

EVALUATING EDEN'S IMPACT ON PARTICIPATION IN LOCAL E-GOVERNMENT

Angus WHYTE¹ and Ann MACINTOSH²

¹ A.whyte@napier.ac.uk, Research Fellow; ² A.Macintosh@napier.ac.uk Professor of E-Governance
International Teledemocracy Centre, Napier University, 10 Colinton Rd, Edinburgh EH10 5DT, UK

ABSTRACT

The EDEN project is funded through the European Commission's Fifth Framework Programme under the thematic programme 'Systems and Services for the Citizen' which includes R&D projects aimed at e-democracy. The overarching objective of the EDEN project is to stimulate and support citizens' participation in the decision-making process. Urban planning was selected as a test case, since in Europe it has a longer history of citizen participation than most areas of public administration, with widespread adoption of statutes that formalise when it should be undertaken. EDEN integrates a set of Natural Language Processing (NLP) tools into urban planning web-based services, aiming to enhance the accessibility and comprehension of online planning information by citizens and planning professionals, and their acceptance of online consultation procedures. The paper describes the evaluation framework developed in collaboration with public administrations in Antwerp, Bologna, Bremen and Vienna. We describe issues in developing relevant methods and measures of participation, and outline early results from the ongoing evaluation. We discuss the applicability of the OECD's framework for engaging citizens in policy-making, which suggests that effective information provision is a pre-requisite for effective consultation models, which are in turn the basis for e-participation tools based on partnership between citizens and administrations.

INTRODUCTION

Governments are increasingly recognising a need to develop new methods to provide easier and wider access to government information and to achieve broader and deeper involvement of citizens in decision-making. The terms e-government and e-democracy have become widely used in policy-making and academic circles. E-government is typically used as the *more general* term, referring to the use of ICT (Information and Communication Technology) to improve government, with 'e-democracy' used more specifically to mean the use of ICT's to promote citizen engagement, through technology-enabled information dissemination, consultation and participation in the policy-making process (OECD, 2003). The work of the OECD (Organisation for Economic Co-operation and Development) to promote frameworks for developing e-government and e-democracy is by no means unique, and similar work is undertaken in many non-OECD countries. Nevertheless it provides a useful framework for discussing the EDEN project, and we will return to it after first outlining the project.

The EDEN (Electronic Democracy European Network) project is funded through the European Commission's Fifth Framework Programme under the thematic programme 'Systems and Services for the Citizen' which specifically includes R&D projects aimed at e-democracy. EDEN is a collaborative project with public administrations (PAs): Bologna, Antwerp, Bremen, Nisko, and Vienna, along with the Bologna based Osvaldo Piacentini Archive, and with research partners: Omega Generation, International Teledemocracy Centre, Public Voice Lab - PVL, Digipolis, TZI - Center for Computing Technology at the University of Bremen, and Yana Research.

The overarching objective of the EDEN project is to stimulate and support citizens' participation in the decision-making process, specifically in the area of urban planning. The project objectives do not explicitly define 'participation', but they do so implicitly. The objectives frame the expected impact of the tools on participation in terms of more accessible and easier to comprehend and navigate information, provided by the public administrations to enable citizens who may be affected by plans, but do not

normally get involved, to actively use it to make ‘more informed’ questions, comments, requests, or complaints.

EDEN focuses on urban planning partly because it is an area of public administration that has a longer history of citizen participation than most, and widespread adoption of statutes that formalise when it should be undertaken. Urban planning is the basis for a case study, the results of which may be applied in other administrative domains. That is, the tools being developed are not urban planning applications but are intended for broader use.

The project objective is being addressed by developing a set of Natural Language Processing (NLP) tools, that will be integrated into the PA’s (Public Administration) infrastructure for communicating with citizens on matters related to urban planning. The functions of these tools are:-

- Automatic routing to offices of citizens’ messages according to their content
- Automated support for PA staff to manage Frequently Asked Question (FAQ) lists, for citizens to search them, and for their feedback to be used to update the FAQs.
- Style checking tools for planning professionals to make documents easier for citizens to understand, by identifying “difficult” expressions and technical terms and suggesting alternatives from an organisation-wide glossary, including translations of terms into foreign languages.
- Natural language access to databases containing maps and planning documents relating to the city neighbourhoods.
- Discussion forums with opinion polling and notification tools to disseminate information according to a match between discussion themes and user profiles.

The tools, some of which we outline in a little more detail later, are being deployed on a pilot basis in various combinations in the participating cities. At time of writing the pilots are underway, and the authors’ role in the project is in coordinating the evaluation of the pilot’s impact, according to a framework developed in collaboration with the project partners.

It is this evaluation, which draws on previous work (Whyte and Macintosh, 2002) that is the main theme of the paper. It is perhaps surprising that published evaluation frameworks and studies of e-democracy impact are notable for their rarity, indeed the OECD recently commented that “No OECD country currently conducts a systematic evaluation of government performance in providing information, conducting consultation and engaging citizens in policy-making” (OECD, 2001, p.4).

