IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# DEMO-net: D 5.2

## eParticipation: The potential of new and emerging technologies


**Editor :** **Claudia Soria, CNR-ILC**

**Asta Thorleifsdottir, UI**

**Revision :** Jacob Nörbjerg, Clive Sanford, Asta Thorleifsdottir and Maria Wimmer


**Dissemination Level :** [TA p. 63]


**Author(s) :** Dimitris Apostolou*, Frantisek Babic°, Georgia Bafoutsou*, Peter Butka°, Spyros Dioudis*, Marian Mach°, Ann Macintosh°°, Tom Gordon[++], Christos Halaris°, Kostas Kafentzis°, Gregoris Mentzas*, Marek Paralic°, Jan Paralic°, Alastair Renton°° Andreas Rosendahl[#], Tomas Sabol°, Christian Schneider[#], Asta Thorleifsdottir[+], Maria Wimmer[#]

* TUK - Technical University of Kosice (SK)

° ICCS -Institute of Communication and Computer Systems (GR)

[+] UI- University of Iceland (IS)

[#] IWVI - University of Koblenz-Landau (DE)

°°Napier University Scotland

[++] Frauenhofer institute, Germany



**Due date of deliverable :** 31st December 2006

**Actual submission date :** 10th of February 2007

**Start date of project :** 01 January 2006

**Duration :** 4 years

**WP no.:** 5

**Organisation name of lead contractor for this deliverable :** CNR-ILC

**Abstract:** *The public sector is dealing with information and knowledge resources by large. The challenges facing increased public participation in societal decision making are to a great extend due to the large scale of information flow, the various formats of input and the organisation of the information and data into useful knowledge that will ensure feedback and more public involvement. The aim of this report is to cast a light on the potential of new and emerging technologies to lower the barriers and face the challenges of eParticipation.*

**Project funded by the European Community under the FP6 IST Programme**

**The DEMO-net Consortium consists of:**

| | | |
|---|---|---|
| County of North Jutland – Digital North Denmark | Coordinator | Denmark |
| University of Leeds | Partner | United Kingdom |
| Örebro University | Partner | Sweden |
| University of Koblenz-Landau | Partner | Germany |
| Fraunhofer Gesellschaft zur Förderung der angewandten Forschung e.V. | Partner | Germany |
| Institut für Informationsmanagment Bremen GmbH | Partner | Germany |
| University of Macedonia | Partner | Greece |
| Institute of Communication and Computer Systems | Partner | Greece |
| Copenhagen Business School | Partner | Denmark |
| Aalborg University | Partner | Denmark |
| Fondation National des Sciences Politiques | Partner | France |
| Technical University of Košice | Partner | Slovakia |
| Consiglio Nazionale delle Ricerche | Partner | Italy |
| University of Bergamo | Partner | Italy |
| Yorkshire and Humber Assembly | Partner | United Kingdom |
| European Projects and Management Agency (EPMA) | Partner | Czech Republic |
| Napier University | Partner | United Kingdom |
| University of Iceland | Partner | Iceland |
| University of Helsinki | Partner | Finland |

# History

| *Version* | *Date* | *Modification reason* | *Modified by* |
|---|---|---|---|
| | 01.07 | Internal review of whole document | Clive Sanford |
| | 01.07 | Internal review of whole document | Jacob Nörbjerg |
| | 02.07 | Insertion of reviewers comments<br>Writing of executive summary<br>Writing of Introduction in Umbrella document<br>Compilation of conclusion section | Asta Thorleifsdottir |
| | 11/02/07 | Insertion of updates in sub-del. on ontologies, SWS, and KM based on reviewers comments;<br>Insertion of the chapter on devices and channels;<br>Updates in Introduction of the umbrella document<br>Insertion of executive summary on devices and channels in chapter 2 of umbrella doc;<br>Updates in executive summaries on KM, ontologies and SWS. | Maria A. Wimmer |
| | 20/02/07 | Substitution of D5.2.2 with newer part | Claudia Soria |

# Table of contents

# Executive Summary

The public sector is dealing with information and knowledge resources by large. The challenges facing increased public participation in societal decision making are to a great extend due to the large scale of information flow, the various formats of input and the organisation of the information and data into useful knowledge that will ensure feedback and more public involvement. The aim of this report is to cast a light on the potential of new and emerging technologies to lower the barriers and face the challenges of eParticipation to enhance public participation in policy and decision-making.

The report is aimed at socio-technical researchers. This is the group that will most likely make use of the information we have collected for future development of eParticipation tools.

The report tackles the following areas as sub-deliverables that are considered of importance to the advancement of eParticipation:

- Collaborative Environments
- Argumentation Support Systems
- Ontologies
- Semantic Web Services
- Knowledge Management and Knowledge Engineering
- Devices, Channels and Mobile Technologies

All of the sub-deliverables cover a broad range of research within their own particular domains, and serve as a good foundation for future endeavours to apply them to eParticipation.

Every section is built up in a similar manner for extended readability. An overview of each sub-deliverable is given in chapter 2.

# 1 Introduction and Overview of the deliverable

The use of information and communication technology (ICT) based tools is increasingly being explored to enhance participation in decision-making processes in Europe. The tools and technologies used so far have mainly depended on text-based documents rather than making ultimate use of newly introduced technology.

To discuss the potential of new and emerging technologies to eParticipation is no easy task. There is a myriad of interesting innovation and development taking place in the ICT world all of which could and can have an impact on eParticipation.
The joint research activity 5 (see TA, p. 25f) aims at investigating eParticipation tools and technologies. Task 5.1 investigated existing eParticipation tools. The aim of the D5.1 report was to identify and describe the current use of ICT tools and technologies to promote and enhance participation. By assessing the current ICT use in eParticipation contexts, a baseline was formed for further advances in the area by developing an assessment framework to describe ICT methods, tool categories and technologies currently used in eParticipation. These tools can involve a considerable variety of underpinning current and emerging technologies (see deliverable D 5.1).

The deliverable at hand aims at investigating current and future emerging technologies to support eParticipation tools and applications.
In p. 25f, the TA stresses that especially the following questions should be investigated:

- How and to what extent can natural language technologies support eParticipation?
- How and to what extent can speech technologies support eParticipation?
- How and to what extent can text mining technologies support eParticipation?
- Are embodied conversational agents appropriate interfaces for eParticipation?
- How can multi-agent systems support eParticipation?
- How and to what extent can ontologies and semantic web services support eParticipation?
- Is Computer Supported Argument Visualisation relevant and if so how?
- Is there a need for CCSW technology and if so where?

Depending on the tool and its use, eParticipation services are being delivered through a variety of channels and devices such as PCs, digital TV and mobile phones. In this report we have selected a few areas of investigation and questions listed in the TA[1] to start with, leaving out others to encounter at a later state. The reader may ask him- or herself why new technologies relating to GIS based systems, three dimensional representation, visualisation; areas of artificial intelligence and so on have been left out. WP 5's task 5.2 is on identifying and assessing emerging technologies and their potential to support eParticipation, Research is required on emerging technologies to better

[1] The TA at p. 25f addresses the whole time-span of the Demo-net project, while the deliverable D 5.2 reports first results of investigations.

understand their potential to promote and support eParticipation. The technological challenges are related to aspects of scale, inclusion, understandability and management.

Under the given timeframe and resources available, the partners in the workpackage decided to select the following themes for emerging technologies:

- Collaborative environments
- Argumentation Support Systems
- Ontologies
- Semantic Web Services
- Knowledge Management and Knowledge engineering
- Channels, Devices and Mobile Technologies.

These are only a few areas for this first report, the other wait for the 2nd time around. The selection is not based on prioritisation but rather to clarify the potential of the technologies already emerging in eParticipation.

In the following, the structure of the document is detailed.

## 1.1 Structure of the document

This reports concern is identifying emerging technologies and their potential to promote and better understand eParticipation. The method chosen is the production of *detailed reports* for each of the technologies identified. The result is a draft policy documents on each emerging technology area following the structure suggested below.

D5.2 consists of six reports or sub-deliverables, in their own right labelled D5.2.1, D5.2.2, etc. The sub-deliverables are introduced by a short overview in chapter 2 of the umbrella document. The six executive reports highlight the findings of each sub-deliverable. The umbrella report is closed by an analytical conclusion chapter focusing on the relationships and overlaps between the technologies, as well as pointing other technologies that need further analysis.

The report deals with the following areas and organisations based on the survey results presented in Demo-Net's Workpackage 4.

- Collaborative environments (D 5.2.1)
- Argumentation Support Systems (D 5.2.2)
- Ontology (D 5.2.3)
- Semantic Web Services (D 5.2.4)
- Knowledge Management and Engineering (D 5.2.5)
- Devices, Channels and Mobile Technologies (D 5.2.6)

These areas are covered in sub-deliverables almost all following the same structure described below:

**1. Overall Description of the Technology**

Includes a detailed description of what the technology is, how it has evolved and key research questions which this technology needs to solve in order to progress further. This is a comprehensive account of the technology in question, not specifically addressing eParticipation.

**2. Examples of research groups and what they are researching**

Detailing the key international research groups in the particular technology area, this section will list the specific subfields which particular institutions concern themselves with and their general levels of expertise.

**3. Current Applications of Technology. Includes status of applications and current constraints – what can and cannot be done. This applies to any industry sector.**

Presents specific examples of the technology being used in particular applications. It details for each instance to what extent the application is constrained by current technology. Where the technology is quite limited, an account of why it may be limited may be given.

**4. Current Applications of technology in eGovernment, including eParticipation.**

A technology may not have reached the same level of maturity in eGovernment as it has done in other application areas. This section deals with to what status the technology is being deployed in eGovernment applications, along the lines of section 3.

**5. Future potential uses of existing technology for eParticipation.**

This section addresses the potential to which the technology may be used to enhance eParticipation. Linked to the original outline in section 1, it identifies where the current state of the art may be used in particular eParticipation activities. It also details what research is required to achieve these scenarios, and how DEMO-net should harness its resources towards these goals.

## 1.2 Executive Reports

This chapter gives an overview of the key results of each sub-deliverable and the importance and potential of the specific technology discussed.

### 1.2.1 D5.2.1 Collaborative Environments

Electronic collaboration has a strong potential to support distinct participation areas and different stakeholders in the various stages of eParticipation. This potential is being examined in the context of this report.

The structure of the section report is as follows:

Chapter 1 is a brief introduction in the e-collaboration technology, including definitions found in the literature.

Chapter 2 provides an overview of the electronic collaboration technology, identifying the typical functions of collaborative systems, presenting classifications of systems and defining collaboration dimensions.

Chapter 3 gives a brief description of research groups dealing with electronic collaboration.

Chapter 4 presents applications of the technology in the business world and e-government. Electronic collaboration can be applied inside the enterprise as well as in the context of inter-organizational relationships, potential that has also been explored by several European projects.

In the context of electronic government, e-collaboration helps to implement electronic solutions as far as all three interfaces of government are concerned: government to government (G2G), government to business (G2B), and government to citizens (G2C).

Finally, Chapter 5 identifies applications of electronic collaboration in the field of eParticipation. Collaborative environments have a strong potential to support e-participation, as they offer a variety of tools for synchronous and asynchronous communication and collaboration among several participants.

A number of projects launched under the fifth Framework Programme (FP5) will try to promote and enable the online participation of all stakeholders in decision making, and improve the interaction between citizens and public administrations, through e-collaboration functionality.

Some of the eParticipation areas, described in Deliverable 5.1, are considered suitable for applying e-Collaboration technology. Section 2 of Chapter 5 presents some use scenarios in the Community building / Collaborative Environments, Electioneering, Consultation and Discourse areas of eParticipation.

### 1.2.2 D5.2.2 Argumentation Support Systems

Argumentation Support Systems are computer software for helping people to participate in various kinds of goal-directed dialogues in which arguments are exchanged. Their potential relevance for eParticipation should be readily apparent, since the goal of eParticipation is to engage citizens in dialogues with government about such matters as public policy, plans, or legislation. Surely argumentation plays a central role in this process. In a public consultation, for example, citizens are given an opportunity to not only make suggestions, but also support these suggestions with arguments.

Typically eParticipation projects make use of generic groupware systems, such as discussion forums and online surveys. These generic groupware systems, however, do not provide specific technical support for argumentation. For example, they provide no way for a citizen to obtain a quick overview of the issues which have been raised, to list ideas which may have been proposed for resolving such issues, to see in one place the pro and con arguments of these proposals, or to get an idea about which positions currently have the best support given the arguments put forward thus far in the dialogue. These are just a few of the kinds of services offered by argumentation support systems.

This report provides an introduction to the theory of argumentation; summarizes prior work of the leading research groups on modelling argumentation and supporting argumentation with software tools; describes various prior applications of argument support systems, mostly in research pilot projects; and presents a number of eParticipation application scenarios for argumentation support systems, as a source of ideas for future pilot projects.

A number of argumentation support systems and associated tools are presented. Some of them focus on the visualization of arguments and here the graphical notation and user interface are important features. Others focus on providing analysis of the situation but typically with a more limited graphical user interface. A number of underlying argumentation models are used including those based on Issue-Based Information Systems (IBIS) and the diagramming method developed by Wigmore for mapping evidence in legal cases. In considering their relevance to eParticipation, we need to consider the features needed to support informed debate in order to support evidence-based policy-making. The systems presented allow users to access various levels of information, to be able to focus on specific information and to have the ability to organize the gathered data to construct an effective argument – all of which are required for eParticipation.

In eParticipation, there is a clear requirement to better understand how technology can support informed debate on issues but there are two main obstacles in achieving this. The first is that the deliberation is typically on complex issues and therefore there are typically a large number of arguments and counter arguments to consider which when presented in linear text can be confusing for the public at large. Secondly, it is not obvious that many people actually have the necessary critical thinking skills to deliberate on issues. In can be seen that the type of argumentation support systems and tools described in this report have the potential to add value to current eParticipation methods. This is explored further in the section on eParticipation scenarios.

As governments seek to consult their citizens over matters of policy, it becomes increasingly important that citizens receive the relevant information in a medium that they can, and will, want to use in forming their opinion upon consultative issues. This report presents sample eParticipation application scenarios of argumentation support systems in order to assess the potential contribution these systems can make to the consultation process. They cover techniques for the presentation of complex information in a thematically arranged format, for identifying those issues that generate a significant response, for collating consultation responses and representing them within an argument structure, and for checking upon the consistency of contributions to a debate. As such, they have something valuable to offer both government and civil society.

### 1.2.3 D5.2.3 Ontologies

During the last decade, ontologies and Ontological Engineering have gained increased attention. The concept of ontology is not a new concept as such. It has been used by philosophers (e.g. Aristotle) since ancient times to analyse and categorise what exists. With the increasing use of sophisticated information and communication technology, ontologies have become a concept of interest for structuring information in a way which is close to the human understanding.

In areas such as Artificial Intelligence, ontology became a powerful conceptual tool for Knowledge Modelling. It provides a coherent base to build on and a shared reference to align with, in the form of a consensual conceptual vocabulary, on which one can build descriptions and communication acts. Accordingly, Ontology Engineering refers to the set of activities that concern the ontology development process, the ontology lifecycle and the methodologies, tools and languages for building ontologies.

Ontologies are now widely used in Knowledge Engineering, Artificial Intelligence and Computer Science in applications related to knowledge management, e-commerce, intelligent information integration and retrieval, Semantic Web and many more.

In respect to e-participation, ontologies can help to structure the complex area thereby creating the natural links among application of ICT and the context of citizen engagement during their discourses with politicians and governments. This way, a proper understanding of the field can be provided, which is at the same time machine-readable and computable. In more advanced e-participation implementations, ontologies can represent the basic underlying concept of structuring domains, lines of argumentation etc. where intelligent reasoning and knowledge extraction may be facilitated. The recent technologies and digital ontology descriptions even enable the exploitation of reasoning and inference mechanisms, consequently providing innovative means for knowledge management and personalised and customised tools and services in a wide range of e-participation.

In the report, we provide a detailed description of the concept and the technology of Ontologies, as well as of the new field of Ontological Engineering. We cover a number of issues such as: definition of ontology in the context of Computer Science, establishment of the ontology engineering principle and coverage of ontology development methodologies and ontology languages. We provide an overview of key research centres of the field, and investigate the application of ontologies in general, and in e-government. We conclude by addressing key research questions in order to further out knowledge in the domain in e-participation by presenting future scenarios of ontology applications in the field.

### 1.2.4 D5.2.4 Semantic Web Services

Semantic Web Services is a technology that extends the very popular computing paradigms of Web Services and Service Oriented Computing by facilitating semantic annotation of web services though the use of ontologies. The ultimate goal is to enable automatic semantic-based discovery, composition and execution of web services across heterogeneous users and domains.

The present document aims to provide a thorough coverage of the field of Semantic Web Services and its potential role in the eParticipation domain and particularly within the Demonet project.

More specifically, in the introductory chapter of this document the notions of Web Services and Semantic Web Services are introduced.

Chapter 2 deals with semantic web services in much more detail by providing a thorough description of the fields of Web Services and Semantic Web, by explaining how these two are combined towards Semantic Web Services and by describing the most common technologies related to the field.

In chapter 3, an overview of the key research centres in the areas of Semantic Web and Semantic Web Services is given while chapter 4 investigates the various application fields of semantic web services.

In chapter 5, the important role of semantic web services in the area of e-government is illustrated through a number of application examples.

Finally, chapter 6 concludes by addressing the key research questions that the technology of semantic web services needs to cope with in order to be effectively applied in the domain of e-participation.

### 1.2.5 D5.2.5 Knowledge Management and Knowledge Engineering

The public sector is dealing with a significant amount of information and knowledge resources. This knowledge has to be appropriately managed and smoothly integrated. Especially in policy formulation, i.e. in various e-participation areas, the activities and results of action are of information and knowledge by nature. Yet, we still lack a clear understanding of what kind of knowledge and information we are dealing with in e-participation, what purposes and rationale lays behind investigations and activities and which tools and technologies of data and knowledge engineering can support e-participation in its various forms.

This sub-deliverable posits answers to several questions:

First, the introduction sets the scope and basis of understanding for information and knowledge in e-government and e-participation. It further raises four key challenges of knowledge management in e-participation.

Chapter 2 is an introduction to the types of information and knowledge in e-government and e-participation. A holistic framework of understanding and specific aspects of knowledge in governmental processes is discussed. An example of democratic deliberation demonstrates the knowledge aspects in this process.

In chapter 3, methodologies for knowledge management are presented. The discussion distinguishes among concepts describing KM processes, methodologies for knowledge engineering, a concept for knowledge distribution and the knowledge spiral of Nonaka and Takeuchi in order to understand the knowledge creation process.

Subsequently, KM tools and technologies are introduced. Chapter 4 covers comprehensive approaches such as corporate memories, knowledge portals, workflow management systems, or case-based reasoning. Furthermore, concepts for structuring information and knowledge, for information retrieval, and for knowledge analysis are discussed. In addition, agent technologies, and alternative concepts such as individualised feeds, recommender systems, social bookmarking and the like are introduced.

Chapter 5 investigates the potential use scenarios of existing KM technologies for e-participation.

Since KM technologies and solutions are not yet widely used in e-participation contexts, chapter 6 concludes with a number of research questions that are pertinent to future e-participation research.

### 1.2.6 D5.2.6 –Devices, Channels and Mobile Technologies

An inclusive European information society and the accessibility of services through a range of communication channels are crucial to enable e-participation for all (cf. D5.1 section 6.4). To allow access for various user types and social groups, their specific means and channels for communication should be addressed. The terminus “channel” in

D5.1 made no distinction between communication channels, application technology or devices.

The sub-deliverable at hand tries to clarify the different understandings of technology and devices on the one hand. On the other hand, their dependencies among each other will be shown. Consequently, features, requirements, user preferences, means of service delivery as well as public value of devices and their specific technologies will be analyzed. The importance and impact of limitations, advantages, conditions, business models and the public value have to be taken into account.

The report is organized as follows.

First, the introduction sets the scope and ground of understanding for devices and channels in e-participation.

Chapter 2 focuses on communication channels for e-participation. Aspects such as flexibility of users, mode of transmission, infrastructures for bearer services, mobile provider services and data access, internet application and services, and general aspects of availability of channels and cots will be discussed.

In chapter 3, device classes will be analysed along their interaction and usability aspects, the primary channels used for transmission, and the types of applications these devices are used for. PCs, mobile phones and digital TV will be investigated.

Subsequently, the importance of mobile technologies and digital TV channels is discussed in view of e-participation. Chapter 4 discusses therefore issues of diffusion, personalization and localization for mobile technologies and devices, as well as digital TV.

Chapter 5 investigates the impact of devices and channels on e-participation

Chapter 6 concludes the report with reflections and an outlook.

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D5.2.1 - Collaborative Environments

| | |
|---|---|
| **Editor :** | **Gregoris Mentzas** |
| **Revision :** | 01 |
| **Dissemination Level :** | [TA p. 63] |
| **Author(s) :** | Georgia Bafoutsou, Christos Halaris, Gregoris Mentzas, Peter Butka |
| **Due date of deliverable :** | 30th December 2006 |
| **Actual submission date :** | 15th January 2007 |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | 5 |
| **lead contractor for this sub-deliverable :** | ICCS |

**Abstract:** Collaborative environments have a high potential for supporting e-participation as they offer a variety of tools for synchronous and asynchronous communication and collaboration among different stakeholders in the various policy stages of eParticipation. This report gives an overview of electronic collaboration technology, presents several research groups activated in the e-collaboration field and identifies applications of the technology in the business world as well as in electronic government. Finally, is presented the potential of the technology to enhance e-participation applications and are described some application scenarios.

**Project funded by the European Community under the FP6 IST Programme**

# Executive Summary

Electronic collaboration has a strong potential to support distinct participation areas and different stakeholders in the various stages of eParticipation. This potential is being examined in the context of this report.

The structure of the report is as follows:

Chapter 1 is a brief introduction in the e-collaboration technology, including definitions found in the literature.

Chapter 2 provides an overview of the electronic collaboration technology, identifying the typical functions of collaborative systems, presenting classifications of systems and defining collaboration dimensions.

Chapter 3 gives a brief description of research groups dealing with electronic collaboration.

Chapter 4 presents applications of the technology in the business world and e-government. Electronic collaboration can be applied inside the enterprise as well as in the context of inter-organizational relationships. This potential has also been explored and put to use in several European projects.

In the context of electronic government, e-collaboration helps to implement electronic solutions as far as all three interfaces of government are concerned: government to government (G2G), government to business (G2B), and government to citizens (G2C).

Finally, Chapter 5 identifies applications of electronic collaboration in the field of eParticipation. Collaborative environments have a strong possibility for supporting e-participation, as they offer a variety of tools for synchronous and asynchronous communication and collaboration among several participants.

A number of projects launched under the fifth Framework Programme (FP5) will try to promote and enable the online participation of all stakeholders in decision making, and improve the interaction between citizens and public administrations, through e-collaboration functionality.

Some of the eParticipation areas, described in Deliverable 5.1, are considered suitable for applying e-Collaboration technology. Section 2 of Chapter 5 presents some use scenarios in the Community building / Collaborative Environments, Electioneering, Consultation and Discourse areas of eParticipation.

# 1 Introduction

When discussing cross-organizational electronic collaboration most people think of applications most commonly displaying the capabilities of today's collaboration tools: e-mail, bulletin boards, chat sessions, and virtual rooms. We think of online collaborative workspaces for virtual corporative teams and e-learning groups, where work is being shared in an e-Room or the equivalent.

E-collaboration and collaborative systems bring geographically dispersed groups together, enhancing communication, coordination and cooperation. This results in significant time and cost savings, decreased travel requirements, faster and better decision-making and improved communication flow throughout the organization. In simple terms collaboration means distributed computing environment (or intranets) accessible by authorised users.

A generic definition for e-collaboration comes from Bowers [Bowers, 1991], who points out that in its most general form, collaboration using computers examines the possibilities and results of the support through the use of technology of people who work together and are involved in communication intensive processes. Borghoff [1995] stresses that Cooperative Work (CW) includes; information, coordination and cooperation and that supported Cooperative Work (SCW) should offer structure, control and motives.

Other researchers emphasize on the aspect of enabling group work or group activities. For example, Greif [Greif, 1988] defines collaboration using computers as coordinated activities supported by computers. Such activities are communication and problem solving by groups of co-working people or people that need to share files or cases.

Broadly defined, the term “Electronic Collaboration” encompasses the support of communication and coordination of two or more people through the use of software programs, in an effort to fulfil an assignment or solve a problem together [Borenstein, 1992; Schooler, 1996].

# 2 Overall Description of the Technology

## 2.1 Functions of Collaborative Systems

There is a large number of systems, either commercially available or research prototypes, which satisfy in some respect the requirement for collaboration over the web and enable sharing of information over the web. Many of these where formally described as document management systems, case management systems and the alike but have now expanded to real-time information sharing. Such systems are often linked to quality and project management and therefore often include a communication tool with common functions such as:

*Electronic mail:* The most common and widespread communication tool. It allows wide contact over the Internet and its primary use is for text messages, normally relatively brief. Often the messages are accompanied by file attachments.

*Chat:* real-time text talk, where messages appear on both users screens. Usually, a split screen is used, where the local typing appears in one part and the remote in the other. There is no particular subject set and it does not scale to more than a very few users.

*Discussion*: a subject is set, which constitutes the fuel for a discussion that progresses over time. Participants can be either online or express their opinion anytime, under the condition that the subject remains open. Users can pick a topic and see a “thread” of messages and replies about it and post their own message.

*Bulletin Board:* a message board, where a conversation can be carried on over time. The user can leave a message for someone, and they can answer it and the initiator can respond back to them later.

*Whiteboard:* whiteboards allow two or more people to view and draw on a shared drawing surface. This may be used for discussing or describing objects, which are difficult to verbalise. Most shared whiteboards are designed for informal conversation, but they may also serve structured communications, or more sophisticated drawing tasks, such as collaborative graphic design, publishing, or engineering applications. Shared whiteboards can indicate where each person is drawing or pointing, by showing tele-pointers, which are colour-coded or labelled to identify each person.

*File & Document management and sharing:* this function includes the possibility of viewing and editing shared files. Files are stored in a central server and users can work on them, either using their local applications, or the tool’s functionality. Occasionally, there is possibility for version control, search, electronic signing and access control.

*Synchronous Work on Files:* files can be edited simultaneously by a number of users, either on each other’s screen, or on a whiteboard.

*Screen Sharing:* both people have the same view of the screen and possibly the remote user can take control of the other user’s system. Screen sharing can mean that either only the view of the screen is shared -essentially a graphic representation of one screen is passed to the other screen- or applications can be shared, in which case events from the remote keyboard and mouse are used to drive the local input and pointer.

*Presentation Capability:* users can conduct presentations, i.e. show and annotate PowerPoint slides.

*Task List:* lists of actions to be performed, pending activities, unresolved problems and scheduled meetings are kept and the user is notified for new items in the list.

*Meeting Scheduling Tools*: meeting scheduling tools include creating meeting agendas and lists of issues or using calendars for organizing meetings.

*Electronic Calendars:* the electronic calendar supports the enhanced collaboration of group members, providing common access to meeting schedules. The members not only have the possibility to register information about their personal appointments, but also have access to similar information involving other users. In several occasions users receive notifications about future, scheduled meetings.

*Workflow or Case Management:* A workflow is defined as a collection of tasks organized in such a way to form a business process.

## 2.2 Classifications of Collaborative Systems

Researchers have identified at an early stage the need for providing a means for classification of the systems supporting e-collaboration. Therefore classification efforts exist since the early 80's, and their number continues to grow.

A classification criterion of collaborative systems defines a dimension of these systems, or a set of possible values that a characteristic of these systems can assume. Criteria are usually presented in taxonomies. A taxonomy creates a relationship between the classification criteria, and therefore can be considered as a multidimensional space, where each criterion corresponds to a dimension.

A first approach to provide a taxonomy of collaborative systems, is to distinguish them by *when* and *where* the interaction takes place (time/space taxonomy, see [DeSanctis & Gallupe, 1987; Ellis et al., 1991; Johansen, 1988]. In this context, two primary dimensions are identified (see Table 1).

In the horizontal dimension we order collaborative tools by the location of participants: they can be either at the same place (also referred to as co-located) or at different places (remote). Similarly, the vertical dimension makes the distinction, whether the interaction happens at the same time (synchronous) or at different times (asynchronous). These dimensions provide four communication scenarios: synchronous, co-located; asynchronous, co-located; synchronous, remote and asynchronous, remote.

| | **Same Time** | **Different Time** |
|---|---|---|
| Same Space | *Face-to-face interaction*<br>• Conference Tables<br>• Public screens<br>• Tools for voting and exchanging of ideas | *Tasks that range over time*<br>• Spaces that belong to groups<br>• Screens accessed by groups<br>*Tasks with different time schedules*<br>• Project management<br>• Tools for voting and exchanging of ideas |
| | *Remote, real-time interactions* | *Communication & Coordination* |

**Table 1: Time/Space Classification**

Grudin [Grudin, 1994] also provides a classification of collaborative systems in terms of *Time and Space*. Time and space settings in collaborative software can be classified as *same, different but predictable and different but unpredictable*. Accordingly, nine different categories of collaborative systems emerge (see Table 2).

A review of the literature reveals several other classifications of systems that support group work. DeSanctis & Gallupe [DeSanctis & Gallupe, 1987] discuss a taxonomy based on *group size* (smaller, larger), and *task type* (planning, creativity, intellective, preference, cognitive, conflict, mixed motive).

Jarczyk et al. [Jarczyk et al., 1992] developed a taxonomy to characterize collaborative systems, where five major classes of criteria are defined: *functional, technical, application, usability and ergonomics and scalability*. The functional criteria describe the features of systems, the technical characterize the platform, the environment and the system architecture, the application criteria help to define the application domain, usability and ergonomics are important for the acceptance of a tool and finally, orthogonality and scalability are meta-criteria, which focus on the flexibility of the system with respect to the other criteria.

| | **Same Time** | **Different Time Predictable** | **Different Time Unpredictable** |
|---|---|---|---|
| **Same Space** | Electronic meetings | Work with different schedules | Spaces belonging to groups |
| **Different Space Predictable** | Whiteboards Conferences with the use of multimedia | Voice mail | Collaborative writing |
| **Different Space Unpredictable** | Broadcast seminars | Asynchronous conferences | Workflow Management |

**Table 2: Collaborative systems** *Adapted from [Grudin, 1994])*

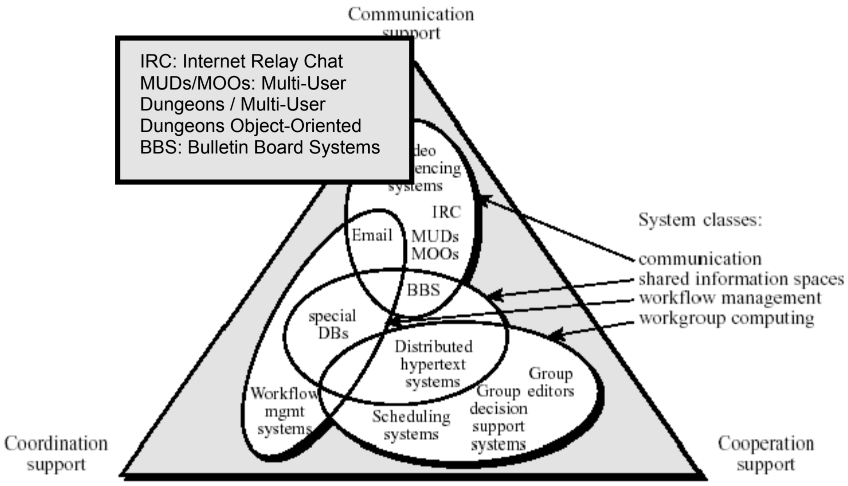
Mentzas [Mentzas, 1993] classifies collaborative software based on four major criteria: *coordination model characteristics*, *type of processing*, *decision support issues* and *organizational environment*.

McGrath & Hollingshead [McGrath & Hollingshead, 1994] deal with a task framework, where group tasks are classified in four quadrants. Each quadrant is characterised by a general performance process (action of a group): generate (alternatives), choose (alternatives), negotiate and execute. The quadrants are then subdivided in two types of tasks each, and as result eight different types of tasks arise. The task circumplex is a two dimensional representation. The horizontal dimension shows a contrast between behavioural or action tasks to the right and conceptual or intellectual tasks to the left. The vertical dimension reflects a contrast between cooperation or facilitative compliance at the top and conflict at the bottom.

Malone & Crowston [Malone & Crowston, 1994] define a taxonomy based on a collaboration/coordination model. According to their framework, four levels of processes are defined: *collaboration/coordination, group decision-making, communication among the collaborators* and *perception of common artifacts.*

Teufel [Teufel, 1995] in an effort to categorize the collaboration systems, distinguish three possibilities of electronic support for collaborative processes: *Communication support*, *Cooperation support* and *Coordination support*. The various systems are placed in a triangle, according to the basic functionality of each one and in relation to the three possibilities for electronic support. The systems are further grouped in four categories: communication systems, shared information spaces, workflow management and workgroup computing (see Figure 1).

In the Groupware Bible of Lotus Corporation are identified three classes of software supporting electronic collaboration: *Communication systems*, *Collaboration systems* and *Coordination systems* [Lotus, 1996]. Communication systems are means that passively transmit information. The complexity of those systems ranges from simple tools supporting same time, same place, one-to-one interaction, to sophisticated software capable of handling same as well as different time and space situations including a large number of participants.




.

**Figure 1: Classification of Collaborative Systems by Teufel** *(Source: [Teufel et al., 1995])*

Collaboration systems are common workspaces, which contribute to the diminution of time and space constraints. Examples of such systems are the electronic conferencing systems and the shared databases. Finally, Coordination systems combine structured communication and collaboration actions and also support informal conversations. Figure 2 presents a graphical representation of this categorization. It should be noted though, that the three cycles overlap, which leads to the conclusion that several systems can be part of more than one category.

Ellis [Ellis, 2000] provides a categorization of collaborative systems according to the underlying technology. Thus, four aspects are determined: *keepers*, *coordinators*, *communicators* and *team-agents*. Briefly, the first aspect, keepers, groups all functionality related to storage and access to shared data. The second aspect, coordinators, is related to the ordering and synchronization of individual activities that make up the whole process. The third aspect, communicators, groups all functionality related to unconstrained and explicit communication among the participants. Finally, the fourth aspect, team agents, refers to intelligent or semi-intelligent software components that perform specialized functions and help the dynamics of a group.

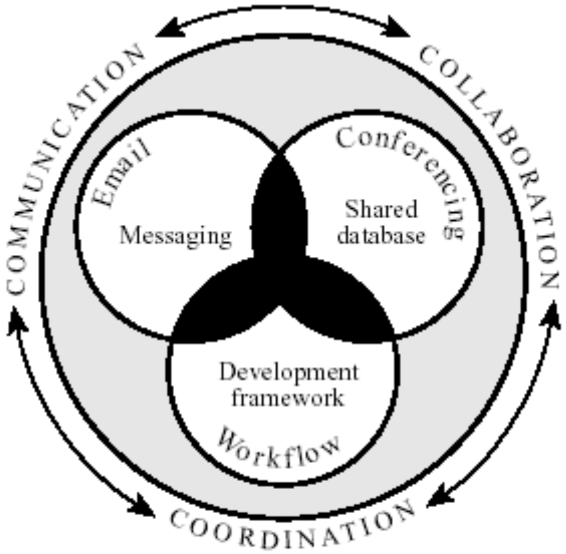



**Figure 2: Classification of Collaborative Systems by Lotus *(Adapted from [Lotus, 1996])***

Meier [Meier, 2002] distinguishes three dimensions in the area of Collaboration and Cooperative work: *coordination*, *communication* and *common ground*. Collaboration support systems are also classified based on whether they provide *synchronous* or *asynchronous communication and collaboration support* and whether they address the needs of *Individuals, Teams,* or *Organizations / Networks / Communities*. Figure 3 presents the above-described classification.

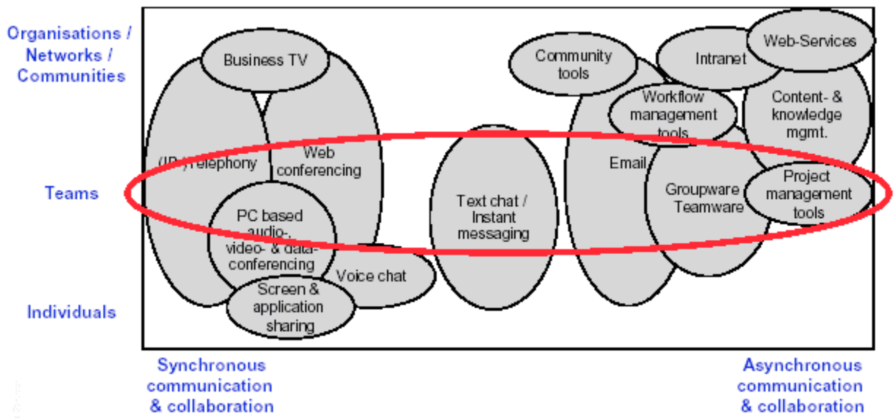



**Figure 3: Classification of collaboration support systems by Meier*(Adapted from [Meier, 2002])***

As mentioned in the previous paragraphs, various researchers have already pointed out the three basic dimensions of e-Collaboration: Communication, Cooperation and Coordination. In the next paragraphs we analyse the basic features and identify the most usual functions of IT platforms that electronically support each of the three dimensions.

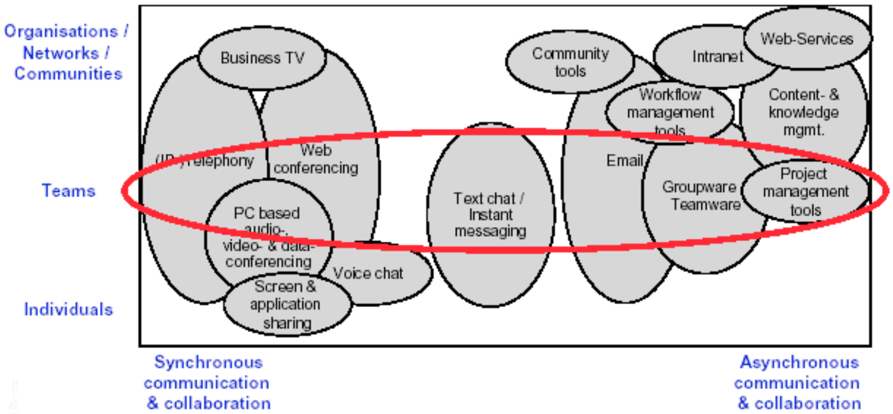
***Communication***

The term Communication includes basic information exchange among the involved in a collaborative situation parties. Emphasis is given on the explicit interaction between two or more people, either in the context of a discussion, or during the exchange of an electronic message.

Communication processes do not usually have structure or a specific sequence of steps. They can take place either randomly or on a pre-defined schedule. There are possibilities for bilateral (one-to-one) or multilateral (one-to-many, many-to-many) communication and real-time or asynchronous interaction.

Communication support has been the primary focus of many software systems. The simple, text-based communication with the use of electronic mail has now been enhanced with multimedia (voice contact and electronic conferences with the use of video).

The software for electronic mail is still the most common and widely used for communication and information exchange. There is need for low cost extended software, offering speed and easiness of application and use.




### *Cooperation*

Under the term Cooperation we group the possibilities for work on shared documents and files of various formats. In essence, cooperation is about the actual collaboration of groups, aiming at the generation of artefacts.

The interaction in this case is implicit and takes place through the reference on the shared artifact, can occur at the same time or asynchronously and the use of multimedia is usually not included. Group work is stored in repositories as data and information and is accessible by all interested parties according to their access rights. The user interface is usually simple and no special programming knowledge is requested.

### *Coordination*

The concept of Coordination focuses on the programming and scheduling of activities performed by the involved actors in a collaboration process.

Simple coordination capabilities involve electronic calendaring tools. Electronic calendars can be used either for personal or for group scheduling.

Also, elementary coordination is accomplished when interfacing with some electronic conferencing products, electronic meeting and electronic workspace systems, through the use of task lists and meeting scheduling tools.

Basically though, the dimension of Coordination is supported by the Workflow Management Systems. These systems offer assistance for strictly structured actions, which happen at a specific order, as well as for semi-structured processes, which require intellectual work and whose parts are insufficiently defined and changeable. In both cases, interaction is implicit and of different time and space.

The functions described in paragraph 1.2 are classified in relation to the time and space dimensions [Bafoutsou & Mentzas, 2002] (see Table 3).

| Function | Temporal Dimension | | Spatial Dimension | |
|---|---|---|---|---|
| | Synchronous | Asynchronous | Co-Located | Remote |
| E-mail | | ■ | | ■ |
| Chat | ■ | | ■ | ■ |
| Bulletin Board | | ■ | | ■ |
| Discussion | ■ | ■ | ■ | ■ |
| Whiteboard | ■ | | ■ | ■ |
| File & Document sharing | | ■ | | ■ |
| Synchronous work on files | ■ | | ■ | ■ |
| Screen Sharing | ■ | | ■ | ■ |
| Presentation Capability | ■ | ■ | ■ | ■ |
| Task list | | ■ | | ■ |
| Meeting scheduling tools | | ■ | | ■ |
| Electronic Calendar | | ■ | | ■ |
| Workflow Management | | ■ | | ■ |

**Table 3: Temporal and Spatial Dimensions of Collaborative Functions** *(Source: [Mentzas & Bafoutsou, 2002])*

As one can conclude studying Table 3, File & Document Sharing is mostly remote and asynchronous, while real-time cooperation takes place in the case of Synchronous Work on Files and Screen Sharing, where both dimensions of space are also included. Presentations can be conducted either synchronously or asynchronously and interacting users can be either in remote locations or at the same place. Finally, asynchronous and remote are the functions of Task lists, Meeting Scheduling, Electronic Calendars and Workflow Management.

In Table 4, all functions are presented according to the degree of Communication, Cooperation and Coordination they support. We consider three levels of support: Low (*), Medium (**), and High (***).

We make the following observations: real-time interaction offers good opportunities for Communication, the support for cooperation is considered high with functions involving shared, real-time editing of files, and finally, Workflow Management corresponds to the highest level of Coordination. Table 4 incorporates the time/space classification shown in Table 3.

In Table 4, are listed functions, typical for each dimension. We usually select the functions with "***" at the corresponding column. In some cases, as is the electronic mail, the incorporation of the function in a dimension is obvious, even with "**" at the required column.

| Collaboration dimensions<br>Function | Communication | Cooperation | Coordination |
|---|---|---|---|
| E-mail | ** | * | * |
| Chat | *** | * | * |
| Bulletin Board | ** | * | * |

| | | | |
|---|---|---|---|
| Discussion | *** | * | * |
| Whiteboard | ** | * | * |
| File management | * | ** | * |
| Synchronous work on files | ** | *** | * |
| Screen Sharing | ** | *** | * |
| Presentation Capability | ** | *** | * |
| Task list | * | * | ** |
| Meeting scheduling tools | ** | * | ** |
| Electronic Calendar | * | * | ** |
| Workflow Management | ** | ** | *** |

| | |
|---|---|
| *** | High support of the dimension |
| ** | Medium support of the dimension |
| * | Low support of the dimension |

**Table 4: Functions and Collaborative Dimensions** *(Source: [Bafoutsou & Mentzas, 2002])*

Table 5 describes briefly the three collaboration dimensions, including a short definition and a list of the basic functions of each dimension.

Concluding, the electronic support of collaboration has primarily focused on two axes: either on the Coordination of business processes performed asynchronously by different actors (using for example a workflow management system), or on the automation of Communication and Cooperation for groups involved in more loose processes (i.e. Electronic Mail, Electronic Workspaces, etc.).

| Collaboration Dimensions | Definition | Functions |
|---|---|---|
| *Communication* | **Explicit interaction of two or more people aiming at the exchange of information of any kind** | E-mail<br>Chat<br>Bulletin Board<br>Synchronous Discussion<br>Asynchronous Discussion |
| *Cooperation* | **Implicit interaction taking place through reference to a common artefact** | File & Document Sharing<br>Synchronous work on files<br>Whiteboard<br>Screen sharing<br>Presentation capability |
| *Coordination* | **Programming and settlement of activities performed by the parties involved in a collaboration process** | Task list<br>Meeting scheduling tools<br>Electronic Calendars<br>Workflow Management |

**Table 5: Brief Description of Collaborative Dimensions** *(Source: [Bafoutsou & Mentzas, 2002])*

Collaboration systems that specialize either in Communication or in Cooperation or even combine both dimensions provide limited support for Coordination during teamwork. One can locate shortcomings in the functions of business process definition, where neither the workflow automation nor workflow monitoring is possible.

However, there are several research efforts towards the integration of workflow management systems that develop tools supporting Communication and Cooperation.

(see [Kreifelts et al., 1999], [Haake & Wang, 1999], [Agostini & De Michelis, 2000], [Kammer & McDonald, 1999], [Bussler, 2000] and [Araujo et al., 2001], [Bafoutsou & Mentzas, 2002].

# 2 Examples of Research Groups

## 2.3 Fraunhofer Institute for Applied Information Technology (Fraunhofer FIT)

http://www.fit.fraunhofer.de

Fraunhofer FIT works to enhance human abilities through flexible, context-adaptive information and cooperation systems.

Systems designed and built in FIT support designers and engineers in real and virtual work environments, provide Internet-based platforms for social as well as task-related interaction of learning communities and virtual teams.

Fraunhofer FIT is organized in three departments with about 30 researchers each, plus a service unit with administrative, service and support functions for the institute: Cooperation Systems Department, Information in Context Department and Life Science Informatics Department.

The Cooperation Systems (CSCW) department develops and evaluates groupware and community systems for virtual teams and organizations. Their work on hardware and software of Mixed and Augmented Reality systems focuses on support for cooperative planning tasks.

The CSCW research department was set up in the 1970's. Among the early results is one of the first electronic mail systems in Europe. The group has since done research on workflow systems, directory systems, organizational knowledge modelling, task and coordination management, and multi-user 3D environments.

In the framework of the Social Web research programme, the department aims to extend Internet-based groupware and mobile technologies to cooperative spaces for people, enabling them to expand their web of relationships to workgroups and communities on a global scale. Social, physical and virtual components are combined to ease cooperative handling of complex information. In the business sector, Social Web systems support distributed groups and virtual organizations in their ability to foster globalisation and to deal with the consequences. In particular, the CSCW research department concentrates on four research issues:

- Development of flexible Internet-based groupware solutions for e-learning and self-organized document management in organizations and teams.
- Development of community ware solutions to support social awareness and collaborative construction of community knowledge.
- Support of cooperative planning processes through mixed reality applications.
- Evaluation and design of groupware and knowledge management systems.

## 2.4 Distributed and Mobile Collaboration Lab, Distributed Systems Group, Information Systems Institute, Technical University of Vienna

http://www.infosys.tuwien.ac.at/dmc/index.html

Business processes and distributed collaboration have been changing radically over the last years. Business environments demand increased flexibility, interconnectivity, and autonomy of involved systems as well as new coordination and interaction styles for collaboration between people. The Distributed and Mobile Collaboration Lab performs research on the latest trends in distributed and mobile collaboration technologies that allow people to move across organizational boundaries and to collaborate with others within/between organizations and communities.

## 2.5 CTRG Groupware and Workflow Research, University of Colorado

http://www.cs.colorado.edu/~skip/ctrg.html

The Collaboration Technology Research Group (CTRG) was formed in January of 1992. Its research involves topics of human collaboration and computer support for cooperative work. CTRG operates on the principle that the state of the art can be advanced by close coupling of theoretical and applied research. The research group is concerned with theories, models, architectures, implementations, and studies/evaluations of computer supported group activity. This is an inherently interdisciplinary topic, so CTRG draws together faculty and students from the Computer Science Department, the Business School, and the Institute for Cognitive Science at the University of Colorado. The main areas of interest are:

- workflow architectures and models
- groupware frameworks and functionalities
- group cognition and group user interfaces
- dynamic change in workflow systems
- social computing
- very large scale collaboration.

## 2.6 GroupLab, The Computer-Supported Cooperative Work and Groupware Research Laboratory, University of Calgary, Canada

http://grouplab.cpsc.ucalgary.ca

The GroupLab at the University of Calgary is investigating CSCW, groupware and Human Computer Interaction. CSCW projects have wide coverage, from asynchronous to real time conferencing, and from same-place to geographically distributed meetings. HCI projects concern usability of World Wide Web browsers and personal information management.

