A SEMANTICS-BASED CONSULTATIONS WORKBENCH

Costas Vassilakis, Dimitris Gouscos, Panagiotis Georgiadis
E-Gov Lab, Department of Informatics and Telecommunications, University of Athens, Greece
{costas, d.gouscos, p.georgiadis}@e-gov.gr

Abstract: eConsultations constitute an effective means to inclusive and informed participation of citizens and society in policy, decision and law formation processes, and an answer to the democratic deficit. eConsultation platforms need to support all stages of consultation processes including agenda setting and topics raising, legislation proposal publicity, notification of developments, proposal debate and commentary, collection, analysis and synthesis of views. In this paper we present the design of an open platform assisting policy makers and the civil society in the set-up, enactment, management and federation of inclusive and informed consultations. The proposed platform employs semantic techniques, such as semantics, content annotation and summarization to support the consultation processes and provide targeted and digested information to participants, and facilitates tailoring of eConsultation procedures by offering basic eConsultation activities as building blocks, which can be combined according to contextual needs. The platform also enables distinct eConsultation processes to be federated, allowing the exchange of information, which may be subject to different semantic annotation and classification, according to the rules of each eConsultation process.

Keywords: eGovernment, eConsultations, semantic web, federation, platform architecture

1. INTRODUCTION

Consultations are processes undertaken by empowered institutions (or agent) to inform itself on one or more as-yet unmade decisions, or to review previous or existing decisions, policies, or programs [1]. Consultations are an indispensable tool for identifying potential impacts of decisions on involved stakeholder groups, which are given the opportunity to express their views before decisions are finalised. The importance of this aspect has been verified in a significant number of cases, such as the EU's 2003 legislation on chemicals [2], where 6500 contributions were received in the context of an internet-enabled consultation that was held for over two months. The consultation process revealed flaws in the legislation, whose correction saved several billions of Euros, in particular for smaller companies. Consultations may also significantly contribute to addressing the democratic deficit in decision-making [3] and to leveraging the consensus around decisions, through the increased sense of participation. The scope of consultations may range from neighbourhood level to national or even international level, depending on the issue under discussion: issues such as pet control or road conditions may be best addressed by local societies, while issues of more global nature, such as educational policy or social policy reforms are bound to be discussed in a wider audience.

Information and communication technologies (ICT) can play a significant role in enabling consultations. An important such use of ICT technologies would be for promoting the on-line provision of information, which is an essential precondition for citizen engagement in the democratic process [4]. Furthermore, ICT technologies can support the various activities related to e-consultations, including raising of issues, collecting and commenting opinions (e.g. through internet-based discussion forums), annotating statements with supporting documents (through references to digital libraries and/or attachment of documents), and conducting of e-voting to draw the final decision.

While ICT technologies can thus be used to assist the "traditional" process of conducting e-consultations, they present new opportunities for further enhancing the consultation process. Firstly, through ICT technologies the on-line provision of information can be more targeted, forwarding to each stakeholder only the information that is relevant to her interests, effectively tackling the issue of information overload [5]. Secondly, ICT technologies can greatly facilitate "bottom-up" democratic process whose agenda is set by the democratic constituency, possibly structured around political groups or NGOs, rather than leaving the initiative to the decision makers (governments, public agencies, etc). Through such initiatives momentum can be built and the decision-makers' attention can be drawn to specific issues. An example of such a bottom-up process is the "I Decide Today" web page of the Estonian Government [6]. Thirdly, semantics-based techniques from the ICT domain can be exploited to improve the efficiency and effectiveness