However the lack of established frameworks applies to e-government generally. The evaluation task is often hindered by lack of clarity in objectives, lack of definitions and indicators of success, the complexity of the relationships between stakeholders, and barriers to reporting both failure and success (OECD, 2003). Yet there is widespread agreement of the need for sound evaluation, given the potential impacts of ICT in alignment with organisational change. As Fountain remarks, those impacts raise fundamental and important questions for central concepts of governance such as accountability, task specialization, and jurisdiction (Fountain, 2002).

EDEN is a relatively small-scale project in e-government terms but, as the list of tool functions above suggests, the project potentially impacts on administrative functions and services, particularly the handling of enquiries, that extend beyond the use of discussion fora that are commonly associated with the term e-democracy. We will discuss how the e-government (service related) and e-democracy (citizen engagement) aspects of EDEN are inter-related, through changes in role for the stakeholders involved, and through issues of how citizens represent themselves. We also discuss some of the complexities of the trajectory from information provision through consultation to participation, a central theme of the OECD’s framework.

EDEN as a Trajectory from Information to Participation

The OECD define three types of interaction associated with citizen engagement (OECD, 2001) that have become widely referred to, namely:

Information: a one-way relationship in which government produces and delivers information for use by citizens. It covers both “passive” access to information upon demand and delivers information for use by citizens and “active” measures by government to disseminate information to citizens.

Consultation: a two-way relationship in which citizens provide feedback to government. It is based on the prior definition of information. Governments define the issues for consultation, set the questions and manage the process, while citizens are invited to contribute their views and opinions.

Active participation: a relationship based on partnership with government in which citizens actively engage in defining the process and content of policy-making. It acknowledges equal standing for citizens in setting the agenda, proposing policy options and shaping the policy dialogue – although the responsibility for the final decision or policy formulation rests with government.

These distinctions indicate a scale of ‘engagement’ in policy-making along which government initiatives could be plotted. That is how its authors use it, in reporting that “efforts to engage citizens in policy-making on a partnership basis are rare, undertaken on a pilot basis only and confined to a very few OECD countries” (ibid.). The scale echoes older ‘ladders of participation’ (Arnstein, 1969), and the principle continues to be applied when characterising participation, especially in urban planning (eg. Nobre, 1999).

Applying this principle to the OECD definitions above, Macintosh (2004) describes three levels of participation that can be used to characterise e-democracy initiatives: -

E-enabling is about supporting those who would not typically access the internet and take advantage of the large amount of information available. The objectives we are concerned with are how technology can be used to reach the wider audience by providing a range of technologies to cater for the diverse technical and communicative skills of citizens. The technology also needs to provide relevant information in a format that is both more accessible and more understandable. These two aspects of accessibility and understandability of information are addressed by e-enabling.

E-engaging with citizens is concerned with consulting a wider audience to enable deeper contributions and support deliberative debate on policy issues. The use of the term ‘to engage’ in this context refers to the top-down consultation of citizens by government or parliament.

E-empowering citizens is concerned with supporting active participation and facilitating bottom-up ideas to influence the political agenda. The previous top-down perspectives of democracy are characterized in terms of user access to information and reaction to government led initiatives. From the bottom-up perspective, citizens are producers rather than just consumers of policy (Macintosh et.al, 2002). Here there is recognition that there is a need to allow citizens to influence and participate in policy formulation.

These terms are helpful since the term ‘participation’ can be applied to *any* of the levels of engagement-to refer to the extent that citizens make active use of information, in relation to their use of e-engagement or e-empowering tools (such as online citizens juries, or online petitioning). On this scale EDEN’s city administrations *declared objectives* for the tools relate to ‘e-enabling’ and ‘e-engagement’ - by providing citizens with tools to better prepare them to take part in consultations, and providing administrations with improved information on the citizens’ questions and concerns. If it meets its objectives EDEN may also build capacity for e-empowerment, by demonstrating a potential to sustain collaborative policy-making through its feedback mechanisms. However the policy-making environment for EDEN does not involve the project in any deliberate efforts by the administrations to encourage bottom-up policy-making.

Despite EDEN’s focus on the ‘lower’ end of the scale, we show in later sections that active participation supported by technology, the ‘high’ end of the scale, has in fact been a pre-requisite for the successes in e-enabling and e-engaging. In other words the scale might be interpreted too narrowly as a simple linear one.

THE EDEN TOOLS: AIMS AND ARCHITECTURE

In the Introduction we outlined the functions of the EDEN tools. We describe them in more detail here, focusing in particular on three of seven that are deployed by the pilot sites in various combinations. Two of the three tools, *Answer Tree* and *Address Guesser* entail the *processing* of citizens’ enquiries by computational linguistics technology, or Natural Language Processing (NLP). The *Guided Fora* tool does not use NLP directly but takes the more conventional form of ‘threaded conversation’ found in discussion fora. Both tools formalise the management of texts through which citizens and administration officers represent themselves, but in different ways. Later we will return to the implied change in the roles of actors that this entails, but in this section we focus on the options for communication that the tools offer when integrated into the web sites of the cities involved, and the benefits sought. These are depicted in Figure 1 below.