## 2.7 Computer Mediated Communication (CMC) Research Group, Stockholm University and KTH Technical University

http://dsv.su.se/jpalme/cmc-research-at-DSV.html

This group performs research both of a social science and technical nature in the area of Computer Mediated Communication (CMC), i.e. the use of computers for communication of information between humans.

The group has special interest in:

- Use of CMC for scientific communication
- Use of CMC as an educational tool
- CMC as an organisational, societal, and cultural phenomenon
- Design of software for e-mail and non-simultaneous group communication
- Standards for e-mail and non-simultaneous group communication
- Information quality and filtering in CMC software
- Quality on the Internet
- Electronic mail and other networked application standards

## 2.8 Cooperative Systems Engineering Group (CSEG), Computing Department, Lancaster University

http://www.comp.lancs.ac.uk/research

The Cooperative Systems Engineering Group (CSEG) researches into all aspects of Systems Engineering and Cooperative Systems. Their work ranges from fundamental research in cooperative working through systems requirements engineering and systems development techniques, to innovative ways of interacting with computer systems.

## 2.9 CITO - Centre for Innovation, Technology & Organisation (CITO)

http://mis.ucd.ie/CITO

The research activities of CITO are focused on understanding the relationship between innovation, technology and organisation, and the associated policy and management implications. Their approach is guided by the assumption that a sophisticated understanding of information and technology, and their role in the constitution of social and organisational life, should be based on an appreciation of how such artefacts come to be embedded within broader institutional (organisational, cultural, economic, political) contexts. As such, their research work is concerned with in-depth empirical studies of information systems implementation and use that are especially attentive to the underlying social relations within which such systems are embedded. They are committed to developing appropriate theoretical perspectives for illuminating such processes by drawing from a variety of intellectual traditions, including philosophy, sociology, political science, psychology, economics and organisation theory. The emphasis is on the pragmatic use of theory to make tangible and insightful contributions to management practice.

## 2.10 Collaboration Technology Laboratory (CTL), Auburn University

http://www.eng.auburn.edu/~kchang/CTL/affec.htm

The objective of the Collaboration Technology Laboratory (CTL) is to study and broaden technologies that foster computer-based environments for collaborative work. This area of study is also known as Computer Supported Cooperative Work (CSCW) or Groupware in the literature. Many subject areas are involved in the pursuit of this exciting study. These include computer networks, database, formal methods, graphical user interfaces, human computer interaction, psychology, programming languages, and sociology.

The goal of this research is to establish a general framework for Computer Supported Cooperative Work (CSCW) applications to support formal electronic collaboration (FEC) work via electronic means such as Web browsers. The framework will accept the formal requirements description of a desired collaboration application, verify the validity of the requirements, and compile the requirements into an intermediate level “byte-code” that can be interpreted as a virtual machine.

## 2.11 E-collaboration group, Comm Tech Lab, Michigan State University

http://commtechlab.msu.edu/randd/collaboration/intro.htm

The Comm Tech Lab e-collaboration group studies existing tools and experiments with the integration of e-collaboration and traditional web site design.

Email was the first collaborative application made possible by the Internet. Newsgroups and Listserves connect strangers with common interests, as do text and graphical chat rooms. Instant Messaging is the latest online collaboration tool to achieve widespread adoption, connecting family, friends and co-workers.

In contrast to these collaboration applications, exploring the web is usually a solitary user experience, void of other human presence even if a lot of other people are concurrently reading the same web page. Browsing doesn't need to be solitary. Online learning, ecommerce, public relations, and museum/collection sites might all benefit from integrating e-collaboration in different ways.

# 3 Current Applications of the Technologies

The wide deployment of the Internet, combined with the development of information technology and telecommunications creates an increasing need for electronic support of small groups and medium sized enterprises as well as of remote partnerships of much larger scale.

Collaboration requirements are as intense inside the enterprise as in the context of inter-organizational relationships.

## 3.1 General

The forces of globalization and ubiquitous digital networking are making the traditional B2B (business-to-business), B2C (business-to-consumer), and B2E (business-to-employee) distinctions less and less relevant. Within the extended enterprise, entities switch between taking the role of vendor, supplier, partner, and customer. Partners are brought into the corporation's extended intranet, while CRM initiatives bring customers into the design and development phase of new products.

The agile virtual enterprise requires the ability to communicate with a variety of people – but increasingly in the same ways. As the distinctions between B2B, B2C, and B2E fade, the emerging requirement is for virtual collaboration that are flexible enough to meet a variety of demands.

### 3.1.1 Home Office Connection

Working from home constitutes one of the trends of the twenty first century. The home workers need access to calendars, important files and addresses, in other words need to transfer the office organization onto the Internet. Electronic collaboration gives remote co-workers the possibility to access and edit common files, exchange ideas and deliver day-to-day tasks faster and more effectively.

### 3.1.2 Project Planning and Execution

The project management capabilities of collaborative software enable the project manager to divide projects into phases and further into workpackages. They are able to apply data including expenditure, co-workers, times and priorities. Project members can be made aware of everything happening within the team. They can look up the current status, fill in their working hours and outline the progress made to date on the project.

### 3.1.3 Preparation and reworking of seminars

Electronic Collaboration is successfully applied in the preparation and execution of training seminars. Training course delegates, who are not co-located, are able to participate in the seminar through obtaining access in common workspaces, where they can attend presentations, exchange documents, and solve problems. Also, using the calendaring and scheduling possibilities of electronic collaboration tools, seminar participants get informed of new seminars and can plan and update their schedule accordingly.

### 3.1.4 Coordination of software development

Software development is a complex iterative process, in which many persons are involved. Requirements must be collected and transferred into functions. During the development, dependencies regarding many different factors must be taken into consideration. Errors arising during tests must be documented and fixed immediately.

Electronic Collaboration is applied throughout the whole software development life-cycle, from planning to implementation and testing.

### 3.1.5 Collaborative Commerce

In the business-to-business sphere, online collaboration is growing very rapidly across the supply chain. Such collaborative services enable corporations to go beyond transactional e-commerce exchanges to use the Web to exchange intellectual capital, facilitate problem troubleshooting and resolution, enable new ideas to cross corporate boundaries, provide rich feedback on goods, services and customer satisfaction, and foster a much closer bond between business partners.

### 3.1.6 Virtual Consortia

The new forms of network business organizations, grouped under the term of “virtual consortia” hold the following characteristics [Halaris et al., 2001]:

- They are created by organizations remotely located, whose fundamental competencies are complementary and are oriented towards the same business opportunity and
- They use the Internet for the exchange of data and information between them.

Working effectively in virtual consortia requires electronic sharing and editing of documents, online discussions and exchange of ideas, preparation and attendance of virtual meetings and precise task sharing and coordination through workflow management. As far as project management is concerned, distributed co-workers should be able to record their respective tasks, times and budgets on a decentralized database, providing responsible managers with clear and up-to-date data regarding completion and cost of tasks.

Electronic collaboration tools provide the virtual consortia members with all required functionality.

### 3.1.7 Virtual Tendering and Bidding

In the case of project-centric business environments, as for instance in the construction sector, a critical business process is tendering/bidding for a new project, where timely opportunity identification and adequate consortium formation are the key factors for winning a contract [Halaris et al., 2001].

Managing virtually the tender/bidding process consists of supporting electronically the execution, partially or in total, of all necessary activities.

Tender documents are made directly available to interested parties and thus lead times are eliminated. Since very often after the publication of the tender, clarifications are made, it is also possible to amend clarifications to tender documents after they are uploaded and to accordingly notify providers who have already downloaded the tender documents.

Having made all the tender documents electronically accessible, a sophisticated search engine is used to access them. This engine should enable quick multi-parameter search of tenders and flexible presentation of results. Additionally, e-mails are automatically sent informing the user of any new tenders that match a predefined profile(s). This profile is defined by the user and contains priorities and interests, which are the basis for the screening of new tenders. Another important issue is the ability to search for and gather information about potential partners, subcontractors and suppliers as well as to have a secure environment ensuring on-time and quick communication with them.

Virtual support has also to deal with the need of exchanging documents and messages within the virtual consortium after its formation, when the preparation of the bid begins.

Towards the end of the tender/bid process, virtual management supports the electronic submission of the bid, the communication between client and provider and the electronic dispatching of the results. If this is accomplished in a way that does not endanger confidentiality of the bids, then substantial advantages are gained from the minimisation of the response times to tenders.

## 3.2 Projects Concerning e-collaboration

***GENESIS - Generalised ENnvironment for procesS management In cooperative Software engineering***

http://www.genesis-ist.org/

GENESIS is an IST project (IST-2000-29380) intending to develop an Open Source environment that supports the co-operation and communication between software engineers belonging to distributed development teams involved in modelling, controlling, and measuring software development and maintenance processes. Moreover, it includes an artifact management module to store and manage software artifacts produced by different teams.

The GENESIS project proposal stemmed from an increasing interest showed by Academics and IT professionals in large organizations toward the communication and coordination problems found in software development activities, especially when software engineering team members are geographically distributed and speak different languages.

The co-ordination aspects are supported by workflow-like engines, which implement the software process models used in an enterprise (a non-invasive approach). The communication involves both formal (e.g., the release of specification documents) and informal (e.g., annotations describing personal considerations about a design choice) aspects, which will be provided through data and control integration with an artifacts management module. The environment supports modelling, controlling and measurement of software design, development and maintenance processes, as well as the co-operation and communication between software engineers belonging to collaborating development teams.

***InContext - INteraction and Context based technologies for collaborative teams*** *(EU Framework 6 STREP research project IST-034718)*

http://www.vitalab.tuwien.ac.at/projects/incontext/

Knowledge workers are increasingly engaged in a number of projects at the same time and thus require a flexible form of collaboration. New team forms emerge spontaneously, featuring dynamic interaction patterns which are currently not supported by existing software services. Team types such as virtual, nimble, or mobile/nomadic evolve and merge to reflect the vibrant nature of human interaction. To enable efficient collaboration among team members and effective participation of individuals in multiple teams at the same time, collaboration environments need to exhibit capabilities for large-scale interaction, peer-to-peer communication, and loose coupling in a trusted serviced-oriented way. inContext strives to enable dynamic collaboration by exploring new techniques and algorithms for pro-active service aggregation, context-aware service adaptation and service provisioning.

***LABORANOVA - Collaboration Environment for Strategic Innovation***

http://www.laboranova.com/

The goal of Laboranova is to create next generation Collaborative Tools which will change existing technological and social infrastructures for collaborating and support knowledge workers and eProfessionals in sharing, improving and evaluating ideas systematically across teams, companies and networks. Laboranova will do research to develop and integrate models and tools in three specific areas, the three pillars in the project: ideation, connection and evaluation. By integrating these efforts the results will be innovative collaboration approaches and organisational models for managing early innovation processes, software prototypes and the integration of the isolated models and tools into a Collaborative Innovation Platform.

***FUSION - Semantic Business Process Fusion***

http://www.fusionweb.org

The FUSION project aims to promote efficient business collaboration and interconnection between enterprises (including SMEs) by developing a framework and innovative technologies for the semantic fusion of heterogeneous service-oriented business applications. Such applications may exist either within an enterprise or in several collaborating companies within the enlarged Europe. Led by SAP AG, the FUSION consortium consists of 14 partners from five European countries (Germany, Poland, Greece, Hungary, Bulgaria).

***SUPPLYPOINT - B2B Electronic Marketplace in the Construction Sector***

Supplypoint was a research project funded by the ESPRIT programme of the European Commission. The project addressed the issues of pan-European electronic trade links for business-to-public electronic commerce, where it supported electronic procurement. This were to be achieved by using supply chains and covering a life-cycle from contract identification to completion, including supply chain management and electronic payments. Thus, it was provided a one-stop shopping service for companies to purchase goods and services from SMEs co-operating in virtual and dynamic supply chains.

## 3.3 Applications in eGovernment

Organizational hierarchies or bureaucracies were once the familiar representation of government, built to organize decision making and communication through their multiple layers and departmental structures. These bureaucracies, characterized by centralized authority, controlled information, and differentiated talents and functions, led to complexity and the development of specialized units. The effectiveness of hierarchies in today's environment is being challenged by multi-dimensional issues, the overload of information available via the Internet, and the proliferation of human networks of communication and action. Communication technologies fostering informal communities are fundamentally changing the way businesses, universities, government agencies, and other organizations operate because they bypass the laws of bureaucracy and allow individuals and organizations to connect across boundaries. For the first time, information can be shared in an open environment, one that does not acknowledge hierarchy or rank. Networks provide a vehicle for quick response and are flexible and adaptable, providing the capability for rapid deployment and just-in-time decision making [Mc Daniel & Carr, 2005].

The challenge for government is to implement electronic solutions across the three elements of eGovernment: government to government (G2G), government to business (G2B), and government to citizens (G2C).

Regarding the G2B and G2C relationships, eGovernment is currently targeting at realizing the necessary infrastructure for offering citizens and enterprises the capability to perform electronically their transactions with the Public Administration (e.g. declarations, applications, etc.), through the electronic provisions of the necessary public services over the Internet [Karacapilidis et al., 2004].

The asynchronous mode of electronic collaboration (e.g. e-mail, bulletin boards) allows citizens and businesses to have the same round-the-clock access to government services and agencies that they are enjoying in their dealings with private-sector institutions such as banks and airlines. Because much of the communication between governments and citizens involves the transmission of sensitive information, however, it is considered imperative to ensure the security and privacy of the transactions in order to realize the full potential of this mode of interaction.

Citizen-centred government comprises integrating at the point of delivery services and information offered to the public. The implementation of eGovernment regarding the achievement of inter-agency collaboration (G2G relationships) presents challenges such as

- realizing the management of Public Administration Business processes (capturing, modelling, redesigning and implementing)
- achieving efficient group collaboration and group working among Public Administration workgroups that participate in common business processes,
- handling group decision–making, concerning difficult and complex social problems, or granting licences and permissions with high social impact and
- addressing interoperability issues among the infrastructures that support different Public Authorities in various countries.

The cooperation of government agencies is crucial to the development of an internetworked government that provides a vehicle for transforming the functions of government and for gaining efficiencies and improvements in coordination. Ensuring effective collaboration is essential to support the technical process of developing integrated service delivery; it provides governments with the potential to develop integrated applications, share resources, adapt to new environments and enhance organisational learning [Scott & al., 2004].

Several European projects are dealing with the application of electronic collaboration in the field of eGovernment.

***ICTE-PAN – Methodologies and Tools for Building Intelligent Collaboration and Transaction Environments for Public Administration Networks*** *(IST-2001-35120)*

http://www.eurodyn.com/icte-pan/

The ICTE-PAN Project, which as mentioned above is implemented in the context of the European Union IST Programme, has been initiated to address the G2G collaboration needs of Public Organizations. The main objectives of this project are:

- to develop a methodology for modelling collaboration among POs, and also for redesigning it based on the state-of-the-art ICTs,
- to develop a complete electronic platform with all the required meta-tools for creating high quality G2G collaborative environments,
- to elaborate sustainable measurement algorithms for evaluating such environments.

The project is implemented by a well-balanced consortium of technology providers and users, consisting of European Dynamics (Greece), University of the Aegean (Greece), TXT Solutions (Italy), National Environment Research Institute (Denmark), Ministry of Environment of Lower Saxony (Germany) and Province of Genoa (Italy).

***EU-publi.com - Facilitating Co-operation amongst European Public Administration employees through a Unitary European Network Architecture and the use of Interoperable Middleware Components*** *(IST-2001-35217)*

The EU-Publi.com project introduces information technology in order to facilitate inter-European collaboration amongst Public Administration employees.

The project attempts to introduce changes in the way that Public Administration employees execute processes in order to provide services to end customers. Currently, typical processes executed by Public Administration employees are fragments of a larger whole that can be identified through the service it provides to its customer. Relating these processes in terms of causal relationships (e.g. who requires what from whom?), macroprocesses can be established that typically extend over several Public Administration organizations.

If a European Citizen requests a service that triggers a macro-process, which cuts across more than one national European administration, then coordinated execution at different administrative level is required. The question of interoperability amongst diverging and heterogeneous information sources scattered across various organisations in several countries is therefore prominent. In order to enable the capacity of civil servants to provide such a service, a unifying architecture on the top of the existing legacy systems is built, into which the collection of distributed, autonomous systems of each Public Administration can be brought together into a common cooperative environment.

EU-Publi.com allows the Public Administration employee to cooperate more easily with other Public Administration employees at the intra-organization, inter-organization as well as at the European level. By enabling interoperability amongst currently heterogeneous Public Administrations, EU-Publi.com will bring about new e-services for Public Administration.

***ONESTOPGOV- A life-event oriented framework and platform for one-stop government*** *(FP6-2004-IST-4-26965)*

http://www.onestopgov-project.org/

Online one-stop government enables 24 hour, single point access to public services that are integrated around citizens needs (usually life-events). Currently however, online one-stop government projects do not care about citizens needs and do not provide integrated services from different back-offices. The OneStopGov project aims to specify, develop and evaluate a life-event oriented, integrated, interoperable platform for online one-stop government. This platform will be accompanied by a coherent framework for realising and exploiting online one-stop government at all levels.

The guiding vision, challenge, innovation and unique selling proposition for the OneStopGov platform involve: the inherent support of life-events; the active, citizen-centric approach; and the definition and use of generic models (e.g. generic workflows, generic reference models). The OneStopGov platform will be based on a number of scientific and technological innovations. First, the life-event ontology will be specified to enable proper representation of the life-event concept. Second, the active life-event portal will be implemented to care for citizens' needs and circumstances. Third, a complete set of life-event reference models will be specified to allow implementing virtually any life-event. Four, these reference models will be implemented using generic workflow Web technologies.

The OneStopGov platform and framework will be deployed in three new Member States (Slovenia, Hungary and Poland) and one Accession country (Romania). The platform will be used for modelling, implementing and deploying 16 life-events. The consortium includes two organisations responsible for eGovernment at national level (Slovenian ministry of Public Administration and the company owned by the Hungarian Prime Minister's Office), one at regional level (the region that includes Bucharest), and one at local level (Polish municipality) thus ensuring maximum visibility and take-up of the project results.

***SAKE - Semantic-enabled Agile Knowledge-based e-Government*** *(FP6, Start date: 01-03-2006)*

http://www.sake-project.org/

Permanent changes in the environment (political, economical and ecological) cause frequent changes in the governments' regulations that may affect public administration processes and systems. To reduce time-to-market with regards to new decisions, regulations, and law, it is necessary to equip public administration with tools supporting the agile response to changes. A change in one activity in a process or in one part of an e-government system (front and back office) may cause many problems in other parts of the same process or system.

Therefore, there is a need for resolving changes in a systematic manner, ensuring overall consistency. Furthermore, these changes impose the need for updating the knowledge needed to perform the administrative process or use the e-government system, which is heterogeneous and fragmented. These needs are more prominent in the case of New Member States, since their full integration heavily depends on the possibility to adapt their public administrations to the existing EU regulations in a very short period of time. The SAKE project addresses the afore-mentioned needs by specifying, developing and deploying a holistic framework and supporting tools for an agile knowledge-based e-government that will be sufficiently flexible to adapt to changing and diverse environments and needs.

***ITAIDE -Information Technology for Adoption and Intelligent Design for E-Government*** *(FP6, Start date: 01-01-2006)*

http://www.itaide.org/

One of the great challenges for European governments is solving the paradox of increasing security of international trade, while at the same time reducing the administrative overhead for commercial as well as public administration organisation. It is vital to have timely information about business transactions. This information gathering is very costly for businesses and public administrations. Finding the right balance between control and cost of information gathering is the key to increase competitiveness of European businesses locally, nationally and internationally.

ITAIDE develops a Common Information Model for electronic documents and document mapping software to improve the pan-European interoperability of taxation and customs systems. This interoperability is an essential prerequisite to achieve strategic goals for e-customs such as the introduction of Authorised Economic Operator and Single Window Access service provisioning for businesses. ITAIDE develops a procedure redesign methodology, supported by an intelligent software tool, to improve the efficiency and simplification of e-customs procedures. To encourage the adoption of these redesigned procedures by Taxation and Customs offices, businesses and technology providers, it should be the result of a truly collaborative co-design process that creates win-win benefits for all these stakeholders.

This requires new public-private partnerships between Taxation and Customs offices and businesses. We also develop organizational network collaboration models to build these new public-private partnerships between Taxation and Customs offices and businesses. ITAIDE integrates and strengthens European research for innovative government by enhancing service offerings and disseminating good governance practice through increased security and controls, while commensurately employing intelligent software tools to reduce administrative load burden.

***eGov-Bus: Advanced eGovernment information service Bus***

http://www.egov-bus.org/web/guest/home

The objective of eGov-Bus is to integrate and extend research and standards in the area of process and content management for government and cross-government systems, with the capability of creating advanced applications of electronic signature enhancing acceptance of the technology and establishing trusted system validity and non-repudiation, relying on web services, process and repository management platforms based on a highly secure, highly available, scalable and distributed architecture providing data access abstraction. A key downstream effect is the reduction of integration costs of many of eGovernment projects.

It will research advanced infrastructure level technologies on which future developments of IDA will be enabled. Specifically, the project will:

- Create adaptable process management technologies by enabling virtual services to be combined dynamically from the available set of e-Gov functions, personalizing preferences and supporting the rules of the specified life event.

- Improve effective usage of advanced web services technologies by e-Government functions with: Service Level Agreements; Audit trail; Semantic representations; Availability and performance.

- Exploit and integrate current and ongoing research results in the area of natural language processing to provide user-friendly personalizable interfaces to the eGov-Bus.

- Orchestrate the available web services according to the specific life-event requirements, creating a comprehensive workflow process and providing explanation to the end-user.

- Support a virtual repository of data structures required by life-event processes, representing declarative (i.e. rules governing life-events categories) and procedural knowledge.
- Research a secure, non-reputable audit trail for composed web services by advancing qualified electronic signature technology.

# 4 eParticipation Applications of the Technology

## 4.1 Current Applications in eParticipation

The term participation means taking part in joint activities for the purpose of reaching a common goal. This encompasses both trivial situations in which participation mainly has a technical meaning ”doing things together”. But participation, even in trivial situations, also has a goal-oriented aspect which means decision making and control are involved. E-participation is a term meaning the support through information and communication technologies of processes involved in government and governance. The e-participation method employed depends to a certain extent on whether the participation concerns service delivery or policy scrutiny. It may also depend on where in the policy or service delivery life cycle the engagement exercise is to take place.

One can consider policy making to comprise 5 high level stages [Macintosh et al., 2005]. These are: **Agenda setting; Analysis; Formulating the policy; Implementing the policy**: **Monitoring the policy.** Similarly, with regard to defining and delivering services, one can consider 4 high level stages: **Need for a service; Design of service; Implementation of service; Monitoring of service.** Depending on what extent or “how far” citizens are involved in policy formulation and new service delivery, four levels are distinguished: **eInforming, eConsulting, eCollaborating** and **eEmpowering.**

Collaborative environments have a high potential for supporting e-participation as they offer a variety of tools for synchronous and asynchronous communication and collaboration among several participants, such as:

***Chat***

Chat rooms allow users to freely interact with one another. Participants post messages to others in shared ‘*chat spaces’*, for normally an hour at most. Each participant can see all other’s responses and these often overlap.

Chat is time-specific and limited. Some of the challenges of *chat* are that it can be somewhat difficult to arrange; timing things to suit all concerned, and ensuring that all people are aware of the opportunity.

***Online Discussions***

A discussion forum is a website for an online discussion group where users, usually with common interests, can exchange open messages. It typically shows a list of topics people are concerned about. Users can pick a topic and see a “thread” of messages and replies about it and post their own message [Macintosh et al., 2005].

Discussion fora are well suited to following similar threads of online discussion when used for e-participation, supporting the exchange of points of view. All users can typically read all comments, however, in certain instances, users have to be registered in order to post and reply to comments. Careful design is required to ensure users can easily navigate through the different threads. Their main advantages are that they have the potential to support interaction, thought, deliberation, debate and allow for a full discussion.

Online discussion fora can be pre-moderated or post-moderated:

- *Pre-moderated*: this means all responses contributed by participants are vetted before they appear online to ensure conditions of use are met. The risk is that participants are discouraged from contributing since they do not see their contribution immediately. The value is the greater degree of control over what is disclosed in online discussion.

- *Post-moderated*: this means that all responses contributed by participants are vetted within a defined period (e.g. 24 hours). The risk is that participants may use abusive language which might be read by others before the message is removed. The value is that they can immediately see that their contribution has been included in the online dialogue.

***Virtual Communities***

Virtual Communities are online spaces in which users with a shared interest can gather to communicate and build relationships [Macintosh et al., 2005].

There is typically a web site organised specifically to support an issue, a range of connected issues or a geographical area.

***Group Decision Making***

Collaborative environments, enhanced with structured group decision making capabilities, can facilitate all argumentative discourses in the context of both policy making and service delivery processes.

A number of projects were launched under the fifth Framework Programme (FP5) in order to promote and enable the online participation of all stakeholders in decision making. The projects cover a wide range of topics within the eParticipation field, like, for instance, the improvement of the interaction between citizens and public administrations, through e-collaboration functionality.

***DEMOS** - Delphi Mediation Online System (IST-1999-20530)*

http://www.demos-project.org/

DEMOS is designed to support and encourage ‘online democracy'. Its vision and long-term goal is to make political processes more democratic by motivating and enabling all citizens, whatever their interests, technical skills or income, to take an active and effective part in debate and decision-making. To achieve this goal, DEMOS exploits the full potential of the Internet as a mass communication medium. More specifically, the DEMOS project builds an online platform supporting large-scale citizen participation and result-oriented public debate on political topics on the Internet – on the local, national or European level.

***WEBOCRACY** - Web Technologies Supporting Direct Participation in Democratic Processes*

http://www.wolverhampton.gov.uk/government_democracy/council/modernisation/egovernment/webocracy.htm

The project relies on Internet-based software components of the Java platform and combines three proven methods of social research and organisation development: the Delphi, the survey and the mediation method. At the heart of DEMOS is forum software, Zeno, with special support for moderators and mediators of consensus oriented discourses. DEMOS has also an attractive, user-friendly interface and tool to conduct online surveys and categorize contributions.

The project will design and develop a Webocrat system - a Web-based system comprising functions of content management, computer-mediated discussion, organisational memory, information retrieval, data mining, and knowledge modelling. The system will support communication and discussion, publication of documents, browsing and navigation, opinion polling on questions of public interest, intelligent retrieval, user alerting, and convenient access to information based on individual needs. The aim of the project at the organisational level is to facilitate discussion between citizens and representatives of local governments, to enable a user-friendly access to information, databases and knowledge repositories for citizens, public servants and elected representatives, to support public discussion and to provide citizens with opportunity to express their opinion.

## 4.2 eParticipation Application Scenarios

Some of the eParticipation areas, described in Deliverable 5.1, are considered suitable for applying e-Collaboration technology. These are

- Community building / Collaborative Environments,
- Electioneering
- Consultation
- Discourse

In the following paragraphs are described application scenarios in the e-Participation field, using e-Collaboration technology.

***Community building / Collaborative Environments,***

Users, become members of virtual communities and may contribute online, using a combination of the e-participation tools, and even initiate debate on specific issues in all stages of the policy formulation and new service delivery processes. Particularly, when the need for a major change in policy or service has been highlighted, the contributions of citizens can be effectively exploited.

***Consultation, Electioneering***

E-Participation Chat Rooms are offered for a specific time-horizon, normally an hour at most. Often, participants need to register in advance. Citizens can take part in live question-answer panels with elected representatives, government staff or experts at pre-arranged times. In case there are young people (under 16's) that wish to take part in such a chat space, moderation is available to control any disruptive behaviour. Participants are given the opportunity to appreciate other perspectives, while transparency in decision-making is enhanced as questions and opinions can be directly corresponded.

In the context of election campaigns, politicians can have direct contact with their voters, listen to their problems and express their intended course of action towards the problem solution.

In the consultation area of eParticipation discussion forums are also useful, especially for the development of a complex policy. However, since they do potentially allow for a broad and deep debate, staff time and skills are required to moderate (at least for legal reasons), support and facilitate such discussion, as well as the content analysis skills to analyse contributions and produce reports to embed the results into the political process and to give feed-back to the users.

Specific e-participation *discussion fora* are the:

- Issue-based fora, i.e. organised around policy issues that have been formulated by policy-makers, interest groups or 'experts', and presented as the heading of one or more discussion 'threads'. Responses are sought in order to gauge opinion or solicit ideas. Position statements, links to topic-related websites and other background information may also be presented, although they are often lacking.

- Policy-based fora, i.e. organised around themes/issues that relate directly to a draft policy that is meant to address these, and where discussion threads are intended to solicit responses from those affected. Participants might be encouraged to submit alternative ideas and suggestions but the format implies that what is being sought is an indication of how far the participants agree (or not) with the proposals, and why.

A clear 'conditions of use' statement is required which can be followed by both moderators and participants, and in some cases registration is necessary. Besides it should be described (and planned before starting) how the results will be used for the further political process, a commitment of the political authorities is helpful.

Discussion forums provide structured interaction that extends normally over a period of days or weeks. When used for consultations, each forum should last between 4 to 12 weeks to increase participation and also to re-visit the forum and reply to others with a definite time restriction to keep discussion focussed as potential users know that they have to make their statement within a specific time-horizon.

For discussion forum-based participation initiatives it is important to have a clear statement of what can and cannot be typed as comments into the forum. A Conditions of Use statement may require legal advice, but the outcomes must be clearly visible and understandable by all. There will also be a need for discussion moderation.

### *Discourse*

Collaborative environments enhanced with decision making capabilities can help citizen groups in reaching a shared understanding of the issues set, as knowledge elicitation, sharing and construction are supported. Taking into account the input provided by government agencies and citizens, illustrative discourse-based knowledge graphs can be constructed, composed of the ideas expressed so far as well as their supporting documents. Moreover through the integrated decision support mechanisms, discussants are continuously informed about the status of each discourse item asserted so far and further contemplate on them according to their beliefs and interests on the outcome of the discussion. Storage of documents and messages being asserted in an ongoing discussion should take place in an automatic way that is upon their insertion in the knowledge graph. On the other hand, retrieval of knowledge can be performed through appropriate interfaces, which aid users explore the contents of the knowledge base and exploit previously stored or generated knowledge for their current needs [Karacapilidis et al, 2004].

# 5 References


Agostini, A., De Michelis, G., (2000) A Light Workflow Management System Using Simple Process Models. *Computer Supported Cooperative Work*, 9, 335-363

Araujo, R., Borges, M., (2001). Extending the Software Process Culture – An Approach based on Groupware and Workflow. *F. Bomarius and S.Komi-Sirvio (Eds): PROFES 2001, LNCS 2188, 297-311.*

Bafoutsou, G., Mentzas, G., (2001). A Comparative Analysis of Web-based Collaborative Systems, Proceedings of the t*welfth international workshop on Database and Expert Systems Applications (DEXA 2001), 496-500. (3-7 September, Munich, Germany).*

Bafoutsou, G. and G. Mentzas (2002) Review and functional classification of collaborative systems, *International Journal of Information Management*, Volume 22, Issue 4, August 2002, pp. 281-305.

Borenstein, N. S., (1992). Computational mail as network infrastructure for computer supported cooperative work. *Proceedings from the ACM Conference on CSCW*, Toronto, Canada, ACM Press, *67-73.*

Borghoff, U.M. and Schlichter, (1995) J.H. *Rechnergestuetzte Gruppenarbeit:Eine Einfuehrung in Verteilte Anwendungen.Springer-Verlag, Heidelberg, 1995.*

Bowers, J. and Churcher, J., (1988). Local and Global Structuring of Computer Medicated Communication: Developing Linguistic Perspectives on CSCW in COSMOS. *Proceedings of the Conference on Computer-Supported Cooperative Work -CSCW'88 (Sept. 26-28, Portland, OR). ACM, N.Y., 1988. 125-139.*

Bussler, C., (2001). The Role of B2B Protocols in Inter-Enterprise Process Execution. *F.Casati, D. Georgakopoulos, M-C.Shan (Eds): TES 2001, LNCS2193, 16-29.*

DeSanctis, G., Gallupe, R.B., (1987**)**. A Foundation for the Study of Group Decision Support Systems , *Management Science, 23(5), 589-609.*

Ellis, L., Gibbs, S.J., Rein, G.L., (1991) Groupware: Some Issues and Experiences, *Communications of the ACM, 34,(1), p. 38-58.*

Ellis, L., (2000). An Evaluation Framework for Collaborative Systems*, Colorado University Technical Report CU-CS-901-00.*

Greif, I., (1988) Computer-Supported Cooperative Work: A Book of Readings. In Computer and People Series, *Morgan Kaufmann Publishers, San Mateo, CA, 1988.*

Grudin, J., (1994). Computer-Supported Cooperative Work: History and Focus. *Communications of the ACM, 37, 1, 92-105.*

Haake, J., Wang, W., (1999). Flexible support for business processes: extending cooperative hypermedia with process support. *Information and Software Technology 41(1999) 355-366.*

Halaris, C., Kerridge, S., Bafoutsou, G., Mentzas, G., Kerridge, Susan, (2001). An integrated system supporting virtual consortia in the construction sector, *Journal of Organizational Computing and Electronic Commerce – Special Issue on "B2B eCommerce and e-Supply Chain Management*.

Jarczyk, A., Loffler, P., Volksen, G., (1992). Computer Supported Cooperative Work (CSCW) - State of the Art and Suggestions for Future Work. *Internal Report*, *Version 1.0, Siemens AG, Corporate Research.*

Johansen, R., (1988). Groupware. Computer-Support for Business Teams. *The Free Press, New York.*

Kammer, P. & McDonald, D. (1999). Putting Words to Work: Integrating Conversation with Workflow Modeling. *Technical Report, UCI-ICS-99-30, Department of Information and Computer Science, University of California, Irvine, August 1, 1999.*

Karacapilidis N., Loukis E., and Dimopoulos S., (2004). A Web-based System for Supporting Structured Collaboration in the Public Sector, *Third International Conference EGOV2004, Zaragoza, Spain, 2004.*

Kreifelts, T., Hinrichs, E., Woetzel, G., (1999). BSCW-Flow: Workflow in Web-based Shared Workspaces. *Proceedings of the Workshop on Cross-Organizational Workflow Management and Co-ordination of the Work Activities Coordination and Collaboration Conference (WACC'99)*, *San Francisco*.

Lotus, (1996). Groupware - Communication, Collaboration, Coordination*, (Lotus Development Corporation, http://www.lotus.com/bible/*

Characterizing E-Participation in Policy-Making (2004). *Proceedings of the 37th Hawaii International Conference on System Sciences, 2004.*

Macintosh, A. Coleman, S. and Lalljee, M., (2005). eMethods for Public Engagement Published by Bristol City Council for The Local eDemocracy National Project. *Available at: http://www.e-democracy. gov.uk/ knowledgepool/ default.htm?mode=1&pk_document=466.*

Malone, T and Crowston, K., (1994). The interdisciplinary study of coordination. *ACM Computing Surveys, 26(1): 87—119.*

Mc Daniel, A., Carr, J., (2005). Cross-Boundary Leadership: A New Challenge for eGovernment, *18th Bled eConference, eIntegration in Action*, *Bled, Slovenia, June 2005*

McGrath, J., & Hollingshead, A.B., (1994). Groups Interacting with Technology. *Thousand Oaks, CA: Sage*.

Meier, C. (2002). *UNITE and the state of art in collaboration support systems*. *UNITE Project Forum, Darmstadt, Germany*.

Mentzas, G. (1993). Coordination of joint tasks in organizational processes, *Journal of Information Technology, v. 8, p. 139-150*.

Mentzas, G. and G. Bafoutsou (2004) Electronic Collaboration, Communication and Cooperation: A Taxonomy and a Prototype, Chapter II in E. Y. Li and T. C.

Du (eds) *Advances in Electronic Business, Vol. 1, Theme: "Collaborative Commerce"*, Idea Group Publishing, pp. 19-52.

Schooler, Eve (1996). Conferencing and collaborative computing. *Multimedia Systems (1996) 4:210-225.*

Scott, M., Golden, W., Hughes, M., (2004). A Click and Bricks Strategy For eGovernment, *17th Bled eCommerce Conference, eGlobal*, *Bled, Slovenia, June 2004.*

Teufel, S., Sauter, C., Muehlherr, T. and Bauknecht, K., (1995). *Computerunterstuetzung fuer di Gruppenarbeit. Addison-Wesley, Bonn*.

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D5.2.2 - Argumentation Support Systems

| | |
|---|---|
| **Editor :** | **Tom Gordon** |
| **Revision :** | Version 1.6 |
| **Dissemination Level :** | [TA p. 63] |
| **Author(s) :** | Tom Gordon, Ann Macintosh, Alastair Renton |
| **Due date of deliverable :** | 31 December 2006 |
| **Actual submission date :** | |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | WP5 |
| **Organisation name of lead contractor for this sub-deliverable :** | Fraunhofer |

**Abstract**. As governments seek to consult their citizens over matters of policy, it becomes increasingly important that citizens receive the relevant information in a medium that they can, and will, want to use in forming their opinion upon consultative issues. This report presents argumentation support systems and sample eParticipation application scenarios of these systems, in order to assess their potential contribution to the consultation process. The systems presented cover techniques for the presentation of complex information in a thematically arranged format, for identifying those issues that generate a significant response, for collating consultation responses and representing them within an argument structure, and for checking upon the consistency of contributions to a debate. As such, argumentation support systems have something valuable to offer both government and civil society.

**Project funded by the European Community under the FP6 IST Programme**

# Executive Summary

Argumentation Support Systems are computer software for helping people to participate in various kinds of goal-directed dialogues in which arguments are exchanged. Their potential relevance for eParticipation should be readily apparent, since the goal of eParticipation is to engage citizens in dialogues with government about such matters as public policy, plans, or legislation. Surely argumentation plays a central role in this process. In a public consultation, for example, citizens are given an opportunity to not only make suggestions, but also support these suggestions with arguments.

Typically eParticipation projects make use of generic groupware systems, such as discussion forums and online surveys. These generic groupware systems, however, do not provide specific technical support for argumentation. For example, they provide no way for a citizen to obtain a quick overview of the issues which have been raised, to list ideas which may have been proposed for resolving such issues, to see in one place the arguments pro and con these proposals, or to get an idea about which positions currently have the best support given the arguments put forward thus far in the dialogue. These are just a few of the kinds of services offered by argumentation support systems.

This report provides an introduction to the theory of argumentation; summarizes prior work of the leading research groups on modelling argumentation and supporting argumentation with software tools; describes various prior applications of argument support systems, mostly in research pilot projects; and presents a number of eParticipation application scenarios for argumentation support systems, as a source of ideas for future pilot projects.

A number of argumentation support systems and associated tools are presented. Some of these focus on the visualization of arguments and here the graphical notation and user interface are important features. Others focus on providing analysis of the situation but typically with a more limited graphical user interface. A number of underlying argumentation models are used including those based on Issue-Based Information Systems (IBIS) and the diagramming method developed by Wigmore for mapping evidence in legal cases. In considering their relevance to eParticipation, we need to consider the features needed to support informed debate to support evidence-based policy-making. The systems presented allow users to access various levels of information, to be able to focus on specific information and to have the ability to organize the gathered data to construct an effective argument – all of which are required for eParticipation.

In eParticipation, there is a clear requirement to better understand how technology can support informed debate on issues but there are two main obstacles in achieving this. The first is that the deliberation is typically on complex issues and therefore there are typically a large number of arguments and counter arguments to consider which when presented in linear text can be confusing for the public at large. Secondly, it is not obvious that many people actually have the necessary critical thinking skills to deliberate on issues. In can be seen that the type of argumentation support systems and tools described in this report have the potential to add value to current eParticipation methods. This is explored further in the section on eParticipation scenarios.

As governments seek to consult their citizens over matters of policy, it becomes increasingly important that citizens receive the relevant information in a medium that they can, and will, want to use in forming their opinion upon consultative issues. This report presents sample eParticipation application scenarios of argumentation support systems in order to assess the potential contribution these systems can make to the consultation process. They cover techniques for the presentation of complex information in a thematically arranged format, for identifying those issues that generate a significant response, for collating consultation responses and representing them within an argument structure, and for checking upon the consistency of contributions to a debate. As such, they have something valuable to offer both government and civil society.

# 1 Introduction

Argumentation Support Systems are computer software for helping people to participate in various kinds of goal-directed dialogues in which arguments are exchanged. Their potential relevance for eParticipation should be readily apparent, since the goal of eParticipation is to engage citizens in dialogues with government about such matters as public policy, plans, or legislation. Surely argumentation plays a central role in this process. In a public consultation, for example, citizens are given an opportunity to comment on draft legislation. These comments will not only contain suggestions for changes, but also support these suggestions with arguments. In some other forms of eParticipation, such as those founded on the ideal of deliberative democracy, other participants are offered an opportunity to view and respond to such arguments with further arguments of their own.

Typically eParticipation projects make use of generic groupware systems, such as discussion forums and online surveys. These generic groupware systems, however, do not provide specific technical support for argumentation. For example, they provide no way for a citizen to obtain a quick overview of the issues which have been raised, to list ideas which may have been proposed for resolving such issues, to see in one place the arguments pro and con these proposals, or to get an idea about which positions currently have the best support given the arguments put forward thus far in the dialogue. These are just a few of the kinds of services offered by argumentation support systems.

The idea of using argumentation support systems for eParticipation is not entirely new. Arguably the idea can be traced back at least to Horst Rittel's pioneering work in the early 1970s on Issue-Based Information Systems (Rittel 1973). Rittel was not a computer scientist but rather a city planner. His idea of an Issue-Based Information System (IBIS) is essentially a visual map of arguments, to help people to collaborate to find solutions to what he called "wicked problems", by which he meant problems which have no algorithmic, scientific or objectively optimal solutions for a variety of reasons, including the lack of consensus among stakeholders about such things as utilities and values. He recognized that city planning, like public policy and legislative development in general, was essentially a social, dialectical process of trying to resolve conflicting goals, values, interests and positions.

One of the first European eParticipation research projects, GeoMed (Geographical Mediation Systems, IE2037), which began in 1996, long before the term "eParticipation" had been coined, aimed to help citizens to participate in city planning by integrating an IBIS-based argumentation support system, Zeno, with a web-based geographical information system (Gordon 1995, Gordon 1996, Gordon 1997). A later version of Zeno served as the technical foundation of the eParticipation platform developed in another European project, DEMOS (Delphi Mediation Online System, IST-1999-20530), which ran from 2000-2004 and was successfully piloted in the cities of Hamburg and Bologna (Gordon 2002, Richter 2002).

The remainder of this report is organized as follows. The next section provides an introduction to the theory of argumentation and an overview of the prior work of leading research groups on modelling argumentation and supporting argumentation with software tools. Next is a section describing various prior applications of argument support systems, mostly in research pilot projects. We then return to the subject of eParticipation by presenting a number of eParticipation application scenarios for argumentation support systems, as a source of ideas for future pilot projects. Finally, there is a section recapitulating the main conclusions.

# 2 Overall Description of the Technology

Argumentation Support Systems cannot be understood or evaluated without some appreciation of the theory of argumentation. Moreover, software tools should be based on carefully considered computational models of the application domain and its tasks, according to the principles of good software engineering. For these reasons, our description of the technology of argumentation support systems has two parts: the first part outlines the theory of argumentation, primarily from the perspective of the field of philosophy, and introduces various efforts to develop formal, computational models of argumentation within computer science; the following section focuses on more applied computer science research by presenting software tools which have been developed for supporting various argumentation tasks, such as argument visualization and mediation systems.

## 2.1 Argumentation Theory

In 1962, Carl Adam Petri, the renowned German computer scientist and inventor of Petri Nets, said (Petri 1962): "Now is the time to shift our view of computers from communications medium to negotiation medium, from knowledge processing to interest processing". Considering that the first email systems had just been invented in 1961 and that the ARPANET computer network, the predecessor of the Internet, did not appear until 1969, this was quite a remarkable statement for the time. Petri anticipated that computer networks would not only be used as a communications medium, for transferring data from place to place, but also provide some kind of intelligent support for helping people to resolve conflicts of interest when confronting practical problems.

Practical problems are problems requiring some action to be taken to achieve goals and promote values. Such problems range from the trivial, such as deciding what to cook for dinner, to the global issues of our time, such as how to preserve the environment or prevent the further proliferation of nuclear weapons. Theoretical problems, in contrast, are concerned with how best to acquire and organize our knowledge of the way the world works. Whereas theories can be revised or replaced at any time, practical decisions typically have consequences, once they have been acted upon, which cannot be undone.

Existing information and communications technology is of limited use for helping people to solve practical problems. Algorithms require the problem to be "well-defined" and perfect input data to produce correct results, following the principle of "garbage in, garbage out". Automatic theorem provers are similar; they may be able to tell us if some premises are inconsistent or what conclusions are entailed by the premises, but they provide no support per se for constructing or challenging the premises. Large databases, particularly loosely coupled and distributed databases such as the World-Wide Web, can provide access to enormous amounts of data, but the informativeness of this data can be questionable and it may be practically impossible to find relevant information or determine its quality (c.f. "information overload"). Knowledge-based systems can provide useful support in narrowly-defined technical domains, but are too expensive to build and maintain for helping with everyday problems requiring common sense ("knowledge acquisition bottleneck"). Also, support systems based on decision theory make strong assumptions about knowledge of the dimensions of the problem space and consensus about the utility curves which are unrealistic for most practical problems.

Typically, when confronted with a practical problem, there is both too much and not enough information, the decision must be made within a limited period of time and other resources such as personnel and money can also be scarce. The expected value of the outcome is usually not high enough to warrant the development of special purpose software. Opinions will differ about the truth, relevance or value of the available information. Arguments can and will be made both for and against and proposed solutions. Reasoning is "defeasible", i.e. further information may require some conclusions to be retracted or make some other solution appear more promising. Value judgments about ethical, legal, political, business or even aesthetic issues are at least as important as objective facts or knowledge about the problem domain. Various stakeholders, with divergent interests, may be affected by the decision. Negotiation may be necessary.

The purpose of Argumentation Support Systems is to support and facilitate the making of practical decisions under such circumstances. The aim is to help assure that the decision-making process is efficient, transparent, open, fair and rational. Not coincidentally, these goals have much in common with the goals of "good governance" and e-participation (Malkia 2004).

The theoretical subfield of computer science which studies the foundations of Argumentation Support Systems is young and goes by many names, such as Computational Models of (Natural) Argumentation or Computational Dialectics. Much work has been conducted as part of Artificial Intelligence, especially in the interdisciplinary field of Artificial Intelligence and Law.

The concepts of dialectic and argumentation are closely related. The ancient Greeks recognized and studied three normative sciences: logic, rhetoric and dialectic. In modern terms, logic is the study of consequence and inference relations between declarative sentences; rhetoric is the study of effective communication and dialectic is the study of norms and methods for resolving conflicting views, ideas and opinions. Argumentation straddles rhetoric and dialectic: whereas rhetoric is concerned with how to select and present arguments, dialectic addresses the question of how to organize the process of exchanging and evaluating arguments in goal-directed dialogues. Whereas the term "argumentation" emphasizes the process of exchanging and evaluating arguments in dialogues, the term "dialectic" emphasizes the process of resolving conflicting arguments (pro v. con), interests (proponent v. opponent) and ideas (thesis v. antithesis). The conflict of interests between two parties can be generalized to dialogues with more than two stakeholders, as is often the case in the context of e-participation.

Thus far our aim has been to introduce the topic of argumentation support systems and demonstrate its relevance for e-participation. In the remainder of this section we present an overview of the modern theory of argumentation, from the field of philosophy, and a summary of computer science research on computational models of argumentation.

This brief overview of the modern philosophy of argumentation is based on Douglas Walton's recent textbook, "Fundamentals of Critical Argumentation" (Walton 2006), beginning with the concept of an argument. An argument links a set of statements, the premises, to another statement, the conclusion. The premises may be labelled with additional information, about their role in the argument. Aristotle's theory of syllogism, for example, distinguished major premises from minor premises. The basic idea is that the premises provide some kind of support for the conclusion. If the premises are accepted, then the argument, if it is a good one, lends some weight to the conclusion.