of the opinion collection and commenting phase: opinions and comments may be linked together and possibly be automatically associated with supporting documents; automatic or semi-automatic classification of opinions and comments to topic realms can promote more targeted discussions; automated document summarisation can support both the discussion phase and the conclusions drawing phase, since the volume of the text than needs to be read is drastically reduced. A final opportunity offered by ICT technologies for the organisation of consultations and discussions is the *content federation* between otherwise independent consultations. Through content federation an e-consultation forum may publish selected portions of its content (topics, comments, supported documents etc), which can then be imported by (properly authorised) affiliate sites. Consultation forums may participate in a federation at a peer-level basis (e.g. two environmental NGOs may exchange views on selected subjects), in a hierarchical fashion (e.g. consultations within university communities may publish their content related to non-local issues to a nation-wide higher education consultation forum) or even in an "adverse party" setup (e.g. a hunter's association may exchange content with an environmental NGO for idea cross-fertilisation).

An important aspect that must be taken into account for the design of ICT systems that will support public consultations, is that no single consultation process model is suitable for all consultation processes: the distinct steps that will be taken within a consultation process (topic raising, proposal submission, commenting, voting etc), their order, the number of iterations etc. may need to be adapted, taking into account the audience's political culture, the issue at hand, the time constraints, or other pertinent parameters [7] [8] [9]. Parameters of various stages, such as registered vs. anonymous commenting, moderated vs. unmoderated discussions, privacy, anonymity, security, uniqueness of voting may also vary depending on the aforementioned parameters. In this sense, a public consultations ICT system must provide a flexible toolkit allowing the consultation initiator to tailor the consultation process to the needs of the specific situation. Such a toolkit can be complemented with checklists, best practises guides and/or automated consultation process verification tools, to assist consultation process designers to formulating robust and effective consultation models, and help them avoid common errors and pitfalls. A final parameter that should be taken into account for public consultation ICT systems is the ability to setup, operate and maintain such a system with low cost and limited technological expertise, so as not to place barriers for stakeholders with restricted resources. The ability for the platform to operate in an application service provision [10] model is also desirable, enabling stakeholders that cannot directly host such a platform to conduct consultations.

In this work we discuss the eConWork platform, which is targeted at facilitating electronically enacted consultation processes, set up and

managed either by the decision makers or by the decision constituency, according to process models varying as outlined above. The rest of this paper is organised as follows: section 2 presents background information, related work and the state of the art related to e-consultations. In section 3 the key aspects of the eConWork platform are presented, while section 4 concludes the paper and outlines future work.

2. RELATED WORK

Consultations are used more and more widely by governments, public organisations as well as by user communities that face common issues or challenges. The European Union has launched the "Your voice in Europe" portal, through which EU citizens have access to consultations and discussions, as well as to a forum for exchanging experiences (http://europa.eu.int/yourvoice/). Consultation portals have also been developed at governmental level (e.g. the "Your opinion counts" portal by the government of Singapore [http://www.feedback.gov.sg/] and the "Consulting with Canadians" of the government of Canada [http://www.consultingcanadians.gc.ca/]), or county level (e.g. the "Ask Me!" portal of the Worcestershire county council [http://www.worcestershire.gov.uk/home/cs-con/cs-con-ask.htm]).

The interest shown in consultations has triggered related research activities, especially in the area of best practices for conducting consultations. The UK Cabinet Office has published a guide on how users should be consulted [11] as well as a comprehensive "top ten tips" list on issues to pay attention to and common errors to avoid [12]. [13] provides guidance on how consultations should be conducted, including information on consultation evaluation and performance measurement and tips on how consultations can be improved. Guidelines for conducting consultations have been published by other administrations as well e.g. [14] and [15], while a number of publications are entitled "consultation toolkits" [16] [17], being however guides on how consultations should be conducted rather than specialised software assisting consultation designers to plan their consultations.