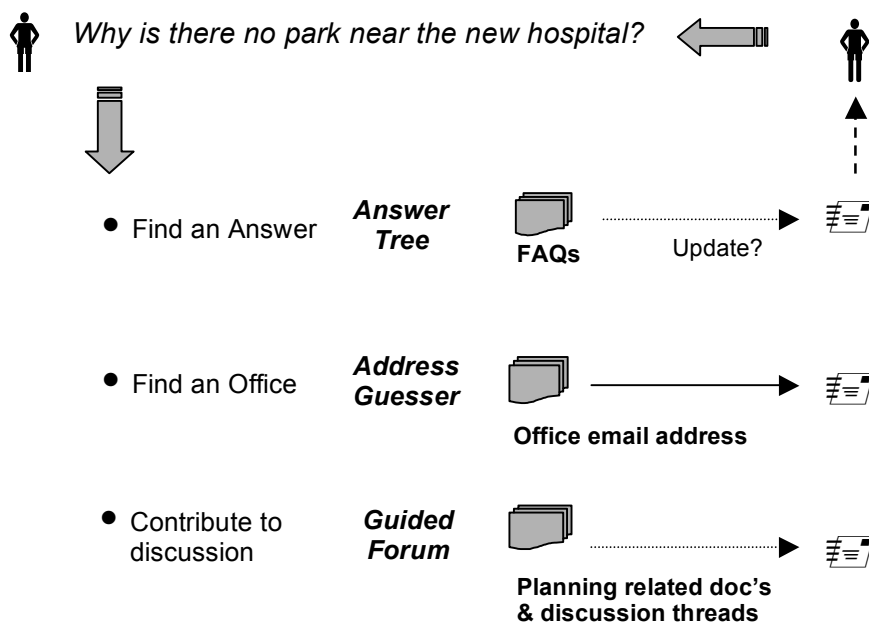


Figure 1. Options for communication between citizen and administrations

EDEN envisages that citizens with a question, comment or complaint (etc.) can choose to go online and send a private message to the administration (via *Address Guesser*), perhaps first checking to see whether their concern is already addressed by a FAQ (Frequently Asked Question) in *Answer Tree*. As the name suggests, this allows a topic tree to be browsed to find an answer. It can also be searched however, and *Answer Tree* can respond to questions in natural language (as in Figure 1) without the user needing to know how to use keywords effectively, or use Boolean operators to combine them.

If a suitable answer is not there, users may email the Administrator of the *Answer Tree* to suggest that it be added, and receive a reply – although probably after the message has been routed to the office specialising in the matter concerned. If the citizen already knows which office that is they may send an email directly to it instead of via the *Answer Tree* Administrator. However many citizens neither know nor care about the administrative structures. In that case *Address Guesser* provides suggestions of the correct office(s), based on a comparison between the content of the message, and that of a compiled set of previous emails answered by each office.

The citizen may however not want, need, or be satisfied with a personal message, or may simply want to know what other citizens think. The *Guided Forum* tool provides threaded discussion, typically focused

on district or neighbourhood-level topics (which may be added to on users' suggestions), and supported by related planning documents.

EDEN thus inter-relates the handling of *private* enquiries, a service typically associated with e-government, with the more *public* discussion normally associated with e-democracy. Benefits are sought for both parties to the communication- citizens (as individuals or civic groups) and for the 'back office' functions and officers, whose role is to respond or proactively seek citizens' feedback. These benefits can be summarised as follows:-

Address Guesser should support enquirers by providing faster and more accurate targeting of their enquiry when they do not know whom to contact. It should support PA offices by providing 'front office' enquiry forwarding functions, by increasing the cost-effectiveness of responding (reduced time per query); and by providing a timelier overview of 'burning issues' according to localities and target groups.

Answer Tree should stimulate an increase in the number of people accessing information on decision-making procedures and outcomes. Its use should allow answers to 'frequently asked' questions about planning to be more easily published on the City website. This should result in more people finding relevant answers to their questions on the website, and more people making better informed choices on whether to contribute their own views about planning matters.

Guided Forum. The PAs aim to increase citizens' engagement in planning, leading to avoidance of (long-term) planning mistakes arising from inadequate participation of those affected by decisions. Online consultations, i.e. discussion fora that are focused on particular planning decisions and which guide citizens through the relevant decision-making phases and background information, should support this in several ways. Firstly in providing additional channels for public discussion between the PA and the citizens, enhancing the capability to involve those legally entitled to a view. Secondly in enhancing the transparency of the PA's response to the citizens' views, by providing summary reports of these that officers/representatives can use in decision-making, and providing citizens with timely feedback on the outcomes. Thirdly, some cost efficiencies are expected to arise from deployment, from decreases in the marginal cost of providing consultation materials on demand.