The goal of argumentation is often described as discovering or determining the "truth" of some claim, where a claim is a statement which has been asserted by some party in the dialogue. When the claim is about a factual or theoretical issue, this may make sense, at least as an ideal. However, when the issue being discussed is about what action to take in order to solve some practical problem, this characterization of the goal of argumentation is more problematical. If for example, in an e-participation context, the plan of a city to build the airport is being subjected to public review, one would not ordinarily characterize this as being an issue of truth or falsity. The question is not whether the plan is true, but whether it is good, acceptable or well-advised.

For this reason, among others, the goal of argumentation is to determine the acceptability of claims, rather than their truth. In the case of factual claims, ideally only true claims would be acceptable. Given unlimited resources, the argumentation should conclude that a factual statement is acceptable if and only if it is true. But in practice, resources will typically be limited and we will often have to decide whether or not to accept claims with less than complete certainty about their truth. Consider criminal cases, to take a familiar example, where a person can be convicted of having committed a crime when the evidence is conclusive "beyond all reasonable doubt". Although this is a high standard of proof, it does not require complete certainty.

Good arguments provide reasons for accepting their conclusion, the conclusion need not be a logical consequence of the premises. Logical consequences are necessary, by virtue of their form, irrespective of their content. Arguments, in contrast, are substantive and "defeasible". They are substantive because they depend not only on the form of the premises, but also their content and acceptability. And they are defeasible because their conclusions are only plausible, not certain, and may be defeated in various ways by additional information, for example by revealing implicit premises which turn out to be untenable or by bringing forward better counterarguments. In the field of Artificial Intelligence, this property of argumentation is known as "nonmonotonicity", a term borrowed from mathematics.

As just suggested, some premises of arguments may be implicit. For the sake of efficiency, the norms of argumentation do not require all premises to be made explicit, at least not immediately. For example, premises which are thought to be common knowledge, or otherwise already accepted by the other participants, are typically left implicit. "Socrates is a man, therefore Socrates is mortal" to use a standard example, is a perfectly understandable argument, even though the major premise "All men are mortal." has been omitted. Implicit premises can be revealed and possibly challenged during the dialogue as necessary.

There are many different kinds of arguments and much research has gone into discovering and classifying various patterns of argument, based on an analysis of the structure and content of arguments reconstructed from natural language texts. These patterns of argument have come to be called "argumentation schemes". Although they are the result of empirical case studies, they also have a normative side. They are a useful tool both for guiding the reconstructing of arguments put forward by other parties, so as to open them up to critical analysis and evaluation, as well supporting the construction ("invention") of new arguments to put forward in support of ones own claims or to counter the arguments of others.

Argumentation schemes generalize the concept of an inference rule to cover plausible as well as deductive and inductive forms of argument. Argumentation schemes are conventional patterns of argument, historically rooted in Aristotle's "Topics" (Slomkowski 1997). Unlike inference rules, argumentation schemes may be domain dependent. Each scheme comes with a set of "critical questions" for evaluating and challenging arguments which use the scheme. For example, the scheme for argument from expert opinion includes a critical question about whether the expert is biased. Argumentation schemes are useful for several purposes, including reconstructing and classifying arguments, criticizing arguments, and as templates for making new arguments.

Since argumentation schemes may be domain dependent, there are an unlimited number of such schemes. Domain dependent schemes, in fields such as the law, may evolve along with the knowledge of some domain. Many schemes, however, are general purpose. Walton and his colleagues have taken on the project of collecting and classifying general purpose schemes. To date their collection contains about 60 schemes. Examples include Argument from Expert Opinion, mentioned previously, Argument from Popular Opinion, Argument from Analogy, Argument from Correlation to Cause, Argument from Consequence, Argument from Sign and Argument from Verbal Classification.

When evaluating arguments put forth in a dialogue, one issue is the "validity" of the argument. An invalid argument has no weight, i.e. provides no support for its conclusion. But how shall validity be defined? In classical deductive logic, an inference is valid if and only if the conclusion must logically be true if the premises are true. This conception of validity is too stringent for arguments, since these only provide plausible support for their conclusions. Nonmonotonic logics strengthen the consequence relation to support consequences which are only plausible. Consequences in nonmonotonic logics are defeasible: it may be that some consequence of a set of premises is not a consequence of some superset of these premises. That is, additional information may require plausible conclusions to have to be retracted.

Nonmonotonic logics retain however the relational approach of argument validity of classical logic: whether or not an argument is valid depends only on the relationship between the set of premises and the conclusion. Walton's theory of argumentation, however, takes a more contextual, procedural view of argument validity: an argument is "valid" if and only it furthers the goals of the dialogue in which it is put forward. From this perspective, the validity of an argument can depend on the state and history of the dialogue. To give a practical example: an argument in favor of some proposal made during the brainstorming phase of a deliberation might be valid during the process of selecting some of these brainstorming ideas for a more in-depth evaluation in the next phase of the deliberation, but not valid in this later phase if this particular proposal had not been selected. To sum up: from a dialectical perspective, whether or not a argument is valid depends on how it is used in a dialogue, not merely on the relation between its premises and conclusion.

Whether or not an argument has been used properly or furthers the goals of the dialogue, depends also on the type of dialogue. Walton has developed a taxonomy or "ontology" of dialogue types, as illustrated in Figure 1.

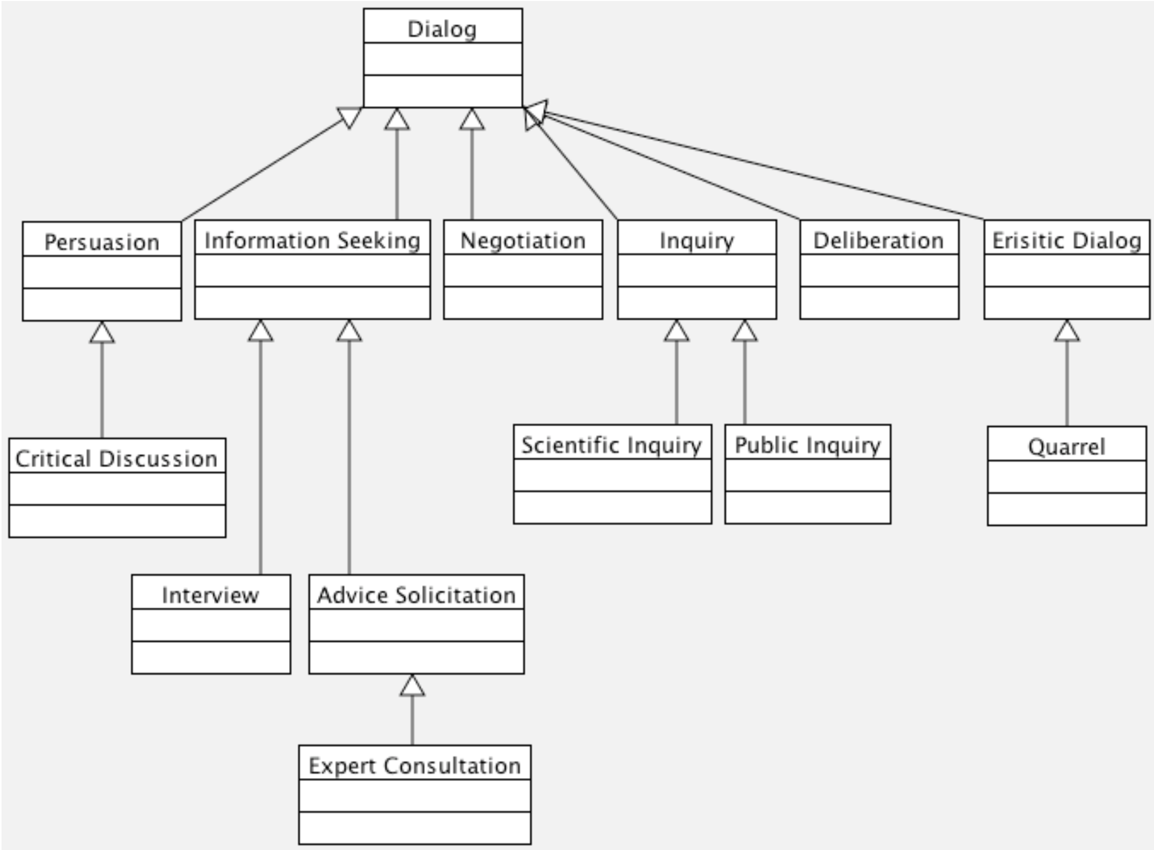



**Figure 1: Taxonomy of Dialogue Types**

Persuasion dialogues debate the truth of some statement. One party, the proponent, claims that some statement is true. The other party, called the respondent, challenges this claim. There are several subtypes of persuasion dialogues. In a "dispute", the respondent not only challenges the proponent's claim, but also claims some opposing, contradictory statement to be true. The roles in a dispute are symmetric. The proponent and respondent each have a burden of proof, for their respective claims. More common, however, is the "dissent" form of persuasion dialogue, in which the respondent only doubts the proponent's claim, but makes no claim of his own. In a dissent, the proponent has the burden of proof and must produce the stronger arguments. The arguments of the respondent need only be strong enough to cast doubt on the proponent's claim.

Although the dialogue types are usually described as involving two parties, they can be generalized to any number of parties. More important than the number of participants is their roles in the dialogue. Several participants could share a role.

An information seeking dialogue has the goal of seeking advice. The starting point is not the assertion of some claim, as in persuasion dialogue, but rather the asking of a question. Expert consultations, for example with medical doctors or lawyers, are a subtype of information seeking dialogues.

The goal of negotiation dialogues is to make a "deal", i.e. to reach an agreement on how to exchange such things as goods, services or money. The starting point is neither a question nor a claim, but rather an offer. This can be accepted by the other party or modified in a counteroffer.

An inquiry is a methodical investigation of some matter, to explain or understand some observations or data. Scientific inquiries try to explain natural phenomena by developing hypotheses and constructing, evaluating and comparing scientific theories. Public inquiries investigate such things as accidents or crimes. The starting point of an inquiry consists of the observations in need of explanation. These observations are not being called into question, unlike the claim of a persuasion dialogue. The question is not whether these observations are true, but how best to explain them.

Deliberation dialogues are about choosing some course of action which takes into account the interests of multiple stakeholders. In a deliberation, one of the first tasks is to identify the stakeholders and their interests. They may not all be participants in the dialogue, at least not initially. And it may not be practical for every stakeholder to take part in the dialogue personally. Stakeholders may need to be represented by others. A common mistake in deliberation is for participants to make and try to defend specific proposals at too early a stage in the dialogue. It is usually better to first spend time trying to identify the stakeholders and understand their interests. Brainstorming may come next, in which ideas are freely collected but participants are not supposed to commit themselves yet to particular proposals.

So-called "eristic" dialogues, from the ancient Greek word meaning wrangle or strife, is an emotional kind of dialogue in which the participants vent their anger, frustration or other deep feelings. Eristic dialogues are considered by some to be irrational and to have no other goal than to "argue for the sake of argument". Walton's view, however, is that such dialogues can serve a positive, "cathartic" function and that they are, like the other kinds of dialogues in his typology, guided by norms, even if these norms are quite relaxed compared to the other dialogue types. For example, the basic civility norms requiring participants to do such things as take turns and give each other a fair opportunity to express their views, remain in force.

Actual dialogues may be mixtures of these various types and a dialogue may shift from one type to the other and back. For example, during a negotiation a salesman may make some claims about the product that might be called into question by the customer, causing a temporary shift to a persuasion dialogue. Similarly, in a deliberation, once the stage has been met to evaluate specific proposals, each such evaluation could take the form of a persuasion dialogue.

What kinds of dialogues are relevant for e-participation? The field of e-participation distinguishes various forms or degrees of "citizen-engagement", such as consultation and deliberation. In a consultation, the government publishes draft plans or legislation and provides citizens with an opportunity to submit comments, but not an opportunity to view or discuss each others comments or to engage the government in a true dialogue. These comments may range from merely casting doubt on the government's draft, criticizing it with arguments against the proposal or, at the other extreme, contain proposals for changes to the draft, supported by arguments. Such consultations are probably best classified as information seeking dialogues, in terms of Walton's typology. The government is seeking information from citizens. Deliberative democracy is some form of deliberation dialogue, but the particular characteristics of deliberative democracy, which distinguish it from general-purpose deliberation dialogues, require further study. As for eristic dialogues, surely when open, unmoderated discussion forums are made available for e-participation, there is a risk that some dialogues will be of this type.

Dialogue types are defined along several dimensions: the purpose or goal of the dialogue, the roles of the participants, the speech acts available, the termination criteria, a process model and a "protocol" for regulating this process. Dialogue types in argumentation theory are normative models of communication. If argumentation dialogues are viewed as games, then the participants are its players, the speech acts its moves, and the protocol defines it rules.

Speech acts are uses of natural language in dialogues, such as asking questions, making claims, putting forward arguments or counterarguments, making concessions or retracting claims. The protocol defines the pre- and postconditions of these speech acts, to regulate when a speech act may be made and, if it is allowed, with what effect. This may depend on the stage of the process and the state of the dialogue, taking into consideration the prior history of the dialogue, i.e. what has already been said.

In addition to defining the preconditions and postconditions of speech acts, the protocol will include rules regulating such things as termination conditions (When is the dialogue finished?), commitments rules (When does a party become committed to some statement?), proof standards (How are the arguments pro and con some statement to be balanced, weighed or otherwise aggregated for each issue?), and finally the distribution of the "burden of proof". There are various kinds of proof burdens to consider: the "burden of questioning" regulates whether some statement can be assumed to be true so long as it has not been called into question; the "burden of production" regulates which party is responsible for producing arguments or evidence suggesting that some presumption may not hold; and the "burden of persuasion" regulates which party must have the stronger arguments when the time comes to make a decision. Usually the same party will have both the burden of production and the burden of persuasion. But this is not always the case. In criminal law, for example, the defense has the burden of production for any exceptions to crimes, such as self-defense in murder cases, but the prosecution has the burden of persuasion, even for such exceptions. Thus, to continue with the murder example, the prosecution has the burden of persuading the court that the killing was not done in self-defense, once the defendant has produced sufficient evidence to meet his burden of production.

## 2.2 Computational Models of Argumentation

This section provides an overview of computer science research on modeling argumentation. Computational models of argumentation are formal models designed for use in specifications of argumentation support systems. These are mathematical models, using such mathematical tools as set theory and formal logic. What makes such a mathematical model "computational" is its intended use as a foundation for computer applications. Thus, computational properties, such as decidability and computational complexity are relevant. Computational models can also themselves be represented in software, using high-level functional or logic programming languages. These "executable specifications" facilitate the empirical testing and evaluation of the models.

It will be helpful for structuring this presentation of computational models to first take a look at the various kinds of argumentation tasks we would like these models to support. Based on prior analyses of argumentation tasks and their interrelationships (Brewka 1994, Prakken 1995, Bench-Capon 2003), we distinguish the following four layers:

- The "logical layer" is responsible for representing statements and argumentation schemes into order to construct or generate arguments by applying argumentation schemes to a "knowledge base" of statements.
- The "dialectical layer" is responsible for structuring, evaluating and comparing arguments which have been put forward during the dialogue, and informing

participants about the status of statements and arguments given these arguments. We also will include the task of "reconstructing" arguments from natural language texts in this level.

- The "procedural layer" is responsible for supporting the process of argumentation, facilitating and guiding the dialogue, to help assure its achieves its normative goals. This layers includes the facilitation tasks of moderators and mediators. One of these tasks is to help participants to obey procedural rules, i.e. the argumentation protocol for the applicable dialogue type. This task in turn requires keeping track of the commitments of the participants in the dialogue.
- Finally, the "rhetorical layer" is responsible for helping participants to "play the game" well. Whereas the procedural layer facilitates the normative goals of the dialogue, this layer provides a private advisor to each participant, analogous to an attorney, to help participants protect and further their own interests. Tasks here include selecting among arguments which could be made and presenting these arguments clearly and persuasively, taking into consideration the intended audience, perhaps using argument visualization techniques. We have also placed the decision-making task of the authority with power to make decisions at this layer.

Figure 2 is a "use case" diagram showing these tasks, divided into the above layers, together with the abstract roles responsible for each task. In concrete situations, one person may have more than one role, some roles may be combined or some roles may need to be distinguished further. For example, in lawsuits the judge may have the moderator role and share the authority role with a jury.

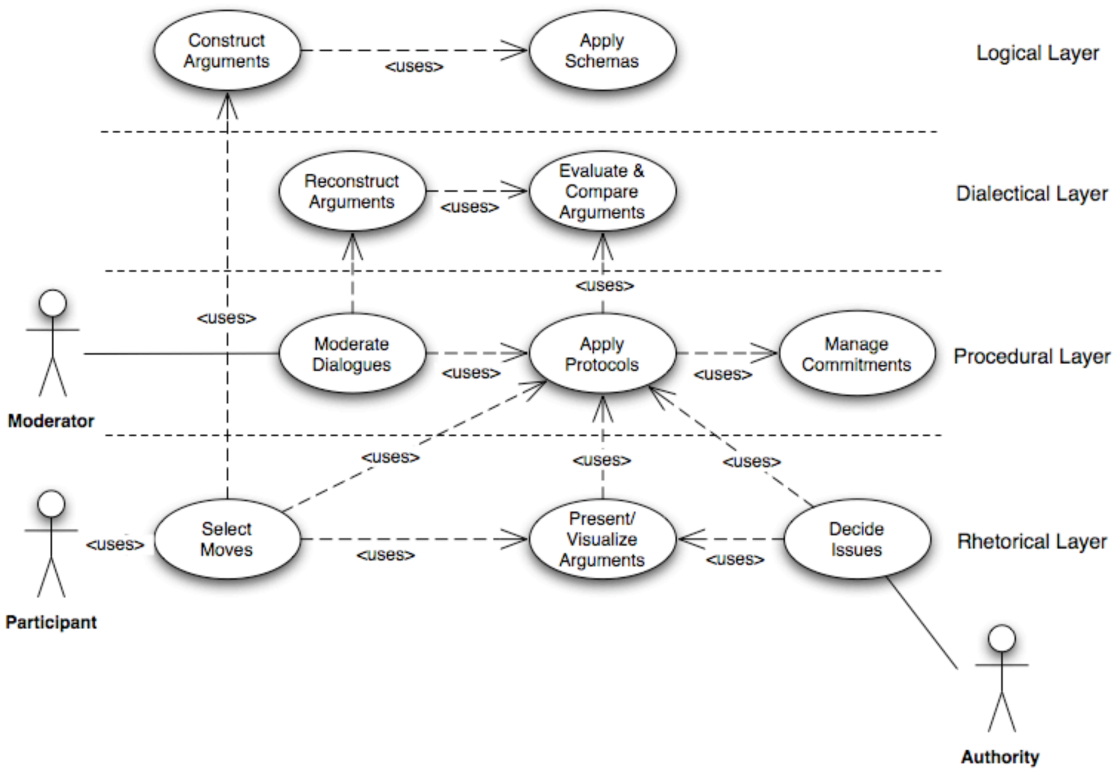



**Figure 2: Argumentation Use Cases**

Let us now begin our review of computational models of argument, starting with the logical layer. Again, here the task, broadly stated, is to construct arguments by applying argumentation schemas to some representation of evidence, facts or knowledge of the domain. The field of Artificial Intelligence (AI) is relevant here (Russell 2003). In mainstream AI, argumentation schemes have not typically been studied as such, explicitly. But AI research on such topics as knowledge representation, nonmonotonic logics, case-based reasoning, reasoning under uncertainty, and machine learning can all be understood, retroactively, as efforts to construct computational models of various argumentation schemes. Moreover, the theory of argumentation schemes provides a framework for understanding how the seemingly diverse forms of reasoning studied by AI can be combined and integrated. For example, research on computational models of legal reasoning in the field of Artificial Intelligence and Law was long divided between case-based and rule-based approaches. But, increasingly, argumentation theory is seen within AI and Law as a way to synthesize these approaches. Prakken has produced a survey of computational models of various argumentation schemes from the field of Artificial Intelligence and Law (Prakken 2005), which was a helpful reference for the work presented in this section. See also (Bench-Capon 2003; Bench-Capon 2006).

The topic of case-based reasoning in AI can be understood as attempts to construct computational models of the scheme for arguments from analogy and related schemes. The first research on case-based reasoning was probably within the interdisciplinary field of AI and Law, at around the time this field was forming in the late 1970s. McCarty's TAXMAN model (McCarty 1977) was one of the seminal works in the field. McCarty's approach to case-based reasoning, based on the idea of constructing and comparing theories of a line of cases, was much ahead of its time. According to this theory-construction model, the better arguments from cases are the ones based on the better, i.e. more "coherent", explanatory theory of those cases.

Probably the most influential computational model of case-based reasoning is Ashley's HYPO model (Ashley 1990). In HYPO, cases are represented as a set of "dimensions", where each dimension includes information about which party is favoured in each direction of the dimension. For example, in the trade secrets domain, the dimension of disclosure favours the defendant, i.e. the party who allegedly violated a trade secret, the more the plaintiff company has disclosed the (so-called) secret to third parties. HYPO formalized the relation of "on-pointedness". One precedent case is more "on-point" than another precedent case if the first case has more dimensions in common with the current case. Arguments were constructed in HYPO by searching for analogous cases, cases with dimensions in common with the current case, which had been decided in favour of the desired party, plaintiff or defendant. (This depended of course on the role of the party trying to construct the argument.) The other party can then try to construct counterarguments, either by distinguishing the current case from precedent case, i.e. by pointing out differences between the two cases, or by searching for more on-point cases in his favour. HYPO is named after its model of reasoning with hypothetical cases. Hypotheticals are imaginary cases, constructed for the sake of argument. Typically, they are variations of the current case, constructed to test proposed interpretations of legal rules or principles. For example, if a party proposes some rule, perhaps by generalizing the decision of some precedent case, the other party could try to construct a hypothetical case showing that this proposed rule leads to some unintuitive or otherwise undesirable result.

The CATO model of case-based reasoning (Aleven 1997), both simplified and extended the HYPO model. It simplifies HYPO by replacing dimensions by boolean "factors", i.e.

propositions which are either true or false in a case. But CATO extends HYPO by organizing these factors into a hierarchy and using this hierarchy to support additional case-based argumentation schemes, in particular schemes for arguments from "downplaying" and "emphasizing" distinctions. A distinction between a precedent case and the current case is downplayed by showing that factors present in both cases have a common ancestor in the hierarchy and arguing that the precedent case is more general, applying to all cases in which this more abstract, common factor is present. For example, if the precedent case involved deception but the current case bribery, one might downplay the distinction between deception and bribery by noting they are both illegal means of obtaining information and arguing the precedent applies to all such illegal means, not just deception.

Other influential models of case-based reasoning in the AI and Law field include GREBE (Branting 2000) and CABARET (Skalak & Rissland 1992). GREBE used semantic networks, the forerunner of ontologies modelled using Description Logic, currently popular in the context of the Semantic Web, in its model of case comparison. CABARET modelled the use of cases to construct arguments about open-textured concepts and included models of argumentation schemes for broadening and narrowing the application of legal rules using cases. Both GREBE and CABARET were early attempts to model an argumentation framework in which argumentation schemes for arguments from both rules and cases could be used together, in an integrated fashion. See also (Prakken & Sartor 1998). Gardner's early model of legal reasoning (Gardner 1987) also needs to be mentioned in this context. Although it was primarily a model of schemes for arguments from rules, it also included scheme for arguments from cases, called interpretation rules, which were applied to open-textured concepts, "when the rules ran out". Loui and Norman (1995) developed a computational model of another case-based argumentation scheme, for arguments from the "rationale" of the case. The scheme exposes presuppositions of the rationale of a case and then argues that these presuppositions do not apply in the present case. For example, in a precedent case which decided that vehicles are not allowed in public parks, there may be a presupposition that the vehicles in question are privately owned. If in the current case the vehicle is not privately owned, this argumentation scheme could be applied to construct an argument that the precedent does not apply. Finally, Bram Roth developed a model of case-based argumentation which also makes use of rationales, represented as reconstructions of the dialectical structure of the arguments in the published opinions of the cases (Roth 2003). In Roth's account, the arguments in a precedent case are applied to the facts of the current case. If the current facts provide at least as much support for the conclusion of the precedent case, considering its arguments, then the conclusion of the precedent case presumptively also applies to the current case. The scheme modelled by Roth is known as argument "a fortiori" (from the stronger argument).

Next we want to address computational models of schemes for arguments from defeasible rules or, as Walton calls them, defeasible generalizations. In the law, the idea of applying rules, by trying to "subsume" the facts of the case under the legal terms of the rule, is quite basic. In the legal philosophy known as "mechanical jurisprudence", this process was thought to be purely deductive. In some early work in the field of AI and Law, this same insight led to experiments with using theorem provers or rule-based systems to build legal expert systems based on first-order logic (Sergot 1986). This approach is adequate for some application scenarios, especially in public administration, and is the basis for most commercial legal knowledge systems today. But models of rules based on classical logic are not well suited for capturing the defeasibility of arguments from rules, since rules can be subject to exceptions, overridden by others rules, or invalid. Nonmonotonic logics have been developed in AI to model reasoning with defeasible rules, but typically these logics do not address the issue of how to integrate reasoning with defeasible rules with other forms of plausible or presumptive reasoning, such as case-based reasoning. Argumentation-theoretic models of reasoning with defeasible rules can overcome these limitations. One influential argumentation-theoretic model of arguments from rules is Hage and Verheij's "Reason-Based Logic" (Hage 1997; Verheij 1996). Other models of defeasible arguments from rules were developed by Gordon (1995), as part of his Pleadings Game model of legal argumentation, and Prakken and Sartor (1996).

When arguments conflict, some way is needed to resolve these conflicts. Some models of argumentation include a "built-in" method for resolving these conflicts. For example, several models always prefer arguments from cases to arguments from rules (Gardner 1987, Branting 2000, Skalak & Rissland 1992). Similarly, some nonmonotonic logics, such as Conditional Entailment (Geffner 1993) always prefer the more specific argument. A more general solution is to support argumentation about argument priorities or strengths (Gordon 1995; Prakken & Sartor 1996; Hage 1996; Verheij 1996; Kowalski & Toni 1996). These priority arguments may apply higher-level principles, such as lex superior (prefer the rule from the higher authority), lex posterior (prefer the newer rule) and lex specialis (prefer the more specific rule), which may themselves be defeasible.

Computational models of argumentation schemes for reasoning with evidence have long been neglected. One of the first models (Lutomoski 1989), represented a number of argumentation schemes for arguments from statistical evidence in the domain of employment discrimination law, including critical questions. Chris Reed, Douglas Walton and Henry Prakken have more recently been working together on computational models of arguments from evidence (Prakken 2003, Bex et al, 2003, Prakken 2004), based on John Pollock's work on a scheme for argument from perception and related schemes (Pollock 1987).

In (Bench-Capon 2002), Bench-Capon analyzed the role of purpose ("teleology") when interpreting a body of case law, motivated by the seminal paper by Berman and Hafner (1993), which identified limitations of the HYPO approach to case-based reasoning in the law. Bench-Capon's central idea is that the rules and rule preferences cannot be derived solely from factors in precedent cases, but must also be informed by the purposes of the rules, i.e. by the values promoted by the rules. Shortly thereafter, Bench-Capon, in collaboration with Sartor, developed this basic idea into a theory-construction model of legal argument (Bench-Capon & Sartor 2003). In this model, legal theories are constructed from precedent cases in a process which takes values and value preferences into consideration to derive and order rules, which may then be applied to the facts of cases to reach decisions. This theory construction approach, first advocated in AI and Law by McCarty (1977), can be viewed as a complex argumentation scheme. The scheme is complex compared to other case-based schemes, because it depends not only on the features of a single case, but rather an analysis of a whole line of cases. The idea of the scheme is to construct a theory capable of explaining the decisions in this line of cases and then to apply this theory to the facts of the current case. Of course several competing theories are possible. Thus the scheme can produce several competing arguments. The conflict between these arguments is resolved by comparing these theories: the better the theory, the better the argument. Which theory is better, or most "coherent" can be debatable and thus an issue to be addressed by further argumentation. There are different criteria for evaluating the quality of theories. Bench-Capon and Sartor address the issue of how to define and model coherence in (Bench-Capon & Sartor 2003). And in (Bench-Capon & Sartor 2001), they present quantitative metrics of theory coherence. See also (Hage 2001). Atkinson (formerly Greenwood), Bench-Capon and McBurney (Atkinson 2005) did further work on modelling teleological reasoning in the law, in which they develop a formal model of the argument scheme for practical reasoning, based on Bench-Capon's Value-Based Argumentation Framework (Bench-Capon 2003). Also relevant for the question of how to model arguments from purpose and values, is Bruce McLaren's thesis (McLaren 1999), in which he develops a formal model of ethical arguments from cases.

Before moving on to models of the dialectical layer, let us say a few words about the role of "classical" knowledge representation in AI for argument construction. The mainstream approach to modelling knowledge uses various subsets of first-order logic, sometimes a couple of subsets in combination. Currently popular is to use a decidable subset of first-order logic, such as Description Logic (Baader, 2003), to model terminological knowledge ("ontologies") and to use some complementary, rule-based, subset of first-order logic to handle knowledge which cannot be expressed in the decidable subset used to model terminology (Russell 2003). This is the approach taken by the World Wide Web Consortium with the Web Ontology Language (McGuinness 2004) and the Semantic Web Rule Language (Horrocks 2004). Since argumentation schemes generalize inference rules of deductive logic, in principle inference engines for knowledge bases expressed in first-order logic, or some subset thereof, may be used to construct arguments, where the arguments represents a formal, deductive proof. Although such reasoners, being limited to first-order logic, are not sufficient for modelling defeasible argumentation, they can be used to construct arguments which can then be compared with, and perhaps defeated by, arguments constructed using other argumentation schemes. For example, arguments from two different OWL ontologies could be pitted against each other. If they are two different versions of the same ontology, the principal of lex posterior could be applied to prefer the argument from the newer version. Or some model of theory coherence could be used to prefer the argument from the more coherent ontology.

Recall that the "dialectical layer" is responsible for structuring, evaluating and comparing arguments which have been put forward during the dialogue, and informing participants about the status of statements and arguments given these arguments. The idea of developing a computer model for managing support and justification relationships between propositions goes back to research on truth and reason maintenance systems in AI, beginning with Jon Doyle's Truth Maintenance System (Doyle 1979). Probably the most famous system of this kind is Johann de Kleer's Assumption-Based Truth Maintenance System (de Kleer 1986). Some nonmonotonic logics, those with an argumentation-theoretic semantics, can be viewed as providing the services of the dialectical layer. Examples include Loui's model of defeat among arguments (Loui 1987), Pollock's OSCAR system (Pollock 1987), which includes an explicit model of relationships between propositions and arguments called "inference graphs", Vreeswijk's work on defeasible dialectics (Vreeswijk 1993), the assumption-based model of defeasible argumentation of (Bondarenko 1997), Prakken and Sartors argumentation-based logic with defeasible priorities (Prakken & Sartor 1997), Verheij's DefLog system (Verheij 2003), and the argumentation semantics for Nute's Defeasible Logic (Nute 1994) developed in (Governatori et al. 2004). An overview of logics for defeasible argumentation is provided by (Prakken & Vreeswijk 2002). Dung's abstract model of argumentation frameworks, which defines the acceptability of arguments solely in terms of an attack relation among arguments, has been extremely influential (Dung 1995), in part because he was able to prove how many prior nonmonotonic logics could be reconstructed as instances of his abstract model. Prakken, however, has argued that Dung's abstract model is not capable of modelling distributions of the burden of proof (Prakken 2001). Prakken and Sartor (2006) have shown that it is important to distinguish between three kinds of burdens (the burden of questioning, the burden of production and the burden of persuasion). And the question of who has some burden must be distinguished from the proof standard used to evaluate whether this burden has been met. Freeman and Farley (1996), were the first to model proof standards, based on such legal proof standards as scintilla of evidence, preponderance of the evidence and beyond reasonable doubt. The Zeno Argumentation Framework (Gordon & Karacapilidis 1997) included a model of argument graphs which used such proof standards to evaluate the dialectical status of statements. Zeno, however, did not distinguish the three kinds of burdens of proof. A recently developed successor of Zeno, called Carneades, in addition to supporting variable proof standards, on an issue-by-issue basis, uses three kinds of premises (ordinary premises, assumptions and exceptions) and information about the dialectical status of statements (undisputed, at issue, accepted or rejected) to allow the three kinds of burden of proof to be allocated (Gordon & Walton 2006).

The first formal models of the "procedural layer" in philosophy were by Hamblin (1970), Rescher (1977) and Mackenzie (1979), but these were not computational. They formally defined protocols for various kinds of argumentation dialogues, in the form of games. One could also mention Lorenzen and Lorenz's Dialog Logic, which is a formal dialogue game for constructing proofs in intuitionistic logic (1978). (From the viewpoint of argumentation theory, this is rather ironic, since intuitionistic logic is even more strict than classical logic about the inferences it allows.) Krabbe (1985) provides a survey of formal systems of dialogue rules up until 1985. Walton and Krabbe (1995) developed a formal model of commitment rules for dialogues. Commitment is one of the fundamental concepts which needs to be handled by a model of dialogue. The basic idea is that a party becomes committed to the premises and conclusions of any arguments he puts forward, as well as to any claims of the other party he concedes. In many formal models of dialogue, these commitments are managed in a so-called "commitment store". One issue to be addressed by the model is to what extent a party should become committed to logical consequences of his explicit commitments. One of the first computational models of argumentation dialogues was Gordon's Pleadings Game (Gordon 1995), which is an idealized model of the process of pleading in civil law cases in common law jurisdictions. The pleading phase is the first phase of a law suit, before trial. Essentially, the goal of pleading is to identify the legal and factual issues to be resolved by the court at trial. Other computational models of dialogue followed shortly thereafter, including Hage's procedural model about how to decide hard cases (Hage et al. 1994), Dialaw (Lodder 1999) which is based on Reason-Based Logic (Hage 1997; Verheij 1996), the Toulmin Dialogue Game (Bench-Capon 1998), which as its name suggests is based on Toulmin's argumentation scheme (Toulmin 1958), and Prakken's formal model of Dutch civil procedure (Prakken 2001), which focuses on modeling the allocation of burden of proof and the role of the judge. Also worth mentioning in this context is the Prakken and Gordon's computational model (1999) of Robert's Rules of Order (Robert 1915) for parliamentary assemblies.

Except for the topic of argument visualization, relatively little research has been done on computational models of the rhetorical layer, which is responsible for selecting arguments to put forward and other moves, and presenting arguments clearly and persuasively, taking into consideration the standpoints, values, commitments and beliefs of the intended audience. However, two chapters of the book "Argumentation Machines" (Reed and Norman 2003) address this issue. The first chapter, entitled "The Persuasion Machine" (Gilbert et al. 2003), presents a high-level description of an argumentation support system, based on insights from computational linguistics, which focuses on rhetorical tasks. Although informed by computational linguistics, this work is too abstract to be considered a computational model, and this presumably was not the authors' intention. Rather, it is a high-level sketch of various rhetorical tasks, i.e. a use-case analysis, together with some initial ideas about how to support these tasks using computer systems. The second chapter, entitled "Computational Models of Rhetorical Argument" (Crosswhite et al. 2003) sounds like it might present a survey of prior research on this subject, but actually presents a new computational model of the rhetorical level, based upon the philosophy of argument in "The New Rhetoric" (Perelman & Olbrechts-Tyteca 1969) and using McCarthy and Buvac's Context Logic (1998) to model audiences. This model is contrasted with prior work by Das et al. (1997), which uses rhetorical argument schemas to select arguments. Interestingly, both of these models are from the multi-agent systems community. Bench-Capon's Value-Based Argumentation Framework (2003) also needs to be mentioned here, since it uses a model of the value preferences of an audience in its evaluation of the acceptability of arguments.

One of the first argument visualization methods was developed by Wigmore, for visualizing the evidence in legal cases (Wigmore 1940). The diagramming method Toulmin used in his "Uses of Argument" (Toulmin 1958) has been very influential. But the argument diagramming method developed by Beardsley (1950) and refined by Freeman (1991) has become the de facto standard in the humanities. Conklin's gIBIS system (Conklin 1988), based on Rittel's idea of an issue-based information system (Rittel & Webber 1973), was perhaps the first computational model designed for visualizing arguments. More recently, a number of software applications for visualizing arguments have been developed, such as Araucaria (Rowe & Reed 2003), some of them as commercial products. Araucaria and other tools for visualizing arguments are covered in more depth in the next section of this report.

## 2.3 Argumentation Support Tools and Associated Research Groups

Douglas Engelbart, inventor in the 1960s of much of today's interactive personal computing tools, draws attention to the need for tools to tackle the "complex, urgent problems" facing society. Forty years on, he has concluded that central to meeting this challenge are argumentation support systems to help clarify the nature of the problems, and scaffold dialogical negotiation of ways forward (Engelbart, 2003). In this section we describe various examples of argumentation support tools.

Some have been developed as an educational resource, both as a means of delivering information but also as a means of teaching critical thinking skills. The legal domain requires its students to develop critical thinking skills and make effective use of argument, therefore it is not surprising that a large number of tools have their roots in this domain, being developed as 'argumentation assistants' for the legal profession. Others have grown within a commercial domain in response to the demands of arriving at, and presenting, strategic decisions within a large, dispersed business community. However, that is not to say that background determines suitability; for instance, 'Reason!Able' has been employed to resolve a dispute, but is also used for the instruction of critical thinking;

similarly, 'Compendium' has used in an informative role, even though its roots are firmly in commercial real-time problem solving.

Bex, Prakken, Reed and Walton (2003), although focusing on the legal domain, usefully consider two distinct types of argumentation support tools. That is those which contain knowledge about a problem domain and can perform reasoning to suggest solutions to the problem, for example dialogue and mediation tools, and those they term 'sense-making' systems (Kirschner et al, 2003) which do not support reasoning but rather structure the problem typically using visualization and may also support logical computation and communication between users of the system, i.e. argument mapping tools. Graphical visualization, through various forms of argument maps, has the potential to help people to create better arguments and analyses. The majority of the systems in this section can be considered to be of the sense-making type, however, some for the legal domain e.g. CATO and PLAID have associated case bases which can be interrogated.

We do not claim that this is an exhaustive survey but it does indicate the breadth of work being undertaken in the development of argumentation support tools. For each argumentation support tool we provide a general description of the system and, if available, the URL where either the tool can be downloaded from or where further information is available. We then briefly describe each tool by considering: the underlying argumentation model it uses, the type of user interface it presents, the domain it has been predominantly used in.

The systems are listed in alphabetical order.

### 2.3.1 Argue! and ArguMed

http://www.ai.rug.nl/~verheij/aaa/index.htm

Argue! was developed in 1998 by B. Verheij at the University of Groningen, The Netherlands. It can be considered as a sense-making tool. Verheij himself describes it as an 'argument-assistance system' which is meant to support, rather than replace, the reasoning process of the user (Verheij 1998). The system was further developed into ArguMed (Verheij 2000). It is an aid in the drafting and generation of arguments, performing such tasks as: administration of the argument process; tracking issues raised and assumptions made; tracking of reasons, conclusions and counterarguments; evaluating the extent to which statements are justified; and checking that users comply with the argument rules.

ArguMed provides graphical structuring for argumentation with a user interface supporting a click and drag metaphor to allow the user to decide whether their input is an assumption, issue, reason or attack. The system then decides whether an issue is justified, not justified or neither. Further details of this work can be found in Verheij (2005).

### 2.3.2 Araucaria

http://araucaria.computing.dundee.ac.uk/

This is an argument mapping tool developed by the University of Dundee, UK (Rowe et al 2003). It is used for analysing arguments where the user is supported in reconstructing and diagramming an argument. There is a simple point-and-click interface. The software supports several different diagramming methods, including Toulmin diagrams and the Beardsley/Freeman "standard" diagramming method.

It provides a user-customisable set of schemes with which to analyse arguments. The latest version of the tool supports Wigmore diagrams, a technique of presenting legal arguments in a diagrammatic form which was introduced into the legal academies in the 1930’s (Wigmore 1931). This argumentation scheme provides for ‘propositions’ and ‘assertions’, or relations, such as ‘supports’ and ‘challenges’.

A more recent publication discusses the use of the Araucaria to support the teaching of philosophy students (Rowe et al 2006).

### 2.3.3 Belvedere

http://lilt.ics.hawaii.edu/lilt/software/belvedere/index.html

This is an argument mapping system that has been designed to support problem-based collaborative learning scenarios, using evidence and concept maps, to teach middle and high school students critical enquiry skills. It was originally developed by Dan Suthers while at the University of Pittsburgh. He is now at the University of Hawaii at Manoa where the system has been further enhanced.

Belvedere is an issue-based argumentation system and supports multiple representational views (tables, hierarchies and graphs) on evidence models and provides support for concept maps and causal models. Users can construct ‘inquiry diagrams’ from a ‘palette’ of icons that represent different types of statements – such as ‘hypotheses’ and ‘data’ – and different types of linkages to indicate relationships between statements – such as ‘for’ and ‘against’. The linkages are colour-coded (green indicating a ‘for’ linkage, red indicating ‘against’) and their thickness can be altered to represent level of belief.

In the process of diagram construction, students working together develop social skills necessary for group problem-solving. They can also compare their maps to ‘model solutions’ provided by their teachers.

### 2.3.4 CATO and CATO-Dial

Both systems were developed by University of Pittsburgh to teach law students about the use of case-based legal argument. The first version of the system, CATO (Case Argument TutOrial), used didactic explanatory dialogue. Students have to develop a position choosing cases presented to them by the system. If they choose a poor example, the system alerts them to the fact and provides an explanation for why that particular case is of no value.

The newer version, CATO-Dial, takes a courtroom simulation approach using dialectic argument, the idea being that students would acquire skills more effectively when engaged in rôle-play, since the learning context would then be more appropriate to those hoping to practice law. The student acts as advocate in a case, selecting argument moves from a menu. The system acts as judge to mediate the proceedings, and as opposing counsel to expose any weaknesses in the student’s argument. It also provides an online help facility for the student to access when their argument goes badly.

CATO is perhaps one of the most popular systems to teach legal argument skills. It is based on previous work of Ashley, i.e. the HYPO model of legal argumentation, which provides an overall framework, argument forms: citation, response and rebuttal and a set of argument moves that can be made within the framework. It has been trailed extensively with law students (Aleven and Ashley, 1994).

### 2.3.5 Compendium

http://www.compendiuminstitute.org

Compendium is an argument mapping system that uses the issue-based information system for indexing and structuring discussions. It has been used for a number of years for commercial real-time problem-solving; originally, applications were concerned with business process re-design. The Compendium tool was designed to overcome some of the known limitations of the QuestMap tool (see below), though it has now grown substantially in scope to include integration with other tools, open source development and generally be more focussed towards use in research..

The system allows for considerable customization of the argument maps by the users and supports outputs in multiple document formats. Elements of a discussion are represented as ‘queries’ and ‘responses’, to which qualifying remarks can be attached indicating ‘support for’, or ‘criticism of’ that contention. Using hyperlinks, users can associate relevant documents with particular nodes to back-up any references. It is also possible to partition the discussion into a series of linked maps, which has the advantage of breaking-down large amounts of data into manageable portions.

Being based upon a MySql database, users can perform searches upon the information contained in the nodes, thereby facilitating the extraction of information contained in the maps. Describing the full functionality of Compendium is beyond the scope of this report, however the Compendium website provides extensive information (Selvin, 2003).

### 2.3.6 Dialaw

http://cli.vu/~lodder/dialaw/ and http://cedire.org/

This is a dialogue game for two players. It is an issue-based system developed by Arno Lodder who is currently at the Computer/Law Institute of the Vrije University Amsterdam (Lodder and Herczog 1995). The idea behind the game is to allow two people to state what they believe about a particular issue under discussion and then see where they agree and disagree. It helps users to understand how to construct logical arguments against opposing claims and also how to defend their own claims. The system allows for users to exchange statements and arguments and this dialogue is then stored and represented in a tree structure. The system supports a procedural model of legal justification.

Basically the game proceeds as follows. A player starts a dialogue (a game) with a claim and then the other player can challenge, make a new claim or concede. The game continues in this fashion with the opportunity for each player to also retract a claim. The system identifies when one player’s statements logically imply a claim of the other player, in which case the player then has to concede to the claim or retract one of his statements that led to the implied commitment. The game finishes when no disagreement remains.

Dialaw is discussed, along with Gordon’s Pleading Game and other dialogues and mediation systems, by Bench-Capon and Prakken (2006).

### 2.3.7 Hermes

http://www.mech.upatras.gr/~nikos/index.html

This argumentation support tool was developed under the European Commission ICTE-PAN project (Karacapilidis 2005). It is based on the Zeno system (Gordon and Karacapilidis, 1997). Hermes is aimed at supporting online group facilitation between government agencies. The developers argue that the majority of existing collaborative argumentation support systems have been designed to support face to face meetings with a human facilitator whereas what is needed for government to government collaboration is virtual support. Therefore the tool has an issue-based discussion forum with special support for argumentation.

The tool allows for the construction of a diagram of the discourse that is composed of the ideas so far expressed during the discussion. The basic elements are: 'issues' - corresponding to decisions to be made or targets to be met; 'alternatives' - corresponding to potential choices; 'positions' - these are assertions associated with an 'alternative', that provide grounds for following or avoiding that choice; and 'constraints' - these represent preference relations. Users can input their preferences to courses of action through a "position, relation, position" tuple, where an example of a relation is "less important than" or "more important then". Not only does Hermes record the users' arguments, but it also checks for inconsistencies among users' preferences, and automatically updates the discourse status according to the entire set of user input.

### 2.3.8 GEOMED

http://cordis.europa.eu/search/index.cfm?fuseaction=proj.simpledocumentlucene&HD_ID=1783331&CFID=169260&CFTOKEN=54309717

GeoMed (Geographical Mediation System, IE2037) was a 4th Framework Telematics European project to develop and validate an web-based groupware system to engage citizens in regional and urban planning. GeoMed integrated support for sharing documents, arguing planning issues and accessing geographical information (Schmidt-Belz et al., 1999).

GeoMed began in 1996 and was thus one of the first European eParticipation research projects, long before the term "eParticipation" had been coined. The project aimed to help citizens to participate in city planning by integrating an IBIS-based argumentation support system, Zeno, with a web-based geographical information system (Gordon 1995, Gordon 1996, Gordon 1997). A later version of Zeno served as the technical foundation of the eParticipation platform developed in another European project, DEMOS (Delphi Mediation Online System, IST-1999-20530), which ran from 2000-2004 and was successfully piloted in the cities of Hamburg and Bologna (Gordon 2002, Richter 2002). More information on Zeno can be found in Section 2.3.15 of this report.

### 2.3.9 Parmenides

http://cgi.csc.liv.ac.uk/~katie/Parmenides1.html

The 'Parmenides' system (Atkinson, Bench-Capon and McBurney 2004 and Atkinson 2006) supports consultation. It has a web-based interface to an argumentation support tool designed to facilitate dialogue between government and individuals. The system uses argumentation theory to support deliberation dialogues and helps users to apply the argumentation scheme for practical reasoning within a discussion.

Presently, it features a debate upon the invasion of Iraq as an illustration of its capabilities. The user is presented with a justification of the invasion in the form of a structured argument. They then have the opportunity either to accept the argument, in which case they are taken to a 'farewell' screen, or they are presented with a series of six possible attacks on the argument with which they can agree or disagree. The user is also able to enter a free text comment summarising their view of the debate.

By storing the users' comments on a database, it is possible to identify the strengths and weaknesses of the issue under scrutiny, thereby affording the policy makers an insight into where their views need bolstering, as well as where they can rely upon public support.

### 2.3.10 PLAID

This argumentation support tool, Proactive Legal Assistance, was developed at the University of Liverpool, UK to teach law students how to develop argument-based briefs as answers to policy questions. (Bench-Capon and Staniford, 1995 and Bench-Capon et al, 1998). PLAID is based on a modified form of the argument schema developed by Toulmin and uses a dialogue game structure. (For more information on dialogue games see Gordon, 1995).

The key features of PLAID are: access to legal sources of information which can be used with minimum adaptation for use by the system; multi-agent based architecture; and a knowledge base to support the development of an argument. The dialogue game is between the user (a law student) who asks for information as part of an argument graph and the computer which holds the entire graph. The system generates a 'brief' for the user, compiled from a number of sources comprising the system's knowledge base; these sources include statutes, leading cases, commentaries, and 'birth, marriage and death' records.