In the software domain, existing consultation portals have been customly developed for the respective administration, as specialised software solutions. The open-source community has contributed various isolated tools that may be used as parts of a consultation process, such as Drupal (http://freshmeat.net/projects/drupal/), Referendum (http://freshmeat.net/projects/referendum/) and Yet Another Community System (http://freshmeat.net/projects/yacs/) that may be used as community tools, phpBB (http://sourceforge.net/projects/phpbb/) and w-agora (http://sourceforge.net/projects/w-agora/) which implement forum functionalities, the demexp (http://freshmeat.net/projects/demexp/) and GNU.FREE (http://freshmeat.net/projects/free/) voting systems and so forth. However, each of these tools supports only a small subset of the consultation

process, no communication between the phases supported by different tools is provided (either via common information repositories of through data exchange procedures) and substantial technical know-how is required to install, configure and operate all the required components, making it thus infeasible for many organisations to undertake the task of organizing public consultations through the Internet.

3. THE ECONWORK PLATFORM

In order to meet the requirements for an effective public consultation platform, the eConWork platform is structured into functional layers, as depicted in *Figure 1*. A single eConWork platform installation may host multiple consultations, either on related topics or on totally diverse subjects. This feature of the eConWork platform enables it to be used as a centralised environment for performing consultations, constituting thus a single point of reference for consultation stakeholders. Users of the platform, however, are provided with tools and facilities for filtering the information available on the platform, so as to avoid information overwhelming. The eConWork platform layers as well as the functionality and the modules incorporated in each layer are discussed in the following paragraphs.

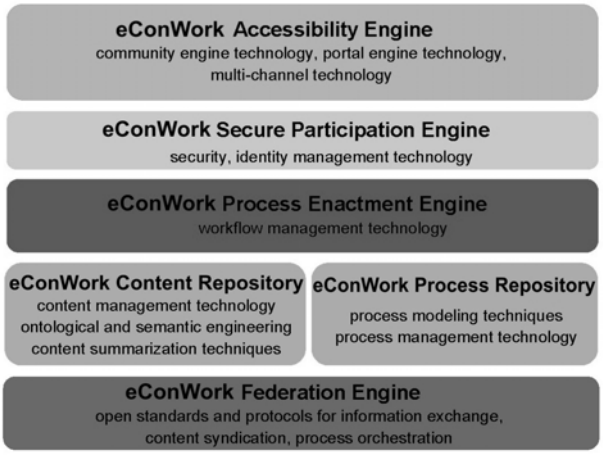

***Figure 1*. Overall eConWork Architecture**

The top layer of the architecture hosts the essential components for users to interact with the eConWork platform. These components include:

- facilities for community collaboration, such as forums and message boards, on-line dialogues, referenda, questionnaire-based surveys, calendars and shared document repositories. Effectively, these are the instruments that will be available to the eConWork platform users to exchange views and access related documentation throughout the consultation processes.

- tools to assist users locate the desired content, including classification taxonomies, search engine interfaces, “what’s related” lists, directories and so forth. These tools are not considered as core elements for performing consultations, however they are essential in an environment with large volumes of possibly diverse content (due to the platform capability to conduct multiple consultations), since they enable users to reach the desired content either by navigation or through querying.
- provisions for user profiling. Users of the eConWork platform are able to specify preferences regarding their interaction with the system as well as designate the content and procedures they are interested in. By specifying their interests, users can effectively avoid information overloading, since the eConWork platform will limit the presented information to the relevant content; user preferences are also taken into account in the notification posting procedures, through which the platform informs its users regarding important developments or deadlines. Besides content characteristics, user preferences may include procedures and processes as well, since certain users may be only interested in participating in certain stages of the consultation (e.g. only in the voting stage). In order to facilitate user profiling, the eConWork platform includes pre-configured user templates (e.g. doctor, farmer, student, professor, environment activist etc), which are linked to content categories related to the interests implied by each template. Preferences inherited through templates may be then further refined by the users.
- adapters for the delivery of platform content and facilities through different dissemination channels. The eConWork platform does not consider only the web channel, but exploits the capabilities offered by other technological developments, including WAP, SMS, telephone centres, teletext and interactive TV, as well as other prominent communication channels. Naturally, the nature of certain communication channels imposes certain restrictions -for instance it is infeasible to communicate a public discussion transcript through the SMS channel. Differentiations may also appear regarding the capabilities of client access devices served through the same communication channel, e.g. a PC connected to the internet offers better capabilities for text presentation and user input as compared to an i-mode enabled cellular phone. The design of the eConWork platform allows for specifying the activities and the content that may be delivered through a specific dissemination channel (for example, the SMS channel may only be used for notifying users regarding major developments as well as for voting), while adapters may be configured to take into account the peculiarities and special features of various client access devices that may be attached to them (e.g. segmenting long texts into page sequences for small displays, using alternative input methods and lexicons to support devices with limited keyboards and so forth).