METHODOLOGY

Since EDEN is a research and development project it remains to be seen to what extent the software tools will be fully deployed by the PAs. Some prototypes have been deployed on an experimental basis for periods of one or two months, and others will continue to be for several months further. Their impact on participation will therefore be relatively time-limited, so it is more appropriate to think in terms of *probable* benefits and risks rather than consequences. Nevertheless, our framework is intended to detect any early impact, as well as to assess the likely benefits for citizens and administrations and the likely barriers to accomplishing improved communication and participation. In this section we describe the methodology used to identify the needs and benefits outlined in the previous section, and how these have been tied to more specific targets, and to risks that may represent barriers to the desired impact.

Key Assumptions

The evaluation framework rests on several methodological assumptions that we should make clear. One of the more fundamental is our preference for a case study approach rather than a statistical experiment. We discuss the reasons for this elsewhere (Whyte and Macintosh, 2002), but briefly the social research in EDEN is exploratory, e-democracy evaluation methods are in their infancy, and there are few previous evaluations that can attest to the validity or reliability of relationships between quantifiable variables. Also, with the EDEN tools deployed on a pilot basis, it is neither desirable nor feasible to try to control the factors that influence their use, especially as the cities energetically promote them both internally and externally. Our aim is rather to add to understanding of those factors, and to use case study strategies for maximising validity, namely the use of multiple sources and methods, and a traceable path from evidence to conclusions (e.g. Yin, 1994)

Our approach is a hybrid of action research, ethnography and evaluation approaches, inspired partly by Suchman and Trigg (1991). The need in our case is to ground our intervention to 'improve'

communication practices, in an understanding of how citizens, their representatives, and planning and communications officers collaboratively accomplish their tasks. Our *action research* role has been to work with some of those actors in each city to elaborate on and clarify the rationale for change, and to coordinate their work with citizens to understand the requirements for the tools. This work draws on Soft Systems Methodology (SSM) (Checkland and Scholes, 1990) and scenario-based methods (Carroll and Rosson, 1992; Erskine et al, 1997). SSM is an approach to collaborative enquiry in the action research tradition, and intended for use in any ‘problematic situation’ that involves ‘would-be problem-solvers’ in reaching an accommodation about feasible and desirable action. The methodology involves describing the ‘relevant systems of human activity’ that systems may support while reflecting on issues that arise from intervening in the situation. It is beyond our scope here to describe SSM, but in general terms it adopts techniques that are typical of participatory design; structured discussion around current tasks and proposed changes to them, seeking accommodation between views of feasible and desirable change.

That involved using scenarios to provide a concrete description of the to-be-designed system in use and the projects’ claims about the feasibility and desirability of change. Scenarios provided *stereotypical* depictions of tasks and interaction, in narrative form, including the social settings, intentions and resources (Nardi, 1991) that we envisaged our would-be actors employing. They were based on prior discussion to identify target groups: for example young people, members of civic associations, architects, call centre officers and other intermediaries the city authority partners envisaged using the EDEN tools. The scenarios then portrayed a fictional urban development undergoing early public consultation. Two versions were prepared, the first depicting events from the perspective of fictional planners and officers dealing with enquiries and managing consultation procedures, and the second depicting those events from the perspective of the various ‘citizens’.

We then carried out a claims analysis of the responses to the scenarios (Erskine et al, 1997), i.e. we related the benefits and consequences that the scenario claimed EDEN tools could have, to how the scenario readers supported or criticised them, or offered alternative proposals. In parallel with the discussions this involved, we surveyed citizens about communications issues and their interests in EDEN. These materials provided a broad understanding of what was thought feasible and desirable about the design proposals, and why.

The role of *ethnographic* methods in the project has been relatively small, but important to understanding how communications are currently accomplished. Ethnography, the descriptive method of anthropology, has been increasingly adopted in studies of collaborative work for systems design (see e.g. Crabtree et al, 2000). It’s role is based on the assumption that, although people may well be capable of identifying the improvements to current practice that they want from a system, the day-to-day practices of collaboration that the system would change are often not given enough attention, and critical design constraints and opportunities may be missed. In systems design contexts ethnography normally involves observations, semi-structured interviewing, and detailed analysis of recorded interactions between the actors involved in accomplishing work in its everyday setting. In our case we have relied on samples of emails and contributions to discussion fora, alongside the accounts given by planning professionals and communications officers in semi-structured interviews about their work. These have informed our development of the system requirements, together with the evaluation indicators and the deployment risks and issues that we describe below.

EVALUATION QUESTIONS, CRITERIA AND METHODS

The main evaluation questions then are:-

1. Do the NLP-based information retrieval tools provide relevant answers to citizens’ questions?
2. Do the pilots demonstrate the anticipated improvements to online access, navigation, and comprehension?
3. To what extent are the tools accepted by citizen and PA users, and why?
4. Do EDEN tools better enable citizens to contribute views on their neighbourhood?