The system agents assist the user by finding information to fill the roles of required by the Toulmin schema – such as 'claim', 'data', 'backing', 'warrant' and 'rebuttal'. When the user is satisfied with their choice, a 'Rapporteur' agent generates a document in English from these arguments. This document can then be edited using a text editor, or by using Plaid's purpose-built editor that allows hypertext documents to be created and manipulated, thereby enabling the cooperative editing of texts.

### 2.3.11 QuestMap

http://www.cognexus.org/id17.htm

QuestMap was based on the gIBIS system (Conklin and Begemann, 1988) and (Conklin, Selvin, Buckingham Shum, and Sierhuis, 2003). Originally QuestMap was developed as an organizational memory and information management tool for collaborative working within a large utilities company in California. It was the company's idea to use it to support group facilitation/deliberation. Therefore, the system supported two different types of applications, supporting asynchronous collaborative information management and supporting group deliberation in face-to-face meetings.

It was based on the IBIS argument notation and provided hypertext and groupware functionality by allowing the user to create IBIS maps and lists. QuestMap used icons, or 'nodes', to represent the IBIS method elements of 'Issues', 'Positions' and 'Arguments' (supporting or contesting statements relative to a position). It was powered by a hypertext engine whose functions were accessed via an interface. The chief features were as follows: the creation of hyperlinks between maps through the copying of one node into another map; a list display of all maps or lists in which a particular node features – clicking on a list element takes the user to the particular instance of that node; additional information could be added to each node by placing text in a 'contents window' – including keyword search terms; and a search engine that could produce lists of nodes containing keywords, where those lists were themselves sets of hyperlinks. A case study on its use is provided by Conklin (2003).

This tool is no longer distributed or supported.

### 2.3.12 Reason!Able and Rational

http://www.goreason.com/ and http://www.philosophy.unimelb.edu.au/reason/

Both Reason!Able and Rationale are argument mapping tools from the University of Melbourne. Reason!Able supports the development of simple diagrams of complex reasoning, so that the evolving argument can be visualized. The tool was developed to support deliberation through the visualization of arguments. Initially the system was intended to help undergraduate students develop their critical thinking skills, then later progressed to support group deliberation in the workplace. The system itself does not support any analysis of the arguments but rather supports the construction and modification of argument visualizations. The Reason! project is developing a method for improving reasoning skills which is centred around Computer-Assisted Argument Mapping (CAAM) using the Reason! software learning environment.

Rationale uses colour and position to represent arguments; 'position' boxes, representing the conclusion, are white and placed at the top of the map; 'reason' and 'objection' boxes are green and red respectively, and are positioned beneath the position they support; 'rebuttal' boxes are orange, and represent an objection to an already existing objection. The argument is laid out on three levels; the top level provides the position being debated; the second level presents the reasons and objections that support or refute the position; the third level provides support to second level reasons and objections, thereby reinforcing them but not directly responding to the initial position. Users can judge the strength of an argument by evaluating its elements (whether they think the case is strong, weak or ambivalent, and whether they agree, disagree or are undecided about the position), and these judgements are represented on the map through the thickness of the lines connecting the various boxes.

For further information see van Gelder (2002 and 2003). An application using Reason!Able is presented in the following section.

### 2.3.13 Risk Agora

http://www.csc.liv.ac.uk/~peter/downloads/may01/rehgmcb.doc

This system supports deliberation about potential health and environmental risks of new chemicals among the scientific community (Rehg et al 2004). It is not intended as a real-time tool but rather to formally model and represent debates in the risk domain, as users posit, assert, contest, justify, qualify and retract claims. This activity is represented using Toulmin's model of argumentation within a dialectical framework.

Using a knowledge-base complied from scientific data, the Agora represents debates for the following purposes: to point-out the logical implications of current scientific belief relating to a particular issue, and the consequences of alternative options; to compare the arguments for and against a particular claim, according to their respective degrees of certainty and cogency; to combine arguments for and against a claim, thereby constructing a case for it; to provide an overview of the debate for the benefit of interested observers; to support group deliberation; and to support government agencies in risk assessment and regulatory determination.

Since regulatory decisions have to be taken regardless of the completeness of the scientific knowledge of a particular issue, it is desirable for regulatory agencies to have a snap-shot of the relevant debate at any time. To enable this, the Agora defines claims according to the arguments presented for and against them. Thus 'Probable' claims are those for which no arguments have been presented that rebut or undercut the claim. Clearly, at any particular time, it is these 'probable' claims that will be of most interest to the relevant agencies.

### 2.3.14 Room 5

http://www.cs.wustl.edu/~room5/

Room 5 is a sense-making tool developed as a game to support computer-mediated defeasible argumentation which is issued-based (Loui et al 1997). It was developed at Washington University in the US, as a testbed for a semi-formal legal argumentation system that could be used by members of the public. Room 5 separates each claim into three parts: the authority for the claim, such as a legal precedent; a paraphrasing of the claim; and a formal statement of the logic behind the claim.

It supports the graphical structuring of argumentation with a rather unusual user interface in that it uses colour codes to distinguish arguments and counter arguments. It also specifically does not use arrows to show linkages but rather uses horizontal and vertical text boxes to represent argument and counter-arguments horizontally and support for arguments vertically.

The tool is primarily aimed at supporting law students and its development is based on past Supreme Court cases. It includes a data mining component to provide access to online legal texts. The law student then has to decide whether the information from previous legal text supports, attacks or re-states evidence in the current case.

### 2.3.15 Zeno, Dito and Diaglo

http://zeno8.gmd.de/zeno/

Dito and its predecessor Zeno provide advanced support for collaborative decision-making using a moderated issue-based discussion forum with special support for argumentation. Diaglo provides a graphical user interface to the systems.

Zeno supports computer-mediated defeasible argumentation (Gordon, 1996). The argumentation model is based on IBIS and aspects of the overall tool are specific to the urban planning domain, e.g. it is integrated with a geographical information system. The basic elements are issues, positions, arguments and preferences. The elements of the IBIS model can be linked together to form various argumentation graphs, for example a dialectical graph. A novel feature of the system is its ability to support inference through 'semantic' labelling of the graphs. As Gordon and Karacapilidis (1997) state:

"It transforms IBIS from a lifeless method to organize and index information into a playing field for stimulating debate. The interested parties can see immediately whether their positions are currently "winning" or "losing", given the arguments which have been made so far, motivating them to marshall still better arguments in favor of their positions" (p17)

Zeno extended the idea of threaded discussions, in which messages are organized in an outline or tree, to the collaborative construction of more general labelled graphs. Both the nodes and links could be labelled, with labels configured by the moderator. And the graphs were not restricted to trees. Other extensions included the assignment of user defined properties to nodes and attachments, as in email attachments. Gordon and Richter (2002) describe the implemented system towards the end of the DEMOS project.

## 2.4 Summary

To conclude, in this section we have presented a number of argumentation support systems and associated tools. With regard to the tools, some of these focus on the visualization of arguments and here the graphical notation and user interface are important features. Others focus on providing analysis of the situation but typically with a more limited graphical user interface. A number of underlying argumentation models are used including those based on IBIS and Wigmore diagrams. In considering their relevance to eParticipation we need to consider the features needed to support informed debate to support evidence-based policy-making. The systems we have presented here allow the users to have access to various levels of information, to be able to focus on specific information and to have the ability to organize the gathered data to construct an effective argument – all of which are required for eParticipation.

In eParticipation there is a clear requirement to better understand how technology can support informed debate on issues but there are two main obstacles in achieving this. The first is that the deliberation is typically on complex issues and therefore there are typically a large number of arguments and counter arguments to consider which when presented in linear text can be confusing for the public at large. Secondly, it is not obvious that many people actually have the necessary critical thinking skills to deliberate on issues. In can be seen that the type of argumentation support systems and tools described in this section have the potential to add value to current eParticipation methods. This will be explored further in the section on eParticipation scenarios.

# 3 Applications of Argumentation Support Systems

This section provides examples of the practical use of argumentation support systems in five specific domains: business and commerce, education, law, urban planning, and conflict resolution. This is not an exhaustive description of the application of such tools but rather specific applications that have direct relevance to eParticipation.

The first two examples use the Reason!able tool. As discussed in the last section, this can be classed as a sense-making tool that supports the development of argument maps. The tool supports group deliberation in face-to-face meetings and also helps students develop their critical thinking skills.

## 3.1 Business and Commerce

This is concerned with the pressure to resolve a difference of policy within a company and involves the use of argument mapping system in a face-to-face setting. Van Gelder (2003, pp. 108-114) recounts a case in Australia involving a dispute at a factory. From pursuing a policy of 'one person, one job', the company decided to switch to training their workforce to be able to perform more than one role. This change in working practice divided the employees, and, although discussions were vigorously undertaken, no consensus upon the matter could be reached.

Representatives from throughout the factory participated in a facilitated argument mapping exercise using Reason!Able in an attempt to resolve the impasse. The group were conducted through a map developing exercise based upon the premise that they should continue with the current policy. Once all the representatives' views had been satisfactorily recorded, the resulting map was reviewed. The visualisation clearly presented the group with an irresistible reason for rejecting this premise; if one person is wholly and uniquely responsible for a task, then when that person is unavailable, the job cannot be performed. Consequently, work dependent upon the performance of that task will be unable to continue. Hence the case for multi-skilling appeared unanswerable.The significance of this result lies in the fact that although the group were aware of this reason all along, visualising the entire debate presented this reason within the context of all other considerations for the first time. With such a complex mesh of reasons for and against, it is too taxing to keep all the points in one's mind together, and too easy to avail oneself of a handy counter argument.

## 3.2 Education

Argument mapping tools have been used widely to support the teaching of philosophy students. They aim to improve the student's ability to follow arguments in a logical manner and to develop the student's critical thinking.

This example is based on research conducted at Monash University, Australia (Twardy, 2004). Twardy was concerned that the critical thinking skills of undergraduate students at Monash University were not as good as those of students at University of Melbourne. The Mebourne students were taught to map arguments using the Reason!able tool. Therefore

Twardy undertook a comparative study of the methods used and to ensure an even comparison of the methods, he himself did the teaching at the two universities. The results provided strong empirical data in favour of the argument mapping method. He concludes: “Despite my own training in analytical philosophy, I feel that mapping helps me with my own thinking”.

Similar research from the Department of Philosophy at Carnegie Mellon University investigated whether using argument mapping could enhance the critical thinking skills of 139 students in an introductory philosophy course. Her research concluded that learning how to construct argument maps significantly improved the student’s critical thinking skills (Harrell 2004).

Critical thinking involves understanding an argument and being able to analyse and evaluate it. With regard to eParticipation, if we wish to develop online deliberation tools to support citizen engagement then we must also ensure that the users have the necessary skills to deliberate, both individually and in groups. Therefore the type of tools that support critical thinking have the potential to also support eParticipation.

## 3.3 Law

The use of symbols to represent arguments has a long history in the legal domain, going back to Wigmore in 1931 (Wigmore 1931). One important motive for thus representing legal cases was not so much to show the reasoning that led to any particular verdict, but to highlight weaknesses in a chain of argument, thereby making the verdict more or less doubtful. This critical use of argument representation is shared by current researchers in the domain of legal argumentation (for example, Bex et al 2003).

However, the application of Argument Support Systems in law is hampered by the necessity to provide any system with sufficient amounts of information for the task it is expected to perform; in the legal domain, there is simply too much knowledge required for any but the simplest routine jobs. Whilst there appears to be little prospect of overcoming this problem in the near future, it is possible to make the size of the knowledge base more tractable, yet large enough to find useful application. Thus, the ability to present legal argument clearly will be of great benefit in preliminary fact investigation, case management and mediation. It also proves its worth in legal training, not only by helping students familiarise themselves with the structure of cases, but also in teaching them the discipline of legal reasoning. The application to be considered in this sub-section relates to a study of law students using such techniques.

A research project was conducted with the aim of discovering whether or not the use of argument visualisation techniques affects the quality and type of arguments produced by second year law students, both in their course work and a final exam (Carr 2003). In the study, a test group of thirty-three students were given training in and access to a Computer Supported Argument Visualization (CSAV) tool called QuestMap™ – which was a commercial version of Conklin's gIBIS system (Conklin et. al, 1988) – whilst a control group of forty students prepared their work using traditional methods. The students were set five problems to solve at intervals throughout the semester. A ‘model’ narrative answer was prepared, along with a mapped version that had been approved by the course professor. Using these documents, the students were able to compare their answers with the ‘model’ solution, with the control group using the narrative version and the test group using the map. At the end of the semester all students sat a practice examination.

Assessment of the impact of using QuestMap™ was based upon the measurement of three indicators: the number and types of argument structures present in the student's answers, as defined by Toulmin's Model of Argument; the professor's judgement of the students' performance in the final exam; the number of nodes created using QuestMap™ throughout the semester, indicating the extent to which skills improved with time (the higher the node count, the greater the depth of the arguments). Tests on the students indicated that there were no significant pre-existing differences between the control and test groups.

A summary of the results are as follows: the arguments of the test group did not get significantly more elaborate over time; the test group did not have a significantly different score from the control group in the practice exam. Two features of this case should be borne in mind; that the students in question were in their second year of Legal Studies and therefore had acquired experience in legal argumentation (which goes to explain why the test group quickly became proficient in using QuestMap™); that the practice exam was held a fortnight before their finals, suggesting that the students may not have spent much effort in preparation.

Whilst these results do not confirm the belief that using CSAV will improve a student's ability to analyse legal arguments, they do provide a number of valuable insights into the effect such tools can have when used in this context. It is felt that were the same test conducted with first year students, who would be expected to lack argumentation skills, then there would be a marked improvement in those using argument support compared with those who lacked it. As it is, there is arguably a benefit to experienced students, in the sense that the tool provides support to their work, allowing them to create answers more efficiently, as well as serving as a focus for discussion. This facility is not to be underestimated; as noted by Suthers whilst researching CSAV as an aid to teaching (Suthers 1999), visual representations can promote discussion amongst students by showing clearly where openings for counter arguments occur, where the fruits of these discussions appear on the map rather than the discussions themselves. Thus the map's function is less to record than to stimulate. As Carr says: "The software then becomes a support for the process of argumentation, rather than a representation of it." (Carr 2003, p.92)

## 3.4 Urban Planning

In 1997, there was a proposal to create a residential area and high-technology 'park' between the cities of Bonn and Sankt Augustin. The land to be developed was reserved for agricultural use, so a change in its status was required by law before any work could proceed. Such a change meant that the plans had to be made available to the general public, and their comments taken into consideration. This provided an opportunity to implement an internet-based support system, 'GeoMed' (Geographical Mediation) to facilitate the process (Schmidt-Belz, Gordon & Voss 1999).

Urban and Regional planning may be characterised as having the following features: effective communication and collaboration between interested parties is essential throughout the planning process; representations of space, such as maps, are a necessary feature of the process; negotiation and decision making are crucial phases of each project.

With this in mind, GeoMed was designed to make planning processes more transparent; to encourage and oversee public participation; to assist in the resolution or avoidance of conflict; and to support co-operation between planners, experts and communities. Success in these aims would have the beneficial result of making urban planning more efficient, less time-consuming and less-expensive.

To perform this function, GeoMed consists of six components:

- Shared workspaces to which owners define access rights, where members can view or upload documents to share with other members.
- A GIS viewer that allows users to pan, zoom and select layers, as well as add new layers, perform simple editing functions to graphics and annotations.
- A service whereby GIS data can be offered for sale, ordered and paid for.
- Software agents that perform notification services to users of shared workspaces and discussion fora.
- Knowledge-based system applications that allow the plans to be analysed with respect to any special regulations that apply.
- A discussion forum to provide a space for users to present their comments, queries and responses.

These components are represented as a single system that provides a number of integrated services. Thus, users of the discussion forum are able to link their comments to plans made available via the GIS viewer, thereby making any debate about features of the plans easier to comprehend.

The discussion forum employs the ‘Zeno Argumentation Framework’ (Gordon & Karacapilidis 1997) in order to be able to offer its users more than the simple functions of viewing and responding to messages. ‘Zeno’ is designed to show dependences between arguments as they emerge in the process of debate, to direct discussion onto solutions that appear to be the most promising. It also assists the moderators of the forum to monitor the propriety of comments by providing them with information on the rights and obligations of participants. Typically, an issue will be raised in the forum – for instance, the benefits of demolishing certain buildings to accommodate a new development – to which contributors will respond, providing comments for, or against. Using Zeno, it is possible to provide diagrammatic views of the issues and their associated positions, providing a clearer view of the relationship between comments. In addition, it shows preference rankings between positions thereby making it possible to judge the relative strengths of the contending solutions. This provides contributors with an immediate view on which positions are currently favoured, thereby motivating those whose opinions are ‘losing’ to strengthen their arguments, and guaranteeing a robust debate of the issue.

GeoMed was subjected to a two-day validation process, in which two groups were given a scenario, rôles to play and tasks to perform, followed by a discussion on the system’s prospective benefits and potentials. Some months later, the opportunity arose to offer it for public use in participating in a planning project. Over a two week period, a GeoMed workspace was made available to citizens containing plans and information; a discussion forum was created where citizens could input their comments.

Some cautious conclusions can be drawn from this experience, always bearing in mind that when this project was conducted in 1997, citizens’ familiarity with ICTs was a pale shadow of what it is to-day. An encouraging number of people used the workspace for information, comparing favourably with the numbers attending a public meeting. However, no-one left a comment on the discussion forum or provided feedback on the system to the project team. It is possible that the lack of communication was due to difficulties with the user interface, coupled with the novelty of the system.

Yet, this project provided many valuable points for future work, not only in the field of planning, but any system supporting group co-operation, internet mapping or public participation. These include: introducing complex systems like GeoMed into organisations will be difficult since they will not only have to accommodate novel processes, but do so within the constraints posed by the legal regulations to which regional and urban planning is subject; planning issues involve people performing different rôles, with distinct interests to promote, but if the discussion is to be of any value it not only has to be available to all but also contributions have to be made from all parties in order to provide a balanced and informed debate; systems like GeoMed run in conjunction with traditional methods, which leads to the administrative problems associated with using paper documents and electronic data.

From the experience of the GeoMed project it is evident that the relationships between and within organisations are highly complex. In this respect it is worth briefly describing a recent project, using argumentation systems to facilitate G2G collaboration for public-policy and decision-making (Karacapilidis, Loukis & Dimopoulos 2005). There are four basic discourse elements in the system; issues, alternatives, positions and preferences. ‘Issues’ correspond to the problems, decisions and goals. Users propose ‘alternatives’ for each ‘issue’, which represent potential choices. These choices are supported or contested by ‘positions’, which may also refer to other ‘positions’ raised in the debate. ‘Preferences’ provide a qualitative means by which users can assess the relative strengths of particular courses of action, and consist of the tuple (position, relation, position), where ‘relation’ can be either ‘more important than’, ‘of equal importance to’, or ‘less important than’. Any expression of a preference may also be subject to support or criticism. User input is used by the system to construct an illustrative discourse-based knowledge graph, representing the user’s arguments and any documents they wish to include supporting their opinion.

The project centred upon the question of whether or not to allow non-state universities in Greece, the outcome of which has significant implications for such institutions as government, education, the municipalities, Chambers of Industry and Commerce, the Church, the private sector, not to mention all the potential students. Four groups were chosen from this collection of interested parties, representing the Ministry of National Education, the University Professors, the Chambers of Industry and Commerce, and owners of existing private educational institutions. All received training in the system and were familiar with using the internet, including the use of electronic fora. The argumentation session was conducted synchronously amongst the fourteen participants.

Evaluation of the session showed that the participants felt that the system was useful and that its basic functions were easy to master. This positive attitude extends to their being prepared to use it again in similar discussions and contexts, with the reservation that some difficulties arose from their being unfamiliar with conducting arguments over the internet rather than face to face. However, whilst online argumentation requires greater effort in the construction and comprehension of arguments when compared with traditional discussion sessions, it is felt that this will ease significantly as users become familiar with the system; allowing longer training sessions and conducting the debate asynchronously would also alleviate the burden. Overall, this project provides evidence that argumentation systems can support the collaborative understanding of social problems and the development of potential solutions, both within and between organisations, as well as with the citizens who will be affected by policy decisions. The authors conclude that: "…it can contribute to the transparency and openness of the whole public policy making and implementation process, by making the relevant information accessible at a very low cost." (Karacapilidis et al 2005, p. 620)

## 3.5 Conflict Resolution

The use of argumentation tools to resolve conflicts cuts across three of the applications named above, and warrants a brief entry in this section. The case to be described (Papadopoulos 2004) conjoins the features of dispute, legal intricacy and planning issues.

The Californian community of Graton faced the problem posed by Mexican day labourers using outlying areas of the town as a base for finding employment. Their presence caused concerns about such matters as traffic, litter, personal safety and contamination of the local creek. The Mexicans too were frustrated at there being few dependable job opportunities and a lack of affordable housing. The construction of a day centre for the labourers was proposed as a solution to this problem, but community opinion was divided upon the merits of this scheme.

In November 2002, the North Bay Consensus Council (NBCC), using a technique called 'Conflict Cartography', was hired to bring the various stakeholder groups together in an effort to build consensus and resolve this issue. Argument visualisation tools played two important roles in this task. Maps representing the state of the project were produced at each significant step in the conflict resolution process. These provided an accessible medium for interested parties, by organising the immense amount of data supplied by the stakeholders, which included such diverse material as aerial photographs of suitable sites, planning regulations, expert analysis, and suggestions concerning the layout of the day centre. Maps were also used to stimulate and represent in real time the community feedback upon the progress of the project and its findings. Through a process of negotiation, involving argumentation support systems, the NBCC managed to bring representatives for the stakeholders from an impasse, to the creation of a set of draft agreements and recommendations.

# 4 eParticipation Application Scenarios of Argumentation Support Systems

This section focuses on eParticipation application scenarios for argumentation support systems and investigates to what extent such systems can be designed to encourage debate and deliberation by citizens on public issues.

Policy making is an iterative process where options to follow have to be discovered over time. The policy solution is such that there is no clear cut right or wrong approach, but instead there are better or worse solutions that need to be debated and where stakeholders hold conflicting views to such an extent that some do not even agree that there is a problem to be solved. The domain involves a large amount of knowledge that must be made explicit in different formats at each stage of the policy-making life cycle. This includes knowledge from many different sources and channels. Where the government has an interest in seeking the public's views on policy, there is an obvious need to supply suitable information upon which particular opinions can be based. As most of this will need to be extracted from this large amount of information, the public are faced with a time-consuming, and thereby off-putting, activity in order to prepare themselves for an informed debate. Additionally, problems of political policy are highly complex, admitting many opportunities for confusion and frustration. Taking all these facts together, they fall within that class of problems classified as 'wicked' (Rittel 1973). As discussed in earlier sections, 'wicked' problems have a number of characteristics that make them both difficult to analyse and resolve.

As we have shown, argumentation support systems are useful both for guiding the reconstruction of arguments put forward by other parties, so as to open them up to critical analysis and evaluation, as well supporting the construction ("invention") of new arguments to put forward in support or one's own claims or to counter the arguments of others. Given that argument maps use icons and arrows to represent the structure of a series of related viewpoints, thereby clarifying the issue under consideration, they have the potential to provide a readily accessible medium by which citizens can follow and join in public debates on policy issues.

In this section we consider the use of argumentation support systems to support the provision of information, consultation and deliberation,- three eParticipation activities identified in D4.1 and D5.1. We particularly focus on the sense-making systems.

## 4.1 Sense-Making Application Scenarios for eParticipation

Napier University has been using the Compendium tool, as an example of a sense-making tool, to investigate how they can be used within a political context to support eParticipation (Renton and Macintosh 2005 and 2006).

### 4.1.1 Supporting provision of information - Representing political debates.

Figure 3 demonstrates how argumentation support tools have the potential to support the provision of information. The map represents a fragment of a parliamentary debate on radio masts from the Scottish Parliament. It concerns the issue of safe usage guidelines. Typically such debates are recorded and where made available electronically on websites are reproduced verbatim without any analysis or structuring, making it very difficult for users to read and comprehend the issues. In the case of this map, the two bulb icons represent the substance of an issue, at a general level (top centre) and as a specific statement (beneath, enclosed in a light blue box). Contributions supporting and contesting this statement are indicated by green ‘pluses’ and red ‘minuses’ respectively, attached by arrows. The text beneath these icons is a summary of the comment taken from the verbatim report of the debate. The blue asterisk to the right of the icon indicates the presence of information that can be revealed when the cursor is rolled-over. In the above instance, the text in the mauve box contains the name, political party and constituency of the individuals making the comment.

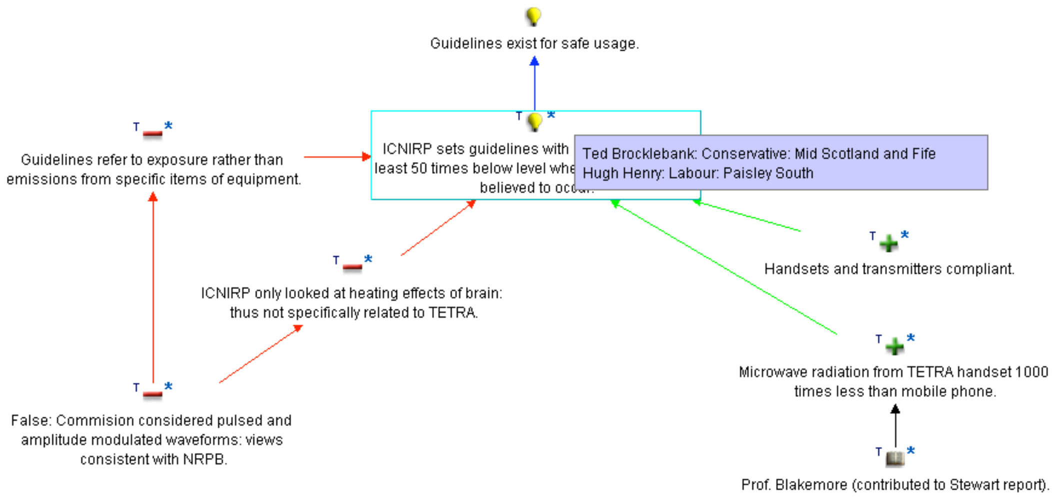



**Figure 3: Arrangement of icons representing part of a political debate**

### 4.1.2 Supporting Consultations

Figure 4 shows an alternative way of setting out the responses to an online consultation on a published draft policy document. The globe icon on the left indicates a hyperlink; in this instance, linked to the site containing the consultation paper. The intention is to improve clarity by making all the section topics visible at once. The blue icons on the far right provide links to further visualisations that provide the user with greater detail. Embedding maps permits information to be organised clearly and efficiently over a number of connected pages, rather than attempting to place all the data on one page. The deepest map contains a representation of the replies made to a particular question.

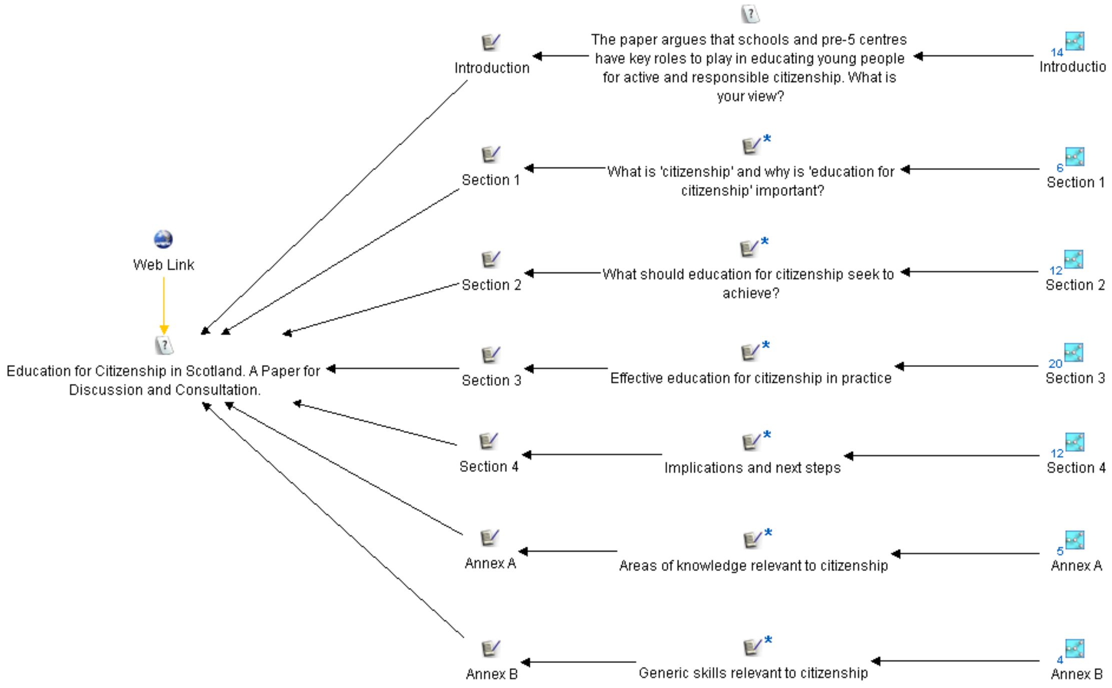



**Figure 4: Visualisation of consultation data and responses**

### 4.1.3 Supporting deliberation

Figure 5 represents a simplified version of consultation responses in the form of an inverted tree. It is designed to allow users to deliberate before making their own conclusions. This process should assist users to see how their convictions on one issue may conflict with other beliefs; thus, one might realise that the principle of ‘freedom of choice’ clashes with a belief in the duty of employers to protect their employees from harm.

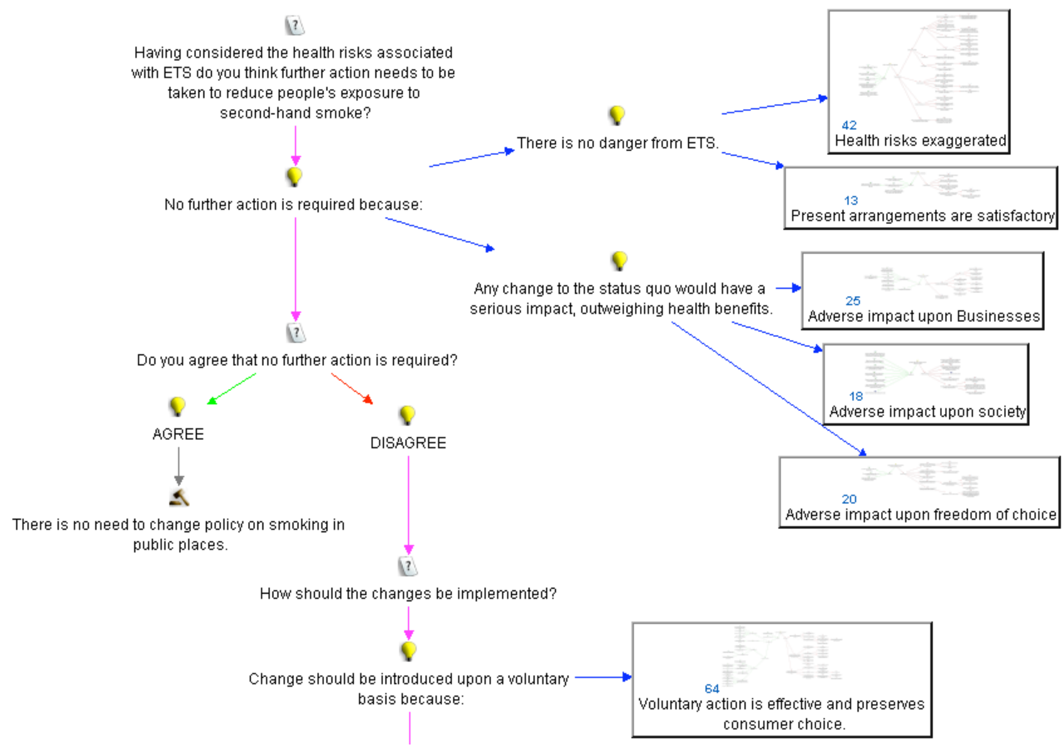



**Figure 5: Supporting deliberation during a consultation**

## 4.1.4 Supporting Analysis of a discussion forum

Figure 6 shows an alternative way of displaying the responses to an online consultation. The contributions to an online discussion forum were taken and reproduced verbatim. Although the map provides information in the same way as those above, it was also designed to establish whether or not individual contributors had remained consistent throughout the debate, and therefore could be used to support the analysis and evaluation of the consultation process.

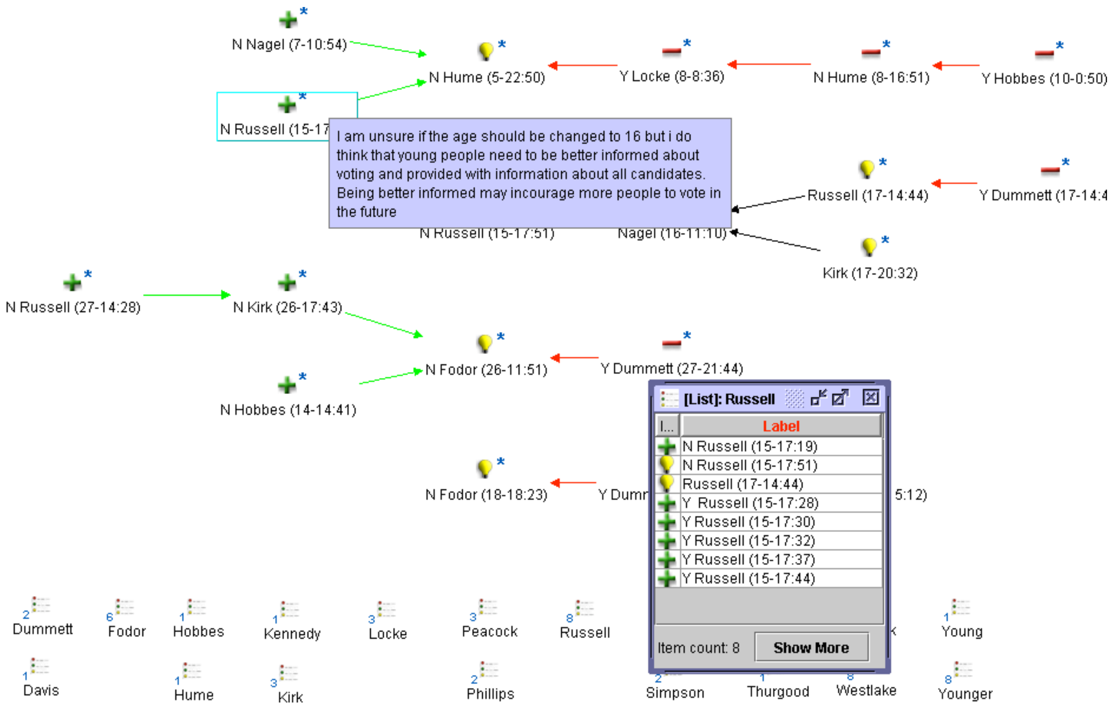



**Figure 6: Supporting analysis of a discussion forum and example of use of the 'list' function to check on consistency.** [2]

## 4.2 Summary

Currently the creation of maps is largely done manually and thereby is quite time consuming. The corollary of this is that the maps will be expensive to produce as well as there being a lag period between the end of the debate and the appearance of the map. Whilst experiments are underway in using semantic searching to extract text for the maps, the results are unclear and further research involving semantic search and text mining is required.

As governments seek to consult their citizens over matters of policy, it becomes increasingly important that citizens receive the relevant information in a medium that they can, and will, want to use in forming their opinion upon consultative issues. This section has presented sample scenarios in order to assess the potential contribution argument support systems can make to the consultation process. They cover techniques for the presentation of complex information in a thematically arranged format, for identifying those issues that generate a significant response, for collating consultation responses and representing them within an argument structure, and for checking upon the consistency of contributions to a debate. As such, they have something valuable to offer both government and civil society.

---

[2] Map created using responses to: '*Should the voting age for the Scottish Parliament be lowered to 16?*' Highland Youth Voice: 28.05.04 – 19.07.04.

# 5 Conclusions

Argumentation Support Systems are computer software for helping people to participate in various kinds of goal-directed dialogues in which arguments are exchanged. Their potential relevance for eParticipation should be readily apparent, since the goal of eParticipation is to engage citizens in dialogues with government about such matters as public policy, plans, or legislation. Surely argumentation plays a central role in this process. In a public consultation, for example, citizens are given an opportunity to not only make suggestions, but also support these suggestions with arguments.

Typically eParticipation projects make use of generic groupware systems, such as discussion forums and online surveys. These generic groupware systems, however, do not provide specific technical support for argumentation. For example, they provide no way for a citizen to obtain a quick overview of the issues which have been raised, to list ideas which may have been proposed for resolving such issues, to see in one place the arguments pro and con these proposals, or to get an idea about which positions currently have the best support given the arguments put forward thus far in the dialogue. These are just a few of the kinds of services offered by argumentation support systems.

This report introduced the theory of argumentation; summarized prior work of the leading research groups on modelling argumentation and supporting argumentation with software tools; described various prior applications of argument support systems, mostly in research pilot projects; and presented a number of eParticipation application scenarios for argumentation support systems, as a source of ideas for future pilot projects.

A number of argumentation support systems and associated tools were presented. Some of these focus on the visualization of arguments and here the graphical notation and user interface are important features. Others focus on providing analysis of the situation but typically with a more limited graphical user interface. A number of underlying argumentation models are used by these systems, including Issue-Based Information Systems (IBIS) and the method developed by Wigmore for mapping evidence in legal cases. In considering their relevance to eParticipation, we need to consider the features needed to support informed debate to support evidence-based policy-making. The systems presented allow users to access various levels of information, to be able to focus on specific information and to have the ability to organize the gathered data to construct an effective argument – all of which are required for eParticipation.

In eParticipation, there is a clear requirement to better understand how technology can support informed debate on issues but there are two main obstacles in achieving this. The first is that the deliberation is typically on complex issues and therefore there are typically a large number of arguments and counter arguments to consider which when presented in linear text can be confusing for the public at large. Secondly, it is not obvious that many people actually have the necessary critical thinking skills to deliberate on issues. In can be seen that the type of argumentation support systems and tools described in this report have the potential to add value to current eParticipation methods. This was explored further in the section on eParticipation scenarios.

As governments seek to consult their citizens over matters of policy, it becomes increasingly important that citizens receive the relevant information in a medium that they can, and will, want to use in forming their opinion upon consultative issues. This report presented sample eParticipation application scenarios of argumentation support systems in order to assess the potential contribution these systems can make to the consultation process. They cover techniques for the presentation of complex information in a thematically arranged format, for identifying those issues that generate a significant response, for collating consultation responses and representing them within an argument structure, and for checking upon the consistency of contributions to a debate. As such, they have something valuable to offer both government and civil society.

# References


[1] Aleven, V. Teaching Case-Based Argumentation Through a Model and Examples. Ph.d., University of Pittsburgh, 1997.

[2] Aleven, V. and Ashley, K. D. (1994) An Instructural Environment for practicing argumentation skills. *Proceedings of AAI 1994*. Cambridge, Mass: MIT press. pp485-492.

[3] Ashley, K. D. Modeling Legal Argument: Reasoning with Cases and Hypotheticals. Artificial Intelligence and Legal Reasoning Series. MIT Press, Bradford Books, 1990.

[4] Atkinson, K. (2006): Value-based argumentation for democratic decision support. In P. E. Dunne and T. J. M. Bench-Capon (editors): *Computational Models of Natural Argument, Proceedings of the First International Conference on Computational Models of Natural Argument* (COMMA 2006), pp. 47-58. Liverpool, UK. Frontiers in Artificial Intelligence and Applications, Vol. 144. IOS Press, Amsterdam, The Netherlands

[5] Atkinson, K., Bench-Capon, T., and McBurney, P. (2005). Persuasive political argument. In F. Grasso, C. Reed and R. Kibble (editors): Proceedings of the Fifth International Workshop on Computational Models of Natural Argument (CMNA 2005), pp. 44-51. IJCAI 2005, Edinburgh, Scotland.

[6] Atkinson, K., Bench-Capon, T., and McBurney, P. (2004). PARMENIDES: Facilitating democratic debate. In R. Traunmüller (editor): Electronic Government, pp. 313-316. Lecture Notes in Computer Science (LNCS) 3183, Springer, Berlin. Third International Conference on eGovernment (EGOV 2004), DEXA 2004, Zaragoza, Spain

[7] Baader, F., Calvanese, D., McGuinness, D., Nardi, D., and Patel- Schneider, P., Eds. The Description Logic Handbook – Theory, Implementation and Applications. Cambridge University Press, 2003.

[8] Beardsley, M. C. Practical Logic. Prentice Hall, New York, 1950.

[9] Bench-Capon, T. Specification and implementation of Toulmin dialogue games. In Proceedings of JURIX 98 (1998), pp. 5–20.

[10] Bench-Capon, T., and Sartor, G. A quantitative approach to theory coherence. In Legal Knowledge and Information Systems (Jurix 2001) (Amsterdam, 2001), B. Verheij, A. R. Lodder, R. P. Loui, and A. J. Muntjewerff, Eds., IOS Press, pp. 53–62.

[11] Bench-Capon, T. The missing link revisited: The role of teleology in representing legal argument. Artificial Intelligence and Law 10, 1-3 (September 2002), 79–94.

[12] Bench-Capon, T. Persuasion in practical argument using value-based argumentation frameworks. Journal of Logic and Computation 13, 3 (2003), 429–448.

[13] Bench-Capon, T., Freeman, J., Hohmann, H., and Prakken, H. Computational models, argumentation theories and legal practice. In Argumentation Machines; New Frontiers in Argument and Computation, C. Reed and T. J. Norman, Eds. Kluwer Academic Publishers, 2003, pp. pp. 85–120.

[14] Bench-Capon, T., and Sartor, G. A model of legal reasoning with cases incorporating theories and values. Artificial Intelligence 150, 1–2 (November 2003), 97–143.

[15] Bench-Capon, T., and Prakken, H. (2006) Argumentation. In Information Technology and Lawyers: Advanced Technology in the Legal Domain, from Challenges to Daily Routine, A. R. Lodder and A. Oskamp, Eds. Springer Verlag, 2006, pp. 61–80.

[16] Bench-Capon, T and G. Staniford. (1995), *PLAID - Proactive Legal Assistance* in Proceedings of the Fifth International Conference on AI and Law, University of Maryland,

ACM Press New York, 1995, pp 81-88.

[17] Bench-Capon, T., Leng, P. H. and Staniford, G. (1998) A computer –supported environment for the teaching of legal argument. *The Journal of Information, Law and Technology* (JLIT) 1998(3).

[18] Berman, D. H., and Hafner, C. D. Representing teleological structure in case-based legal reasoning: The missing link. In Proceedings of the 4th International Conference on Artificial Intelligence and Law (New York, 1993), ACM Press, pp. 50–60.

[19] Bex, F., Prakken, H., Reed, C., and Walton, D. Towards a formal account of reasoning with evidence: Argumentation schemes and generalizations. Artificial Intelligence and Law 11, 2-3 (2003), 125–165.

[20] Bondarenko, A., Dung, P. M., Kowalski, R. A., and Toni, F. An abstract, argumentation-theoretic approach to default reasoning. Artificial Intelligence 93, 1-2 (1997), 63–101.

[21] Branting, L. K. Reasoning with Rules and Precedents: A Computational Model of Legal Analysis. Kluwer Academic Publishers, Dordrecht, 2000. Book version of 1991 PhD Thesis.

[22] Brewka, G., and Gordon, T. F. How to buy a Porsche: An approach to defeasible decision making. In Working Notes of the AAAI-94 Workshop on Computational Dialectics (Seattle, Washington, 1994), pp. 28–38.

[23] Carr, C.S., (2003). Using Computer Supported Argument Visualization to Teach Legal Argumentation. In Kirshchner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.), *Visualizing Argumentation*, pp. 75-96. London: Springer-Verlag.

[24] Conklin, J., (2003). Dialog Mapping: Reflections on an industrial strength case study. In Kirshchner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.), *Visualizing Argumentation*, pp. 117-136. London: Springer-Verlag.

[25] Conklin, Jeff and Michael L. Begemann (1988). gIBIS: A Hypertext Tool for Exploratory Policy Discussion. Conference on Computer-Supported Cooperative Work CSCW, Portland (Oregon)

[26] Conklin, J., Selvin, A., Buckingham Shum, S. and Sierhuis, M. (2003). Facilitated Hypertext for Collective Sensemaking: 15 Years on from gIBIS. Keynote Address, *Proceedings LAP'03: 8th International Working Conference on the Language-Action Perspective on Communication Modelling*, (Eds.) H. Weigand, G. Goldkuhl and A. de Moor. July 1-2, 2003, Tilburg, The Netherlands http://eprints.aktors.org/262

[27] Crosswhite, J., Fox, J., Reed, C., Scaltsas, T., and Stumpf, S. Computational models of rhetorical argument. In Argumentation Machines — New Frontiers in Argument and Computation, C. Reed and T. J. Norman, Eds. Kluwer Academic Publishers, 2003, ch. 6, pp. 175–209.

[28] Das, S. K., Fox, J., Eldson, D., and Hammond, P. A flexible architecture for autonomous agents. Journal of Experimental and Theoretical Artificial Intelligence 9 (1997), 407–440. de Kleer, J. An assumption-based TMS. Artificial Intelligence 28 (1986).

[29] Doyle, J. A truth maintenance system. Artificial Intelligence 12 (1979), 231–272.

[30] Dung, P. M. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. Artificial Intelligence 77, 2 (1995), 321–357.

[31] Engelbart, D. C., (2003). Afterword. In Kirshchner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.), Visualizing Argumentation, pp. 205-208. London: Springer-Verlag.

[32] Felscher, W. Dialogues as a foundation for intuitionistic logic. In Handbook of Philosophical Logic; Vol. III: Alternatives in Classical Logic, D. G. Günthner and F., Eds.

1986, pp. 341–372.

[33] Freeman, J. B. Dialectics and the Macrostructure of Arguments: A Theory of Argument Structure. Foris Publications, 1991.

[34] Freeman, K., and Farley, A. M. A model of argumentation and its application to legal reasoning. Artificial Intelligence and Law 4, 3-4 (1996), 163–197.

[35] Gardner, A. An Artificial Intelligence Approach to Legal Reasoning. MIT Press, 1987.

[36] Geffner, H., and Pearl, J. Conditional entailment: Bridging two approaches to default reasoning. Artificial Intelligence 53, 2-3 (1992), 209–244.

[37] Gilbert, M. A., Grasso, F., Groarke, L., Gurr, C., and Gerlofs, J. M. The persuasion machine — argumentation and computational linguistics. In Argumentation Machines — New Frontiers in Argument and Computation, C. Reed and T. J. Norman, Eds. Kluwer Academic Publishers, 2003, ch. 5, pp. 121–174.

[38] Gordon, T. F. The Pleadings Game; An Artificial Intelligence Model of Procedural Justice. Kluwer, Dordrecht, 1995. Book version of 1993 Ph.D. Thesis; University of Darmstadt.

[39] Gordon, T. F. Zeno: A WWW system for geographical mediation. In Collaborative Spatial Decision-Making, Scientific Report of the Initiative 17 Specialist Meeting, P. J. Densham, M. P. Armstrong, and K. K. Kemp, Eds., Technical Report. Santa Barbara, California, 1995, pp. 77–89.

[40] Gordon, T. F., Karacapilidis, N., and Voss, H. Zeno — a mediation system for spatial planning. In CSCW and the Web - Proceedings of the 5th ERCIM/W4G Workshop, U. Busbach, D. Kerr, and K. Sikkel, Eds., GMD Technical Reports. Sankt Augustin, Germany, 1996, pp. 55–61.

[41] Gordon, T. F., and Karacapilidis, N. The Zeno argumentation framework. In Proceedings of the Sixth International Conference on Artificial Intelligence and Law (Melbourne, Australia, 1997), pp. 10–18.

[42] Gordon, T. F., and Richter, G. (2002) Discourse support systems for deliberative democracy. In eGovernment: State of the Art and Perspectives (EGOV02) (Aix-en-Provence, 2002), R. Traunmüller and K. Lenk, Eds., Springer Verlag, pp. 248–255.