The second functional layer of the eConWork platform manages issues related to security and identity management. This layer exploits information provided by the users (e.g. username-password combinations, certificates, data from physical tokens such as smart cards, etc) as well as information provided by the dissemination channel adapters (IP address of connecting access device for computer networks, caller ID for telephone networks and so on), to determine (a) the user identity with a *degree of confidence* and (b) the most prominent methods for establishing secure communication with the user, and compute the maximum level of security that these communication methods may guarantee. These two metrics are used by the platform to control access to various functionalities within consultation processes, in order to meet the uniqueness and security specifications set by consultation process designers. For example, if within the voting stage of some consultation process the uniqueness of the vote is not vital, a medium degree of confidence for the user identity (e.g. through nickname and password) may be adequate to ensure the voting right; if, however, the uniqueness of voting is considered crucial, more strict criteria may apply. Similar provisions are made for the security level, since expressing an opinion regarding pet control within a neighbourhood is bound to require less protection than casting a vote in the national elections.

The third functional layer of the eConWork platform is addressed to organisers of public consultations, empowering them to set up, roll out, monitor and manage consultations, as well as designate the content that will be published or imported through the federation mechanism (discussed below). In order to set up a consultation, organisers may select the pertinent stages from the *consultation process toolchest*, and combine them in a sequential, or even an iterative manner. The consultation process toolchest includes all individual stages that may appear within a consultation, such as publishing a proposed agenda, collecting comments on it, agenda refinement and finalisation, requesting contributions for the initial proposal, collecting comments on the proposal, evaluating the comments, publishing a revised proposal, voting and so on. Not all possible consultation stages will be required to be included in every consultation; for instance, if the designers of a consultation estimate that the consultation agenda is straightforward, then the stages associated with agenda setting may be omitted from the consultation design. Each consultation stage may be supported by a number of *instruments*, e.g. comment collecting may be supported through discussion forums, on-line synchronous discussions, e-mail and SMS messages, citizen panels, focus groups and so forth, while user identity and security parameters may be set for each of the stages. Time parameters may be also set for consultation stages and stakeholder notifications may be scheduled. Note that instruments are not necessarily limited to on-line tools, e.g. notification of stakeholders may be supported via newspaper publications or TV broadcasts, while the voting procedure may include

"traditional" voting with ballots. When such instruments are used, it is possible for the eConWork platform to send alerts to designated recipients, who should take the necessary actions in the physical world. *Figure 2* illustrates a screenshot of the consultation process designer, where the process toolchest displays the components available to be used in consultations, whereas the editor window (on the right) displays the components included in the consultation currently being edited and their sequence.