Answers to these key questions will tell us whether or not EDEN meets its aims, but there are also important issues that we need to address to give an account of how and why it does so:-

- What comparisons can be made between the traditional offline methods and online methods with EDEN?
- Are the skills & capacity available to use the tools effectively, from citizens' and PA perspectives?
- For the PA, what are the conditions that need to be met for full deployment of the EDEN tools?

The third of these issues is currently being addressed in the final stages of the project and we do not have the scope to discuss it here. However the first two have turned out to be key ones, affecting the answers to each of the evaluation questions listed above, which we now describe in a little more detail along with the criteria and methods we use to establish whether the project is indeed 'enhancing participation'.

Evaluating whether NLP tools give relevant results

To answer our first evaluation question we can draw on established practice in assessing the performance of information retrieval systems. Both *Answer Tree* and *Address Guesser* can be considered as question answering systems, and the standard measures used to assess these are *precision* (the number of relevant answers in the results set, divided by the total number of answers returned), and *recall* (the number of relevant answers in the results set, divided by the number of relevant answers present in the system). A variant of the precision measure is the *reciprocal rank score* (Radev et al, 2002). This takes account of the position of relevant answers in the results set, i.e. it gives more weight to a results set that has a relevant answer shown first than shown second (etc.). The reciprocal rank score is the sum of the reciprocal ranks of all the relevant responses. For example, if a query gives a set of 3 question-answer pairs, and only the second and third are relevant, the score is 0 for the first, 1/2 (0.5) for the second, and 1/3 (0.33) for the third, giving a reciprocal rank score of 0.83. Normally the tests for reciprocal rank (and for recall) are carried out for 50-100 test questions, and a mean score obtained.

These tests were conducted, first of all, as an 'internal' quality check to ensure that the data (FAQs for *Answer Tree*, office email samples for *Address Guesser*) had been properly set up to get the best results from the NLP software resources. The latter include software representations of grammatical rules (for Italian, German and Dutch/Flemish) and glossaries covering urban planning terms that, when used to parse the users' message texts, should distinguish the relevant terms and syntactic structures (e.g. 'noun phrases'). Since the rationale for using NLP in the project is that retrieval should produce good results with natural language questions like that in Figure 1, the validation targets were high. For example for *Answer Tree* we specified 95% for recall, and 0.87 for reciprocal rank, which translates as 75% of results for test queries should show a relevant FAQ first, and the rest should show one in second place.

Two important points about these measures are, firstly, that they are dependent on a more or less subjective judgement about the relevance of the results to the question that was asked, so the test results are dependent on the testers and the system being tested. Secondly, the tests are normally carried out by information retrieval specialists in laboratory settings rather than by ICT staff in city councils, and normally using questions that are worded to match the capabilities of the system and the information being tested.

This last point is a critical issue, since it highlighted differences in expectation between the testers (city partners) and the software suppliers (technical partners). For reasons that we return to in the next section, the targets for evaluation with real users' questions were lower, even though the internal validation gave excellent results.

Evaluating whether information provision leads to improved acceptability

Acceptable information retrieval performance is a necessary but insufficient condition for the aims of EDEN. Whether or not an information retrieval system achieves high precision and recall in relation to the test of a query, what users (citizens) are interested in is whether or not the results help them meet their overall information-seeking needs (Spink, 2002). For EDEN, what we are interested in is whether those needs coincide with the public administrations' requirement to get more informed questions and comments on matters related to urban planning.

The EDEN tools are intended, as we have said, to enhance access, navigation and comprehension. These can be considered as intervening conditions for meeting the larger goal of increased acceptance of online tools by citizens wanting a say in decision-making, and by policy-makers responsible for those decisions. These criteria need more exact working definitions, and to arrive at them criteria and targets were drafted and discussed among the city partners, and importantly with citizens. As a result of these discussions, indicators and targets were developed based on the following basic definitions:-

Access refers to how much information is made available, who accesses it, when and how often, and how much time is needed to do this. When assessing this, we emphasise access by people who say they normally do not get involved in commenting on matters affecting their neighbourhoods.

Navigation and Comprehension are both concerned with active use of the EDEN tools after access is made. That is, once users have gained access, do they manage to find documents that help, or offices to contact? And what do they make of the outcomes; in terms of being able to comprehend and act on what is said?

Acceptance is concerned with the perceived *legitimacy* of the tools and their content as media for online participation in urban planning.

The indicators and targets were expressed as statements, broken down from each of these criteria and related to each of the tools. One example for the *Guided Fora* was “More people making better informed choices on whether to give views on planning”. These statements were then re-worded to relate to each of the overall criteria above, and so that they could be assessed quantitatively and qualitatively from the available sources of data. This resulted in between 4 and 8 indicators per tool, examples of which are shown below in Table 1.

