development – OECD (2001), *Citizens as Partners – Information, Consultation and Public Participation in Policy-Making*, OECD Publication Service, Paris.
- Organization for Economic Co-operation and Development (OECD)(2003), *Engaging Citizens Online for Better Policy-making-Policy Brief*, OECD Publication Service, Paris.
- Organization for Economic Co-operation & Development – OECD (2004a), *Evaluating Public Participation in Policy Making*, OECD Publication Service, Paris.
- Organization for Economic Co-operation and Development (OECD) (2004b), *Promise and Problems of e-Democracy: Challenges of Online Citizen Engagement*, OECD Publication Service, Paris.
- Panopoulou, E., Tambouris, E., Sanchez-Nielsen, E., Zotoú, M. and Tarabanis, K. (2011), "Learning from eParticipation initiatives of regional and local level authorities in Greece and Spain", *Journal of Balkan and Near Eastern Studies*, Vol. 3 No. 1, pp. 13-35.
- Pedersen, J., Kocsis, D., Tripathi, A., Tarrell, A., Weerakoon, A., Tahmasbi, N., Xiong, J., Deng, W., Oh, O. and Vreede, G.J.D. (2013), "Conceptual foundations of crowdsourcing: a review of IS research", *Proceedings of the 46th Annual Hawaii International Conference on System Sciences*, Maui, HI.
- Rentoumi, V., Giannakopoulos, G., Karkaletsis, V. and Vouros, G. (2009), "Sentiment analysis of figurative language using a word sense disambiguation approach", *Proceedings of RANLP 2009*, Borovets.
- Rentoumi, V., Petrakis, S., Karkaletsis, V. and Vouros, G. (2010), "A collaborative system for sentiment analysis", *Proceedings of SETN 2010*, Athens.
- Rittel, H.W.J. and Weber, M.M. (1973), "Dilemmas in a general theory of planning", *Policy Sciences*, Vol. 4, pp. 155-169.
- Rogers, E. (2003), *Diffusion of Innovations*, The Free Press, New York, NY.
- Saebo, O., Rose, J. and Flak, L.S. (2008), "The shape of eParticipation: characterizing an emerging research area", *Government Information Quarterly*, Vol. 25, pp. 400-428.
- Santucci, D. (2007), "Studying e-petitions: state of the art and challenges", *ESF-LIU Conference on Electronic Democracy, Vadstena, Sweden*.
- Snead, J.T. (2013), "Social media use in the U.S. executive branch", *Government Information Quarterly*, Vol. 30 No. 1, pp. 56-63.
- Stewart, O., Huerta, J. and Sader, M. (2009), "Designing crowdsourcing community for the enterprise", *Proceedings of the ACM SIGKDD Workshop on Human Computation*, ACM, New York, NY.
- Surowiecki, J. (2004), *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, NY.
- Tarrell, A., Tahmasbi, N., Kocsis, D., Tripathi, A. Pedersen, J., Xiong, J., Oh, O. and Vreede, G.J.D. (2013), "Crowdsourcing: a snapshot of published research", *Proceedings of the Nineteenth Americas Conference on Information Systems (AMCIS) 2013*, Chicago, IL.

-
- Thomas, J. and Cook, K. (Eds) (2005), *Illuminating the Path: Research and Development Agenda for Visual Analytics*, IEEE Press, Los Alamitos, CA.
- Tiwana, A., Konsynski, B. and Bush, A. (2010), "Platform evolution: coevolution of platform architecture, governance, and environmental dynamics", *Information Systems Research*, Vol. 21 No. 4, pp. 675-687.
- Torres, L.H. (2007), "Citizen sourcing in the public interest", *Knowledge Management for Development Journal*, Vol. 3 No. 1, pp. 134-145.
- United Nations (2008), *E-Government Survey 2008: From eGovernment to Connected Governance*, United Nations Publications, New York, NY.
- Wandhöfer, T., Taylor, S., Alani, H., Joshi, S., Sizov, S., Walland, P., Thamm, M., Bleier, A. and Mutschke, P. (2012), "Engaging politicians with citizens on social networking sites: the wegov toolbox", *International Journal of Electronic Government Research*, Vol. 8 No. 3, pp. 22-43.
- Wong, P.C. and Thomas, J. (2004), "Visual analytics", *IEEE Computer Graphics and Applications*, Vol. 24 No. 5, pp. 20-21.

Appendix

Part A: Personal information

- A1. Your age is between (18-25, 25-35, 35-50, over 50 years).
- A2. Your gender is (male or female).
- A3. Your educational level is (high school, bachelor's/college degree, master's degree, PhD/postgrad research degree or other – please specify).