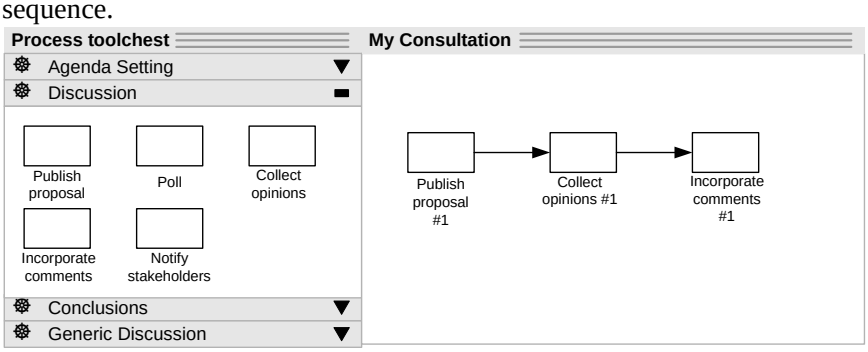

Figure 2. **Consultation process designer**

While designing a consultation process, designers are supported by reference process models and best practice guidelines. The eConWork platform encompasses a comprehensive list of best practises, success stories, checklists and common pitfalls, which consultation process designers may review to find techniques and practices best suited to the case at hand. All material is organised under a number of classification schemes (e.g. by consultation topic type, by number of expected participants, by communication channel), to ease the task of locating pertinent information.

Best practice guidelines are not only provided as reference documents, but a number of consultation parameters are also expressed in formal models, enabling the use of the *consultation process verifier,* an automated tool that ascertains that recommendations are followed in the designed consultation processes and provide advise to designers on how the consultations can be improved. The consultation process verifier may check various parameters of the consultation, such as inclusion of all mandatory phases (e.g. expression of views and voting), the order of phases (e.g. discussion of the agenda, if included, should be the first item in the consultation), duration of individual phases or the consultation as a whole (for instance, the phase of expressing views cannot last less than 10 days or more than three months; the duration of the whole consultation would be suggested not to exceed eight months), the number of iterations and so forth. All recommendations of the consultation process verifier are optional for the consultation designers to follow, providing flexibility for creating consultation models that are considered best for particular situations.

Complementary to the above tools and aids, *template consultations* are also provided, modelling the most commonly found types of consultations. Consultation process designers may instantiate a template consultation and use it directly, or appropriately customise it to fine-tune it for the specific situation.

The fourth functional layer of the eConWork platform hosts the actual platform content and provides the necessary infrastructure for semantics-related operations. The platform content is split into two distinct subsets: the first subset is comprised of consultation process models, best practices and other information and knowledge pertaining to the platform use, constituting thus effectively *meta-information* on the platform. The second content subset includes the actual contents of the consultations (messages, documents, transcripts etc), together with ontologies and semantic artefacts allowing for the semantic annotation of the information contents of consultation processes. When new consultation content is stored into this area, it is automatically matched against the semantic structures defined for the specific consultation, to create facilities that will prove useful to the eConWork platform users for handling the specific content. The facilities created for each new piece of content are dependent on the content type and the relevant consultation stage: for instance, ontologies are used for content classification [18] [19], and for generation of appropriate taxonomies, through which users may locate the content; content summarisation [20] [21] can be employed for dialogue transcripts and for lengthy supporting documents; document similarity [22] can be utilised to cluster opinions and comments into groups of similar or "almost tantamount" stakeholder views and so on. Ontologies should be specifically developed for each consultation (or set of consultations) to reflect the semantics that consultation organisers consider important as well as the interrelations between them, while other mechanisms are generic and may be used in any consultation. The results of the semantic processing within this layer are made available to the eConWork platform users either through explicit provisions on the top functional layer (classification taxonomies, search engine interfaces) or as hyperlinks added within the affected documents (e.g. a document summary contains a hyperlink to the full document and vice versa).

The fifth functional layer of the eConWork platform is a federation engine, enabling the semantics-based content federation between multiple independent platform installations. The federation engine extends the RSS content syndication mechanism [23] by allowing the exploitation of the semantic structures and mechanisms built into the eConWork platform. In this sense, consultation organisers may designate the content to be published by selecting nodes in an ontology or branches in a taxonomy, effectively defining that content related to the specific ontology node(s) or classified under a certain branch in the taxonomy will be made available for affiliate sites to import. Since an eConWork platform may host multiple (possibly

non-related) consultations, within each export procedure the affiliate sites that may import the published content are listed (e.g. if a platform hosts a consultation regarding the environment and one about national defence, the affiliate sites to which each consultation content should be exported are bound to be different). Export options also include whether only consultation contents will be exported or supporting documents will be provided as well, rules for excluding content from publishing (e.g. content categorised as *restricted* may be chosen to remain local only).