Tool	Criteria	Indicator	Sources
Address Guesser	Comprehension	Officers can answer routed messages as easily as direct e-mails to office	Message samples, questionnaire, discussion
Answer Tree	Navigation	Fewer than 25% of questions forwarded to the Administrator are already in the Answer Tree	Questionnaire; discussion
Guided Forum	Acceptability	Moderator ratings of contribution quality higher than previous fora Majority of citizen users feel more involved in decision-making	Ratings Questionnaire; discussion

Table 1: Examples of Evaluation Indicators

The indicators need some further explanation, firstly about the sources and secondly about the basis for the indicators themselves. The main sources for the evaluation are:-

- Samples of messages between citizens and officers; and contributions to discussion fora.
- Discussion with officers and citizens invited to ‘user panel’ group interviews, and individually.
- Observation of users trying the tools, and analysis of usability problems and their severity.
- Questionnaires: both in print form, with user panel participants, and online to allow any user to respond.
- Log files: the tools log all queries to *Answer Tree* and *Address Guesser*, and the responses provided by the system (FAQs and office addresses, respectively). Web server log files also provide details of page requests and visits to the

When we established the requirements for *Address Guesser* we analysed samples of e-mails between citizens and officers responsible for responding to queries to generic office email addresses. It was clear from these, and substantiated from discussions with the officers, that such emails were often not formulated so that what the citizen really wanted to know was stated in their first message. This was

significant because an assumption of the *Address Guesser* use of NLP technology is that messages are worded sufficiently clearly to identify the key terms that represent what the user wants to know, and which identify the office to which their message should be routed. When a correct office cannot be automatically determined, the Address Guesser sends it to a ‘default office’, responsible for handling general enquiries, who can then manually forward it. Even then, however, there is a risk that the recipient office may not know how to help without asking the user for clarification.

Answer Tree has a similar route to a general enquiry handling office, the “Administrator”, as shown in Figure 1. A key design assumption is that the messages that citizens send are not those that the user could find by browsing the FAQ ‘tree’, but instead are non-routine questions that may be useful for new FAQs, and alleviate the enquiry-handling function of the more routine ones. But if the search function has low performance, or if users cannot easily browse to find an answer, there is a risk the Administrator could be inundated – in effect a risk of success on the ‘access’ criterion at the expense of failure on ‘acceptability’.

For the *Guided Forum* tool, the criteria are consistent with the interest of the Bremen city partners who are piloting it, in promoting *deliberative* discussion of local plans and related issues. Concepts of deliberative democracy underlie much of the benefits sought for e-democracy, and in EDEN’s case are translated into a rating scale used by the discussion moderators to assess the contributions that citizens make online, as follows below (Westholm, 2003).

- ++ opinion and constructive proposal, reply of an expert
- + opinion, proposal, information regarding a fault/problem
- 0 opinion, wish
- opinion and unproductive proposal
- box for complaints, swearing, insults

These ratings allow an assessment of whether the forum outcomes are acceptable from the perspective of the administration. Our interest is also however in whether, from the citizens’ perspective, the experience of contributing leads them to feel more involved in local policy making, or whether they see other (met or unmet) benefits in participating. These are assessed partly from the results of an online questionnaire, and partly from discussion with users willing to take part in user panels.

EARLY RESULTS AND CRITICAL ISSUES

Since the pilots are ongoing until January 2004 we cannot yet report results that answer all of the evaluation questions. Nevertheless some critical issues have emerged that, as we discuss in the conclusions, affect how we place EDEN in the OECD’s framework for citizen engagement. Both for the NLP-based tools and the discussion forum, the issues revolve around the capacity of the administration to respond to citizens’ input, and in turn on how citizens represent themselves online.

Finding Answers and Offices: Gaps between NLP and language-in-use

The *Answer Tree* and *Address Guesser* tools have both, as we remarked earlier, been validated to assess their performance using the standard criteria of recall and precision (or reciprocal rank). In each case they achieved results on test query sets approaching the high targets set for them. However the tests proved controversial, not because of the results but because of the criteria set by the technical partners to compile the test queries. This resulted from differences in perspective about what constitutes ‘natural language’.

For the technical partners, as computational linguists, ‘natural language’ is a theoretical concept representing the set of all ‘well-formed’ sentences of a human language (English, Italian, etc.). The notion of ‘well formedness’ depends on the notion that a phrase or sentence can be recognised as conforming to a set of grammatical rules that specify the possible syntactic structures of that language (Moreschini et al, 2003). To take an example “one cats rans over the carpet” would fail this test, even although a human would have little trouble interpreting it sufficiently well to act competently – even if only to ask for clarification.

For the city partners however, ‘natural language’ was taken to mean the language actually used by citizens. So when required to re-write the test FAQs (for *Answer Tree*) and sample email messages (for *Address Guesser*) to conform to a ‘well formed’ sentence structure their concern was that the system would not perform adequately with the queries submitted online by users. To take a further example contrasting with that in Figure 1, consider a message stating “Park on my street? You must be joking! But your plan permits hundreds of people to do that”. This statement can reasonably be understood by a human to be about ‘parking’ rather than ‘a park’, but not about ‘parking permits’. It remains to be seen whether the NLP tools will cope with such ambiguities sufficiently well to provide the relevant FAQs or office addresses.