[43] Gordon, T. F., and Walton, D. The Carneades argumentation framework — using presumptions and exceptions to model critical questions. In Proceedings of the First International Conference on Computational Models of Argument (COMMA 06) (Liverpool, September 2006), P. E. Dunne, Ed., pp. 195–207.

[44] Governatori, G., Maher, M., Antoniou, G., and Billington, D. Argumentation semantics for defeasible logic. Journal of Logic and Computation 14 (2004), 675–702.
Hage, J. A theory of legal reasoning and a logic to match. Artificial Intelligence and Law 4, 3-4 (1996), 199–273.

[45] Hamblin, C. Fallacies. Methuen, London, 1970.

[46] Hage, J., Leenes, R., and Lodder, A. Hard cases: A procedural approach. Artificial Intelligence and Law 2 (1994), 113–167.

[47] Hage, J. C. Reasoning with Rules – An Essay on Legal Reasoning and its Underlying Logic. Kluwer Academic Publishers, Dordrecht, 1997.

[48] Hage, J. C. Formalising legal coherence. In Proceedings of the Eighth International Conference on Artificial Intelligence and Law (New York, 2001), ACM Press, pp. 22–31.

[49] Harrell, M. (2004). Using Argument Diagrams to Teach Critical Thinking Skills. http://www.hss.cmu.edu/philosophy/harrell/Using_Argument_Diagrams.pdf

[50] Horrocks, I., Patel-Schneider, P., Boley, H., Tabet, S., Grosof, B. N., and Dean, M. SWRL:

A semantic web rule language combining OWL and RuleML.

[51] Karacapilidis, N., E. Loukis, and S. Dimopoulos. 2005. Computer-supported G2G collaboration for public policy and decision making. *Journal of Enterprise Information Management* 18(5):602-624.

[52] Kirschner, Paul A., Simon J. Buckingham Shum and Chad. S. Carr (2003). Visualizing Argumentation. Software Tools for Collaborative and Educational Sense-Making. Computer Supported Cooperative Work. D. Diaper and S. Colston. London, Berlin, Heidelberg, Springer.

[53] Kowalski, R. A., and Toni, F. Abstract argumentation. Artificial Intelligence and Law 4, 3-4 (1996), 275–296.

[54] Krabbe, E. C. W. Formal systems of dialogue rules. Synthese 63, 3 (1985), 295–328.

[55] Kunz, Werner and Horst W.J. Rittel (1970). Issues as elements of information systems. Berkeley, University of Berkeley.

[56] Lodder, A. R. and Herczog, A. (1995), Dialaw – *A Dialogical Framework for Modelling Legal Reasoning*, Proceedings of the 5th International Conference on Artificial Intelligence and Law, College Park Maryland USA, May 21-24 ,1995. pp146-55.

[57] Lorenzen, P., and Lorenz, K. Dialogische Logik. Wissenschaftliche Buchgesellschaft,, Darmstadt, 1978.

[58] Loui, R. Defeat among arguments. Computational Intelligence 3 (1987), 100–106.

[59] Loui, R., and Norman, J. Rationales and argument moves. Artificial Intelligence and Law 3, 3 (1995), 159–190.

[60] Loui, R., P. Norman, J. Altepeter, D. Pinkard, D. Craven, J. Lindsay, and M. Foltz. (1997) *Progress on Room 5: A testbed for public interactive semi-formal legal argumentation*. In Proc. of the Int. Conf. on Artificial Intelligence and Law (ICAIL-97), Melbourne, Australia, pages 207--214.

[61] Lutomski, L. S. The design of an attorney's statistical consultant. In Proceedings of the Second International Conference on Artificial Intelligence and Law (ICAIL-89) (1989), pp. 224–233.

[62] Malkia, M., Anttiroiko, A.-V., and Savolainen, R., Eds. eTransformation in Governance – New Directions in Government and Politics. Idea Group, London, 2004.

[63] Mackenzie, J. D. Question-begging in non-cumulative systems. Journal of Philosophical Logic 8 (1979), 117–133.

[64] McCarthy, J., and Buvac, S. Formalizing context (expanded notes). In Computing Natural Language, A. Aliseda, R. Glabbeek, and D. van Westerstahl, Eds. CSLI Publications, Stanford, Calif., 1998, pp. 13–50.

[65] McGuinness, D. L., and van Harmelen, F. OWL Web Ontology Language overview, 2004.

[66] McLaren, B. Assessing the Relevance of Cases and Principles Using Operationalization Techniques. Ph.d., University of Pittsburgh, 1999.

[67] Nute, D. Defeasible logic. In Handbook of Logic in Artificial Intelligence and Logic Programming, D. Gabbay, C. Hogger, and J. Robinson, Eds. Clarendon Press, Oxford, 1994, pp. 253–395.

[68] Papadopoulos, N., (2004). Conflict Cartography: A Methodology Designed to Support the Efficient and Effective Resolution of Complex, Multi-Stakeholder Conflicts. http://www.viewcraft.com/pdfs/ViewCraft_ConflictCartographyMarch04.pdf Accessed 12.12.06.

[69] Perelman, C., and Olbrechts-Tyteca, L. The New Rhetoric. University of Notre Dame Press, Notre Dame, 1969.

[70] Petri, C. A. Kommunikation mit Automaten. Technical report, GMD, Sankt Augustin, Germany, 1962.

[71] Pollock, J. Defeasible reasoning. Cognitive Science 11, 4 (1987), 481–518.

[72] Prakken, H. From logic to dialectic in legal argument. In Proceedings of the Fifth International Conference on Artificial Intelligence and Law (Maryland, 1995), pp. 165–174.

[73] Prakken, H., and Sartor, G. A dialectical model of assessing conflicting argument in legal reasoning. Artificial Intelligence and Law 4, 3-4 (1996), 331–368.

[74] Prakken, H., and Sartor, G. Argument-based extended logic programming with defeasible priorities. Journal of Applied Non-classical Logics 7 (1997), 25–75.

[75] Prakken, H., and Sartor, G. Modelling reasoning with precedents in a formal dialogue game. Artificial Intelligence and Law 6, 2-4 (1998), 231–287.

[76] Prakken, H., and Gordon, T. F. Rules of order for electronic group decision making — a formal methodology. In Collaboration between Human and Artificial Societies, J. A. Padget, Ed., Springer Lecture Notes in AI. Berlin, 1999, pp. 246–263.

[77] Prakken, H. Modeling defeasibility in law: Logic or procedure? Fundamenta Informaticae 48 (2001), 253–271.

[78] Prakken, H. Modelling reasoning about evidence in legal procedure. In Proceedings of the Eigth International Conference on Artificial Intelligence and Law (New York, 2001), ACM Press, pp. 119–128.

[79] Prakken, H., and Vreeswijk, G. Logics for defeasible argumentation. In Handbook of Philosophical Logic. Kluwer Academic Publishers, 2002, pp. 218–239.

[80] Prakken, H., Reed, C., and Walton, D. Argumentation schemes and generalisations about evidence. In International Conference on Artificial Intelligence and Law (Edinburgh, 2003), pp. 32–41.

[81] Prakken, H. Analysing reasoning about evidence with formal models of argumentation. Law, Probability and Risk 3, 1 (2004), 33–50.

[82] Prakken, H. AI & Law, logic and argument schemes. Argumentation 19, Special Issue on the Toulmin Model Today (2005), 303–320.

[83] Prakken, H., and Sartor, G. Presumptions and burden of proof. In Legal Knowledge and Information Systems (Jurix 2006) (2006). Accepted for publication.

[84] Reed, C., and Norman, T. J., Eds. Argumentation Machines — New Frontiers in Argument and Computation. Kluwer Academic Publishers, Boston/Dordrecht/London, 2003.

[85] Rehg, W., McBurney, P. and Parsons, S. (2004), Computer decision support systems for public argumentation: Assessing Deliberative Legitimacy" AI & Law Society, Nov, 2004 32pp.

[86] Renton, A. & Macintosh, A. (2005). *Exploiting Argument Mapping Techniques to Support Policy-Making*. In K. V. Andersen, A. Gronlund, R. Traunmüller and M. Wimmer (Eds.), Electronic Government: Workshop and Poster Proceedings of the Fourth International Conference, EGOV 2005. Linz: Trauner Verlag.

[87] Renton, A. and Macintosh, A. (2006) Computer Supported Argument Maps as a Policy Memory. Accepted for *The Information Society Journal*.

[88] Rescher, N. Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge. State University of New York Press, 1977.

[89] Richter, G., and Gordon, T. DEMOS – Delphi Mediation Online System. ERCIM News 48 (January 2002), 22–23.

[90] Rittel, H. W., and Webber, M. M. Dilemmas in a general theory of planning. Policy Science 4 (1973), 155–169.

[91] Rittel, Horst W.J. and Melvin M. Webber (1973). "Dilemmas in a General Theory of Planning." Policy Science: 155-169.

[92] Robert, H. M., Robert's rules of order revised for deliberative assemblies. Scott Foresman and Company, Chicago, New York,, 1915.

[93] Rolf and Magnusson (2002). Developing the art of argumentation. A software approach. Paper presented at the 5th international conference on argumentation University of Amsterdam

[94] Roth, B. Case-Based Reasoning in the Law — A Formal Analysis of Reasoning by Case Comparison. University of Maastricht, 2003.

[95] Rowe, G. W. A., Reed, C. A. and Katzav, J. (2006) "Araucaria: Marking up argument" in Working Notes of the European Conference on Computing and Philosophy 2003, Glasgow.

[96] Rowe, G. W. A., Mancagno, F., Reed, C. A. and Walton, D. (2006) "Araucaria as a toll for diagramming arguments in teaching and studying philosophy", *Teaching Philosophy*, 29(2).

[97] Russell, S., and Norvig, P. Artificial Intelligence: A Modern Approach, second ed. Prentice Hall Series in Artificial Intelligence. Prentice Hall, 2003.

[98] Schmidt-Belz, B., Gordon, T. F., and Voss, H. (1999) Urban planning with geomed - first user experiences. In Eurocities, 4th European Digital Cities Conference. Salzburg, 1999, pp. 135–138.

[99] Selvin, A. M., (2003). Fostering Collective Intelligence: Helping groups use visualised argumentation. In Kirshchner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.), *Visualizing Argumentation*, pp. 137-163. London: Springer-Verlag.

[100] Sergot, M., Sadri, F., Kowalski, R., Kriwaczek, F., Hammond, P., and Cory, H. The British Nationality Act as a logic program. Communications of the ACM 29, 5 (1986), 370–386.

[101] Skalak, D. B., and Rissland, E. L. Arguments and cases: An inevitable intertwining. Aritificial Intelligence and Law 1, 1 (1992), 3–45.

[102] Slomkowski, P. Aristotle's Topics. Brill Academic Publishers, 1997.

[103] Suthers, D., (1999). Representational support for collaborative enquiry. Paper presented at the Hawaii International Conference on System Sciences, Maui, Hawaii.

[104] Toulmin, S. E. The Uses of Argument. Cambridge University Press, 1958.

[105] Twardy, C. (2004). Argument Maps Improve Critical Thinking. *Teaching Philosophy*, 27 (2), 95-116.

[106] van Gelder, T., (2003). Enhancing Deliberation Through Computer Supported Argument Mapping. In Kirshchner, P.A., Buckingham Shum, S.J. & Carr, C.S. (Eds.), *Visualizing Argumentation*, pp. 97-115. London: Springer-Verlag.

[107] Van Gelder, T. J. (2002). Argument Mapping with Reason!Able (pdf file). *The American Philosophical Association Newsletter on Philosophy and Computers*, 85-90.

[108] Verheij, B. Rules, Reasons, Arguments. Formal Studies of Argumentation and Defeat. Ph.d., Universiteit Maastricht, 1996.

[109] Verheij, Bart (1998). Argue! - an implemented system for computer-mediated defeasible argumentation. *NAIC '98. Proceedings of the Tenth Netherlands/Belgium Conference on Artificial Intelligence* (eds. H. La Poutré & H.J. van den Herik), pp. 57-66. CWI,

Amsterdam.

[110] Verheij Bart (2000) Dialectical Argumentation as a Heuristic for Courtroom Decision-Making. *Rationality, Information and Progress in Law and Psychology. Liber Amicorum Hans F. Crombag* (eds. Peter J. van Koppen and Nikolas H.M. Roos), pp. 203-226. Metajuridica Publications, Maastricht.

[111] Verheij, B. Deflog: on the logical interpretation of prima facie justified assumptions. Journal of Logic and Computation 13, 3 (2003), 1–28.

[112] Verheij, B. (2005). Virtual Arguments on the Design of Argument Assistant for Lawyers and other Arguers. T.M.C. Asser Press, The Hague

[113] Vreeswijk, G. Defeasible dialectics: A controversy-oriented approach towards defeasible argumentation. Journal of Logic and Computation 3, 3 (1993), 317–334.

[114] Walton, D., and Krabbe, E. C. W. Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning. SUNY series in logic and language. State University of New York Press, Albany, 1995.

[115] Walton, D. Fundamentals of Critical Argumentation. Cambridge University Press, 2006.

[116] Wigmore, J. H. A., 1931, *The principles of judicial proof*, 2nd edition. Little, Brown & Company: Boston.

[117] Wigmore, J. H. A Treatise on the Anglo-American System of Evidence, vol. 1. Little, Brown and Company, 1940.

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D5.2.3 – Ontologies

| | |
|---|---|
| **Editor :** | **Gregoris Mentzas** |
| **Revision :** | |
| **Dissemination Level :** | [TA p. 32] |
| **Author(s) :** | Gregoris Mentzas, Kostas Kafentzis, Christos Halaris, Maria Wimmer, Marian Mach, Peter Butka and Tomas Sabol |
| **Due date of deliverable :** | 30th December 2006 |
| **Actual submission date :** | |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | 5 |
| **Organisation name of lead contractor for this deliverable :** | ICCS |

**Abstract:** The present document provides a thorough coverage of the technology of Ontologies and its potential role within the Demonet project. More specifically, it explains the notions of Ontology and Ontological Engineering, presents the enabling technologies of the field, lists the research groups that conduct research in the area and describes the current applications of the technology as well as its potential application in eGovernment and eParticipation.

**Project funded by the European Community under the FP6 IST Programme**

# 1 Introduction

During the last decade, ontologies and Ontological Engineering have gained increased attention. The concept of ontology is not a new concept as such. It has been used by philosophers (e.g. Aristotle) since ancient times to analyse and categorise what exists. With the increasing use of sophisticated information and communication technology, ontologies have become a concept of interest for structuring information in a way which is close to the human understanding.

In areas such as Artificial Intelligence, ontology became a powerful conceptual tool for Knowledge Modelling. It provides a coherent base to build on and a shared reference to align with, in the form of a consensual conceptual vocabulary, on which one can build descriptions and communication acts. Accordingly, Ontology Engineering refers to the set of activities that concern the ontology development process, the ontology lifecycle and the methodologies, tools and languages for building ontologies.

Ontologies are now widely used in Knowledge Engineering, Artificial Intelligence and Computer Science in applications related to knowledge management, e-commerce, intelligent information integration and retrieval, Semantic Web and many more.

In respect to e-participation, ontologies can help to structure the complex area thereby creating the natural links among application of ICT and the context of citizen engagement in the discourse with politicians and governments. This way, a proper understanding of the field can be provided, which is at the same time machine-readable and computable. In more advanced e-participation implementations, ontologies can represent the basic underlying concept of structuring domains, lines of argumentation etc. where intelligent reasoning and knowledge extraction may be facilitated. The recent technologies to digital ontology descriptions even enable the exploitation of reasoning and inference mechanisms, consequently providing innovative means for knowledge management and personalised and customised tools and services in a wide range of e-participation.

In the following sections, we provide a detailed description of the concept and technology of Ontologies, as well as of the new field of Ontological Engineering covering a number of issues such as: definition of ontology in the context of Computer Science, establishment of the ontology engineering principle and coverage of ontology development methodologies and ontology languages. We provide an overview of key research centres of the field, and investigate the application of ontologies in general, and in e-government. We conclude by addressing the key research questions that the technology needs to cope with in order to advance the domain in e-participation based on potential future scenarios of ontology applications in the field.

# 2 Overall Description of the Technology

## 2.1 Ontology Definition

The word ontology can be used and has been used with very different meanings attached to it. Originating from philosophy, the term 'Ontology' was borrowed by the Knowledge Engineering Community. Since then, many have attempted to provide a comprehensive definition of "Ontology". Setting apart philosophical-related definitions, an appropriate and comprehensive AI-related definition of ontology is provided by Studer and colleagues ([40]) who have merged and explained the definitions given by Gruber ([16]) and Borst ([2]):

*"An ontology is a formal, explicit specification of a shared conceptualization. Conceptualization refers to an abstract model of some phenomenon in the world by having identified the relevant concepts of that phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine-readable. Share reflects the notion that an ontology captures consensual knowledge, that is, it is not private of some individual, but accepted by a group."*

Another definition which highlights the more practical aspects of ontologies is provided by the W3C Recommendation ([47]) according to which "an ontology defines the terms used to describe and represent an area of knowledge. Ontologies are used by people, databases, and applications that need to share domain information[3]. Ontologies include computer-usable definitions of basic concepts in the domain and the relationships among them[4]. They encode knowledge in a domain and also knowledge that spans domains. In this way, they make that knowledge reusable. Ontologies are usually expressed in a logic-based language, so that detailed, accurate, consistent, sound, and meaningful distinctions can be made among the classes, properties, and relations."

## 2.2 Ontology Engineering and Ontology Development Process

During the 1990s and the first years of this century many computer scientists and ontology engineers became interested in formalizing approaches for building ontologies from scratch and for reusing other ontologies in order to minimize the time and effort required for building ontologies. Until the mid-1990s this process was considered an art rather than an engineering activity, and each development team usually followed their own set of methods for manually building ontologies. This lack of common and structured methodologies resulted in increased development times and limited reusability.

[3] The authors (Studer et al, 1998) detail that in their understanding a domain is a specific subject area or area of knowledge, like medicine, tool manufacturing, real estate, automobile repair, financial management, etc.

[4] The authors stress that throughout the W3C recommendation document, the ontology definition is not used in the technical sense understood by logicians

The first attempt to tackle the above issue was made in 1996, in the first workshop on Ontological Engineering that was held in conjunction with the 12th European Conference on Artificial Intelligence. The workshop's goal was to explore and propose a number of principles, design criteria and patterns, and rules of good practice for building ontologies.

A second workshop was held in 1997 with the same topic in Stanford. One of the main aspects dealt with in this workshop was the use of methodologies for designing and evaluating ontologies. Since then, methodological aspects related to different activities of the ontology development process and its lifecycle are included in most of the international conferences on the Ontological Engineering field.

A generic framework that identifies the main activities that should be included in the ontology development process was defined as part of the METHONTOLOGY methodology for ontology construction ([7]). These activities fall into three main categories: ontology management activities, ontology development activities and ontology support activities.

*Ontology management activities* include scheduling, control and quality assurance. The scheduling activity defines and orchestrates the tasks to be performed and estimates the time and the resources needed. This activity is essential in cases of abstract large-scale ontologies. The control activity monitors the executions of the tasks ensuring their appropriate execution and their timely completion. Finally, the quality assurance activity ensures that the quality of each and every product output (ontology, software and documentation) is satisfactory.

*Ontology development activities* are divided into pre-development, development and post-development activities. During the pre-development, the application domain of the ontology is analyzed and the applications where the ontology will be integrated are identified. The reason for this is to determine among others whether it is possible or suitable to build the ontology.

In the development, the specification activity details the reason why the ontology is being built and identifies its intended uses and users. The conceptualization activity is responsible for structuring the knowledge of the application domain transforming it into a meaningful knowledge model represented in a commonly understandable form. This representation is transformed into a formal machine-processable representation during the formalization activity using some formal ontology language. Finally, during the post-development, the maintenance activity updates and corrects the ontology if needed.

*Ontology support activities* consist of a series of activities that are performed in parallel with the development activities and are important to the development process. These activities are:

- Knowledge acquisition: The goal of the knowledge acquisition activity is to acquire the required knowledge for the application domain of the ontology from corresponding experts.
- Evaluation: The evaluation activity is constantly judging from a technical perspective the intermediate and the final results of the ontology development process.
- Integration: The integration activity takes place when already available ontologies are reused for building a new ontology.

- Merging: Merging refers to building a new ontology by unifying concepts, terminology, definitions, constraints, etc. from already available ontologies of the same domain.
- Alignment: The alignment activity maps different ontologies so that they can be used together without having been merged.
- Documentation: The documentation activity describes in a detailed, clear and exhaustive manner the ontology development process and its results.
- Configuration management: The configuration management controls the various versions of the ontology and its documentation.

As it can be recognized, there are some commonalities to the knowledge management processes as being introduced in section 5.2.5 Knowledge Management and Engineering.

## 2.3 Ontological Engineering Methodologies

Classical methods for developing ontologies are:

- Cyc ([27]): The method used for building Cyc, a huge knowledge base with common sense knowledge.
- Uschold and King's method ([46]): The method used for building the Enterprise Ontology.
- Gruninger and Fox's methodology ([17]): The methodology that was used for building the TOVE project ontology.
- KACTUS ([20]): A method used to build an ontology in the domain of electrical networks as part of the Esprit KACTUS project.
- METHONTOLOGY ([14]).
- SENSUS ([42]).
- On-To-Knowledge ([39]).

However, these methods have not been created merely for building ontologies from scratch. Instead, they cover other activities of the ontology development process as well.

For example, in the integration activity it often occurs that the reused ontology is implemented in a language with a underlying knowledge representation paradigm different to the one used by the ontology that reuses it. For dealing with that, METHONTOLOGY includes a re-engineering method ([12]).

Also, ontology learning methods have been thought up in order to decrease the effort during the knowledge acquisition activity. Such methods are used for creating ontologies from scratch as well as for enriching existing ones with new knowledge.

Another important issue for which methods have been created is ontology alignment and ontology merging. Methodologies for ontology merging include the ONIONS methodology ([9]), FCA-Merge method ([41]) Chimarea ([30]) and PROMPT ([35]) while an ontology alignment methodology is AnchorPROMPT ([36]

## 2.4 Ontology languages

A key issue in the ontology development process is the language in which the ontology will be implemented. In the last decade, many ontology implementation languages have been created and other general Knowledge Representation (KR) languages and systems have been used for implementing ontologies despite the fact that they were not specifically created for this purpose.

Generally, we can divide ontology languages into two categories (see Figure). The first category contains a set of AI-based ontology languages, created at the beginning of the 1990s, and based on first order logic, frames and description logics. Some of these languages are **KIF** ([10]), **CycL** ([27],[32]), **Ontolingua** ([5]), **LOOM** ([29]), **OCML** ([31]) and **FLogic** ([23]).

The second category refers to web-based ontology languages, the syntax of which is based on existing markup languages such as HTML and XML. In contrary to the AI-based ontology languages, the purposes of these markup languages are data representation and data exchange respectively. Some of these languages, also known as ontology markup languages, are SHOE ([28]), XOL ([21]), RDF ([26]), and RDF Schema ([3]).

RDF was developed by the W3C (the World Wide Web Consortium) as a semantic network based language to describe Web resources. It is a W3C Recommendation since 1999. The RDF Schema language was also built by the W3C as an extension to RDF with frame-based primitives. The combination of both RDF and RDF Schema is known as RDF(S).

These languages have established the foundations of the Semantic Web ([1]). In this context three more languages have been developed as extensions to RDF(S): OIL ([19]), DAML+OIL ([18]) and OWL ([56]).

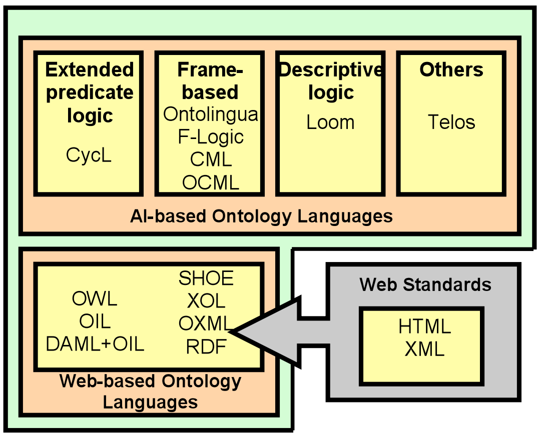



**Figure 1: Overview of description languages for ontologies**

## 2.5 Key research questions and challenges

Building ontologies is difficult for three reasons. First, articulating knowledge in sufficient detail that it can be expressed in computationally effective formalisms is hard. Second, the scope of shared background knowledge underlying interactions of two agents can be enormous. For example, two doctors collaborating to reach a diagnosis might combine common sense conclusions based on a patient's lifestyle with their specialized knowledge. Third, there are unsolved problems in using large bodies of knowledge effectively, including selecting relevant subsets of knowledge, handling incomplete information, and resolving inconsistencies. These problems arise even more in a changing environment.

Thus, research efforts in the ontology domain are concentrating on the fields of ontology learning, ontology merging and large-scale ontology development which try to tackle the above problems respectively. According to [4], the combination of knowledge management (see as well the section 5.2.5 Knowledge Management and Engineering) based on ontologies and Peer-to-peer computing has nowadays become a very popular investigation area.

# 3 Examples of Research Groups

## 3.1 Laboratory for Applied Ontology

http://www.loa-cnr.it/

The Laboratory for Applied Ontology (LOA) belongs to the Institute of Cognitive Sciences and Technology of the Italian National Research Council and performs basic and applied research on the ontological foundations of conceptual modelling, exploring the role of ontologies in different fields, such as: knowledge representation, knowledge engineering, database design, information retrieval, natural language processing, and the semantic web. The group is characterized by a strong interdisciplinary approach that combines Computer Science, Philosophy and Linguistics, and relies on logic as a unifying paradigm. On the application side, special emphasis is given to the use of ontologies for electronic commerce, medical information systems, enterprise modelling, integration of lexical resources, and information access to the Web.

## 3.2 Bremen Ontology Research Group

http://www.fb10.uni-bremen.de/ontology

The Bremen Ontology Research Group is a member of the European Centre for Ontology Research in Saarbrücken established in 2003 in order to support the ontology activities of the Collaborative Research Center on Spatial Cognition and other initiatives requiring formal and linguistic ontologies.

Among the group's goals is the construction of a set of interrelated general common sense ontology modules suitable for all areas of representation, but focusing particularly on spatial representations and information relevant for robotic movement in space: including size, shape and colour representations, routes, and obstacles. Also, they intend to construct a general linguistically motivated ontology for interfacing between all computational components developed within their research centres and natural language technology components for natural dialog.

Finally, they investigate formal modelling and management techniques with varied computational properties as applied to both the ontologies individually and to their mutual relationships, providing for perspectivalism, functional roles and granularity.

## 3.3 DERI Ontology Management Group

http://www.omwg.org

The Ontology Management Working Group's mission is to create the DERI Ontology Management Environment (DOME) aiming at providing an integral solution for the Ontology management problem. The main inspiring principles of their approach are simplicity, completeness (i.e. solving all aspects of the ontology management problem) and reusability.

The Ontology Management suite will comprise tooling support for ontology versioning, merging and alignment that will be integrated as part of the general ontology Editing and Browsing tool. It will make use of an ontology-neutral API that will provide access to different ontology repositories and reasoning engines. Further, a repository and a repository management tool for the administration of the repository will be developed.

The technological mission is to create an efficient and reliable Ontology Management suite enabling application developers to make extensive usage of Ontology technologies for enhanced information processing facilities.

## 3.4 Ontology Engineering Group – Technical University of Madrid

http://www.dia.fi.upm.es/

The Ontology Engineering Group at the Artificial Intelligence Laboratory in the Computer Science School at Universidad Politécnica de Madrid (UPM) carries out research on the Ontological Engineering field and the Semantic Web.

They use ontologies in applications related with e-commerce, knowledge management, natural language processing and the Semantic Web. The most relevant technologies built by the group are WebODE, which is an integrated platform for ontology modeling and management, and ODESeW, an ontology-based application that automatically generates and manages a knowledge portal for Intranets and Extranets.

This technology is based on the widely used and tested methodology for building ontologies named METHONTOLOGY. METHONTOLOGY is the methodology recommended for ontology development by FIPA.

## 3.5 The Information Management Group – University of Manchester

http://img.cs.man.ac.uk/

The Information Management Group (IMG) conducts research into the design, development and use of data and knowledge management systems. Such research activities are broad in nature as well as scope, including basic research on models and languages that underpins activities on algorithms, technologies and architectures.

The group has an international reputation for its work on designing ontology languages, Description Logics, and reasoning systems; it plays a leading role in the development and standardisation of ontology languages, in the design of decision procedures for the underlying Description Logics, in the investigation of the computational complexity of these logics, and in the development of highly optimised reasoning systems based on these procedures. The FaCT and FaCT++ systems developed by the group are widely recognised as having revolutionised the design of description and modal logic reasoners, making the use of languages such as OWL feasible in large scale applications. The design of these systems has been imitated in almost all modern Description Logic reasoners, and has been commercialised in the Cerebra system. The group continues to investigate, design, and build the infrastructure needed for the development and deployment of ontologies in advanced information systems in medical informatics, bioinformatics and hypermedia.

## 3.6 Knowledge Systems Laboratory – Stanford University

http://www.ksl.stanford.edu

KSL conducts research in the areas of knowledge representation and automated reasoning in the Artificial Intelligence Laboratory of the Department of Computer Science at Stanford University. Current work focuses on enabling technology for the Semantic Web, hybrid reasoning, explaining answers from heterogeneous applications, deductive question-answering, representing and reasoning with multiple contexts, knowledge aggregation, ontology engineering, and knowledge-based technology for intelligence analysts and other knowledge workers.

## 3.7 ForschungsZentrum Informatik (FZI – Germany)

http://www.fzi.de

Forschungszentrum Informatik provides the methodological and technological means for formal knowledge representation (ontology engineering, process modelling, service description) and knowledge processing (for service discovery, service composition). It is a non-profit research center with 12 departments headed by professors from the Universities of Karlsruhe, Stuttgart, Tuebingen and Munich. It focuses on novel information technologies for providers of investment and consumer products, of production processes, and of information services.

FZI is Authorized Java Centre (Sun), member of the World Wide Web Consortium (W3C), and member of the Object Management Group (OMG), the standardization body for the Common Object Request Broker Architecture (CORBA). AIFB and FZI are world-wide renowned in the area of Ontology-Based Information Systems and Semantic Web. Through the development of the OntoBroker system, creation of the KAON Open Source Ontology Management suite of tools, and by participation in well-known projects like EU-IST On-To-Knowledge project which established OIL, a Web ontology language building on RDF(S); the DARPA-DAML OntoAgents project; the Onto-logging project which develops methods and tools for ontology-based corporate Knowledge Management; EU-IST project SWWS (Semantic Web Enabled Web Services) which shapes the profile of industrial semantics-based Web Service applications. WIM joined the W3C Web Ontology Working group that establishes future standards for Web ontology representation languages. WIM has grounded experience in the fields of engineering of ontologies and relational metadata based on state-of-the-art Semantic Web methods, tools and standards.

## 3.8 Mindswap – University of Maryland

http://www.mindswap.org

The University of Maryland established the Maryland Information and Network Dynamics (MIND) Lab to provide a focal point for defining, developing, evaluating and deploying new information technologies. In partnership with industry, government agencies and other universities, the MIND Lab addresses key research, education and technology challenges, and creates new opportunities for the market. Its mission is to provide an environment for collaborative research and development efforts among its participants.

The Semantic Web Research Group, or MINDSWAP, is a group of people working with Semantic Web technology inside the MIND LAB at University of Maryland Institute for Advanced Computer Studies. In fact, the group's website is built upon semantic web technologies and specifically OWL.

## 3.9 Knowledge Management Lab - German Research Center for Artificial Intelligence (DFKI GmbH)

http://www.dfki.de

Based in Kaiserslautern and Saarbrücken, the German Research Center for Artificial Intelligence (DFKI GmbH), is the leading industry-geared research institute for innovative software technology in Germany. In the international scientific community the DFKI, which has succeeded in rapidly transforming research into applications, ranks among the important "Centres of Excellence" in the world.

The Knowledge Management department is developing innovative technologies to support the management of data, information and knowledge efficiently, as well as to optimize business working processes within companies. The following topics are primarily being investigated:

- Representation and Management of Data, Information and Knowledge:
- Information Retrieval and Automatic Information Extraction:
- Integration of knowledge-intensive applications into Workflow-Management-Systems
- Document Analysis Techniques for the Optimization of Business Working Processes
- Preservation of knowledge
- Query Assistance

In order to create solutions for these topics it is necessary to develop a multitude of fundamental technologies. These comprise methods for image processing, OCR, lexical post-processing, object recognition, document classification, morphological analysis, syntactical/semantic analysis, pattern matching, benchmarking, ontologies, constraint processing, planning, and configuration.

# 4 Current Applications of the Technology

## 4.1 Knowledge Management

Ontologies are a key concept in the application of knowledge management (KM), thereby providing advanced structures of knowledge which is semantically enriched, machine computable, and which enables automatic reasoning.

A representative application of ontologies in knowledge management is OntoBroker. OntoBroker (http://ontobroker.aifb.uni-karlsruhe.de/index_ob.html) is one of the first implemented tools powered by the use of ontology and Semantic Web technology. It contains three core elements: a query interface for formulating queries, an inference engine used to derive answers, and a webcrawler used to collect the required knowledge from the web. It targets the support for providing a service that can be used more generally for the purpose of knowledge management and for integrating knowledge-based reasoning and semiformal representation of documents. The query formalism is oriented toward a frame-based representation of ontologies that defines the notion of instances, classes, attributes and values ([6]). OntoBroker has been successfully used in several user case scenarios:

- Semantic Community Web Portals: A Portal for the Knowledge Acquisition Community
- Knowledge Annotation Initiative of the Knowledge Acquisition Community (KA)2
- ProPer: Human Resource Knowledge Management.

## 4.2 E-Commerce

Electronic Commerce can be defined as the exchange of information between involved stakeholders using a telecommunication infrastructure. There are two main scenarios: Business-to-Customer (B2C) and Business-to-Business (B2B).

Ontologies in B2C applications enable service providers to advertise their offers, and customers to find offers which match their needs or demands. By providing a single access to a large collection of frequently updated offers and customers, an electronic marketplace can match the demand and supply processes within a commercial mediation environment.

B2B applications have a long history of using electronic messaging to exchange information relating to services previously agreed between two or more businesses. However, a new generation of B2B systems is being developed under the ebXML (electronic business in XML) ([48], [49]) label. These will use classification schemes to identify the context in which messages have been, or should be, exchanged and they will introduce new techniques for the formal recording of business processes, and for the linking of business processes through the exchange of well structured business messages. These schemes, however, usually consist of small-scale ontologies of limited scope.

Therefore, a redesign of these schemes based on state-of-the-art ontology development methods is required, and in several projects under way.

### 4.2.1 Representative applications in E-Commerce

Some representative applications of ontologies in the field of e-commerce are IST-MKBEEM ([50]) and CEN/ISSS MULECO ([51]).

The MKBEEM platform focuses on adding multilinguality to the following stages of the information cycle for multilingual B2C portal services: products or services content and catalogue semi-automated maintenance; automated translation and interpretation of natural language user requests, and natural dialogue interactivity and usability of the service making use of combined navigation and natural language inputs. The main overall goals of MKBEEM were to: develop intelligent knowledge-based multilingual key components (NLP and KRR) for applications in a multilingual electronic commerce platform; validate and assess the prototypes on a pan-European scale (France and Finland) with three basic languages (Finnish, English and French) and two optional languages.

On the other hand, the Multilingual Upper-Level Electronic Commerce Ontology (MULECO) is designed to provide a mechanism whereby relationships between the high-level terms in business ontologies can be inter-connected. Recognizing that most existing ontologies are domain and language specific, and that there is a need to be able to relate terms in one ontology with those in another ontology as part of the semantic translation that is needed to inter-relate applications (as defined within the E-Commerce Integration Meta-Framework (ECIMF) being defined by the CEN/ISSS EC Workshop), the MULECO team are hoping to define techniques that will allow multilingual querying of ontologies based on the relationships between the local ontologies and a set of well defined business and business process classification schemes.

## 4.3 Information Retrieval

In information retrieval (IR) applications, ontologies are used to enhance the relevance of the search results. This is possible because the ontology provides a better representation model of the domain allowing thus the IR system to have a better understanding of the concepts being searched.

Problems in IR systems appear either when the system fails to retrieve relevant answers to the query or when retrieves answers that are irrelevant to the query. The use of ontology-enhanced search and retrieval promises to address this issue by enabling semantic-based retrieval which is more accurate and has more relevant results. In order, however, for this approach to be really effective it needs to take in mind a number of search issues such as context, information quality, and user search mode.

Context is the conceptual framework that determines how relevant a query result is while information quality declares how up to date the information is and whether there are different versions of the same or contradictory information. Finally, user search mode refers to the implicit goal the search has. The search may often differ ranging from browsing, to discovery, to focused search. In each case the relevance of the expected results is different depending exactly on the search’s goal.

### 4.3.1 Relevant applications in Information Retrieval

An ontology-based application for information retrieval is OntoSeek ([52]), a system that adopts a language of limited expressiveness for content representation and uses a large ontology based on Wordnet ([53]) for content matching. Ontoseek combines an ontology-driven content-matching mechanism with a moderately expressive representation formalism. Differently from most of current systems, the user is not assumed to have familiarity with the vocabulary used for component encoding, but the system relies on a large linguistic ontology called Sensus ([42]) to perform the match between queries and data. It assumes that the information encoding and retrieval processes will involve a degree of interactivity with a human user.

Another ontology-based system for information retrieval is SARI ([54]). The SARI (Software Agents for Retrieval of Information) system is intended to act as a broker between human users or other computerized systems (ie applications) needing information at one end, and heterogeneous information sources with different search engines at the other. SARI's architecture reflects the system's role as a broker between its users and information sources. SARI has agents of the following types:

- Application Agents represent the users (humans or other computerized systems) to the SARI system.
- Search Agents mediate information sources. They compile queries coming from Control Agents into the query languages of their information sources, and send the results back to the Control Agents
- Control Agents act as brokers in the SARI system. Each Control Agent receives agent messages containing information retrieval requests from Application Agents, decides to which Search Agents it forwards the requests, sends messages containing the retrieval requests to the appropriate Search Agents, receives messages containing search results from the Search Agents, combines them into information retrieval results, and sends the retrieval results back to the Application Agents
- Ontology Agent contains metadata in the form of ontologies that describe the conceptual structure of the information present in the information sources used by SARI.

In addition, there are also Content Provider Agents that represent content providers to the SARI system. Content providers are organizations or individuals who own one or more information sources that are accessible to the SARI system. Control Agents form the heart of SARI. They make their brokering decisions on the grounds of the user information lying in user profiles, and of the metadata about the information to be retrieved lying in ontologies. Control Agents can form federations with each other, as a rule, but there is just one Control Agent in the present pilot version of SARI. The conceptual structure of the information contained in the information sources available to SARI is described by ontologies. The ontologies describing Web resources are specified as RDF schemas and descriptions for SARI. Ontologies can be graphically browsed in SARI.

In SARI the concepts of different ontologies are linked to each other by making use of the notions of viewpoint and bridge. The ontologies interlinked in such a way form the ontological structure that can be viewed from different perspectives. For example, there is a bridge between the concepts Commodity and Product which are respectively the root classes of the classifications under the foreign trade and manufacturing viewpoints. Future goals with SARI include making the formation of bridges between the concepts of different ontologies semiautomatic, and also semiautomatic generation of RDF metadata from Web resources.

## 4.4 Portals and Web Communities

The widely-agreed core idea of the Semantic Web is the delivery of data on a semantic basis. Intuitively the delivery of semantically processable data should help with establishing a higher quality of communication between the information provider and the consumer. The vision of the Semantic Web is closely related to ontologies as a sound semantic basis that is used to define the meaning of terms and hence to support intelligent providing and access to information for Web communities.

Navigating through Web portals based on a topic thesaurus, like http://dmoz.org or http://www.yahoo.com6, is almost equivalent to browsing a static hierarchy. Those with a richer semantic model, such as KA2Portal [AIF00] (http://ka2portal.aifb.uni-karlsruhe.de), offer simple navigation through a class hierarchy.

The Ontobroker project [6] lays the technological foundations for the AIFB portal. On top of Ontobroker the portal has been built and organizational structures for developing and maintaining it have been established. The approach closest to Ontobroker is SHOE [28]. In SHOE, HTML pages are annotated via ontologies to support information retrieval based on semantic information.

Besides the use of ontologies and the annotation of web pages the underlying philosophy of both systems differs significantly: SHOE uses description logic as its basic representation formalism, but it offers only very limited inferencing capabilities. Ontobroker relies on Frame-Logic and supports complex inferencing for query answering. Furthermore, the SHOE search tool does not provide means for a semantic ranking of query results. A more detailed comparison to other portal approaches and underlying methods may be found in [39].

A richer semantic model, such as SEAL Portal [ Stojanovic et al. 2001] or C-Web [Saglio et al. 2002], offer navigation through a class hierarchy for dynamic exploration. Such framework should help users to navigate through portals leading into very large resource spaces and to find quickly many resources but only those of interest for them.

# 5 Applications in eGovernment

In e-government, semantic web technologies aim to achieve information interoperability. In that direction, ontologies, as a key technology for semantic web development, are considered essential to guarantee data and content interoperability in heterogeneous and multilevel knowledge fields (see ontology section). W3C standard technologies foster semantic interoperability as a whole. In this sense, ontologies are important for e-government service integration. Public services are developed through several collaborations of organizations and ontologies contribute to create a common vocabulary to facilitate web resources access and retrieval. If Governments carry out the integration of different levels of e-government services without a common vocabulary, the integration of heterogeneous resources from different administrative levels and the adequate retrieval of relevant resources would not be possible.

Several Governments around the world are using semantic web technologies for e-government services development ([8]). In the European Union, several IST projects propose the use of ontologies for e-government services.

## 5.1 Webocracy

The Webocracy project ([38]) was a project oriented to the usage of ontologies for structuring organizational information in order to improve information representation and retrieval capabilities. A Webocrat system was developed in order to speed up the modernisation of information systems and allow citizens to participate more actively in opinion polls and the discussion of local issues as well as allowing them to access information about local services more easily through the Internet.

## 5.2 OntoGov

The OntoGov project ([43]) was another e-government project that aimed to develop, test and validate an e-government platform semantically enriched through the use of ontologies that will facilitate the consistent composition, re-configuration and evolution of e-government services. More specifically, OntoGov's goals were:

- To define a high-level generic ontology for the e-government service lifecycle (i.e. covering all the phases from definition and design through to implementation and reconfiguration of e-government services) that will provide the basis for designing lower-level domain ontologies specific to the service offerings of the participating public authorities
- To develop a semantically-enriched platform that will enable public administrations to model the semantics and the processes of their e-government service offerings at different levels of abstraction; easily and consistently re-configure their e-government services; and knowledge-enrich the provision of e-government services to citizens and businesses

- To pilot and evaluate the OntoGov platform in three public administrations in three different European countries. The evaluation of the project results will not be limited merely to the technical evaluation; rather, it will take into account both organisational and social aspects of the project.

## 5.3 EU-PUBLI.com

http://www.eu-publi.com

EU-PUBLI.com (IST-2001-35217) aimed at further automation of inter-agency communication through the structured use of integrating Web Services architectures and ontologies. The approach taken resembled the one of OntoGov: complex public macro-processes (e.g. applying for a tax benefit) shall be broken up into atomic micro-processes that are implemented as Web Services. Europe-wide cooperation of public service agencies shall be achieved through a common framework architecture.

## 5.4 E-POWER

E-POWER project was funded by the European Union's Information Society Technologies (IST) programme, and was initiated by the Leibniz Center for Law of the University of Amsterdam in cooperation with Application Engineers and the Dutch Tax and Customs Administration (DTCA). E-POWER developed a method and supporting tools with which legislation can be 'translated' into formal specifications that can be used by computers. Both method and tools decreased the time to market for new/changed legislation and facilitate the maintenance of legislation. It improved the access to the governmental body of knowledge by offering new e-services. Furthermore, the use of this method and tools results in a more efficient use of scarce knowledge resources. By providing easy Internet access to vital information the project contributes to the social inclusion of citizens. Projects main achievement was creation of MetaLex standard (http://www.metalex.nl/). MetaLex is an open format and a generic and extensible framework for the XML and RDF encoding of the structure and contents of legal documents. It aims to be jurisdiction and language-neutral, and is based on modern publishing concepts like XSLT-based transformation pipelines and emerging Semantic Web standards like RDF and OWL. The MetaLex standard intends to provide a generic and easily extensible framework for the XML encoding of the structure and contents of written public decisions, or public legal documents of a general and regulatory nature. The standard makes only minimal requirements on the structure of documents. The standard tries to adopt, as far as is practicable, relevant guidelines and standards of the World Wide Web Consortium (W3C). The standard is also language-independent, supporting multiple language versions of a legal document, and even localization of XML element names.

## 5.5 QUALEG

www.qualeg.eupm.net/my_spip/index.php

QUALEG (IST 507767: Quality of Service and Legitimacy in eGovernment) The project aims at enabling local governments to manage their policies – they should be able to measure the performance of services they offer, to assess the satisfaction of citizens and to re-formulate policy orientation. Among key enabling blocks of the project software solution can be found: a semantic engine for Web services/workflows coupled with an ontology management system. It aims to jointly publish semantically rich Web services interacting with legacy applications and information sources.

## 5.6 SMARTGOV

Finally the SmartGov project ([55]) had as primary objective the specification, development, deployment and evaluation of a knowledge-based platform that would assist public sector employees to generate online transaction services. This was achieved by simplifying their development, maintenance and integration with already installed IT systems. The SmartGov project, through its software platform, aimed to minimise the reliance on IT skills to develop EGovernment services.

SmartGov was intended to benefit any public authority that is planning or already delivering electronic transaction services, whether or not they have access to the SmartGov platform. It was designed to help improve co-operation, effectiveness and efficiency. The framework was underpinned by the E-Government services ontology. This was intended to provide a common understanding of the principles of E-Government services, an understanding from which people can communicate, discuss and build models of their own.

There are also new ongoing projects which are dealing with knowledge models and might deliver new usable ideas as well as real applications in the area of e-government ontologies.

## 5.7 SAKE

http://www.sake-project.org/

One such project is SAKE (Semantic-enabled Agile Knowledge-based e-Government - FP6 027128, Start date: 01-03-2006). The objective of the SAKE project is to specify, develop and deploy a holistic framework and supporting tools for an agile knowledge-based e-government that will be sufficiently flexible to adapt to changing and diverse environments and needs. It will ensure continual improvement of the quality of the decision making process, through the application of semantic technologies and it will enable empowering of public administrators, by providing efficient access to knowledge needed to resolve cases rapidly and accurately. Among others, SAKE will provide:

- a collaborative working environment that will bring every public servant to the same level of effectiveness and productivity and will ensure more efficient knowledge sharing by guarantee at the same the time the reliability and the consistency of the decision making process a change management system that will ensure harmonisation of requests for changes, resolution of changes in a systematic way and their consistent and unified propagation to the collaborative and knowledge space, in order to ensure the high quality of the decision-making process and

- a platform for proactive delivery of knowledge that enables creation of an adaptable knowledge sharing environment through learning from the collaboration between public servants and their interaction with the knowledge repository and supporting in that way full empowerment of public servants.

## 5.8 BRITE

http://www.briteproject.net/

Another project is BRITE (Business Register Interoperability Throughout Europe, FP6 27219, start date 01-03-2006). BRITE aims to build up a Business Register Interoperability platform through the use of advanced technologies such as semantic web technologies, ontologies, web services, etc. A key result of BRITE will be the specification of the BRITE ontology, an agreement among Business Registers on a unique standard for data exchange in the Business Register domain across Europe.

In this context, already implemented domain ontologies based on different geographical, organizational and historical roots will be faced and harmonised resp. linked up with each other. Instead of ”reinventing the wheel”, the re-usability of existing ontologies will be checked and aggregation of existing data, document and process schemata should be aspired on an overarching level. Furthermore, the BRITE ontology has to fulfil criteria such as openness, dynamics and flexibility in order to allow for future changes and integration of laws to come (especially also in respect to EC enlargement policies). The aim of BRITE is to combine Domain Ontologies and Process Ontologies in a way to achieve maximum productivity. This combination is necessary in order to

a) harmonize the vocabulary towards a common upper level conceptual standard,

b) to get an understanding of the individual, national processes and

c) to integrate the corresponding processes correctly.