Once the content to be published is selected and authorisations are given, affiliate installations may proceed to *import* this content. Contrary to the RSS feeds, the import procedure does not simply transfer documents and data from one installation to another: as imported consultation contents are stored within the eConWork content repository, each storage activity triggers the execution of the semantic indexing and annotation mechanisms built into it, producing thus meta-information for navigation, summaries, links to other documents or other pertinent artefacts. We point out here, that the semantic artefacts created in the receiving installation are based on the semantic structures defined by its local consultation designers. Thus, imported items will be linked to local ontological nodes and taxonomy branches, be summarised according to the new hosting site's rules etc, being thus seamlessly integrated with the local content. The eConWork platform enables the use of digital signatures [24] for content that is exported/imported to guarantee the authenticity and integrity of content appearing in installations other than its first origin.

The content federation mechanism integrated in the eConWork platform enables different setups of affiliate sites to be formulated, depending on the affiliation goals, the relationships between affiliate sites in the real world, the consultation semantics etc. Among all possible federation schemes, the following three appear to be of more interest, since they correspond to common cooperation schemes:

- *full replication*. For administrative, organisational or even technical reasons (e.g. lack of on-line network connections or performance improvement) the same consultation is run on multiple installations. The content of the installations is periodically synchronised through the content-federation mechanism. In such a setup, semantic structures defined in affiliate sites are expected to be identical, thus when imported content is semantically annotated according to local semantic structures, the same results (classifications, linkages etc) will be reached.
- *selective, peer-level federation*. Each installation runs its own set of consultations, some of them being of local interest (e.g. in a higher education domain, teaching equipment or student housing problems), while others being of global interest (e.g. national funding for higher education and research). In this setup, each installation may retain consultations of local interest private, and allow for content exchange

within consultations of global interest. Such a setup may be also employed between installations hosted by "adverse parties", e.g. the Ministry of Public Works and environmental NGOs. Under this setup, the semantic structures defined in affiliate sites may be different (in the case of adverse parties this is the most likely scenario), thus the semantic annotation procedure is bound to produce different results.

- *selective, hierarchical federation.* Again, each installation runs its own set of consultations, some of them being of local interest and some of global interest. Contrary to the previous setup, however, in this case content from global interest consultations are published to an entity conducting consultations at a higher hierarchy level (e.g. company branches publish to company headquarters, EU member states publish to European Commission-level consultations and so forth).

  In a hierarchical federation scheme, content flow is also possible downwards the hierarchy, in order to meet two key functional requirements. Firstly, since the higher levels in the hierarchy consolidate the content of the lower-level installations, the consolidated content may be pushed downstream in order to keep each installation up-to-date with the global view of the consultation. Secondly, specific articles of regulations resulting from consultations at higher levels may be subject to *local arrangements*, which must be settled at a lower level. For example, a European-level law on environment protection may dictate that member states should define which bodies are responsible for monitoring and reporting environmental indicators within their jurisdiction. In such a case, a *triggering event* for the relevant legislative bodies should be raised to indicate that the appropriate procedure should commence.

Figure 3 illustrates two eConWork platform installations that are affiliated using the full replication approach, while *Figure 4* depicts two sites that selectively exchange contents. Best practices and guidelines regarding the federation model to employ for various cases are also included in the eConWork process repository.