Since queries that the NLP tools cannot handle are routed by default to the administrations’ general enquiry handling staff, variations in performance of the NLP tools have critical effects on the skill and capacity required from the administrations, and in turn their acceptance of the tools. To achieve NLP performance that is consistent with EDEN tools’ aims of offering an effective alternative to telephone enquiries, itself depends critically on how citizens express their questions and comments.

Informed Participation in Guided Fora

The ‘quality ratings’ used to assess contributions to discussion fora similarly raise issues about how citizens represent themselves, and the administrations’ capacity to respond. The pilots of this tool in neighbourhood consultations in Bremen have been intended to complement the conventional procedures for ‘early consultation’ on plans, which involve public meetings and the option to contribute in writing. Despite political support for the pilots, they came at a difficult time for the planning officials who would be required to assess the contributions.

Partly in anticipation of the constraints on capacity, and partly to increase citizens’ ownership of the process, the Bremen partners adopted a strategy of recruiting local citizens to perform the role of moderator. Citizens were recruited through publicity in local newspapers, and trained in the tool functions and editorial tasks (checking that contributions adhered to an acceptable use policy, checking that they were on topic, etc.).

The collaboration of the citizen volunteers was considered highly successful, not only to reducing the costs for the administration but also to publicising the discussions locally. The level of response has in both of the two pilots carried out been higher than that received by the traditional channels, and the quality of the responses, as assessed according to the criteria for deliberative quality given above, was considered high (further details are given in Westholm, 2003).

It is important to note two things. Firstly, the *Guided Forum* pilots have been conducted without the planned-for integration of the *Answer Tree* tool, which is about to be piloted in Bremen and Bologna. Instead the Bremen team provided a selection of links to documents providing minutes of planning meetings, technical documents (zoning codes) and overviews of urban planning in Bremen. However the web server logs indicated that few of the visitors to the site had downloaded these documents. Secondly, although the response of elected representatives and the city administration was positive, feedback on the action taken to respond to the citizens input is still awaited, and the acceptance of the tool in ‘early stage’ consultation procedures remains to be assessed in the final stages of the pilot phase.

FINAL REMARKS: NO ENGAGEMENT WITHOUT REPRESENTATION

To discuss ‘conclusions’ may seem premature since the evaluation of EDEN is incomplete, but it is already clear that lessons can be drawn from the project that we can summarise in this final section.

Participation, but at what level?

In the Introduction we discussed the OECD’s characterisation of a scale of citizen engagement, from information provision to partnership, and provided a corresponding characterisation of e-democracy initiatives, as e-enabling, e-engaging, and e-empowering. We continue to find this scale a useful conceptual tool. EDEN as we remarked earlier is conceived as a set of e-enabling tools that, when integrated with online discussions for e-engagement, lay the groundwork for closer relationships with the online citizenry that may be described as e-empowering.

There is however a risk that the framework may be taken to imply that administrations should build e-democracy tools from the bottom end of the scale upwards. In other words that information provision must be ‘got right’ before it can enable e-engagement with online consultation models, which in turn must be proved effective before anything as bold as ‘e-empowerment’ can be considered. Our overview of the EDEN *Guided Forum* pilots in Bremen has not tackled many of the evaluation questions described in our framework. Yet what it has illustrated is that an effective *partnership* with citizens, i.e. the volunteer moderators, was critical to achieving credible responses from other citizens in the neighbourhoods concerned. Moreover, this was accomplished without the anticipated integration of highly accessible, easy to understand and navigate background information.

The possibility remains of course that, with the integration of the ‘e-enabling’ NLP tools, the participation in online discussions and the nature of enquiries to the relevant offices will be richer than would be the case otherwise. That is what we are assessing in the remainder of the project. However the scale of engagement portrayed in the OECD and similar frameworks is, we suggest, better understood as the extent to which the individual and collective needs of citizens on the one hand, and administrations on the other, are mutually negotiated and accomplished. The notion that the scale should represent the *intentions* of either administrations or citizen groups to form online partnerships is inadequate, simply because such intentions may not be realisable.

Citizen representation in and through e-government

The second conclusion we draw is that the response of administrations to citizens’ online input depends critically on how citizens are enabled to represent themselves online. This applies both to our *Guided Forum* example, and to the NLP-based *Answer Tree* and *Address Guesser* examples. In the case of the *Guided Forum*, citizens represent themselves through their discussion contributions, and are assessed according to the deliberative nature of those contributions. There are of course issues about how such assessments are made, by whom, and how transparently. However regardless of the criteria used, the effect is to filter the contributions and limit the onus on the administration and elected representatives to act on them. That is, perhaps, the nature of politics whether the dialogue takes place online or offline.