Part B: General citizens' participation information

- B1. Do you use some specific tool for monitoring and measuring citizens' opinion/feedback for the purpose of policymaking? (no, yes – please specify)
- B2. Do you use any specific method for attracting people in order to participate in policymaking procedures? (no, yes – please specify)
- B3. Do you find the idea of getting early feedback about societal reactions useful in your line of work? (no – please justify, yes – please justify)
- B4. During policymaking process, when is the right time to involve citizens? (during idea formulation, once draft is ready, at final stages of legal description, other – please specify)
- B5. Which of the following may influence your attitude during legislation process? (forum discussions, blog posts, reaction in SM, comments in news sites/portals, polls, none, other – please specify)

Part C: Current use of social media in policymaking processes

- C1. What methods do you use for the gathering of arguments and opinions about existing policies? (polls, surveys, internet search, internet applications, other – please specify)
- C2. Have you ever used social media in order to discuss and get feedback from citizens? (yes – please specify which social media did you use, no)
- C3. How often do you discuss policy-related issues in social media? (very frequently, frequently, occasionally, rarely, very rarely or never)
- C4. Which social media do you use? (blogs [e.g. Blogger, Wordpress], microblogs [e.g. Twitter, social networks, e.g. Facebook, Myspace], video/photo-sharing platforms [YouTube, Picassa], news sites with commenting functionalities, discussion forums, other – please specify, none)

- C5. Name each of the above social media that you use for discussing/promoting new policies (blog names, Facebook or Twitter groups, news site columns, etc.)
- C6. How many comments did you have on average from citizens about a new policy proposal? (none, 1-50, 50-100, more than 100)
- C7. How often did you adapt the policy according to citizens' comments? (always, most of the times, sometimes, rarely, never)
- C8. Do you think the use of social media for policymaking is beneficial? (yes, no – please specify why)
- C9. Do you think incorporating citizens' opinions into the policymaking process is helpful? (yes, no – please specify why)
- C10. What is the importance of each of the following positive effects of the use of social media for policymaking? (1 = less important, 5 = most important) (faster and easier citizen feedback, faster dissemination of policy, direct and effective communication between policymakers and citizens, more effective targeted dissemination to citizen groups, other – please specify)
- C11. What is the importance of each of the following possible negative effects of the use of social media for policymaking? (1 = less important, 5 = most important) (possible lack of user authentication and low security level, user anonymity, many irrelevant comments, lack of credibility of statistical results, other – please specify)
- C12. What are the constraints for using social media in policymaking that have to be taken into account? (lack of time needed to search and analyze results, infrastructure deployment difficulties, lack of software for analyzing citizen feedback, organization difficulties, political considerations, other – please specify)
- C13. What category of social media do you consider as more suitable for the policymaking process? (1 = less important, 5 = most important) (blogs, microblogs, social networking sites, wikis, social bookmarking sites, platforms for multimedia and entertainment, platforms for news and information, government organizations forums, online votes/petitions, other – please specify)
- C14. Do you regularly search social media just to get a grasp of citizens' feelings about specific policies? (yes, no) If yes, which social media? (blogs, microblogs, social networks, news sites with commenting functionalities, other – please specify)
- C15. Do you use internet-search applications for this purpose? (yes, no) If yes, which application do you use? (Google search, reputation measurement tools (e.g. socialmention.com, trackur.com), other search tool – please specify)
- C16. Have you ever used decision-making software? (yes, no) If yes, which of the following? (1000Minds, AIRM Online, Analytica, Criterium, Decision Plus, DecideIT, Decision Lab, Decision Lens, D-Sight, Expert Choice, Hiview3, Logical Decisions, MakeItRational, MindDecider, other – please specify)
- C17. What did you like about it?
- C18. Have you ever used any commercial tools for social media monitoring? (yes, no)
- C19. If yes, which of the following? (Alterian SM2, Brandwatch, Converseon, Cymfony Maestro, evolve24 Mirror, Media Metrics socialMeme, Meltwater Buzz, NM Incite My BuzzMetrics, Radian6, Sysomos, Visible Technologies Intelligence, other – please specify)
- C20. If yes, which of the following disadvantages and shortcomings have you encountered with the above tools (technical and non-technical)? (limited access to social media data, poor data analysis, slow processing time, inadequate presentation of results, other – please specify)

Part D: Specific requirements and NOMAD general assessment

- D1. What tools/apps would be useful for you during the formation/validation of a new policy? (data mining, semantic/sentiment analysis, statistical processing, graphic or visual presentation tools, monitoring, collecting and presenting user input, other – please specify)
- D2. What topics would you like to monitor? (health, environment, economy, employment, politics, other – please specify)
- D3. What kind of feedback would you like to collect? (views and opinions, suggestions, criticisms, sentiments, demographics, statistics, other – please specify)
- D4. What kind of content would you like to be searched in social media/Internet by a tool that assists in policymaking? (news, events, articles, comments/opinions, polls, bookmarks, other – please specify)
- D5. How would you value citizens' opinions? (based on socioeconomic background, based on demographics (e.g. age), based on expertise, other – please specify)
- D6. Do you think that NOMAD could provide a decision-support solution to your organization? (it would be very useful, it would probably be useful, indifferent, it would not be useful at all)
- D7. How effective do you think that NOMAD will be in allowing citizens to affect policymaking? (very effective, somehow effective, indifferent or not effective at all)
- D8. How would you assess the degree of innovation required in developing the tools and technologies needed for NOMAD? (state of the art, incremental innovation, radical innovation, systematic innovation)
- D9. Briefly, what do you expect from NOMAD?

Please let us know if we have left anything out or if there are any important subjects that are not covered.

Corresponding author

Euripidis N. Loukis can be contacted at: eloukis@aegean.gr