# 6 eParticipation Applications of the Technologies

Knowledge technologies are recognised as the key enabling technologies in all fields of emerging eParticipation research and potentially commercial directions**.** By its very nature eParticipation is a knowledge intensive process which is incremental and dynamic, requiring meaningful messages to be extracted over time from large assemblages of data and information produced by multiple stakeholders.

Future research should explore the extent to which a novel combination of basic Knowledge technologies like Ontological Engineering, Semantic Web Services and other mainstream and modern technologies, like groupware or Computer Supported Argument Visualisation (CSAV), can be designed to constructively encourage debate and deliberation by citizens on public issues and to support the analysis and management of contributions to inform policy and services over time (rarely less than several months, and often years at regional level and above).

In fact ontologies are expected to unleash the power of eParticipation to a greater extent. As mentioned above, Ontologies may be applied as the knowledge backbone for eParticipation applications, to offer the standard level of commonly agreed understanding. eParticipation ontologies should then be designed and tested in order to promote common understanding of knowledge and standardization. Technologies like the Semantic Web, XML, RDF and OWL for data interchange and ontology publishing (cf. discussions above) shall be applied for the creation of eParticipation applications.

Ontologies in eParticipation environments should cope with the main stakeholders in participatory processes (citizens, PA's, politicians and moderators), the subject matter of political issues and a number of supportive issues related to the above mentioned questions. In the following, we identify the major areas that ontologies can contribute to eParticipation and relate them with th eParticipation areas described in Deliverable 5.1 at chapter 2.1.

**Information and content organization**: This area refers to the content organization and collection mechanisms, which content is necessary to support the various stages of the eParticipation processes.

The area of content organization has a clear meaning in terms of the policy-making process stages aiming to ease the handling of the complexity of great amounts of contributions and eParticipation related documents, and to ensure coherency (OECD 2004).

*Information Provision*

Ontologies in this context could be used for the categorization and codification of knowledge in a standard manner on top of which information services could provide the appropriate information to users in the different stages of public dialogue. Ontologies could be used to structure a content repository as a "knowledge bank" or a "library" that will provide access to the appropriate information and knowledge to citizens and help "contributors", like public servants or moderators, to put content into the repository in a coherent and concise manner.

*Community building*

Open on-going forums can play an important role in public dialogue. Communities are built around a variety of topics and citizens participate by contributing ideas that could be very useful for authorities in order to design their policies and make their decisions. An example is www.droppingknowledge.org, a forum that brilliantly utilizes ontologies to structure more or less the whole world in order to help people make their contributions and access other people questions and opinions.

*Deliberation*

Organizing citizen's contributions in a structured manner, different from the typical sequential way of Internet discussion forums, and providing links among information pieces can help them reflect better on public dialogue and thus gain more perspective. Ontologies could be the backbone for the structuring, re-structuring and transformation of contributions and knowledge of public dialogue among different renditions, e.g. a sequential forum, a knowledge map, a structured report etc (cf. section 5.2.3 Computer Supported Collaborative Argumentation and Computer Supported Argument Visualisation)

*Mediation*

As ontologies provide a shared understanding of a domain their use could be applied for conflict prevention by unveiling concepts and relations whose ambiguity may lead to misunderstanding or dispute or conflict. The capability to structure a domain in a semantically rich and consistent manner with ontologies and make this knowledge of the domain accessible (e.g. using star-trees) allows participants to be optimistic for better understanding, conflict prevention and improved citizen contribution.

**Information and knowledge extraction**: This area includes the use of language techniques in combination with ontologies for the clear interpretation of public opinion as it is captured e.g. in discussion forums, blogs, wikis or other forms of common discussion spaces.

Despite the fact that there is a large number of commercially available front-end engagement tools for government to deploy, there is limited support for analysis of citizens' contributions to facilitate the input that influences the political agenda. Citizens in eParticipation environments post their views, or refer their inconveniences through public governance web sites that could be forums, or specific blogs or even by e-mail, thus creating large volumes of heterogeneous data. Ontologies will offer a common understanding of the heterogeneous data and through the implementation of knowledge extraction and statistical techniques upon the data, we will extract meaningful messages that will represent to the greater extent public opinion and serious concerns or strong arguments in the public life.

*Discourse*

NLP (cf. section 5.2.1 Natural Language technologies) and knowledge extraction technologies may allow specifying and developing of large scale discourse analysis techniques in order to enable facilitators to support citizen deliberation in various participatory processes. For example, policy-making through stakeholder participation articulates one of the fundamental problems of information and knowledge management, that of abstraction of meaningful messages from large volumes of heterogeneous data. A lot of research in the field of ontologies and knowledge and information extraction may be exploited in order to build the appropriate ontologies and design discourse analysis techniques for analysing large-scale information sources of political discourse. This type of research will be significantly useful in the promotion of a democracy of a superior quality.

**Visualization**: This area refers to the use of techniques for visualizing public opinion through the use of ontologies, achieving a common understanding level.

A promising area on ICT-related research in order to cope with accessibility problems and the so far limited scalability of electronic deliberation is the more appropriate visualization of information, participants, interactions and results. Visualization of interactions in (for example) newsgroups or discussion forums ([57], Ericsson, 1999-2006) is a promising approach to providing a more interesting and informative representation of who is currently available to take part, who has taken part, and which topics they have contributed on. The use of ontologies is necessary in order to achieve a common understanding in many aspects that are necessary for their successful application in eParticipation.

Current systems, including to a certain extent rather advanced eConsultation platforms developed in European projects like DEMOS or Webocracy, are neither appealing nor usable enough to engage the attention of non-technical people and to support neutral third parties in facilitating and analysing the ongoing process. The lack of the standard ontologies that will offer the standard basis upon which visualization may be achieved is recognised as crucial. eParticipation systems must cope with increasingly complex information handling tasks, but must also be intuitive and simple to use.

*Deliberation, Consultation*

Ontology-based visualization can provide support in presenting efficiently in discussion forums, newsgroups etc a lot of useful information like

- which issues, positions and arguments have been brought into the discussion so far?
- which knowledge/results have been produced up until now?
- how have the participants' contributions found their way into the summaries/results?
- what are the varying roles, rights, and responsibilities of participants?
- what is the relationship between background information from policy-makers and citizens responses on the other?

The state of the art can be moved forward by using and integrating state of the art technologies and concepts of computer-supported argumentation (CSAV) visualization and ontology engineering. To date, little research has focused on visualising the substance of political consultations and debates. Recent research in CSAV, most notably published in the book by Kirschner, Buckingham Shum and Carr (2003) demonstrates ontology based techniques being used in facilitating multi-stakeholder deliberation in business processes and industrial conflicts, communicating the key ideas in complex public debates, enabling faster assimilation and critical thinking about complex arguments, and supporting strategic goal setting in businesses. For further details see the report 5.2.3 on Computer Supported Collaborative Argumentation and Computer Supported Argument Visualisation.

**Advanced search utilities:** The main utility of ontologies is the one associated with search facilities in heterogeneous data. As already mentioned, coping with large volumes of data is very important in participatory processes.

*Information Provision*

By building upon emerging XML-based standards of the Semantic Web, including RDF and OWL, and Application Programmer Interfaces (APIs) for interacting with eParticipation knowledge systems, we can enable better and more efficient searches in content and data repositories that can subsequently improve access to relevant and better quality information that leads to more informed decisions and to-the-point contributions

Furthermore, ontology based searching on combined resources such as political knowledge bases and legal corpus schematics is a crucial service for the creation of participatory environments that will lead to better policy formulation.

*Deliberation, Consultation*

Through ontologies and their use for searching we can more easily and efficiently gather and categorize in a unified manner information and public opinions submitted to a variety of sources like forums, newsletters, posters etc. This way, an overall view on a single issue (e.g. a very specific theme, like garbage collection) can be achieved by gathering all the relevant information from disparate resources to one place.

A main aim of DEMO-net is to establish a (virtual) centre of excellence, where stakeholders can find relevant information and knowledge on eParticipation. Within different workpackages and a number of deliverables of Demo-net, a wide range of aspects of e-participation is being investigated and collected. The information gathered in this way is very valuable for researchers and practitioners responsible for advancements in e-participation in their field. The virtual centre of excellence of Demo-net should provide access to this knowledge and it should provide a joint platform of exchange of findings. To structure that information and to provide the knowledge of the area in a reasonable way, an e-participation ontology seems to be a proper concept of information structuring, reasoning, retrieval and visualization

With such a domain ontology, an overall structure (i.e. a knowledge map of eParticipation) can be built up in order to create a common understanding, to structure the many aspects of the field and to provide an overview on the initiatives going on in the different areas. A comprehensive eParticipation ontology shall demonstrate the broad scope of the field, and it shall reflect the landscape of expertise and research disciplines involved in certain aspects of eParticipation, or being involved in eParticipation projects. Exploiting the concepts of ontology and semantic web services for information retrieval, the Demo-net virtual centre of excellence, should become the leading knowledge portal of eParticipation in Europe and worldwide which would be consulted by research, practice and development.

# References

[1] Berners-Lee, T., Hendler, J. and Lassila, O. *The Semantic Web. Scientific American 284, 5 (May 2001), 34-43*

[2] Borst WN (1997) *Construction of Engineering Ontologies*. Centre for Telematica and Information Technology, University of Tweenty. Enschede, The Netherlands

[3] Brickley D, Guha RV (2003) *RDF Vocabulary Description Language 1.0: RDF Schema.* W3C Working Draft. http://www.w3.org/TR/PR-rdf-schema

[4] Davies J., Fensel D., van Harmelen F., Towards the semantic web, JOHN WILEY & SONS, LTD, 2003.

[5] Farquhar A, Fikes R, Rice J (1997) *The Ontolingua Server: A Tool for Collaborative Ontology Construction.* International Journal of Human Computer Studies 46(6):707-727

[6] Fensel D, Decker S, Erdmann M, Studer R (1998) *"Ontobroker in a Nutshell" (short paper)*. In: C. Nikolaou et al. (eds.), Research and Advanced Technology for Digital Libraries, Lecture Notes in Computer Science, LNCS 1513, Springer-Verlag Berlin.

[7] Fernandez-Lopez M, Gomez-Perez A, Juristo N (1997) *METHONTOLOGY: From Ontological Art Towards Ontological Engineering.* Spring Symposium on Ontological Engineering of AAAI. Stanford University, California, pp 33-40

[8] Fraser J et al (2003). *Knowledge management applied to e-government services: the use of an ontology.* In: Knowledge management in e-Government. KMGov-2003 Proceedings. Berlin: Springer, 2003. p. 116-126.

[9] Gangemi A, Pisanelli DM, Steve G (1999) *An overview of the ONIONS Project: Applying Ontologies to the Integration of Medical Terminologies.* Data & Knowledge Engineering 31(2):183-220

[10] Genesereth MR, Fikes RE (1992) *Knowledge Interchange Format. Version 3.0. Reference Manual.* Technical Report Logic-92-1. Computer Science Department, Stanford University, California. *http://meta2.stanford.edu/kif/Hypertext/kif-manual.html*

[11] Gomez-Perez A, Fernandez-Lopez M, Corcho O *"Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web* Springer-Verlag London Limited 2004

[12] Gomez-Perez A, Rojas MD(1999) *Ontological Reengineering and Reuse.* In: Fensel D, Stude R (eds) 11$^{th}$ European Workshop on Knowledge Acquisition, Modelling and Management (EKAW'99). Dagstuhk Castle, Germany. (Lecture Notes in Artificial Intelligence LNAI 1621) Springer-Verlag, Berlin, Germany, pp 139-156

[13] Gomez-Perez A (1998) *Knowledge Sharing and Reuse.* In: Liebowitz J (ed) Handbook of Expert Systems. CRC Chapter 10, Boca Raton, Florida

[14] Gomez-Perez A, Fernandez-Lopez M, de Vicente A (1996) *Towards a method to conceptualize domain ontologies.* In: van der Vet P (ed) ECAI'96 Workshop on Ontological Engineering. Budapest, Hungary, pp 41-52

[15] Gomez-Perez A, Juristo N, Pazos J (1995) *Evaluation and assessment of knowledge sharing technology.* In: Mars N (ed) Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing (KBKS'95). University of Twente, Enschede, The Netherlands. IOS Press, Amsterdam, The Netherlands, pp 289-296

[16] Gruber TR (1993a) *A translation approach to portable ontology specification.* Knowledge Acquisition 5(2):1999-220

[17] Gruninger M, Fox MS (1995) *Methodology for the design and evaluation of ontologies* In Skuce D (ed) IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing, pp 6.1-6.10

[18] Horrocks I, van Harmelen F (eds) (2001) *Reference Description of the DAML+OIL (March 2001) Ontology Markup Language.* Technical Report. http://www.daml.org/2001/03/reference.html

[19] Horrocks I, Fensel D, Harmelen F, Decker S, Erdmann M, Klein M (2000) *OIL in a Nutshell.* In: Dieng R, Corby O (eds) 12th International Conference in Knowledge Engineering and Knowledge Management (EKAW'00). Juan-Les-Pins, France. (Lectire Notes in Artificial Intelligence LNAI 1937) Springer Verlag, Berlin, Germany, pp 1-16

[20] KACTUS (1996) *The KACTUS Booklet version 1.0.* Esprit Project 8145 KACTUS. http://www.swi.psy.uva.nk/projects/NewKACTUS/Reports.html

[21] Karp Pd, Chaudri V, Thomere J (1999) *XOL: An XML-Based Ontology Exchange Language.* Version 0.3. Technical Report. http://www.ai.sri.com/~pkarp/xol/xol.html

[22] Kietz JU, Maedche A, Volz R (2000) *A Method for Semi-Automatic Ontology Acquisition from a Corporate Intranet.* In: Aussenac-Gilles N, Biebow B, Szulman S (eds) EKAW'00 Workshop on Ontologies and Texts. Juan-Les-Pins, France. CEUR Workshop Proceedings 51:4.1-4.14. Amsterdam, The Netherlands (http://CEUR-WS.org/Vol-51/)

[23] Kifer M, Lausen GG, Wu J (1995) *Logical Foundations of Object-Oriented and Frame-Based Languages*. Journal of the ACM 42(4): 741-843

[24] Klischewski R (2003a). *Top down or botton up? How to establish a common ground for semantic interoperability within e-government communities*. In: TRAUNMÜLLER, R.; PALMIRANI, M. eds. E-government: modelling norms and conceps as key issues. Proceedings of 1st international workshop on e-government at ICAIL 2003. Bologna: Gedit edizioni , 2003.

[25] Klyne, G., D., Carroll, J.J. (eds.): *Resource Description Framework (RDF):*

*Concepts and Abstract Syntax*. W3C Proposed Recommendation (work in progress). http://www.w3.org/TR/rdf-concepts/. (2003)

[26] Lassila O, Swick R (1999) *Resource Description Framework (RDF) model and Syntax Specification* W3C Recommendation. http://www.w3.org/TR/REC-rdf-syntax

[27] Lenat DB, Guha RV (1990) *Building Large Knowledge-based Systems: Representation and Inference in the Cyc Project.* Addison-Wesley, Boston, Massachusetts.

[28] Luke S, Heflin JD (2000) *SHOE 1.01. Proposed Specification.* Technical Report. Parallel Understanding Systems Group. Department of Computer Science. University of Maryland. *http://www.cs.umd.edu/projects/plus/SHOE/spec1.01.htm*

[29] MacGregor R (1991) *Inside the LOOM classifier.* SIGART bulletin 2(3):70-76

[30] McGuinness D, Fikes R, Rice J, Wilder S (2000)*The Chimaera Ontology Environment.* In: Rosenbloom P, Kautz HA, Porter B, Dechter R, Sutton R, Mittal V (eds) 17th National Conference on Artificial Intelligence (AAAI'00). Austin, Texas, pp 1123-1124

[31] Motta E (1999) *Reusable Components for Knowledge Modelling: Principles and Case Studies in Parametric Design.* IOS Press, Amsterdam, The Netherlands

[32] NCITS (1998) *Draft proposed American National standard for Knowledge Interchange Format.* National Committee for Information Technology Standards, Technical Committee T2 (Information Interchange and Interpretation). *http://logic.stanford.edu/kif/dpans.html*

[33] Newell A (1982) *The Knowledge Level.* Artificial Intelligence 18(1):87-127

[34] Noy NF, Musen MA (1999) *SMART: Automated Support for Ontology Merging and Alignment.* In: Gaines BR, Kremer B, Musen MA (eds) 12th Bannf Workshop on Knowledge Acquisition, Modelling and Management. Banff, Alberta, Canada, 4-7:1-20

[35] Noy NF, Musen MA (2000) *PROMPT: Algorithm and Tool for Automated Ontology Merging and Alignment.* In: Rosenbloom P, Kautz HA, Porter B, Dechter R, Sutton R, Mittal V (eds) 17th National Conference on Artificial Intelligence (AAAI'00). Austin, Texas, pp 450-455

[36] Noy NF, Musen MA (2001) *Anchor-PROMPT: Using Non-Local Context for Semantic Matching.* In: Gomez-Perez A, Gruninger M, Stuckenschmidt H, Uschold M (eds) IJCAI'01 Workshop on Ontologies and Information Sharing. Seattle, Washington, pp 63-70

[37] OntoWeb IST-2000-29243 (2002) Deliverable D2.1 *Successful scenarios for ontology-based applications*

[38] Paralic J, Sabol T (2003). W*ork with knowledge for support of e-government.* In: Proceedings of the Znalosti 2003 conference. Vostrava (Czech Republic): Vojtech Svatek, 2003.

[39] Staab S, Schnurr HP, Studer R, Sure Y (2001) *Knowledge Processes and Ontologies*. IEEE Intelligent Systems 16(1)26-34

[40] Studer R, Benjamins VR, Fensel D (1998) *Knowledge Engineering: Principles and Methods*. IEEE Transactions on Data and Knowledge Engineering 25(1-2):161-197

[41] Stumme G, Maedche A (2001) *FCA-MERGE: Bottom-Up Merging of Ontologies*. Bernhard Nebel (ed) Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI 2001). Seattle, Washington. Morgan Kaufmann Publishers, San Francisco, California, pp 225-234

[42] Swartout B, Ramesh P, Knight K, Russ T (1997) *Toward Distributed Use of Large-Scale Ontologies*. InFarquhar A, Gruninger M, Gomez-Perez A, Uschold M, van der Vet P (eds) AAAI'97 Spring Symposium on Ontological Engineering. Stanford University, California, pp 138-148

[43] Tambouris E et al. (2004). *Ontology-enabled e-gov service configuration: an overview of the OntoGov project*. In: WIMMER, Maria A. ed. Knowledge management in e-Government. KMGov-2004 Proceedings. Berlin: Springer, 2004. p. 106-111.

[44] Uschold M, (1996) *Building Ontologies: Towards A Unified Methodology*. In: Watson I (ed) 16th Annual Conference of the British Computer Society Specialist Group on Expert Systems. Cambridge, United Kingdom. *http://citeseer.nj.nec.com/uschold96building.html*

[45] Uschold M, Gruninger M (1996) *Ontologies: Principles, Methods and Applications*. Knowledge Engineering Review 11(2):93-155

[46] Uschold M, King M (1995) *Towards a Methodology for Building Ontologies*. In: Skuse D (eds) IJCAI'95 Workshop on Basic Ontological Issues in Knowledge Sharing. Montreal, Canada, pp 6.1-6.10

[47] W3C (2004). *OWL Web Ontology Language Use Cases and Requirements*. *http://www.w3.org/TR/2004/REC-webont-req-20040210/*

[48] http://www.ebxml.org

[49] http://www.ibm.com/developerworks/xml/library/x-ebxml/

[50] http://mkbeem.elibel.tm.fr/

[51] http://xml.coverpages.org/muleco.html

[52] N. Guarino, C. Masolo, and G. Vetere, `OntoSeek: Content-Based Access to the Web', IEEE Intelligent Systems, 14(3), 70--80, (May 1999). http://citeseer.ist.psu.edu/guarino99ontoseek.html

[53] http://wordnet.princeton.edu/

[54] SARI : http://www.ercim.org/publication/Ercim_News/enw35/taveter.html

[55] http://www.smartgov-project.org/

[56] Mike Dean, Guus Schreiber (eds.), Sean Bechofer, Frank van Harmelen, Jim Hendler, Ian Horrocks, Deborah McGuinness, Peter Patel-Schneider and Lynn Andrea Stein. OWL Web Ontology Language Reference. W3C Recommendation 10 February 2004. Latest version: http://www.w3.org/TR/owl-ref/

[57] Smith, M. A. and A. T. Fiore (2001). *Visualization components for persistent conversations. In Conference on Human Factors in Computing Systems*, ACM. http://citeseer.ist.psu.edu/smith01visualization.html

[58] http://bibster.semanticweb.org/index.htm)

[59] Kafentzis, K., Georgolios, P., A. Bouras and G. Mentzas (2006) “An Ontology-based Architecture for Knowledge Commerce”, accepted for publication at the Hawaii International Conference on Systems Sciences, HICSS-39, January 4-7, 2006, IEEE Computer Society.

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D 5.2.4 – Semantic Web Services


| | |
|---|---|
| **Editor :** | **Gregoris Mentzas** |
| **Revision :** | |
| **Dissemination Level :** | [TA p. 29] |
| **Author(s) :** | Gregoris Mentzas, Kostas Kafentzis, Christos Halaris, Maria Wimmer, Marian Mach, Peter Butka and Tomas Sabol |
| **Due date of deliverable :** | 30th December 2006 |
| **Actual submission date :** | |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | 5 |
| **Organisation name of lead contractor for this deliverable :** | ICCS |


**Abstract:** The present document provides a thorough coverage of the technology of Semantic Web Services and its potential role within the Demo-net project. More specifically, it explains the notions of Web Services, Semantic Web and Semantic Web Services, presents the enabling technologies of the field, lists the research groups that conduct research in the area and describes the current applications of the technology as well as its potential application in eGovernment and eParticipation.

**Project funded by the European Community under the FP6 IST Programme**

# Executive Summary

Semantic Web Services is a technology that extends the very popular computing paradigms of Web Services and Service Oriented Computing by facilitating semantic annotation of web services though the use of ontologies. The ultimate goal is to enable automatic semantic-based discovery, composition and execution of web services across heterogeneous users and domains.

The present document aims to provide a thorough coverage of the field of Semantic Web Services and its potential role in the eParticipation domain and particularly within the Demonet project.

More specifically, in the introductory chapter of this document the notions of Web Services and Semantic Web Services are introduced.

Chapter 2 deals with semantic web services in much more detail by providing a thorough description of the fields of Web Services and Semantic Web, by explaining how these two are combined towards Semantic Web Services and by describing the most common technologies related to the field.

In chapter 3, an overview of the key research centres in the areas of Semantic Web and Semantic Web Services is given while chapter 4 investigates the various application fields of semantic web services.

In chapter 5, the important role of semantic web services in the area of e-government is illustrated through a number of application examples.

Finally, chapter 6 concludes by addressing the key research questions that the technology of semantic web services needs to cope with in order to be effectively applied in the domain of e-participation.

# 1 Introduction

In recent years, distributed programming paradigms have emerged that allow generic software components to be developed and shared. The basis of these paradigms is object-oriented programming and interoperability architectures such as CORBA, but their wide scale adoption (with TCP/IP networking – CORBA has been in use for some time) came only after the adoption of XML as a common data syntax and through the definition of Web Service standards.

Web services are well defined, reusable, software components that perform specific, encapsulated tasks via standardized Web-oriented mechanisms ([3]). Their real value however is detected in their ability to be automatically discovered, invoked, and composed along with other services through well defined service modelling frameworks.

Web services are considered a promising technology for applications such as eCommerce and enterprise-wide integration. However, standard technologies for Web services provide only syntactic-level descriptions of their functionalities, without any explanation of what these syntactic definitions might mean. This means that fully automated service discovery and composition (i.e. without human intervention) is not possible, limiting thus the usage of Web Services in complex business contexts.

Semantic Web Services (SWS) aim to relax this restriction by annotating Web Services with semantic descriptions of their capabilities, thus facilitating automated composition, discovery, dynamic binding, and invocation of services. Such a semantic interoperability, however, demands a proper infrastructure which is expected to be provided by the Semantic Web.

In the following sections, we provide a detailed description of the fields of Web Services and Semantic Web, we explain how these two are combined towards Semantic Web Services and we describe the most common technologies related to the field.

# 2 Overall Description of the Technology

## 2.1 Web Services

A Web Service is a software program identified by an URI (Uniform Resource Identifier), which can be accessed via the Internet through an interface that is publicly available. The interface description declares the operations which can be performed by the service, the types of messages being exchanged during the interaction with the service, and the physical location of ports, where information should be exchanged ([3]). Web services are usually deployed in Web servers so that they can be invoked by any Web application, Web agent or even Web Service, independently of their implementations.

Key to the interoperation of Web services is the adoption of a set of enabling standard protocols based on XML namely XML schema, SOAP (Simple Object Access Protocol) ([15]), WSDL (Web Services Description Language) ([16]) and UDDI (Universal Description and Discovery and Integration) ([13]). The underlying architectural concept can be a so-called service-oriented architecture (SOA: [11], [12] or a web services architecture concept ([18]).

XML schema (XML-S) ([1]) is the underlying framework for both defining the Web Services Standards, and variables/objects/data types that are exchanged between services. SOAP is the Web Services standard protocol, recommended by W3C that is used for the exchange of XML data over standard web communication protocols such as http. The use of XML-S for defining SOAP messages ensures that different services interpret the exchanged data in the same manner.

WSDL ([4]) on the other hand is the W3C recommended language for describing the service interface. This description includes the definition of atomic method calls, or operations, in terms of input and output messages as well as the mapping of operations and associated messages to physical endpoints, in terms of ports and bindings.

Operations define the way in which messages are handled e.g. whether an operation is a one-way operation, request-response, solicit-response or notification. Ports declare the operations available with corresponding inputs and outputs and the bindings declare the transport mechanism (usually SOAP) being used by each operation.

Finally, UDDI is a standard for building web services registries which can be browsed and queried by other users, services and applications. Service discovery in a UDDI registry is typically human oriented, based on metadata descriptions of service types, or information about the service providers. Additionally, limited automation of service discovery and invocation is possible within UDDI through stored references to WSDL descriptions of the contained services. However, given that no explicit semantic information is defined, automated comprehension of the WSDL description is limited to cases where the provider and requester assume pre-agreed ontologies, protocols and shared knowledge about operations.

1 gives an overview of how the different web services technologies as introduced above are deployed within the concept of service-oriented architectures (SOA), an upcoming

architectural concept that enables the sharing of web services distributed among, and deployed at diverse web servers.

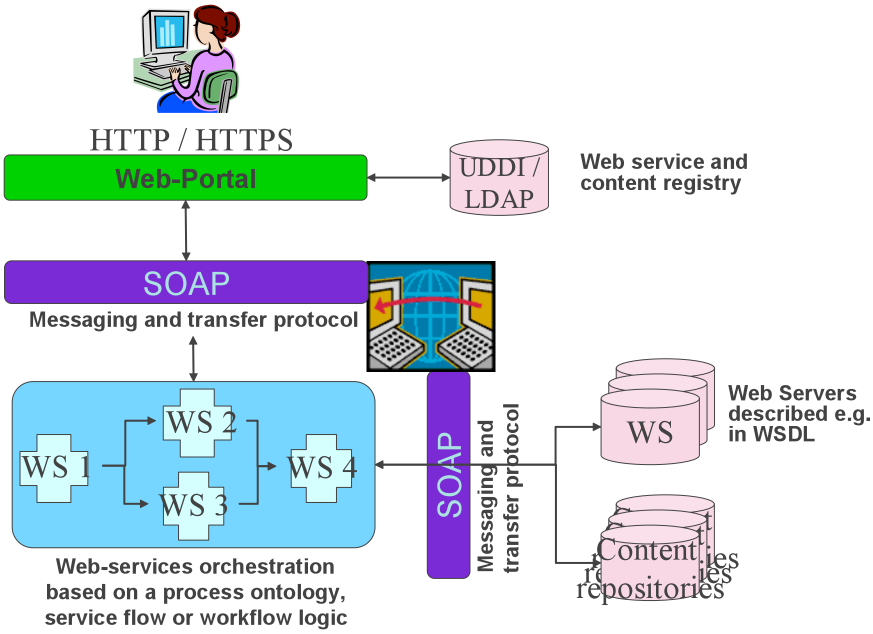



**Figure 1: A general service oriented architecture concept combining distinct semantic web technologies**

The W3C has put the different technologies into a layered reference framework indicating them from the bottom to the top thereby extending the semantic richness and trustworthiness of technology (cf. Figure 2, [5]).

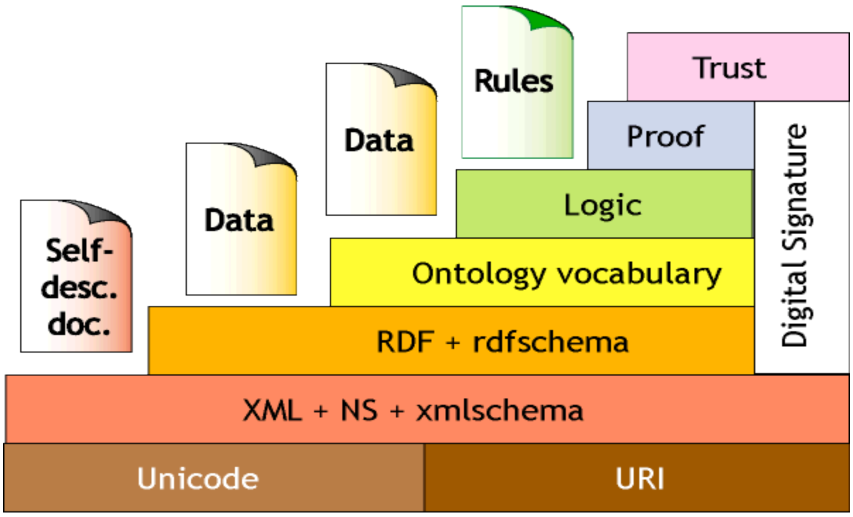



**Figure 2: W3C layers of technologies for semantic web [5]**

## 2.2 Semantic Web Services

The Semantic Web is a vision of a Web as a source of meaningful content and services that can be interpreted by computer programs ([6]). To achieve its goals, it provides the necessary infrastructure for publishing and resolving ontological descriptions of terms and concepts (see paragraph 2.1.1). Additionally, it provides the necessary techniques for reasoning about these concepts, as well as resolving and mapping between ontologies, thus enabling semantic interoperability of Web Services through the identification (and mapping) of semantically similar concepts.

Semantic Web enabling standards are also based on URIs and XML Schema. The current components of the Semantic Web framework are RDF, RDF Schema (RDF-S and the Web Ontology Language – OWL (see paragraph 2.1.4).

Semantic descriptions of Web services enable their automatic discovery, composition and execution across heterogeneous users and domains. These semantic descriptions usually take the form of ontologies as the latter are a rather prevailing standard for describing the semantics of various application domains. Thus, a Semantic Web Service can be defined as a web service whose capabilities and functionalities are semantically annotated (usually through ontologies) and consequently machine interpretable ([3]). Furthermore, a Semantic Web Service incorporates in its description knowledge about the application domain on which it is going to be used. This is done by defining Semantic Web Services through ontologies (cf. sub-chapter on ontology) that enable machine interpretability of the services' capabilities as well as integration with domain knowledge.

Key success factor for the deployment of Semantic Web Services is the creation of the appropriate infrastructures that combine effectively Web Services and Semantic Web enabling technologies.

In general, Semantic Web Service infrastructures can be characterized along three orthogonal dimensions: usage activities, architecture and service ontology (cf. figure 4, [3]). Usage activities define the functional requirements any framework for Semantic Web Services should support while the architecture defines the components needed for accomplishing these activities. Finally, the service ontology is an ontology that aggregates all concept models that describe a Semantic Web Service, thus describing and supporting the usage of the service. An example of such an ontology is OWL-S and it is described in the next section.

The main usage activities required for running an application using SWS include publishing, discovery, selection, composition, invocation, deployment and ontology management. The latter refers to the ontologies that annotate semantically both the SWS and their application domain.

The publishing or advertisement of SWS allows agents or applications to discover web services based on its capabilities. This publishing is performed by storing instances of the web service ontology for each service in a semantic repository. Semantic repositories are engines, similar to the database management systems (DBMS), which allow for storage, querying, and management of structured data. The major differences with the DBMS are can be summarized as follows:

- Semantic repositories use ontologies as semantic schemata. This allows them to automatically reason about the data.

- They work with flexible and generic physical data models (e.g. graphs). This allows them to easily interpret and adopt "on the fly" new ontologies or metadata schemata.

As a result, semantic repositories offer easier integration of diverse data and more analytical power ([19]).

The knowledge modelled by the service ontology is used for service discovery, service composition and service invocation. Discovery of services is nothing else than a semantic matching between the ontology-based description of a web service request and the respective descriptions of published web services. For the matching process any characteristic of a service can be used but the most commonly used attribute is the service's capability. In any case, the service request is expressed, by means of ontologies, in such a way so that it can be directly compared to the candidate services' published description. In case multiple services match the request then these services are ranked according to other non-functional attributes such as cost or quality.

Also, composition is an activity that allows Semantic Web Services to be defined in terms of other simpler services. This allows for minimization of redundancy in the description of SWS and enables more efficient service discovery. The reason the discovery process is more efficient is because it might be that a certain service request cannot be matched with any single SWS description but it can be matched with a proper combination of them (the term service orchestration is also widely used).

The invocation of SWS involves a number of steps, once the required inputs have been provided by the service requester. First, the service and domain ontologies associated with the service must be instantiated. Second, the inputs must be validated against the ontology types. Third, the service can be invoked through the grounding provided.

Finally, the management of service guarantees that semantic service descriptions are created, accessed and reused within the Semantic Web.

From the architecture perspective, SWS are defined as a set of components which realize the activities above, with underlying security and trust mechanisms. These components include a register, a reasoner, a matchmaker, a decomposer and an invoker.

The reasoner is used during all activities and provides the reasoning support for interpreting the semantic descriptions and queries. The register provides the mechanisms for publishing and locating services in a semantic registry as well as functionalities for creating and editing service descriptions. The matchmaker mediates between the requester and the register during the discovery and selection of services. The decomposer is the component required for executing the composition model of composed services. The invoker will mediate between requester and provider or decomposer and provider when invoking services. These components are illustrative of the required roles in the SWS architecture for the discussion here as they can have different names and a complexity of their own in different approaches.

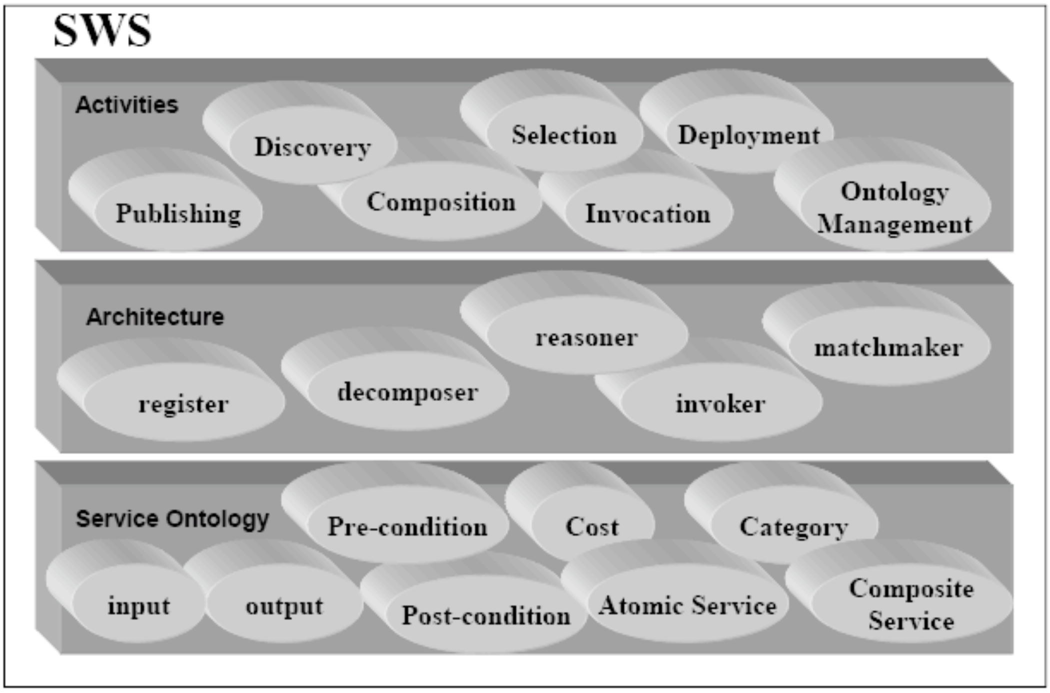



**Figure 3: Semantic Web Services infrastructure dimensions**

## 2.3 Semantic Web Services Technologies

Three main approaches have been driving the development of Semantic Web Service frameworks: IRS-II ([7]), OWL-S ([10]) and WSMF ([6]).

The Internet Reasoning Service - IRS-II is a Semantic Web Services framework, which allows applications to semantically describe and execute Web services. IRS-II is based on the UPML (Unified Problem Solving Method Development Language) framework ([9]), which distinguishes between the following categories of components specified by means of an appropriate ontology that comprises the following elements (figure 4):

- Domain models. These describe the domain of an application (e.g. healthcare, education etc.).
- Task models. These provide a generic description of the task to be solved, specifying the input and output types, the goal to be achieved and applicable preconditions.
- Problem Solving Methods (PSMs). These provide abstract, implementation independent descriptions of reasoning processes which can be applied to solve tasks in a specific domain.
- Bridges. These specify mappings between the different model components within an application.

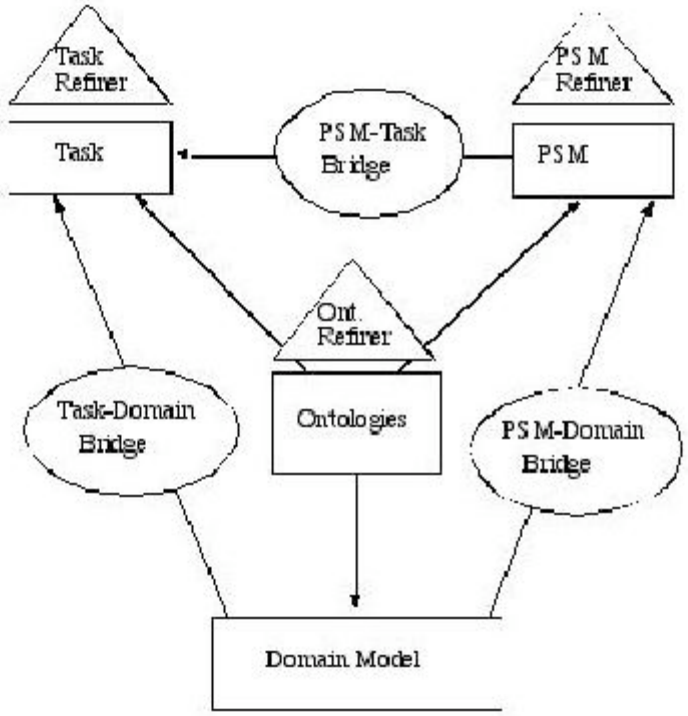



**Figure 4: UPML Architecture**

The main components of the IRS-II architecture are the IRS-II Server, the IRS-II Publisher and the IRS-II Client, which communicate through the SOAP protocol. The IRS-II server holds descriptions of Semantic Web Services at the knowledge level using the UPML framework of tasks, PSMs and domain models. These are currently represented internally in OCML ([8]), an Ontolingua- derived language which provides both the power to express task specifications and service competencies, as well as the operational support to reason about these.

The IRS-II Publisher plays two roles in the IRS-II architecture. Firstly, it links Web services to their semantic descriptions within the IRS-II server and secondly it automatically generates a wrapper which turns the code into a Web service. Once this code is published within the IRS-II it appears as a standard message-based Web service, that is, a Web service endpoint is automatically generated.

A key feature of IRS-II is that Web service invocation is capability driven. The IRS-II supports this by providing a task centric invocation mechanism. An IRS-II user simply asks for a task to be achieved and the IRS-II broker locates an appropriate PSM and then invokes the corresponding Web service.

IRS-II was designed for ease of use. Developers can interact with IRS-II through the IRS-II browser, which facilitates navigation of knowledge models registered in IRS-II as well as the editing of service descriptions, the publishing and the invocation of individual services. Application programs can be integrated with IRS-II by using the Java API. These programs can then combine tasks that can be achieved within an application scenario.

OWL-S ([10]) consists of a set of ontologies designed for describing and reasoning over service descriptions. It combines the expressivity of description logics (in this case OWL) and the pragmatism found in the emerging Web Services Standards, to describe services that can be expressed semantically, and yet grounded within a well defined data type formalism. It consists of three main upper ontologies ([10]): the Profile, Process Model and Grounding.

*The Profile* is used to describe services for the purposes of discovery; service descriptions (and queries) are constructed from a description of functional properties (i.e. inputs, outputs, preconditions, and effects - IOPEs), and non-functional properties (human oriented properties such as service name, etc, and parameters for defining additional meta data about the service itself, such as concept type or quality of service).

The *Process Model* describes the composition or orchestration of one or more services in terms of their constituent processes. This is used both for reasoning about possible compositions (such as validating a possible composition, determining if a model is executable given a specific context, etc) and controlling the invocation of a service.

Three process classes have been defined: the composite, simple and atomic process. The atomic process is a single, black-box process description with exposed IOPEs. Inputs and outputs relate to data channels, where data flows between processes. Preconditions specify facts of the world that must be asserted in order for an agent to execute a service. Effects characterize facts that become asserted given a successful execution of the service, such as the physical side-effects the execution of the service has on the physical world. Simple processes provide a means of describing service or process abstractions – such elements have no specific binding to a physical service, and thus have to be realized by an atomic process (e.g. through service discovery and dynamic binding at run-time), or expanded into a composite process. Composite processes are hierarchically defined workflows, consisting of atomic, simple and other composite processes.

The profile and the process model provide semantic frameworks whereby services can be discovered and invoked, based upon conceptual descriptions defined within Semantic Web ontologies (i.e. OWL).

The *Grounding* provides a pragmatic binding between the concept space and the physical data/machine/port space, thus facilitating service execution. The process model is mapped to a WSDL description of the service, through a thin grounding. Each atomic process is mapped to a WSDL operation, and the OWL-S properties used to represent inputs and outputs are grounded in terms of XML data types. Additional properties pertaining to the binding of the service are also provided (i.e. the IP address of the machine hosting the service, and the ports used to expose the service).

The Web Service Modelling Framework (WSMF ([6])) provides a model for describing the various aspects related to Web services. Its main goal is to fully enable e-commerce by applying Semantic Web technology to Web services. WSMF is the product of research on modelling of reusable knowledge components. WSMF is centered on two complementary principles: a strong de-coupling of the various components that realize an e-commerce application; and a strong mediation service enabling Web services to communicate in a scalable manner. Mediation is applied at several levels: mediation of data structures; mediation of business logics; mediation of message exchange protocols; and mediation of dynamic service invocation.

WSMF consists of four main elements: ontologies that provide the terminology used by other elements; goal repositories that define the problems that should be solved by Web services; Web services descriptions that define various aspects of a Web service; and mediators which bypass interoperability problems. WSMF implementation has been assigned to two main projects: Semantic Web enabled Web Services (SWWS); and WSMO (Web Service Modelling Ontology). SWWS will provide a description framework, a discovery framework and a mediation platform for Web Services, according to a conceptual architecture. WSMO will refine WSMF and develop a formal service ontology and language for SWS.

WSMO ([2]) service ontology includes definitions for goals, mediators and web services. A web service consists of a capability and an interface. The underlying representation language for WSMO is F-logic. The rationale for the choice of F-logic is that it is a full first order logic language that provides second order syntax while staying in the first order logic semantics, and has minimal model semantics. The main characterizing feature of the WSMO architecture is that the goal, web service and ontology components are linked by four types of mediators as follows:

- OO mediators link ontologies to ontologies,
- WW mediators link web services to web services,
- WG mediators link web services to goals, and finally,
- GG mediators link goals to goals.

Since within WSMO all interoperability aspects are concentrated in mediators the provision of different classes of mediators based on the types of components connected facilitates a clean separation of the different mediation functionalities required when creating WSMO based applications.

# 3 Examples of Research Groups

## 3.1 DERI Web Service Execution Environment (WSMX) Working Group

http://www.wsmx.org

The SDK WSMX working group, part of the SDK Cluster aligns the research and development efforts in the areas of Semantic Web Services between the SEKT, DIP and Knowledge Web research projects. Members of this working group include key participants with expertise in Semantic Web-related research areas.

The mission of the SDK WSMX working group is to built up a reference implementation of an execution environment for WSMO. The goal is to provide both a testbed for WSMO and to demonstrate the viability of using WSMO as a means to achieve dynamic interoperability of web services. The development process for WSMX includes defining its conceptual model, defining the execution semantics for the environment, describing an architecture and software design and building a working implementation.

## 3.2 DERI Web Service Modelling Language (WSML) Working Group

http://www.wsmo.org/wsml/

The mission of the SDK WSML working group is to, through alignment between key European research projects in the Semantic Web Service area, work towards further standardization in the area of Semantic Web Service languages and to define a common architecture and platform for Semantic Web Services.

Specifically, the working group aims at developing a language call Web Service Modelling Language (WSML) that formalizes the Web Service Modelling Ontology (WSMO).

## 3.3 DERI Web Service Modelling Ontology (WSMO)

http://www.wsmo.org

The mission of WSMO is to create a Web Service Modelling Ontology, for describing services and its automation process. A world-wide standard will be provided, which will be developed together with industrial partners, research groups, and aligned with many different research projects.

The pillars of the project will be the Web Service Modelling Framework (WSMF), which will provide some basic concepts that will be further developed in the course of the project, and the current available initiatives that try to address similar problems, which drawbacks will be overcome.

## 3.4 Semantic Computing Research Group (SeCo)

http://www.seco.tkk.fi/

The Semantic Computing Research Group (SeCo) of Helsinki University of Technology (TKK) researches machine-processable semantics related to, e.g., the Semantic Web. Its research is focused on semantic media technologies, such as the Semantic Web and intelligent web services. In addition to research and publications, the group also creates prototype applications that demonstrate the new possibilities of semantic technologies, such as semantic portals for end-users, semantic infrastructural services, and ontologies and tools for creating semantic applications.

## 3.5 Intelligent Software Agents Lab – Carnegie Mellon University

http://www.cs.cmu.edu/~softagents/

The Intelligent Software Agents Lab's research on Semantic Web Services concerns the automation of web services discovery and invocation as well as the autonomous interaction with each other reducing thus the need for human mediation.

Therefore the Lab has adapted the RETSINA Discovery mechanism to Web Services, with the development of the DAML-S Matchmaker. The application of its matchmaking technology to the domain of Web Services will help companies reduce the cost of doing e-business, become more agile, deploy faster solutions, and open up new business opportunities.

## 3.6 Semantic Web Technologies Lab – University of Liverpool

http://www.csc.liv.ac.uk/SemanticWeb/

The University of Liverpool Semantic Web Lab is part of the Agent A.R.T. Research Group, located in the Department of Computer Science. The members of the Semantic Web Lab are engaged in a wide range of research topics related to the development of the Semantic Web and the application of Semantic Web technologies. These topics include:

- ontologies for knowledge representation on the Semantic Web and in multi-agent systems
- use of agent research and agent technologies on the Semantic Web
- combining business rules and ontologies to enhance knowledge representation
- indexing and annotation of Semantic Web content
- Semantic Web services, and their intersection with agent services cross-over between Semantic Web and Semantic Grid technologies, and the role of agents in both domains.

The Semantic Web Lab is involved in a number of European and UK funded research projects relating to the development and application of Semantic Web technologies.

## 3.7 Knowledge Media Institute – The Open University

http://kmi.open.ac.uk/

The Knowledge Media Institute's (KMI) research in the Semantic Web research area looks at the potentials of fusing together advances in a range of disciplines, and applying them in a systemic way to simplify the development of intelligent, knowledge-based web services and to facilitate human access and use of knowledge available on the web.