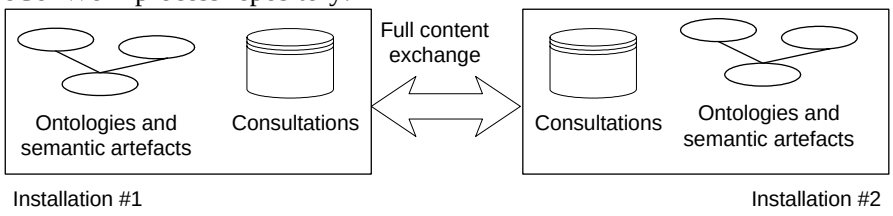

Figure 3. **Full replication with eConWork federation engine**

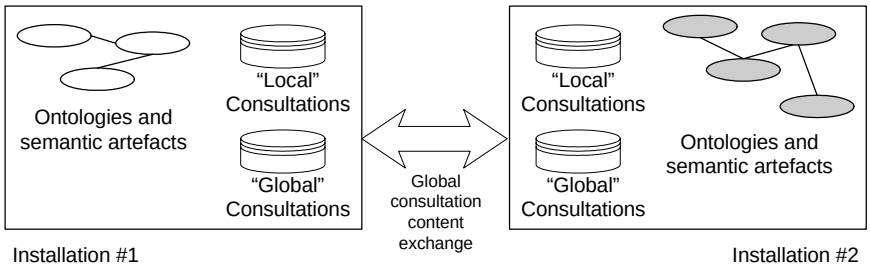

Figure 4. **Selective, peer-level federation**

All layers of the eConWork platform are based on open source software. Integration activities have been scheduled to homogenise the user interface of the various tools, to consolidate all data requirements into a single, shared repository and to provide the necessary hooks to the semantic components, integrated within the eConWork content repository. The eConWork platform will itself constitute an open source project, allowing interested parties to tailor its behaviour to specific needs, substitute functional modules with others that are considered best for the situation at hand, or even integrate commercial components, such as voting systems (e.g. Pnyx [25]).

4. Conclusions – Future work

In this paper we have presented the eConWork platform, which is an integrated environment for hosting public consultations. The eConWork platform uses semantic technologies to support various stages of the consultation process, providing thus targeted notifications, automated or semi-automated document classification, content summarisation, linkage to supporting documents etc, allowing thus stakeholders to participate efficiently in the consultation activities. Consultation designers are provided with tools for tailoring their consultation structure to the specific needs, reference models, consultation templates and best practices to guide them through the design process, plus validation tools that check for common errors in the consultation design. The platform encompasses a federation mechanism, allowing multiple installations to exchange content of mutual interest, while a single installation may host multiple consultations, increasing the platform cost-effectiveness. Future work will include the creation of reference ontologies for selected common consultation topics (education, environment, citizen rights etc), which will be integrated into the platform to further reduce the effort required for installation customisation. The expansion of the provided channel adapter library will also be addressed, since most of the currently used tools, harvested from the Open Source software repository, are primarily addressed to the web channel. Finally, evaluation activities, including all classes of public consultation stakeholders (designers, users, administrators etc), through pilot projects are scheduled.