Discussion fora are the archetypal form of e-democracy initiatives. Yet it is also clear from our discussion of the NLP-based modules that similar ‘politics’ prevail in the shift of enquiry-handling roles from human operators to software. The nature of the gap in performance between our validation, using syntactically ‘well formed’ queries and information, and the evaluation, where citizens will enter queries of their own choosing, remains to be seen. It is clear though that a gap can be expected, and that ‘well-formedness’ represents a threshold that, if it is set too high, has implications for how the administration organises itself to respond, and its’ capacity to do so. Moreover it shows the dependence of such e-enabling tools on citizens’ ability to articulate what they want.

Effective use of online information retrieval tools, for e-government purposes or any other, commonly depends on the skills of users, and we remain confident that the EDEN tools will reduce that burden. Our points are that, firstly, our awareness of the nature of the risks and benefits has depended on access to information about how citizens currently represent themselves in their communications, and to citizens themselves as participants in the evaluation.

Finally and related to that point, the need for citizens to be enabled and engaged in the development of e-governance systems transcends the divide between ‘e-democracy’ and ‘e-government’. We might draw an analogy here with the principle of ‘no taxation without representation’. Taxation, now a typical e-government ‘service’, was resented by the 18th century Bostonians because it was not sufficiently tied to rights of political consent and representation. One of the many challenges of e-governance is to ensure such ties are maintained online.

REFERENCES

- Arnstein, S. (1969) A ladder of citizen participation. *Journal of the American Institute of Planners*, 35, pp. 216-224.
- Carroll, J.M and Rosson, M.B (1992) Getting round the task-artefact cycle: how to make claims and design by scenario. *ACM Transactions on Information Systems*. 10, pp. 181-212
- Checkland, P. and Scholes, J. (1990) *Soft Systems Methodology in Action*. Wiley, Chichester, UK.
- Doubleday, A., Ryan, M., Springett, M., and Sutcliffe, A. (1997) A comparison of usability techniques for evaluating design Proceedings of DIS '97 Amsterdam (NL); ACM, New York.
- Erskine, L., Carter-Tod, D., and Burton, J.K. (1997) Dialogical techniques for the design of web sites. *International Journal of Human-Computer Studies*, 47, pp. 169-195
- Fountain, J. (2002) Toward a Theory of Federal Bureaucracy for the Twenty-First Century. In: *Governance.com: Democracy in the Information Age*, Ed. E. Karmarck and J. Nye, Brookings Institute Press, Washington, D.C. pp. 117-140.
- Macintosh, A. (2004) Characterizing E-Participation in Policy-Making Proceedings of the Hawaii International Conference on Systems Sciences HICSS-37 January 5-8, 2004., Hawaii (to appear).
- Macintosh, A., Davenport, E., Malina, A.; and Whyte A.(2002) Technology to Support Participatory Democracy. In *Electronic Government: Design, Applications, and Management*. Ed. Åke Grönlund. Idea Group Publishing, January 2002; pp. 223-245.
- Moreschini, P. Vanocchi, M. and Carenini, M. (2003) EDEN: The gap between us and them – Introducing NLP in eDemocracy. In. *AI to Support E-Government. Workshop in conjunction with AI-2003*. Peterhouse College, Cambridge, UK. 15 December 2003 (to appear).
- Nardi, B.A. (1991) The use of scenarios in design. *SIGCHI Bulletin*, 24(4). pp. 13-14
- Nobre, J. (1999) Improving community participation in urban planning decision making. In. *Public Participation and Information Technologies 1999* Ed. Pedro Ferraz de Abreu & João Joanaz de Melo, CITIDEP & DCEA-FCT-UNL.
- OECD (2001). *Citizens as Partners: Information, consultation and public participation in policy-making*: OECD, Paris.
- OECD (2003). *The e-Government Imperative*. OECD, Paris.
- Radev, D., Libner, K., and Fan, W. (2002) Getting Answers to Natural Language Questions on the Web *Journal of the American Society for Information Science and Technology* 53(5), pp. 359-364
- Spink, A. (2002) A user-centered approach to evaluating human interaction with Web search engines: an exploratory study. *Information Processing and Management* 38(3): pp. 401-426
- Suchman, L. and Trigg, R. (1991) Understanding Practice: Video as a Medium for Reflection and Design In: *Design at Work: Cooperative Design of Computer Systems*. Ed. Greenbaum, J. and Kyng, M. Lawrence Erlbaum Associates, Hove, U.K.
- Westholm, H. (2003) “Adaptability” in online democratic engagement: a multi-channel strategy to enhance deliberative policies. *Communications*. 28. pp.205-227.
- Whyte, A. and Macintosh, A.; (2002) Analysis and Evaluation of e-consultations; *e-Service Journal*; Volume 2, No 1 “e-democracy in Practice”; pp. 9-34.
- Whyte, A. and Macintosh, A. (2003) Representational Politics in Virtual Urban Places. *Environment & Planning A*. 35(9) pp. 1607 - 1627
- Yin, R.K. (1994) *Case Study Research: Design and Methods*. Sage. Beverly Hills, CA.