Its researchers are exploring ways in which text mining and natural language processing technologies can be harnessed to support smart annotation of web resources and new forms of web navigation, which go beyond following simple hypertextual links. They are developing infrastructures to support rapid development and deployment of semantic web services, which can be used to create web applications on-the-fly. They are also investigating ways in which semantic technology can support learning on the web, through a combination of knowledge representation support, pedagogical theories and intelligent content aggregation mechanisms.

The group's aim is to be at the forefront of both theoretical and practical developments on the Semantic Web not only by developing theories and models, but also by building concrete applications, for a variety of domains and user communities, including KMI and the Open University itself.

## 3.8 Information Management Unit – Institute of Communication and Computer Systems

http://www.imu.iccs.gr

Based in Athens, Greece the Information Management Unit is a multi-disciplinary Unit engaged in research and development activities in information Technology Management. IMU is a research unit of ICCS (Institute of Communication and Computer Systems), which was established in 1989 by the Ministry of Education and the School of Electrical & Computer Engineering of the National Technical University of Athens, Greece. IMU has developed a research programme in "smart e-government", i.e. a government which provides intelligent dynamic and interactive services to spatially and socially mobile citizens and businesses by fully exploiting knowledge management internally and by using semantically enriched, web-enabled technologies. The research challenges IMU addresses concern the semantic-enrichment of e-government service architectures and the linking of effective organizational transformation to public administration, in order to guarantee increased participation of citizens and enable democratic processes.

# 4 Current Applications of the Technology

Applications of the technology of Semantic Web Services cover many business areas, such as Tourism, Healthcare, Learning, etc., in which interoperability is an emerging need. Several IST projects are trying to find ways of integrating Semantic Web Services in these business areas that aim to create the necessary infrastructures and platforms that facilitate the development of such services.

## 4.1 SATINE (Semantic-based Interoperability Infrastructure for Integrating Web Service Platforms to Peer-to-Peer Networks)

http://www.srdc.metu.edu.tr/webpage/projects/satine/

The SATINE Project aims to develop a secure semantic-based interoperability framework for exploiting Web service platforms in conjunction with Peer-to-Peer networks in the tourism industry. This framework will provide tools and mechanisms for discovering and invoking Web Services through their semantics in peer to peer networks, thus exploiting the synergies between these two technologies. Tourism companies, such as hotel chains, rent-a-car agencies and airline companies can use the SATINE tools and infrastructure in order to wrap their applications with Web Services, enrich those services with semantic descriptions and publish them on the P2P network. The same tools can also be used by service requestors, such as travel agencies in order to discover services based on their semantics, invoke them and combine simple services to complex ones.

## 4.2 Artemis (A Semantic Web Service-based P2P Infrastructure for the Interoperability of Medical Information Systems)

http://www.srdc.metu.edu.tr/webpage/projects/artemis/home.html

The Artemis project facilitates the interoperability of medical information systems through semantically enriched Web services. An essential element in defining the semantic of Web services is the domain knowledge. Medical informatics is one of the few domains to have considerable domain knowledge exposed through standards such as the HL7 Clinical Document Architecture. These standards offer significant value in terms of expressing the semantic of Web services in the healthcare domain.

In Artemis project, HL7 is used to semantically annotate Web services by categorizing the events in healthcare domain and considering service functionality which reflects the business logic in this domain. This classification is used as a basis for defining the service action semantics through a "Service Functionality Ontology". In this way, semantic discovery of Web services is facilitated.

## 4.3 DIP (Data, Information, and Process Integration with Semantic Web Services)

http://dip.semanticweb.org/

DIP's mission is to develop and extend Semantic Web and Web Service technologies in order to produce a new technology infrastructure for Semantic Web Services (SWS) - an environment in which different web services can discover and cooperate with each other automatically. DIP's long term mission is to deliver the enormous potential benefits of Semantic Web Services to e-Work and e-Commerce.

The core objectives of DIP are:

- To further develop the vision of the Semantic Web based on machine-processable semantics as a new communication and co-operation infrastructure. Machine-processable semantics enable the automation of information access and processing.
- To combine Semantic Web technology with Web Services for semantics-based services. DIP believes that a combination of Semantic Web and Web Services technology may well deliver the killer application for the Semantic Web. Semantic Web Services can provide an infrastructure that will not only revolutionize information processing but also the way we access computational resources in general. They will provide a completely new infrastructure to facilitate more effective and cost-efficient electronic business and enable people to work together in better and more innovative ways.
- To apply Semantic Web Services as an infrastructure in real world scenarios within an organization and between organizations and its customers, partners and suppliers. DIP will also address one of the critical success factors in the market take-up of Semantic Web Services by creating practical solutions to real-world business challenges. These solutions will be showcased in scenarios within single organizations, and between and across multiple organizations operating along the classic business value chain. DIP aims to develop practical technology, that can be deployed in new methods for eWork, eGovernment, and eCommerce. The main types of applications are:
    - Intelligent Information Management: The Semantic Web, which ranks as one of this decade's most important software developments, has the potential to improve human information access to unstructured and semi-structured information. Through the use of metadata – information about information – the Semantic Web will help us organise and access the vast amount of material on the Web.
    - Enterprise Application Integration: Semantic Web Services hold the promise of moving beyond the simple exchange of information, the dominant mechanism for application integration today, to accessing application services that are encapsulated in both old and new applications. This means organizations will be able to not only move information from application to application, but also will be able to create composite applications by combining services found in any number of different local or remote applications.
    - Dynamic & Smart eCommerce: Semantic Web Services in B2B applications will enable virtual and smart organizations in commercial and non-commercial environments. Here we are talking about the integration of data, processes, and applications between different organizations introducing advanced requirements for openness, heterogeneity, and change.

DIP will strive to develop Semantic Web Services as a scaleable and cost effective solution to the integration problem, thereby dealing with one of the key bottlenecks of modern networked society. According to Gartner analysts, there will be strong and in some cases explosive demand for Web Services and integration technology and services as businesses react to the need for more integration and more agility.

Making disparate systems share information cost-effectively is a perennial problem for companies and represents billions of dollars in technology spending, with an estimated 30% of worldwide IT budgets dedicated to Enterprise Application Integration (EAI) type projects.

## 4.4 FUSION (Semantic Business Process Fusion (2006-2007)

http://www.fusionweb.org/

FUSION aims at a three-fold focus:

- Development of an innovative approach, methodology and integration mechanism for the semantic integration of a heterogeneous set of business applications, platforms and languages within SMEs.
- Integration of research activities carried out in the Enlarged Europe in the areas of Business Process Management, Semantic Web and Web Services.
- Validation of research results by developing proof-of-concept pilots in collaborative commerce growth across semantically-enriched value networks across the Enlarged Europe. In particular, FUSION will facilitate three trans-national cases.

The expected results of FUSION include:

- The FUSION approach for Semantic Service-oriented Business Application integration covering essential business processes between collaborative organizations.
- The FUSION Methodology for Semantic Service-oriented Business Application Integration that will facilitate the integration of business software applications.
- The FUSION integration mechanism will simplify the interconnection of heterogeneous information systems, resource sharing and services provision.
- Three FUSION cases which will prove the concepts and tools of the solution, concerning three different "Enlarged Europe" use cases.

## 4.5 FIT: Self-adaptive e-government Service Improvement with Semantic Technologies

http://www.fit-project.org/

For e-government initiatives to succeed, public services should be organised in away to serve every citizen individually. Since citizens pose different access possibilities, skills and motivation, service delivery should be tailored to the widest possible end-user population. This adaptivity means that all citizens have access to the public services in a manner which is enabling and satisfying. It requires not only the personalized service delivery, but more important the extension of the service description by including dynamically changing citizens' needs. The overall objective of FIT is to develop, test and validate a self-adaptive e-government framework based on semantic technologies that will ensure that the quality of public services is proactively and continually fitted to the changing preferences and increasing expectations of e-citizens

## 4.6 SEKT (FP6-506826, Semantic Enabled Knowledge Technologies

http://www.sekt-project.org/

The vision of SEKT is to develop and exploit the knowledge technologies which underlie Next Generation Knowledge. This vision is (among others) based on the assumption that advanced reasoning capabilities will strongly support the evolution of ontologies and metadata and greatly reduce the overhead for maintenance. Work in advanced reasoning will include the development of techniques for robust reasoning, i.e. reasoning in the presence of inconsistencies, i.e. in order to give meaningful results even when the overall ontology has conflicts. It will also include flexible reasoning which can cope with changes and conflicts in a given model and can fall back to old versions or change the scope of reasoning to a consistent set of statements. The advanced reasoning work will support the evolution of ontologies and meta-data, in order to reduce maintenance overhead. SEKT introduces the notion of knowledge workplaces where the boundaries between document management, content management, and knowledge management are broken down, and where knowledge management is an effortless part of day to day activities. Appropriate knowledge is automatically delivered to the right people at the right time at the right granularity via a range of user devices.

The SEKT concept demonstrator has already been created (see the web page) to show the business potential of SEKT technology. The scenario is set in the financial services sector, but the ideas can be easily extrapolated to a range of sectors and application domains. Please take a few minutes to view this demonstrator and learn about the power of SEKT technology.

## 4.7 NEON (FP6-27595, Lifecycle support for Networked Ontologies)

http://www.neon-project.org/

NeOn is a major European initiative shaping the future infrastructures for semantically aware and Semantic Web applications. The aim of NeOn is to create the first ever service-oriented, open infrastructure, and associated methodology, to support the development life-cycle of such a new generation of semantic applications, with the overall goal of extending the state of the art with economically viable solutions. These applications will rely on a network of contextualized ontologies, exhibiting local but not necessarily global consistency. NEON also aims at aspects of tailoring the human-ontology interaction to users' profiles.

# 5 Applications in eGovernment

Applications of the technology of Semantic Web Services in the area of eGovernment are mainly focused on achieving information interoperability among different e- government services by annotating them semantically. In that way, orchestration and composition of different e-government services is possible thus achieving collaboration and interoperability in an intergovernmental level.

In the European Union level, a number of IST projects propose the use of the Semantic Web Services technology as the main way of modelling e-government services.

## 5.1 TERREGOV

http://www.terregov.eupm.net/my_spip/index.php

One of these projects that address the issue of interoperability of eGovernment services for local and regional governments is TERREGOV.

TERREGOV integrates the dimensions of technological R&D, pilot applications involvement and socio-economic research in order to offer a European reference for the deployment of interoperable eGovernment services in local governments. It deals with the challenge that the governments face in redesigning their business processes in order to implement government processes that invoke services (eProcedures, access to existing legacy information systems and databases) from multiple administrations, make these government processes available to other administrations as eGovernment services and support civil servants involved in such eGovernment processes in getting a clear knowledge of the processes and of the services in order to act as a knowledgeable front-end to citizens (providing advices, identifying the most adequate services, launching the processes for specific citizen cases).

Taking the view that government services are offered by a number of administrations interacting one with each other and that local administrations often act as a front office to the Citizen, TERREGOV's goal is to make it possible for local, intermediate (municipality groupings, districts, ...) and regional administrations to deliver online a large variety of services in a straightforward and transparent manner regardless of the administration(s) actually involved in providing those services.

To promote transparency and responsibility in eGovernment, TERREGOV promotes and supports the creation of centrally controlled orchestrated procedures. In order to allow the most complete use of these procedures, TERREGOV provides means of supporting and helping civil servants in their daily use of the system.

From a technological perspective, TERREGOV focuses on the needs for flexible and interoperable tools to support the change towards eGovernment services, in emerging eGovernment interoperability frameworks. It unfolds in 3 technological R&D Streams

- Web Services and eGovernment Processes to combine flexible eGovernment interoperable services in end-to-end process workflows.
- Semantic enrichment eGovernment Services to enable Web Services to discover each other on a semantic basis.

- Support to Civil Servants to enable civil servants to focus on the added value of the service delivered to Citizens - increasingly acting as advisers.

## 5.2 QUALEG

The QUALEG project implemented a WSDL based workflow management system which brings together the interoperability features of Web services and the business process design and enactment features of workflow management,

## 5.3 EU-PUBLI.com

EU-PUBLI.com's main challenge has been described as proving whether or not Web Services orchestration can be applied in practical e-government scenarios under realistic service load. According to the publications available, this claim has been proved successfully in prototypical simulations. The recommendations are that Web Services can efficiently be used for electronic government services on a large scale with their interfaces directly accessible to service requesters (i.e. the citizens). The DEMO-net project might also use outcomes from the new ongoing projects which are dealing with semantic web services.

## 5.4 SEMANTICGOV

http://www.semantic-gov.org/

The SEMANTICGOV project (FP6-2004-IST-4-027517: Services for Public Administration, starting date: January 1st, 2006) aims at building the infrastructure (software, models, services, etc) necessary for enabling the offering of Semantic Web Services by public administration through the use of the Semantic Web. Through this infrastructure, SemanticGov will address longstanding challenges faced by public administrations such as streamlining cooperation amongst agencies both within a country as well as amongst countries, easing the discovery of public administration services by its customers, facilitating the execution of complex services often involving multiple agencies in interworkflows. The project intends to utilise the infrastructure represented by WSMO, WSML,and WSMX to implement components supporting the aim of the project.

## 5.5 Access-eGov

http://www.access-egov.org/

The FP6-2004-27020 Access-eGov Project (funded by the EC under the Sixth Framework Programme in the Information Society Technologies Programme) aims at development of component-based enhancements of existing e-Government infrastructure based on Semantic Web technologies and distributed architectures (service-oriented and peer-to-peer). These components will enable e-Government service providers (on all levels of public administration - local, regional, national, and European) to easily introduce any (new) service to the world of e-Government interoperability. Once the service is registered in the Access-eGov system, it may be localised, contracted and used (in case of e-service) automatically through agents and other IT components. For service users (citizens as well as businesses) Access-eGov will increase accessibility and facilitate connectivity of the existing e-services across organisational and regional borders, provide more information necessary for the use of traditional PA services and thus enable "integration" of traditional and e-services into "hybrid scenarios". And since not all users feel comfortable when dealing with a myriad of PA services, a virtual personal assistant will guide users through this scenario. the project uses semantic technologies to be able to search for appropriate government services and to ensure their semantic interoperability, All semantic descriptions are expected to be stored in a decentralised semantic directory infrastructure.

## 5.6 R4eGov

http://www.r4egov.info/

R4eGov (FP6 IST-4-26650, running 2006-2008) is another project that aims at the development of concepts and tools for eAdministration in the large. Thereby, the specific foci are interoperability by large and secure cross-organisational collaborative workflows. Web services and semantic web services technologies will be extensively used in a series of service cases.

# 6 eParticipation Applications of the Technology

The description of architectures enabling service delivery in the eParticipation area is quite crucial. Luis Álvarez Sabucedo and Luis Anido Rifón in [20] attempt to propose an architecture that will allow the use of semantic web services in the context of e-government including eParticipation.

Similar approaches were also proposed in the eGov-Interop 05 conference mainly addressing the issue of providing interoperable e-government services with respect to the semantic heterogeneity that they bear. In the proceedings of this conference one can retrieve several architectures proposed in numerous projects like the TERREGOV, eGovernment interoperability on a semantically driven world by Vicente et al 2005.

Nevertheless most of these approaches are schematised by taking into account that most eGovernment projects share certain characteristics like:

- Open Source. Software elements must be available for every user involved in the project and no assumption may be done about the operative system or required tools. Of course, it is not acceptable to force the adoption of some certain programs when there are available free and open alternative.
- Adoption of open standards. By using open standards in every layer of projects, it is possible to guarantee the maintenance and the support for new improvements in the scope of constant researching. Standards used are mainly the OWL-S standard for semantic we services.
- Interaction of any agent. The interaction must be supported even with agents not provided by the administration. Thus, it will be possible to develop agents by anybody that may become part of the system. This feature will largely increase eParticipation as you allow citizens to really take part of Public Administration by mean of their contributions.
- Support for multiplatform devices. Agents in the system may use different network supports, i.e., agents may use wired network, WAP devices, Wi-Fi devices or any other support as they work with open standards.
- Ontology-driven. By using this feature, we will achieve a high level of interoperability as machines will be able to process data with little or no human participation. This also provides interesting advances on server composition and mechanization of procedures by mean of autonomous systems. This feature requires a lot of efforts to implement it in a proper way.

The challenge of creating real life eParticipation applications by using Semantic Web Services has not yet been thoroughly addressed. Obstacles that we identify are:

- The lack of eParticipation protocols that set the border of communication between the stakeholders in eParticipation issues
- The lack of standard eParticipation ontologies
- The lack of the agent based approach, clear identification of agents in deliberation processes that will reveal the requirements of the level of semantic interoperability in Participation issues

The research direction in semantic web services and eParticipation should be directed toward the combination of communication and collaboration protocols in terms of public policy making and policy development. The main areas of eParticipation where semantic web services could contribute the most are the following:

*Information Provision*

Providing access to information and knowledge by publishing repositories using semantic web services [22] can be a main application that will enable co-operation among public authorities in the eParticipation context. Either this information is for only informative purposes (e.g. the content of a fact bank about a theme) or it includes citizens' public opinion (e.g. the content-contributions from a discussion forum), semantic web services can be an interoperable and easy way to publish and distribute it. Therefore, an opportunity arises for public administration to gain access to dispersed public opinions that could provide, after appropriate analysis, a more comprehensive and collective view of citizens' opinion and lead to improved decision making processes.

*The Software as a Service Concept*

Most non-research projects and initiatives in eParticipation have a small budget that provides several constraints on the ICT functionality that will be available to the citizens and the authorities. From the side of authorities, the lack of ICT support for back-office operations can lead to serious delays, increased workload and poor analysis results. For example, managing and analysing the large assemblages of data and information produced by multiple stakeholder participation over time is a very cumbersome task that requires ICT support.

The concept of providing for fee or for free this functionality as a service to the authorities can be a critical success factor of eParticipation initiatives and projects under the typical budget constraints of such projects since a service-based approaches incur decreased costs. The Google and Microsoft Live cases prove that for free service provision is also a viable approach.

# References


[1] Biron, PV, Malhotra, A (2001): *XML Schema Part 2: Datatypes,* W3C Recommendation, 2 May 2001. *http://www.w3.org/TR/xmlschema-2/.*

[2] Bruijn, J. de, Fensel, D., Keller, U. Lara, R. (2005). Using the Web Service Modeling Ontology to enable Semantic E-Business. In CACM, 48 (12), pp. 43 – 47.

[3] Cabral, L., Domingue, J., Motta, E., Payne, T., and Hakimpour, F. (2004) *Approaches to Semantic Web Services: An Overview and Comparisons,* 1st European Semantic Web Symposium, Heraklion, Greece

[4] Christensen, E. Curbera, F., Meredith, G., Weerawarana, S. (2001) *Web Services Description Language (WSDL)*, W3C Note 15. http://www.w3.org/TR/wsdl.

[5] Daconta, M., Obrst, L., Smith, K. (2003). *The Semantic Web. A Guide to the Future of XML, Web Services, and Knowledge Management*. Indianapolis: Wiley.

[6] Fensel, D., Bussler, C. The *Web Service Modeling Framework WSMF*. Electronic Commerce: Research and Applications. Vol. 1. (2002). 113-137

[7] Motta, E., Domingue, J., Cabral, L., Gaspari, M.: *IRS-II: A Framework and Infrastructure for Semantic Web Services*. In: Fensel, D., Sycara, K., Mylopoulos, J. (volume eds.): The SemanticWeb - ISWC 2003. Lecture Notes in Computer Science, Vol. 2870. Springer-Verlag, Heidelberg (2003) 306–318

[8] Motta E (1999) *Reusable Components for Knowledge Modelling: Principles and Case Studies in Parametric Design.* IOS Press, Amsterdam, The Netherlands

[9] Omelayenko, B., Crubezy, M., Fensel, D., Benjamins, R., Wielinga, B., Motta, E., Musen, M., Ding, Y..: *UPML: The language and Tool Support for Making the Semantic Web Alive*. In: Fensel, D. et al. (eds.): Spinning the Semantic Web: Bringing the WWW to its Full Potential. MIT Press (2003) 141–170

[10] OWL-S Coalition: *OWL-S 1.0 Release*. http://www.daml.org/services/owl-s/1.0/. (2003)

[11] Richter, J.P., Haller, H., Schrey, P. (2005). Serviceorientierte Architektur. In: Informatik Spektrum, October 2005, pp 413 - 416

[12] Schwegler, B. (2004): Service-oriented Architecture, Presentation at the e|Gov Days 2004 of the Forum e|Government, Vienna, 2004, http://weblogs.asp.net/beatsch

[13] UDDI Consortium. (2000) *UDDI Specification.* http://www.uddi.org/specification.html

[14] W3C (2003). *SOAP 1.2,* W3C Recommendation. http://www.w3.org/TR/soap12-part0/

[15] D. Box, et al.; *Simple Object Access Protocol (SOAP) 1.1*; W3C Note, 08 May 2000, http://www.w3.org/TR/SOAP/

[16] http://www.w3.org/TR/wsdl

[17] AAAI 2006 Fall Symposium on Semantic Web for Collaborative Knowledge Acquisition http://www.cild.iastate.edu/events/aaai06symposium.html

[18] Booth, D. et al.; Web Services Architecture, W3C Working Group note, p. 61, http://www.w3.org/TR/ws-arch/

[19] http://www.ontotext.com/inference/semantic_repository.html

[20] Luis Álvarez Sabucedo, Luis Anido Rifón: *UDDI Service in eGovernment Environments*, eGov-Interop'05 Annual Conference

[21] http://bibster.semanticweb.org/index.htm)

[22] Kafentzis, K., Georgolios, P., A. Bouras and G. Mentzas (2006) “An Ontology-based Architecture for Knowledge Commerce”, accepted for publication at the Hawaii International Conference on Systems Sciences, HICSS-39, January 4-7, 2006, IEEE Computer Society.

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D 5.2.5 – Knowledge Management and Knowledge Engineering


| | |
|---|---|
| **Editor :** | **Maria Wimmer, Andreas Rosendahl** |
| **Revision :** | D5.2.5-v. 07 [final] |
| **Dissemination Level :** | PU |
| **Author(s) :** | Dimitris Apostolou*, Frantisek Babic°, Spyros Dioudis*, Marian Mach°, Gregoris Mentzas*, Marek Paralic°, Jan Paralic°, Andreas Rosendahl#, Tomas Sabol°, Asta Thorleifsdottir+, Maria Wimmer# [in alphabetical order]<br>* TUK - Technical University of Kosice (SK)<br>° ICCS -Institute of Communication and Computer Systems (GR)<br>+ UI- University of Iceland (IS)<br># IWVI - University of Koblenz-Landau (DE) |


| | |
|---|---|
| **Due date of deliverable :** | 30th December 2006 |
| **Actual submission date :** | 29th January 2007 |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | 5 |
| **Organisation name of lead contractor for this deliverable :** | IWVI |



**Abstract:**

*The public sector is dealing with a vast amount of information and knowledge resources. Many activities and results are of the nature of information and knowledge. However, there is still a lack of a clear and comprehensive understanding of how knowledge and information are being used and supported in e-participation. The purpose and rationale of investigations is to get an overview of which tools and technologies of data and knowledge engineering can support e-participation in its various forms. The sub-*

*deliverable first discusses the various kinds of knowledge in e-government and e-participation. Subsequently, concepts of knowledge management, tools and technologies supporting KM processes are discussed. Among them are knowledge repositories, knowledge structuring concepts, data storage concepts, knowledge extraction concepts (KDD, OLAP, Data Mining), and more advanced concepts of case-based reasoning. The deliverable concludes with future scenarios of applying KM concepts and technologies in different e-participation areas, and are embodied in distinct e-participation tools and processes.*

**Project funded by the European Community under the FP6 IST Programme**

# Executive Summary

The public sector is dealing with a significant amount of information and knowledge resources. This knowledge has to be appropriately managed and smoothly integrated. Especially in policy formulation, i.e. in various e-participation areas, the activities and results of action are of information and knowledge by nature. Yet, we still lack a clear understanding of what kind of knowledge and information we are treating in e-participation, what purposes and rationale lays behind investigations and activities and which tools and technologies of data and knowledge engineering can support e-participation in its various forms.

The sub-deliverable at hand tries to give answers to several questions:

First, the introduction sets the scope and ground of understanding for information and knowledge in e-government and e-participation. It further raises four key challenges of knowledge management in e-participation.

Chapter 2 is an introduction to the types of information and knowledge in e-government and e-participation. Thereby, a holistic framework of understanding, and specific aspects of knowledge in governmental processes are discussed. An example of democratic deliberation demonstrates the knowledge aspects in this process.

In chapter 3, methodologies for knowledge management are presented. Comparisons distinguish between concepts describing KM processes, methodologies for knowledge engineering, a concept for knowledge distribution and the knowledge spiral of Nonaka and Takeuchi to understand the knowledge creation process.

Subsequently, KM tools and technologies are introduced. Chapter 4 covers comprehensive approaches such as corporate memories, knowledge portals, workflow management systems, or case-based reasoning. Furthermore, concepts for structuring information and knowledge, for information retrieval, and for knowledge analysis are discussed. In addition, agent technologies, and alternative concepts such as individualised feeds, recommender systems, social bookmarking and the like are introduced.

Chapter 5 investigates the potential use scenarios of existing KM technologies for e-participation.

Since KM technologies and solutions are not yet widely used in e-participation contexts, chapter 6 concludes with a number of research questions to be investigated in future e-participation research.

# 1 Introduction

## 1.1 Motivation

Entering the 21st century, knowledge has become an objective of utmost interest for individuals and organisations. Although defined as *the* critical success factor for the future, we are far from being aware of the knowledge around us. Accordingly, John Naisbitt remarked: "We drown in information but we are thirsty for knowledge".

To rearrange large volumes of data and information to locate and manage the residing knowledge, a good understanding from different perspectives is required:

- we need to go beyond the technical view when developing ICT support for knowledge management;
- we need to change the perception of knowledge as some kind of information net enriched with some contextual information;
- we need to realise that knowledge is around us, embodied in tools and artefacts, and that it is up to us to benefit from using that knowledge.

For modern organisations, knowledge needs to be perceived as the basic resource, without caring about its physical shape or abstract form. In performing any process, knowledge flows among the components and objects belonging to that process. Such components may be among others: ICT, process and workflow descriptions, physical artefacts such as manuals or signs, and - most importantly – humans, in their skills, know-how and culture. In any productive system knowledge is distributed among many components that belong to that system.

With the hype of the information society a precarious factor arose: ICT is supposed to solve nearly every problem. Yet, ICT is developed by people and in doing so, it is shaped by the knowledge the developer embodied there. What has been neglected is that knowledge dynamically changes in an ever evolving society. So far, ICT cannot migrate and develop on its own - it needs to be maintained by humans. Individuals, however, are capable of adapting quite fast to changing requirements through their intellectual capabilities. So, we need to develop tools where we can continuously elaborate knowledge, where knowledge is in an accessible form everywhere, and where the explicit form of knowledge is not the only one of interest and being dealt with.

## 1.2 Knowledge in the Public Sector

The public sector is dealing with information and knowledge resources in large. This knowledge has to be appropriately managed and smoothly integrated. Hence, *knowledge enhancement and knowledge management* options must be integrated in development from the beginning, in e-government and e-participation alike[5]. As both are complex

[5] Also, e-participation is in many cases reliant on underlying e-government systems.

socio-technical systems, a multitude of aspects need to be considered: people and organizations, ICT systems and architectures, multiple access channels, multiple devices, a variety of process structures, multiple parties (authorities and clients) involved, official procedures and norms, and multiple other physical and abstract knowledge sources.

Many issues are to be addressed for successfully implementing e-participation systems. Hence, a holistic framework is required to understand the knowledge types, the knowledge carriers and their interaction, based on three key characteristics of e-participation:

- different user groups with diverging needs and interaction requirements,
- citizen participation areas (employed in D5.1),
- support of participation with modern ICT.

Many public organizations are chiefly "intelligence organizations" and officials can be considered as knowledge workers. Complex decisions are particularly knowledge demanding. Hence, much knowledge has to be transferred to and discussed by citizens for the purpose of deliberation and participation. In the domain of e-participation, knowledge management has to deal primarily with four challenges as sketched in the following.

## 1.3 Four Challenges of KM in e-Participation

### 1.3.1 Content Integration

Content integration refers to the tough task of connecting the countless existing data collections. Mostly a collection of rather heterogeneous data repositories is entailed that contains data of diverse type formats and that originates from different sources.

Content integration engages all sorts of conventional ways of keeping data: files, databases, legacy information systems, that are based on proper structures of data. Efforts should lead to sophisticated content management. However, many obstacles have to be overcome for such sophisticated content management, such as rendering information visible by use of one browser for all diverse data types and formats involved; or joining different content, where the semantics of data in a particular application are diverging because these have been defined a long time ago. Problems especially accrue in automatic computation such as in data mining[6], when semantic inconsistencies in data may lead to statistical artefacts that cause misinterpretations.

Particular issues of knowledge and knowledge management in e-participation:

- e-participation ontologies[7] and standards enabling data interchange
- Providing knowledge enriched information and intelligent e-participation tools for the stakeholders in e-participation (citizens, administrators, elected representatives, community groups, NGO's, private enterprises etc.)[8].

---

[6] Data mining is treated in more detail in section 4.4.4

[7] Ontologies are being treated in more detail in Deliverable 5.2.3, standards and interoperability needs will be investigated in phase II of Demo-net.

[8] A discussion on the stakeholders of e-participation is provided in deliverable D 5.1.

### 1.3.2 Dissemination of Knowledge

Disseminating knowledge means orientation towards the addressee; information on actual and potential users is necessary for matching of supply and demand. A pro-active approach is needed to ensure a sufficient flow from sources to demand. To promote this idea, the somehow placard-style notion of a knowledge pump[9] has been invented by Borghoff and Pareschi (1998). For data in e-participation areas some additional questions arise, especially supporting the dissemination of knowledge from governments to citizens and from citizens to governments:

- How to prepare knowledge for public display (cf. also visualization below)?
- How to ensure data protection?
- How to secure inspection rights of citizen?
- How to balance content to have it comprehensive and readable?
- How best present geographical databases and environmental information?
- How to handle information delivered to the system by external stakeholders?
- How to ensure feedback on the usefulness of their input to participants?

A variety of options and conditions have to be considered: Potential appearances and forms of design, specific needs of addressees, and legal framing conditions to name but a few. Some parameters that shape a concrete design are:

- Tradeoffs between push- and pull-approaches to knowledge dissemination
- Choice and suitability of the access channel[10]
- Diverse organizational forms and physical settings of demand (office, kiosk, home)
- Balance of human and software mediators/knowledge bearers
- Routing of offer/demand according to competencies of the targeted stakeholder
- Intricacies of the subject matters (legal norms and decisions)
- Translation from administrative/legal jargon to everyday world and vice versa[11].

### 1.3.3 Visualization of Knowledge

The graphical, textual or animated visualization of knowledge is a key feature for communicating content and processes to the users. Especially in e-participation contexts, this is a critical factor: many heterogeneous users with different system platforms and devices should be able to comprehend the information and knowledge contents of e-participation processes in an effective and convenient way. Several challenges of knowledge visualization have to be addressed:

- *Viewing information and knowledge:* the basic requirement for comprehending knowledge is to communicate it through a proper interface. This refers to textual visualization as well as graphics, animations, audio, and interactive comprehension.
- *Behavioural aspects of knowledge visualization:* Knowledge management refers to a large amount to interaction among humans and machines in terms of communicating codified knowledge to a user who elaborates, internalizes and

---

[9] See section 0 for more details.

[10] Devices and access channels for e-participation are explored in more detail in sub-deliverable 5.2.6

[11] A specific requirement for e-participation, since the general public and the expert Government employees and politicians may speak a diverse technical jargon.

contributes knowledge back to the technical system. This cognitive model of knowledge comprehension and knowledge externalization is crucial for successful integration of KM in system applications such as in e-participation.

- *Usability:* A wide range of usability criteria should be implemented for securing smooth work. Examples of such criteria include: malleability of mechanisms for adoption to personal preferences; indicators reminding of the basic status (such as what, where, how) when managing subtasks simultaneously; semantic conformity of notational primitives corresponding to the context of usage.

### 1.3.4 Knowledge Delivered to the Diverse Collaboration and Cooperation Contexts

*C*ollaboration and cooperation among actors are essential in nearly every e-participation scenario. Two basic requests are treated here:

- *Blending different modes of cooperation:* There is need for a wide spectrum of possibilities, depending on whether strictly structured cooperation (workflow) is involved or more informal collaborative modes (message exchange, discussion fora, meeting rooms, blogs, wikis, etc.). A smooth transition between both modes and the inclusion of auxiliary functions such as filtering and calendaring is mandatory.
- *Integrating the knowledge processes into the conventional e-participation solutions*: A key requirement in e-participation is to understand and design the participatory processes thereby defining the proper technical means to support the purpose of participation. Advanced support requires mapping the knowledge process with the participation process in proper way.

## 1.4 Outline of the Report

In order to elaborate the technologies and concepts of knowledge management and knowledge engineering in the settings of e-government and e-participation, we start with a general discussion of what kind of knowledge is to be dealt with in these environments (cf. chapter 2). Chapter 3 deals with the management processes of knowledge. Consequently, concepts of knowledge management are being introduced, ranging from the core KM processes of Borghoff and Pareschi (1998), the Spiral Model of Nonaka and Takeuchi (1995), recent concepts of KM as proposed e.g. by the UN in 2003, and leading to the concept of organizational memories. In Chapter 4, we introduce tools and technologies supporting KM. Among them, knowledge repositories, knowledge structuring concepts, data storage concepts, knowledge extraction concepts (KDD, OLAP, Data Mining), and more advanced concepts of case-based reasoning are being introduced. Chapter 5 discusses future scenarios of applying KM concepts and technologies in different e-participation areas, and embodied in distinct e-participation tools and processes. We conclude with a reflection of future research needs in order to exploit the potential of KM in e-participation contexts.

# 2 Types of Information and Knowledge in e-Government and e-Participation

A starting point for discussion is to get an understanding of what kind of knowledge has to be dealt with in e-participation.

In e-government contexts, we have made an attempt to develop a comprehensive concept for understanding the various kinds of knowledge in the public sector. We investigate first the location of knowledge by sketching a holistic view. Next, a three layers concept of categorizing such knowledge in a more abstract way is introduced. Afterwards, three perspectives on knowledge in e-government are introduced, and finally some attempt is made to categorize knowledge in e-participation contexts.

## 2.1 Promoting a Holistic View on Distributed Knowledge

For locating and redistributing the huge knowledge, comprehension has to include different perspectives and degrees of abstraction. In Figure 1, we introduce a systemic view (derived from the SHEL model (Edwards, 1988)) to demonstrate that any productive process is always performed by a specific combination of system resources.

The SHEL model supports the understanding that knowledge is perceived as the basic system resource of a productive system which is constructed within an organisation. With a productive system, any kind of system where services or goods are produced is meant, be these results information or physical products. Hence, processes (tasks in other disciplines (Diaper, 1989; Kirwan and Ainsworth, 1992; van der Veer et al, 1996) are the core point of interest from many perspectives (e.g. business economics, workflow, organisational management, system development, HCI, etc.). Also, Activity Theory (Engeström et al, 1999; Nardi, 1996) with its focus on activities is in line with that.

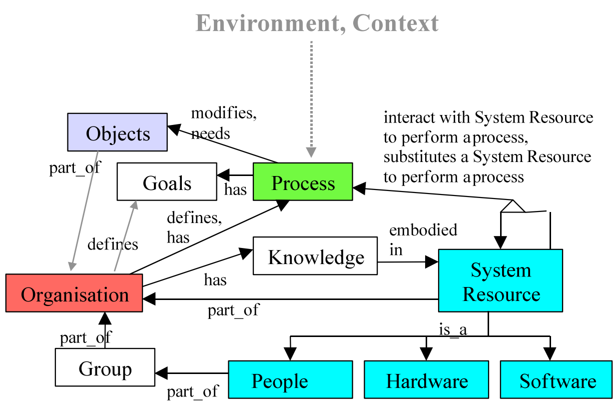



**Figure 1: Systemic view on a productive system**

Processes are always driven by a specific goal to be reached that is defined from the organisation, i.e. from a group of people belonging to the organisation. These people have discretionary power to decide on which strategic, tactical or operative goals, and how these should be achieved. Further, within this holistic view, also the objects serving as input, throughput or output for a process are strictly related to the process and with the organisation and, hence, have to be integrated. The organisational construct of people, other system resources, processes and objects is further embedded in a given system environment which may have a certain impact on how e.g. a process is performed or how a system resource (physical or abstract artefact) is shaped.

From a systemic perspective, the system resources (people and artefacts) are the crucial tools and actors within an organisation reflecting the knowledge that belongs to the organisation. According to this systemic view, in performing a productive process, knowledge flows among the components and objects being part of the process. Here, we recognise that knowledge is distributed among many components and people being part of a productive system which brings us to introduce a framework to better understand the forms of knowledge distributed among distinct system resources.

## 2.2 A Framework to Understand the Different Forms of Knowledge

In an attempt to develop a comprehensive concept for understanding the various kinds of knowledge in the public sector, Wimmer and Traunmüller (2001) suggest a three-layer concept as depicted in Figure 2. This systemic view on a productive system such as e-government reflects many knowledge aspects which must be combined to reach the intended goals.

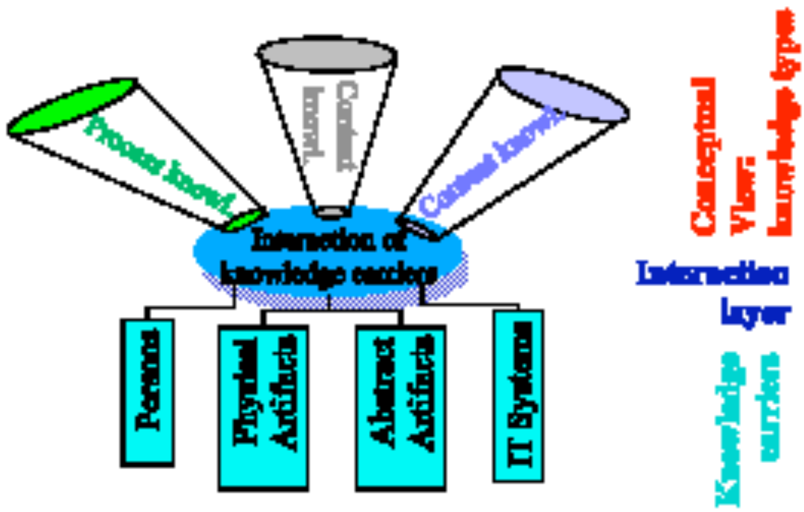



**Figure 2: Three-layers concept of knowledge identification**

The model describes a knowledge transformation process and is based on the following:

1. The conceptual layer (knowledge types) provides a general distinction on three different categories of knowledge: process, context and content knowledge.

2. The interaction layer discusses the combination of different knowledge categories according to the systemic view on a productive process. According to Cole and Engeström (1993), knowledge is embodied in different types of artefacts that interact to perform a specific process.
3. The representation layer treats aspects of knowledge reflected in different knowledge carriers[12] and how these components are interwoven (Wimmer, 2000). Such knowledge carriers may be ICT systems, where knowledge is stored as information and embodied in workflow systems, information systems, specific application systems, databases, controlling systems, management information systems, etc. Knowledge is created in people's minds and as such needs to be transformed into information that can be stored and shared. From the people's point of view the third layer of Figure 2 comprises specific know-how, expertise, skills intuition and information. Thirdly knowledge can be stored in physical and abstract artefacts belonging to the system. Putting focus on material and ideal artefacts, manuals, equipment, formal rules and laws etc. are reflected (cf. Edwards, 1979).

In a holistic view, the layered concept can be considered as a three-dimensional space addressing on one axes the three different layers, on the second axes the different types of knowledge and on the third one the different knowledge stakeholders: people, ICT systems and other material and ideal artefacts belong to the productive system.

It is the interchange between the conceptual layer of knowledge types and the knowledge representations that attracts our attention, because an adequate mapping of knowledge concepts and flow of knowledge is a prerequisite for attaining a proper functionality of a knowledge-enhanced system. Key traits of knowledge types and knowledge stakeholders mark and emboss the interaction in the interaction layer. In that way, a determining influence is exerted with broad effect: on the particular tool, on the interoperability of tools and on the way tools are used in co-operation. A further strong influence between the layers is that a highly achieved functionality has always a history of adapting and tuning, i.e. a dynamic evolution of knowledge is an important matter.

## 2.3 The Knowledge Part in Administrative Processes

Both, the role of the law and special features of information and knowledge have to be dealt within the light of the core task of public administration, i.e. decision-making in non-routine cases (see Lenk et al, 2002). Such decision-making exhibits a big variety, according to the complexity and to the circumstances of both the policies at stake and the individual case. This does not mean that it is not possible to construct a basic process model of administrative action. Lenk and Traunmueller (1999, p. 55ff) have developed a process model proper to public administration, which tries to combine two views: On one hand, processes can be seen as production processes, which obviously is a perspective owing much to Workflow Management Systems (WFMS)[13]. On the other hand, there is a decision-making view. It draws on two sources: the classical account of administrative behaviour by Herbert Simon (Simon, 1957), and administrative procedure as it is

[12] See discussion in section 4.1.1

[13] See also discussion in section 4.1.2

conceptualised in procedural legislation, the administrative acts (e.g. the US (Federal) Administrative Procedure Act or the German Verwaltungsverfahrensgesetze).

### 2.3.1 Stages in Administration's Decision Process

Combining the production and decision-making process views in the pubic sector, the following process elements can be identified (Lenk and Traunmueller, 1999):

1. Observing
2. Policy formulation (planning, taking decisions)
3. Policy implementation (executing).

Taking a closer look, one can derive a circular set of administrative and participatory activities:

- *Observation and Information*: Information is collected on citizens, society, market, behavior, the environment, etc. Such observations can be made for specific purposes (e.g. by the police authority) or for general planning purposes.
- *Substantiating facts*: The material gained from such observations is evaluated in the light of legal and policy premises. In this way, a "case" is constituted.
- *Decision to act*: When enough material is collected and combined with the facts, a decision needs to be taken for action and intervention (either in policy implementation or in the policy formulation).
- *Intervention*: In a typical administrative act, the results of the decision-making process are simply communicated to the addressee. But physical-technical actions can occur as well, such as e.g. arresting a person, paying a sum of money, setting up of roadblocks, closure of a bridge. In rule-making acts, a new law may be drafted and enforced.
- *Execution*: If some addressees do not comply with the orders, an execution of the order may become necessary. A common example is the forcible way of tax collection.
- *Evaluation*: In the last step it has to be checked whether the action taken had the intended effect concerning the influence on the society. The results of this evaluation should be used for improving both administrative decision-making (policy implementation) and the rules guiding it (i.e. policy formulation).

A closer consideration of the stages listed above underlines the prominent role of information and knowledge. Physical action plays only a minor role. Moreover, the information entering government action has a pronounced reflexive character. Every observation made during a process will not only be introduced in the puzzle game of reaching a decision or an agreement, but it will at the same time change the entire texture of the contextual knowledge of the individuals working on the case and hence also of the organisation itself. This creates a reflexive memory of the process which is important for the management of this process. Moreover, the result will not be lost with the process coming to an end. It should contribute to the gradual constitution of a "domain knowledge" - *Dienstwissen*, as Max Weber called it in his studies on bureaucracy. This is perhaps the most important source of government knowledge to be dealt with in respect to knowledge management in e-government and e-participation.

### 2.3.2 Limits of process models

Process models often abstract from many aspects, which are quite important in administrative practice:

- Processes are often not isolated but part of a standing relationship between an agency and citizens.
- Processes involve co-operation among actors and with stakeholders in the field to be regulated. Such a co-operation often transgresses organizational boundaries.
- Many processes are distributed onto several agencies or institutions for reasons of balancing the use of public power. So, the agencies involved (e.g. police, prosecutors, judges and probation officers) have their own way of thinking and action and will follow a logic of their own.
- The knowledge aspect is somewhat under-addressed in process thinking. Many observing and information-gathering activities take place without producing tangible results. The collected pieces of information may not be used directly but these may contribute to organizational learning within an agency. Conversely, actions can be taken without any external information coming in.

Especially the interdependencies of information and knowledge among the administrative system and the participatory processes are not well understood yet. E-government and e-participation are considered each in isolation, whereas both depend on the other especially in the provision of knowledge and substantiating facts.

### 2.3.3 Types of processes in the public sector

In the field of operational administrative action, a huge variety of different processes can be encountered. A tension exists between fully structured production processes and complex decision processes. Most actual processes fall in between these two extremes. Yet, numerous cases exist where at the moment when a process starts, it is far from clear how complex it will eventually become. Since the later stages cannot be anticipated, such processes will loose much of their quality if they are subjected to strictly defined workflow management systems.

Although many ways exist in which different agencies make their distinct interventions into the social fabric (using regulations, services, transfers etc.), we use a very coarse-grained distinction of administrative processes (Wimmer et al, 2001):

1. Recurrent and well-structured processes: processes which are legally controlled in a strict way and formalised to a large extent, and which are characterised by a continuous repetition of mostly homogeneous operation steps that give only minor discretion to the persons in charge of each step.
2. Processing of cases: individualised decision-making which are characterised by a high degree of communication with the citizen and by the need to take special situations into account.
3. Negotiation processes: processes enforcing the law and co-operating with their environment, with a reasonable amount of discretionary power of the decision-maker
4. Weakly structured processes in the field of policy-making including all phases of the policy process, and democratic deliberation. Examples are: bills of parliament, answers to parliamentary inquiries or complex political decisions.

This distinction again emphasises the difference between “production processes” where almost no choices are to be made and which can often be fully automated, and decision processes which can at best be partly standardised.

Most important to policy-making activities are meetings. In order to reach adequate support environments, one has to blend conventional data and decision support with collaborative functions. In this way, a set of highly-modular components may be established to handle particular tasks, which provides all kinds of knowledge as required and in adequate form.

## 2.4 Exemplifying Knowledge Needs in e-Participation via the Process of Democratic Deliberation

Knowledge sharing platforms are of particular importance to democratic decision-making. Instead of dealing with the entire range of policy decisions, we will concentrate here on democratic deliberation in a local environment. Information systems can support and promote citizen participation in public planning as well as in mediation processes in different ways:

- by providing information on a problem and its background also in interactive and multimedia form including "Virtual Reality" techniques;
- by supporting communication processes in different modes and between spatially distant persons ("Telecooperation");
- by structuring debates;
- by directly supporting decision processes, e.g. through voting.

Concerning information provision, the concept of citizen information has to be enlarged to some extent: Basic information needed in executing democratic rights could quite well be provided via the Internet. However, explicit policies are needed to do so in a way which effectively reaches all citizens concerned. One way would be to incorporate it – or provide references to it – in platforms which initiators of a democratic decision process could call upon and make available to all those who want to have a say in the process. We will not deepen this argument here. Rather, for what follows, we concentrate on the second and the third point which are of particular interest for democratic deliberation.

Besides easing access to information, the Internet can help to establish platforms on which democratic debates can happen. The idea of "Issue-Based Information Systems" (IBIS) has been developed by Horst Rittel in 1970, and it has been rediscovered in the late eighties. Such systems can be used to structure debates on controversial issues. The structuring of information is particularly useful in the early stages of the policy process, i.e. for identifying problems and elaborating solutions (Lenk, 1999).

If such platforms are developed, we will soon see innovative systems in which elements of the potential for the supply of information, for the support of communication and for the support of decision-making processes are combined. The design of comprehensive systems for the support of citizen participation can profit especially from developments in the field of CSCW. Access to information, and communication support for the participation of citizens are a precondition for the efficiency of the approach. More complex issues arise at a different stage: structuring debates, allocating rights to raise issues and comment on them. So far, only little experience has been gathered regarding the use of such combined systems which support the structuring of decision-making beyond the level of information supply.

An example of new types of participation-enhancing information systems is provided by the GEOMED project[14], which was funded by the European Union. It comprises three key issues:

- *Information services*: A wide variety of planning tasks require access to geographical information which is typically represented in maps. Thus, the accessibility of geographical information in heterogeneous GIS systems is established over the Internet. Users are able to access, view and manipulate maps embedded in HTML pages from ordinary WWW client PCs. Other information present in the WWW is accessible, too.
- *Documentation services*: A "shared workspace" is established for the elaboration, storage, and retrieval of documents, for messages related to particular geographical planning projects and for contributions to the debates of a planning process or a mediation procedure. So, in a very convenient way, ordinary users can add information to the documents available.
- *Mediation services* provide assistance to human mediators of a round table. In order to structure interventions, an Issue-Based Information System (IBIS) is provided.