5. References

1. O'Maley, P, *Design Criteria For Public Consultation Programs In the Federal Government Setting*, 2004. Available at http://www.omalco.com/consultw.htm
2. Timmers, Paul, "Agenda for eDemocracy - an EU perspective". In *The New Agenda for eDemocracy: Lessons from Initiatives Round the World*, Oxford Internet Institute, 2004. Available at http://europa.eu.int/information_society/activities/egovernment_research/doc/agenda_for_edemocracy.pdf
3. Crombez, Christophe, "The Democratic Deficit in the European Union". European Union Politics, 2003, 4, 1: 101-120
4. OMB Watch, *Engaging Citizens Online For Better Policy-making*, 2003. Available at http://www.oecd.org/dataoecd/62/23/2501856.pdf
5. Cohill, M. A., *The Knowledge Democracy – Challenges and Opportunities*, 2001 Available at http://www.knowledgedemocracy.org/library/docs/amc/kd_concepts.pdf
6. Estonian Ministry of Foreign Affairs. *E-Estonia*, 2005 Available at http://www.vm.ee/estonia/kat_175/pea_175/2972.html
7. Connor M. D., "Adapting Public Consultation to Different Political Cultures". In *Annual meeting of the International Association for Impact Assessment in Christchurch*, New Zealand, 1998. Available at http://www.connor.bc.ca/connor/adapting.html
8. Worcestershire County, *Worcestershire Consultation Toolkit*, 2004 Available at http://www.worcestershire.gov.uk/home/cs-con/cs-con-toolkit.htm
9. WelcomeTrust. *Snap Decisions: Novel approaches to public consultation*, 2004. Available at http://www.wellcome.ac.uk/doc_WTD006112.html
10. Kern, T., Kreijger, J., "An Exploration of the Application Service Provision Outsourcing Option". In "Proceedings of the 34th Hawaii International Conference on System Sciences", 2001 Available at http://www.hicss.hawaii.edu/HICSS_34/PDFs/OSTOI04.pdf
11. UK Cabinet Office, *An introductory guide: How to consult your users*, 1998. Available at http://archive.cabinetoffice.gov.uk/servicefirst/1998/guidance/users/index.htm
12. UK Cabinet Office. *How to consult: top ten tips*, 1998. Available at http://archive.cabinetoffice.gov.uk/servicefirst/1998/guidance/users/topten.htm
13. Elliot, R. and Nichols, J., "It's Good to Talk: lessons in public consultation and feedback". In *Police Research Series*, Paper 22, 1996. Available at http://www.homeoffice.gov.uk/rds/prgpdfs/fprs22.pdf
14. Exeter City Council, *Best Practice in Consultation*, 2005. Available at http://www.exeter.gov.uk/index.aspx?articleid=2105
15. Scottish Executive, *Consultation good practice guidance*, 2005. Available at http://www.scotland.gov.uk/Resource/Doc/1066/0006061.pdf
16. Brent Council, *Brent Consultation Toolkit*, 2002. Available at http://www.brent.gov.uk/consultation.nsf/0/1ba5f40a89659aaa80256b28005e24ff
17. Worcestershire County Council, *Consultation toolkit*, 2004. Available at http://www.worcestershire.gov.uk/home/cs-con/cs-con-toolkit.htm
18. Ruger, S. M. and Gauch S. E., *Feature reduction for document clusteringand classification*. Technical report, Computing Department, Imperial College, London, UK, 2004. Available at http://www.doc.ic.ac.uk/research/technicalreports/2000/DTR00-8.pdf
19. Embley, D. W., Ng Yiu-Kai, Xu, L., "Recognizing Ontology-Applicable Multiple-Record Web Documents". In *ER 2001: 20th International Conference on Conceptual Modeling*, Yokohama, Japan, 2001 pp. 555-570
20. Goldstein, J., Vibhu, M., Carbonell, J., Kantrowitzt, M., "Multi-Document Summarization by Sentence Extraction". In *Proceedings of the ANLP'2000 Workshop on Automatic Summarization*, New Brunswick, New Jersey: Association for Computational Linguistics, 2000, pp. 40-48.
21. Mani, I., "Recent developments in text summarization". In *Proceedings of the tenth international conference on Information and knowledge management*, Atlanta, Georgia, USA, 2001, pp. 529–531.

22. Merlo P., Henderson, J., Schneider, G., Wehrli, E. “Learning Document Similarity Using Natural Language Processing”. *Linguistik online*, 2003, vol. 17, available at http://www.linguistik-online.de/17_03/merlo.pdf
23. Hammersley, B. *Content Syndication with RSS*. O'Reilly, ISBN 0596003838, 2003.
24. Grant, G., *Understanding Digital Signatures: Establishing Trust over the Internet and Other Networks*. Computing Mcgraw-Hill, ISBN: 0070125546, 1997.
25. Scytl. *Pnyx product information*, 2003. Available at http://www.scytl.com/