Depending on their design, information systems embody and represent structures which may help to overcome well-known organisational problems related to democracy: the bringing together of like-minded people, the structuring of debates and the embedding of rules which give a better say to people who have difficulty in expressing themselves. Local democracy like any democracy involving more than a small group of people ("seven plus/minus two") is clearly a problem of organisation. Information technologies, beyond their function of supporting telecommunication and providing access to stored information, are technologies of organisation. It is important, therefore, that we gain a clearer impression of the manifold forms which electronic support of citizen participation could take.

[14] Geographical Mediation Systems, EC project, see http://arti.vub.ac.be/geomed/geomed.html

# 3 Methodologies for Knowledge Management

## 3.1 Definitions for Basic Understanding

Within the knowledge society of the 21st century, industrialised countries perceived a tremendous shift towards global markets where knowledge has become one of the most important assets of modern organisations. This change strongly impacts the nature of work and of workforce itself. The most successful enterprises in this new "world" are those who recognised the importance of knowledge as a resource and who investigate in the management thereof.

Knowledge Management is a field of research and application which is not yet illuminated in full. Many aspects shaping the domain and impacting its success are still to be uncovered, as can also be recognized from above discussions. Furthermore, many experiences and lessons learnt still have to be collected from practical implementations. We need to be aware of the fact that KM is an instrument that requires proper adaptation to the situational context and to the individual facilities of an organisation. Hence, no “one and only” concept of KM may exist, but a diversity of dynamic, lively concepts shapes the field.

It has to be recognised that Knowledge Management may not solely be treated and satisfactorily be implemented within a technical discipline. Instead, KM touches as well psychological, social, organisational and communication issues. Knowledge Management is a multidisciplinary field that calls for know-how and expertise from distinct research fields and applications. Among the research fields to be integrated are knowledge engineering, human factors, communications engineering, psychology, sociology, computer science, artificial intelligence, organisations engineering and management sciences.

Before introducing some state-of-the-art Knowledge Management methodologies and concepts, general definitions of this field are provided, which are settled in the IEEE Standards 610.12 and 1233:

“A **methodology** is a comprehensive, integrated series of techniques or methods creating a general systems theory of how a class of thought-intensive work ought to be performed”. A **method** is a set of “orderly process or procedure used in the engineering of a product or performing a service”. A **technique** is “a technical and managerial procedure used to achieve a given objective”. A **process** is a “function that must be performed in the software life cycle”. A process is composed of activities. An **activity** is “a constituent task of a process”. A **task** “is a well defined work assignment for one or more project members. Related tasks are usually grouped to form activities”.

In the following, a concept of KM processes of Probst et al (2003) is introduced to demonstrate the general steps of knowledge management. Subsequently, a number of methodologies for knowledge engineering (KE) are being introduced:

- The Know-net Method by Mentzas et al (2002),
- The CommonKADS developed by Schreiber et al 1999,
- The DÉCOR Methodology as a follow-up of CommonKADS and IDEF5 (KBSI,

1994),

- The method of knowledge mapping introduced by Plumley (2003), Wexler (2001) or Eppler (cited in Glassey, 2004),
- The Knowledge Management Toolkit as a 10-step approach to capturing and disseminating knowledge introduced in Tiwana (2000),
- The United Nations KM Methodology introduced in (United Nations, 2003),
- The concept of Communities of Practices and CoP Practitioner's Guide described in Navsea (2001), and

Thereafter, the Knowledge Pump elaborated in Borghoff and Pareschi, (1998) is being introduced as a concept for knowledge dissemination.

The spiral model (or SECI model) of Nonaka and Takeuchi (1995) is introduced as a concept for knowledge creation.

Many more approaches to Knowledge Management and Knowledge Engineering exist. However, the set of concepts and methodologies introduced should suffice to give a comprehensive insight into the aims and aspects of KM. As one will recognize when reading through the descriptions, many of them have much in common. The difference is often just the specific focus of methodology, a certain technique applied etc. It can also be stated that many of the concepts may be applied in parallel in order to counterbalance certain weaknesses of methods and to balance these with the strengths of others.

Overall, the knowledge management concepts provide an essential basis to comprehensively understand the processes of managing knowledge, independent of where the concepts are being applied (i.e. public or private sector, e-government or e-participation, etc.).

## 3.2 Concepts describing KM processes

### 3.2.1 Knowledge Management Processes of Probst et al.

In the era of the Information Society, information and knowledge have become key factors of success. Classical concepts of knowledge management as e.g. defined by Probst et al (2003) define a number of core processes of knowledge management as shown in Figure 3. The authors distinguish among the core KM processes and the control loop processes. The KM process should start with the definition of the knowledge objectives, which should be the basis for the core processes of KM. These can also be performed in sequence (main cycle), however, any sequence can be executed and phases may be left out, depending on the specific process needs. Finally, knowledge needs to be assessed in terms of actuality, relevance, completeness, etc. Consequently, Probst et al (2003) introduce a Knowledge Evaluation process which closes the cycle of the core KM processes and the control loop.

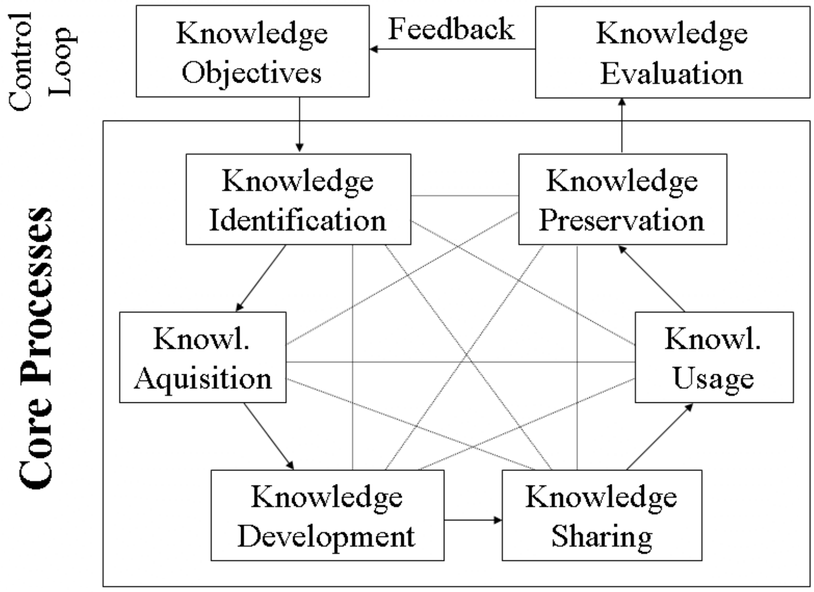



**Figure 3: Modules of knowledge management according to Probst et al (2003)**

## 3.3 Methodologies for Knowledge Engineering

### 3.3.1 Know-Net Method

The Know-Net method (Mentzas et al, 2002)[15] is a phased approach to implement Knowledge Management in an organisational setting. It consists of the following phases:

- Awareness about the benefits of knowledge management and its relationships to strategic as well as operational and day-to-day issues in the corporate environment.
- Stage I - Knowledge Management Strategic Planning phase: In this phase an organisation determines the vision and readiness for a knowledge management initiative and the scope and feasibility of the project.
- Stage II - Develop the KM Project: This is the phase in which an organisation transforms itself to a knowledge intensive company based on the company-specific KM value propositions derived in Stage I.
- Stage III - Operate the KM Organization: In this phase an organisation rolls-out a company-wide implementation plan with a holistic approach to KM.
- Measurement of the level of leveraging of knowledge assets with a KM effort. This phase aims to provide consistent support for measuring the creation, sharing and use of knowledge assets within the company.

---

[15] http://www.know-net.org/

- Training of both the knowledge workers to the new processes and technologies as well as of the staff to take up new knowledge-related roles (e.g. Chief Knowledge Officers, knowledge analysts).

### 3.3.2 CommonKADS

The CommonKADS methodology (see Schreiber et al, 1999) offers a structured approach towards the development of knowledge-based systems. The development of a system in CommonKADS entails constructing a set of models of problem solving behaviour in a concrete organisation or application context. Thereby, CommonKADS engineers a predefined set of models, each of them focusing on a limited aspect, but together providing a comprehensive view:

- *Organisation model*: The organisation model supports the analysis of the major features of an organisation in order to discover problems and opportunities for knowledge systems, to establish their feasibility and to assess the impacts on the organisation of intended knowledge actions.
- *Task model*: Tasks are the relevant subparts of a business process. The task model analyses the global task layout, its inputs and outputs, preconditions and performance criteria, as well as needed resources and competencies.
- *Agent model*: Agents are executors of a task. An agent can be human, an information system, or any other entity capable of carrying out a task (cf. the knowledge carriers in Figure 2). The agent model describes the characteristics of agents, in particular their competencies, authority to act and constraints in this respect. Furthermore it lists the communication links between agents in carrying out a task.
- *Knowledge model*: The purpose of the knowledge model is to explicate in detail the types and structures of the knowledge used in performing a task. It provides an implementation-independent description of the role that different knowledge components play in problem-solving (in a way that is understandable for humans – cf. as well the sub-deliverable 5.2.3 on ontologies). This makes the knowledge model an important vehicle for communication with experts and users about the problem-solving aspects of a knowledge system, during both development and system execution.
- *Communication model*: Since several agents may be involved in a task, it is important to model the communicative transactions between the agents involved. This is done by the communication model in a conceptual and implementation-independent way (just as with the knowledge model).
- *Design model*: The above CommonKADS models together can be seen as constituting the requirements specification for the knowledge system, broken down in different aspects. Based on these requirements, the design model gives the technical system specification in terms of architecture, implementation platform, software modules, representational constructs, and computational mechanisms needed to implement the functions laid down in the knowledge and communication models.

The organization models, task models, and agent models describe the organisational environment and the corresponding critical success factors for a knowledge system. The knowledge and communication models yield the conceptual description of problem solving functions and data that are to be handled and delivered by a knowledge system. The design model converts this into a technical specification that is the basis for software system implementation. However, not always do all models have to be constructed.

CommonKADS[16] distinguishes two basic types of processes:

- Knowledge processes: address the handling of the already set-up KM solution (cf. as well the core processes identified in the model of Probst et al, 2003, - see 3.2.1)
- Knowledge meta processes: address aspects of introducing a new KM solution into an enterprise as well as maintaining it (cf. as well the control loop processes identified in the model of Probst et al, 2003, - see 3.2.1).

From a knowledge creating perspective, knowledge items are integrated into a KM solution via two key knowledge processes, which are supported by some tools based e.g. on ontology:

a) The creation of knowledge (knowledge items are being inserted into document templates, filled in on-line to enable Knowledge import, etc.;

b) knowledge capture, i.e. knowledge is extracted from knowledge items based on their context, for example one tool called OntoAnnotate (Staab et al, 2001) provides building objects and describing them by attributes and their relations to other objects. These identifiers are collected from web pages, tables or text documents. Along this action process, metadata is created.

### 3.3.3 DECOR Methodology

By amalgamating some elements from the CommonKADS and the IDEF5 (see KBSI 1994) methods, the DECOR Methodology (see Abecker, 2003) provides the methodological guidance for running a Business Process oriented Knowledge Management (BPOKM) project and comprises the following steps:

- **Step 1: Business Process Identification**. This activity involves the identification of the most important business process/-es to be supported in the BPOKM project. The business processes with high process complexity and stronger knowledge intensity are those that have a high KM support potential. If they are also central to the company's value creation, they become candidates for a BPOKM project.
- **Step 2: Business Process Analysis**. This activity involves a general description of the selected business process/-es in terms of
    a. tasks constituting the business process;
    b. roles involved; and
    c. key people and source material.
- **Step 3: Task Analysis**. This activity involves a more detailed description of the individual tasks e.g. their input and output objects, control relations between tasks, etc. Moreover, every task in the process is assessed through its contribution to the core activities of Knowledge Management, i.e. generate, store, distribute and apply knowledge. This could lead to characterising some tasks as knowledge-related and/or knowledge intensive.

[16] http://www.sics.se/ktm/projects/kads.html

- **Step 4: Business Process Design**. This activity involves the modelling of the to-be business process using a graphical tool. The output of this step is a business process model enhanced with Knowledge Management tasks for the knowledge flow in the business process.
- **Step 5: Ontology Creation**. In the DECOR Methodology, ontologies are used as a complex model for structuring indices used to describe document content. The Ontology Creation of the DECOR methodology builds upon three central concepts of an ontology[17]:
    d. *Kinds* (an objective category of objects sharing a set of properties),
    e. *Characteristics* (the properties belonging to a Kind) and
    f. *Relations* (the sorts of general features that Kinds exhibit jointly rather than individually).
- **Step 6: Ontology Refinement.** This activity involves the refinement and validation of the developed ontology.

### 3.3.4 Knowledge Management Toolkit

The Knowledge Management Toolkit (Tiwana, 2000) is a four-phase, 10-step approach for capturing and disseminating knowledge, and measuring the impact of an organisation's efforts in the form of Return On Investment (ROI) and other performance metrics. The phases and steps outlined are:

- Phase I: Infrastructure evaluation. This phase is accomplished by the following steps:
    - (Step 1) analysing the organisation's existing infrastructure (the "as-is" part) and
    - (Step 2) aligning the organisation's KM and business strategies (to ensure that one is not devising solutions to "non-problems" and/or one's strategy addresses real business requirements)
- Phase II: KM System Analysis, Design and Development. This second phase encompasses the following steps:
    - (Step 3) designing the knowledge management architecture and integrating existing infrastructure,
    - (Step 4) auditing and analysing existing knowledge,
    - (Step 5) building the knowledge management team,
    - (Step 6) creating the knowledge management blueprint and
    - (Step 7) developing the knowledge management system.
- Phase III: System Deployment. This phase entails the following steps
    - (Step 8) deploying with "results-driven incrementation" methodology and
    - (Step 9) change management and cultural considerations.

---

[17] For a more detailed discussion on ontologies, the reader is referred to sub-deliverable 5.2.3.

- Phase IV: Infrastructural evaluation. The final phase is performed with the last step in the approach:
  - (Step 10) measuring results of knowledge management, devising ROI metrics and evaluating system performance.

### 3.3.5 United Nations Knowledge Management Methodology

The United Nations KM Methodology (see United Nations, 2003) proposes four phases in order to implement any knowledge management initiative. These phases involve the initiation of four projects, namely:

- Phase 1: Knowledge vision and strategy. In order to develop a knowledge vision and strategy, it is important for organizations to identify clearly the direction in which the organization is headed in terms of knowledge, as well as the reason for pursuing that particular course and the means by which it will be pursued. Knowledge management investments become more convincing and tangible once such issues are identified.
- Phase 2: Knowledge architecture and content. The purpose of this phase is to organize and catalogue knowledge into taxonomies, select the technological platforms to be used in order to store knowledge and, subsequently, determine the software tools that will make it possible to add content to and retrieve it from repositories. A knowledge validation procedure needs also to be designed in order to ensure the reliability and coherence of the entire KM system.
- Phase 3: Knowledge infrastructure. The present phase aims to define roles for knowledge providers and users in order to generate and facilitate communities of practice or other knowledge management initiatives. Description of knowledge need to be clearly articulated and documented and properly measured and rewarded.
- Phase 4: Knowledge culture. The aim of this phase is to motivate users to supply and submit their own knowledge voluntarily while also using the knowledge of others. The main activities involved in this phase encompass training programmes, two-way communication and internal marketing tasks, as well as reinforcement programmes such as measurement, reward and recognition. It is essential that such activities are embedded in organizational processes and technologies.

The knowledge vision and strategy phase forms the basis of the remaining phases and must be carried out first although the other three phases can be initiated in parallel with one another.

### 3.3.6 Community of Practice Practitioner's Guide

The Community of Practice Practitioner's Guide is designed as a tool for establishing and sustaining communities of practice (see Navsea, 2001).

Communities of Practice (CoPs) provide a good means for enabling organizations to share knowledge enterprise-wide. Organizations are strengthened through an improved network of contacts and better results. Individuals benefit through peer-group recognition and continuous learning. Communities of Practice promote and strengthen the effective use of knowledge in an organisation comprising therefore one of the cornerstones of successful Knowledge Management.

The Community of Practice Practitioner's Guide proposes a four step approach to creating Communities of Practice:

Step 1: How to create a community.

Key tasks in this step include conducting core planning, preparing for initial community workshop, hosting initial community workshop, checking community progress and building Community Experience Locator. Expected outcomes of this first step are a clear understanding of the roles and responsibilities involved in a community, a community identity and a foundation for community activities.

Step 2: Creating knowledge.

The purpose of this step is to provide the community under development with suggested tools to help it create, capture, and share its knowledge. A tool, in this instance, does not refer to automated systems for transferring information. Rather, these tools are techniques or forums for thinking, e.g., techniques for generating ideas and building relationships, or forums that promote knowledge flow and transfer. An indicative list of tools that can be employed towards that direction are ad-hoc sessions, roadmaps to generating new knowledge (problem-solving and brainstorming), learning history, interviews, action learning, learn from others, guest speakers, relationship building and systems thinking. Expected outcomes of this second step are designs for community forums for critical thinking and knowledge flow and practical techniques for knowledge creation, capturing, and transfer.

Step 3: Building the knowledge base.

The purpose of this step is to provide a framework for building the community's knowledgebase. Key Tasks in this step include:

- Requirements: Map identified collaborative tool functions to business requirements.
- Inventory: Define knowledge assets in a business process context and identify whether created by the community or borrowed from other business owners.
- Taxonomy: Develop a business context classification structure for organizing inventory.
- Flow Model: Model as-is business processes based on the flow of inventory assets to and from customers. Focus on how assets are created and disseminated.
- Migrate: Provide necessary technical support to migrate inventory assets that exist in legacy repositories.
- Map: Identify owners of the inventory folders and designate life-cycle responsibility at a folders structure level.
- Asset Rules: Establish business rules for the use of the groupware to maintain consistency while performing business transactions.
- Transformation: Identify, in priority order, High Value – Low Risk business processes that provide the group with the highest value in terms of Customer Service, Cycle Time Reduction, and Total Ownership Cost.
- Training: Transformed business processes will be simulated in a training environment for user testing and acceptance.

- Help Desk: Enable a functional help desk specifically for community members.

Step 4: Sustaining Communities

This step refers to ways the community leadership can employ in order to assess progress, recognize the natural evolution of community interactions, recognize and reward both individual and community contributions, and continuously foster innovation and growth. Expected outcomes include process adjustments and continuous infusion of new knowledge.

Key tasks include:

- Assessing community progress
- Understanding community evolution
- Recognizing community contribution
- Sparking new knowledge creation and sharing

### 3.3.7 The learning cycle by Neches et al

Many experts have their own models for the KM process. In all these models, many activities and steps are often collateral, sometimes repeated and not in linear sequence. The last concept we herewith introduce is the one presented in Neches et al (1991), and elaborated further on in Raab and Studer (2001) (see Figure 4):

1. Knowledge creation: there are two extremes (one is very formal one and the other one is very informal knowledge)
2. Knowledge import
3. Knowledge capture: when knowledge items have been created, but not yet fully captured from their context, the next step is the capturing of their essential contents
4. Knowledge retrieval and access: many parts of retrieval and access from an ontology-based organizational memory are performed through some GUI.
5. Knowledge use: this part is most often neglected, because some KM systems find some relevant document and pretend that everything is done.

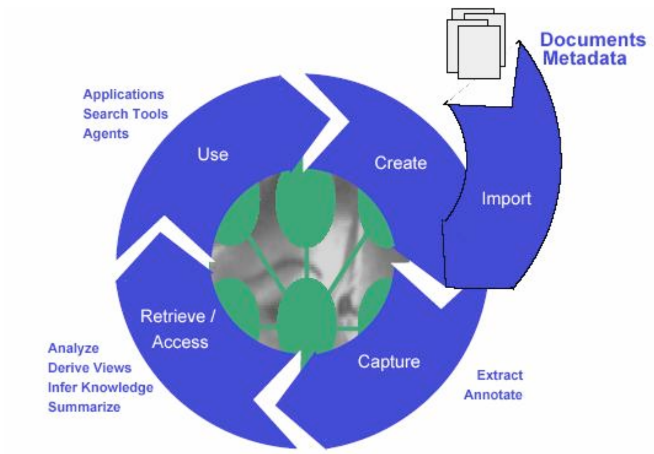



**Figure 4. Learning cycle by Neches et al (1991), graphically elaborated in Raab and Studer (2001)**

## 3.4 A Taxonomy of Knowledge Distribution and Collection

Several types of knowledge repositories do exist. According to Borghoff and Pareschi (1998), knowledge repositories can be divided into four groups (see Table 1). The division is based on the way how knowledge is collected and how it is distributed.

| | Passive collection | Active collection |
|---|---|---|
| Passive distribution | Knowledge Attic | Knowledge Sponge |
| Active distribution | Knowledge Publisher | Knowledge Pump |

**Table 1** Types of knowledge repositories

### Knowledge Attic

This is the simplest form of knowledge repository. The repository is used as an archive which can be consulted when needed. In practice this kind of repository is often the most feasible one. The advantage of this type is that it is not intrusive. It emphasises the bottom-up nature of community learning.

The content is submitted to the repository by community members (when someone finds some knowledge fragment interesting for him/her or when he/she finds it to be potentially useful for other community members, he/she can decide to insert the fragment into the repository). The content is retrieved by members of the community (when someone identifies a lack of knowledge he/she can try to find missing knowledge fragments within the repository). Obviously, in order for the repository to function well it requires a high discipline of the community members – only that knowledge can be retrieved from the repository, which was submitted by someone.

### Knowledge Sponge

Using knowledge sponge knowledge repository tries to address the step dedicated to collecting of knowledge. Community is actively trying to develop a more or less complete knowledge repository, while feeding knowledge into repository is not left to willingness of particular community members to share knowledge they possess or find with other community members. Knowledge acquisition process is clearly defined and supported (mostly by some tolls for automatic knowledge acquisition) in a way independent of particular community members. Use of knowledge is still left to the individual members.

### Knowledge Publisher

In this kind of knowledge repository, asserting knowledge into the repository is left to individual persons. The role of repository maintainers (being humans or not) is to analyse the incoming knowledge, combine the knowledge and forward it to community members. In this way, distribution of knowledge has a proactive character. It is not left to individuals to search for knowledge but the knowledge stored in the repository is filtered and appropriate knowledge fragments are pushed to appropriate persons. This can be done in the forms of briefings, newsletters, various kinds of messages, etc.

### Knowledge Pump

This is the most complex type of knowledge repository. It does neither leave knowledge collection nor knowledge distribution to willingness of individuals. Both knowledge input and output steps are supported by tools and firmly defined processes. In theory, it ensures that knowledge is fully exploited to improve community functioning.

## 3.5 Methodologies for Knowledge Creation

From the methodological point of view the knowledge creation processes have been studied in different contexts:

- Carl Breiter's knowledge building approach has emerged from cognitive studies in the educational context (Breiter and Scardamalia, 1993)
- Yrjö Engeström's theory of expansive learning is based on Activity Theory (AT) (Engeström, 1999, Cole and Engeström, 1993)
- Nonaka and Takeuchi's SECI model of organizational knowledge creation origins from the analysis of work in Japanese companies (Nonaka and Takeuchi, 1995)

According to Paavola and Hakkarainen (2006), these theories can be generalized into a trialogical way of thinking about learning (see Figure 5). Trialogical learning refers to the process where learners are collaboratively developing shared objects of activity (such as conceptual artefacts, practices, products) in systematic way. It concentrates on the interaction through these common objects (or artefacts) of activity, not just among people or within one's mind.

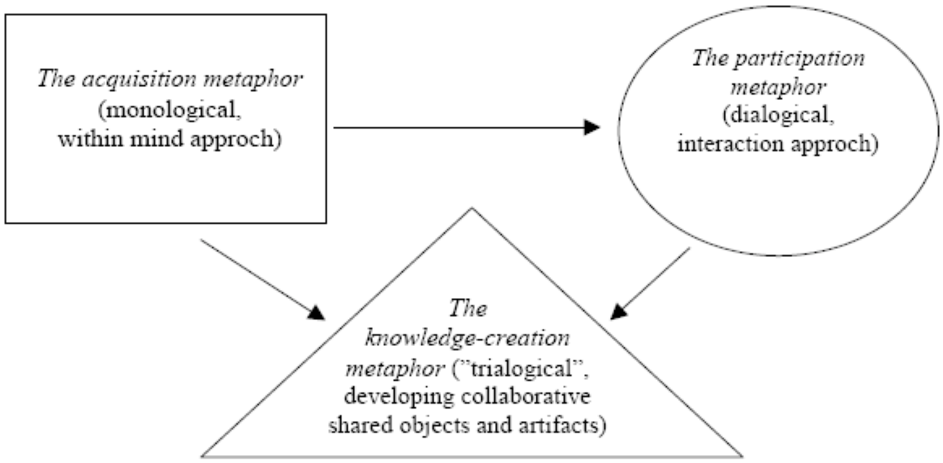



**Figure 5: Three metaphors of learning**

The general challenge for describing knowledge processes are such that: all participants should get an up-to-date understanding of the process; they should be able to make their individual and interconnected contributions to the process; and they should have the possibility to reflect on the course of process and practices of working together.

Knowledge creation processes can be described through two ways: models based on the workflow or models based on the process ontology.

### 3.5.1 Knowledge Spiral (SECI model) of Nonaka/Takeuchi

The creation of knowledge based on the SECI model of Nonaka and Takeuchi (1995) is a continuous process of dynamic interactions between tacit and explicit knowledge. The authors define the organizational knowledge creation process as a basic framework that contains two parts:

- Epistemological - shows that only individuals create knowledge. Therefore, organizational knowledge creation should be understood as a process that organizationally amplifies the knowledge created by individuals and crystallizes it.
- Ontological - interactions between tacit and explicit knowledge.

The aspect is often also addressed as a communication among the external world and the mental internal of individuals, as it is expressed in the spiral model (SECI model) of Nonaka and Takeuchi (1995) The two parts (externalization and internalization) are a base for defining four knowledge creation processes (see Figure 6):

- Socialization: transfers tacit knowledge from one person to tacit knowledge in another person (sympathize / empathize knowledge).
- Externalization: making tacit knowledge explicit among individuals within a group (conceptualize or articulate knowledge – i.e. dialogue).
- Combination: refers to the knowledge transfer once knowledge is explicit (connect, systematize and link knowledge).
- Internalization: understanding and absorbing explicit knowledge into tacit knowledge (operationalize, embody knowledge and put it into action).

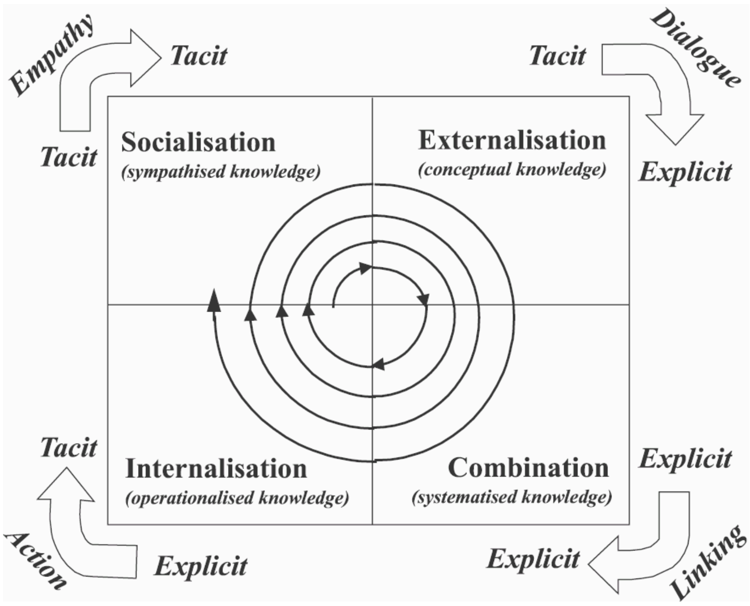



**Figure 6: The knowledge spiral (SECI model) according to Nonaka and Takeuchi (1995)**

The spiral becomes larger in scale as it moves up through organizational levels, and can trigger new spirals of knowledge creation.

# 4 KM Tools and Technologies

## 4.1 Comprehensive Applications

### 4.1.1 Knowledge repositories / Corporate memories

Knowledge is scattered throughout different stores (including heads of people) often in a non-organised way. As a result, it is common on one hand to find contradictory pieces of knowledge or not to find anything useful for a given case on the other hand. Recently, an increased number of people have begun to realise that it is possible to utilise ICT for knowledge organising and sharing. This idea targets building knowledge repositories accompanied by the (continuous) accumulation of knowledge on input side and distribution of relevant knowledge on output side.

Although a knowledge repository can serve everyone and can cover everything (WWW can serve as a representative of such kind of repository), a far more common situation is when a repository serves only a limited community of people – i.e. limiting its attention only to those knowledge fragments, which can be useful and interesting for the members of the community or organization. The community can be based on different principles, e.g. on regional principles (people living together in some geographic area) or on professional principle (people interested in some issue(s)). But the predominant type of community using its own knowledge repository is organisational – people belonging to some organisational (most commonly business) entity.

The aim for organisations is to obtain a tool enabling to catch knowledge in some part of the organisation and to make it available in another part of the organisation (or to catch it in some time and provide it in other time). But the aim is not to just conserve knowledge in a repository. On the contrary, the aim is to keep knowledge flow – distributing right knowledge to right people in right time enables to generate new knowledge which can be captured and distributed again, etc. In this way, knowledge repositories represent a very important means for organisations to function in a competitive environment of today, to adapt continuously to the external environment and to improve their competitive power.

From the technical point of view, knowledge repositories can have different forms and can be based on different technologies – from simple document repositories and database systems through different kinds of information systems up to complex corporate memories.

In a broader sense, knowledge repositories target storing any kind of knowledge. Explicit knowledge is stored in repositories in an explicit way and can be retrieved directly from the repositories. Tacit knowledge is stored in minds of those individuals who are organised around the repositories and the repositories can help to identify a person possessing the required piece of knowledge.

Figure 7 shows the decisive parts of a corporate memory (cf. Borghoff and Pareschi, 1998). It consists of three elements. The knowledge repositories and (data-)bases contain information in the form of documents and electronic bases. The knowledge cartography allows navigating through this information and mapping its contents. Finally the community of knowledge workers handles and uses the gained information. The corporate memory needs central processes that channel flow and transfer of knowledge. These are the processes of knowledge management in the proper sense (cf. parts of the KM processes introduced in chapter 3).

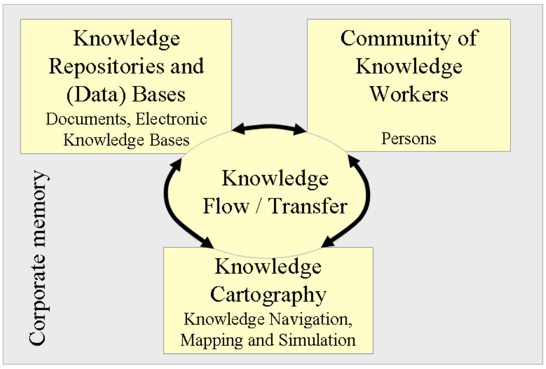



**Figure 7: The framework of a corporate memory**

#### 4.1.1.1 The Webocrat system as an example for a knowledge repository

The Webocrat system (Mach et al 2002) represents tool which can be used as a knowledge repository serving a community of people. The system was developed within the Webocracy project[18] (IST-1999-20364) as a web-based knowledge enabler solution combining a powerful knowledge management technology back-end with a fully customizable web user interface. Basically, it can be used as a knowledge attic system, but its functionality enables it to be used as a knowledge publisher as well, since it provides the possibility to automatically notify users on appearing new knowledge fragment in the repository users can be interested in.

From the point of view of functionality of the system it is possible to break down the system into several parts and/or modules. They can be represented in a layered sandwich-like structure which is depicted in Figure 8.

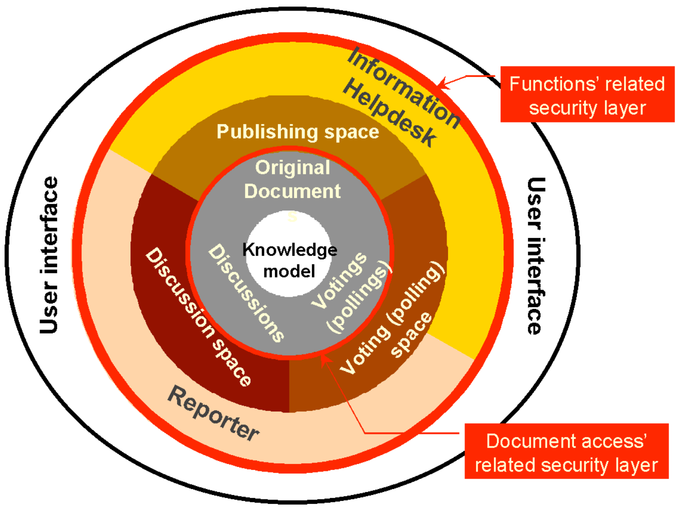



**Figure 8 Structure of the Webocrat system**

[18] http://www.webocrat.sk/webocrat/index.html

The central part of this structure is occupied by a knowledge model module. This system component contains one or more ontological domain models providing a conceptual model of a domain. The purpose of this component is to index all information/knowledge fragments stored in the system in order to describe the context of this information (in terms of domain specific concepts). The central position symbolises that the knowledge model is the core (heart) of the system – all parts of the system use this module in order to deal with knowledge stored in the system (both for organising it and accessing it).

Knowledge stored within the system has the form of documents of different types. Since three main document types are expected to be processed by the system, a document space can be divided into three subspaces – publishing space, discussion space, and opinion polling space. These areas contain published documents expected to be read by users, users' contributions to discussions on different topics of interest, and records of users' opinions about different issues, respectively.

Documents stored in these three document subspaces can be inter-connected with hyper-textual links – they can contain links to other documents – to documents stored in the same subspace, to documents located in another subspace, and to documents from outside of the system. Thus, documents within the system are organised using net-like structure. Moreover, documents located in these subspaces should contain links to elements of a domain model.

Since each document subspace expects different way of manipulating with documents, three system's modules are dedicated to them. Web content management module offers means to manage the publishing space. It enables to prepare documents in order to be published (e.g. to link them to elements of a domain model), to publish them, and to access them after they are published. Discussion space is managed by discussion forum module. The module enables users to contribute to discussions they are interested in and/or to read contributions submitted by other users. Opinion polling room module represents a tool for performing opinion polling on different topics. Users can express their opinions in the form of polling – selecting those alternatives they prefer.

In order to navigate among knowledge stored in the system in an easy and effective way, one more layer has been added to the system. This layer is focused on retrieving relevant knowledge from the system in various ways. It is represented by two modules, each enabling easy access to the stored knowledge in a different way. Citizens' information help-desk module is dedicated to search. It represents a search engine based on the indexing of stored documents. Its purpose is to find all those documents which match user's requirements expressed in the form of a query.

The other module performing information retrieval is the Reporter module. This module is dedicated to providing information of two types. The first type represents knowledge in an aggregated form. It enables to define and generate different reports concerning knowledge stored in the system. The other type is focused on providing particular documents – but unlike the Citizens' information help-desk module it is oriented on off-line mode of operation. It monitors content of the document space on behalf of the user and if knowledge the user may be interested in appears in the system, it sends an alert to him/her.

The upper layer of the presented functional structure of the system is represented by a user interface. It integrates functionality of all the modules accessible to a particular user into one coherent portal to the system and provides access to all functions of the system in a uniform way.

Usability of the Webocrat system has been proven also in the field of e-government. The system enables a user to define, open, and use many communication channels between citizens and a local authority and share knowledge between them. The purpose of these channels can vary from one-to-one communication (e.g. between a citizen and a representative) to many-to-many mode of operation (e.g. informing all citizens about upcoming events). They enable citizens to submit their ideas and proposals to municipalities, ask anything about the municipality, obtain information about the services of local government and events, communicate with representatives and departments of the institution, join public discussions on various local and non-local issues, etc.

The system described by Mach and Stofanik (2006) represents a knowledge sponge (see section 3.4). It tries to collect knowledge (represented in the form of web pages), while focusing on only some limited domain. In principle, to build a collection of web pages, it is possible to employ the following three approaches:

- Manual selection of pages and assignment into a proper class(es)
- Using results submitted by a web search engine and assignment of the returned pages into a class(es)
- Employment of classification techniques to classify pages into a set of classes

All these approaches suffer from some drawbacks. The first approach is able to produce a collection of a very high quality – but due to the required effort it does not pay out. The second solution is much less laborious (although some manual work to query the search engine is still necessary) – but, on the other hand, the results produced by current search engines are not very. And the problem of the last approach is how to obtain pages to feed them into a classification procedure. The proposed solution tries to mix all three above given approaches in order to achieve a method of populating collections with web pages which requires only limited user intervention and simultaneously provides good scalability to build large collections in (semi-)automatic way. The structure of the system is in Figure 9.

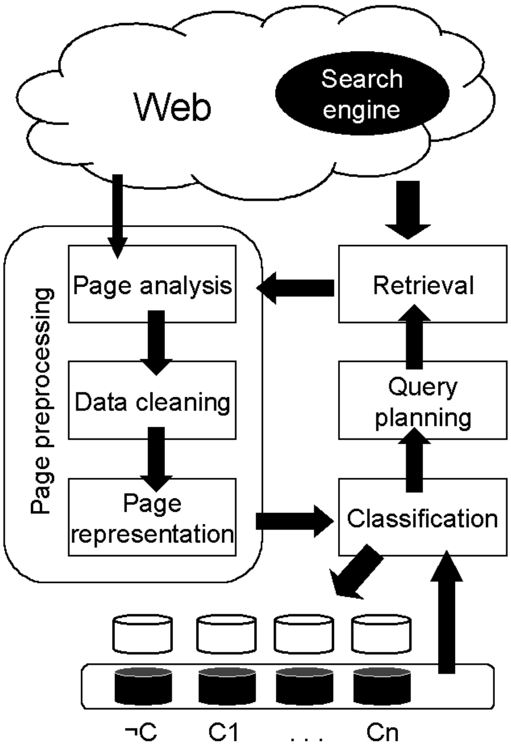



**Figure 9 Structure of the knowledge sponge**

A collection of pages is being created for each class. The collections grow steadily as new pages are retrieved from the Web and distributed into appropriate classes. A part of each collection (marked in black) serves as a set of instances of the given class – it is used to train a classifier for it to be able to recognise those pages which belong to the class.

The approach basically combines technology of web search engines and document classification technology to construct a system providing a steady stream of pages targeting relevant collections. This automatic part is complemented with user involvement necessary to initialise the process of feeding, supervise the overall process, assess the quality of created page collections, and reinitialise the process when necessary to increase its quality.

The proposed structure is composed from several modules providing required functionality. They can be characterised in the following way:

- **Retriever** is responsible for obtaining relevant pages from the Web. It ensures communication with a web search engine (currently, Google search engine was employed) to utilise wealth of knowledge accumulated by the search engine during scanning the Web.
- The aim of **pre-processor** is to convert content a page can deliver into the form suitable for classifier to classify the given page. It covers three basic activities: analysis, cleaning, and representation production.
- **Classifier** is responsible for checking the relevancy of retrieved pages, selecting relevant pages and distributing these pages into appropriate page collections.
- **Planner** represents a glue between web search and classification technologies. Based on results of classification (and the current setting of the classifier), it is

responsible for making decision on which pages should be retrieved from the Web within next iteration.

In order to start it is necessary to form an initial collection of web pages for each class as well as a collection of web pages which are not instances of any given class. These collections must be created by user – he/she can download the pages directly from the Web or can use a web search engine by constructing queries manually. In addition to retrieving the pages, user must classify them into classes. These initial collections represent training set of pages.

Based on the training collections classifiers can be constructed. From now on, any page (downloaded manually by user or automatically) can be classified and inserted into a collection of pages of a class (but not into the training part of the collection).

In order to perform populating of the page collection on some topic, a query to retrieve relevant pages is constructed and executed. To form the query, the classifier is utilised to obtain information on relevant terms associated with the given topic. Pages retrieved using the Google search engine can be inserted into the collection directly or (maybe better) can be classified and only those of them which are classified as really relevant are inserted into the collection of pages on the given topic.

### 4.1.2 Workflow Management Systems

A workflow is defined as the automation of business processes, in which the structure of the tasks and the responsibilities of tasks are predefined. The workflow system takes care of the execution and synchronization of tasks and the information flows to support individual tasks. At present, some workflow systems standards exist, but many of them are still in development. One of the main problems is complexity of workflow systems, which causes the limited adoption of existing standards.

The first effort of standardization was the one performed by the Workflow Management Coalition (WfMC)[19] in 1993. This coalition has developed a framework for the establishment of workflow standards. Two basic specifications from this edge are of interest: the WfMC reference model[20] and XPDL (WfMC-XPDL, 2002). The latter defines an XML schema for specifying the declarative part of a workflow. Some other standards are Business Process Execution Language (BPEL)[21] and Business Process Execution Language for Web Services (BPEL4WS) (IBM, 2005).

### 4.1.3 Extending WfMS with Activity Theory components

Traditional workflow models can be extended for describing knowledge creation processes: so-called dynamic workflow models (DWM). DWM should offer robustness to visualize the coordinated process, and flexibly to reframe a process. In such a way, DWMs enable knowledge practices to be more flexible by combining the process, objects and people (see e.g. Adams et al, 2003).

[19] see http://www.wfmc.org/index.html

[20] For the Workflow Reference Model Diagram see http://www.wfmc.org/standards/model.htm

[21] Business Process Execution Language, see: http://en.wikipedia.org/wiki/BPEL or http://www.theserverside.com/tt/articles/article.tss?l=BPELJava

DWM are based on activity theory. Activity theory (AT) is a powerful and clarifying descriptive theory focusing on understanding of human activity and work practices. It incorporates notions of intentionality, history, mediation, collaboration and development (Nardi, 1996). Based on the AT principles, six functionality criteria for DWM models have been identified by Adams et al (2003):

- Flexibility and reuse means that, at any point in time, several possible sequences that can be followed may exist utilizing a subset of available actions to achieve the objective of the activity. Choices are dependent on the actual circumstances of the activity at that time (context). Thus, the model needs to manage a catalogue of actions that, at runtime, could be chosen based on contextual information.
- Adaptation via reflection means that a model should support evolutionary adaptation of processes based on the experience gained during each execution of the process. Planned adaptations for future instantiations can be achieved e.g. by recording the occurrence of deviations. Thus a plan is an artefact that contains history of its development.
- Dynamic evolution of work practices requires, from the model, support of evolution of processes towards individual specializations without risking the loss of motivation for the overall activity.
- Locality of change means that modifications should be able to be fully applied by changing a minimal number of components, and should impact minimally on associated components. One approach would be to support the definition of a workflow process as a set of sub-processes, each of which is a distinct entity representing a single action. Changes made within one sub-process (activity) will not impact the other sub-processes.
- Comprehensibility of process models to all stakeholders, supporting representation at different levels of granularity. One possible approach is a hierarchical set of linked encapsulated sub-processes, i.e. each sub-process would be a (simpler) workflow model on its own.
- The evaluation of exceptions to “first-class citizens” regards exceptions as events that provide an opportunity for a learning experience. Exception handlings will be needed in both, design as well as execution phases. Selection of suitable actions from a repository of available actions could be made contextually dependent. If no precise action can be found, some approximate one can be selected and adopted to a required context.

An attempt to adopt AT in workflow contexts is described in Bardram (1997), within the project SAIK: Developing Computer Support for Clinical Work. In this project, a planning tool has been designed, implemented and tested in a hospital. This tool supports so-called situated planning, which is based on the activity theory, emphasizing the connection between plans and the contextual conditions for realizing these plans in actual work (Suchman, 1993).

### 4.1.4 Upgrading existing legacy systems to integrate KM functionality

From the technological point of view, most Knowledge Management (KM) projects aim at creating large, homogeneous knowledge repositories, in which corporate knowledge is made explicit, collected, represented and organized, according to a single - supposedly shared - conceptual schema (see also Figure 10). The typical outcome of this kind of project is the creation of an Enterprise Knowledge Portal (EKP), usually with a web-based interface, which provides a unique access point to corporate knowledge. Such an architecture is generally based on technologies like content management tools (text miners, search engines, and so on), which are used to produce a shared view (either implicit – e.g., clusters, neural nets – or explicit – e.g., ontologies, taxonomies[22]) of the entire collection of corporate documents; common formats (such as HTML, XML, PDF), used to overcome the syntactic heterogeneity of documents from different knowledge sources; chats and discussion groups, used to satisfy the need of social interaction.

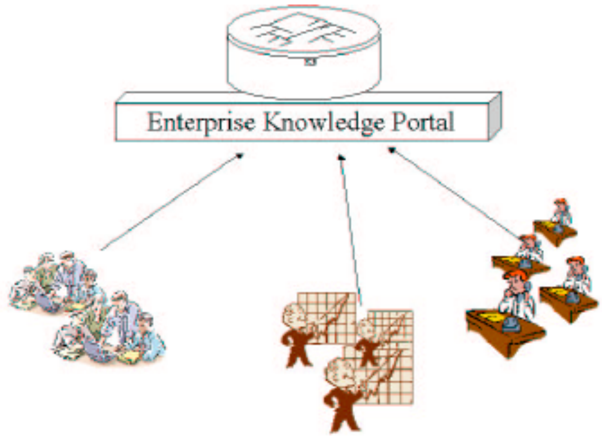



**Figure 10. The traditional KM approach**

Building new knowledge-based systems today usually entails constructing new knowledge bases on existing knowledge in an organization and in systems. Thereby, system developers have to face the big challenge of making existing systems interoperate among each other and with the new knowledge system. Thereby, existing systems have to provide new functionalities such as to perform some reasoning, to make knowledge accessible and to extract knowledge from stored knowledge. Such functionality is usually not implemented in existing legacy systems. However in this way, declarative knowledge, problem-solving techniques, and reasoning services shall be shared among systems.

Therefore the techniques for managing knowledge within the organization are drawn from two distinct areas:

- Techniques used previously for business management, for example, SWOT (Strengths Weaknesses Opportunities Threats) analysis, balanced scorecards, modelling languages such as: IDEF (Process Flow and Object State Description Capture Method)[23] and RADs (Role Activity Diagrams, see e.g. Murdoch and McDermid, 2000); and
- Knowledge techniques used previously for the disciplined development of knowledge-based applications, such as:

[22] See also sub-deliverable 5.2.3 on ontologies for further discussion.

[23] see: http://www.idef.com/

- National Institute for Standards and Technology Process Specification Language (NIST PSL)[24]
- Process Interchange Format (PIF)[25]
- US Defense Advanced Research projects Agency Shared Planning and Activity Language (DARPA SPAR)[26]
- Ontoweb European Union Research Programme - Web Contents Standards[27]
- Semantic Web Services Language[28]
- Workflow Management Coalition (WfMC) Workflow Process Definition[29]

### 4.1.5 Distributed Knowledge Management

Distributed Knowledge Management is proposed within the project EDAMOK (Bonifacio et al, 2002). In this approach, subjectivity and sociality are viewed as a potential source of value, rather than as a problem to overcome. The concept of absolute knowledge, which refers to an ideal, objective picture of the world, leaves the place to the concept of local knowledge, which refers to different, partial and approximate interpretations of the world, generated by individuals and within groups of individuals.

DKM (see Figure 11) is based on the two following principles:

- Principle of Autonomy: each organizational unit should be allowed a large degree of autonomy in managing (creating, representing, organizing, selecting, sharing) its own knowledge ("local" knowledge);
- Principle of Coordination: knowledge sharing across organizational units should be thought of as a form of coordination between multiple autonomous perspectives rather than as a process of creating (and imposing) a supposedly shared knowledge structure.

---

[24] See: http://www.mel.nist.gov/psl/

[25] See: http://ccs.mit.edu/pif/

[26] See: http://www.aiai.ed.ac.uk/project/spar/

[27] See: http://www.ontoweb.org/

[28] See: http://www.w3.org/Submission/SWSF-SWSL/

[29] See: http://www.wfmc.org/standards/standards.htm

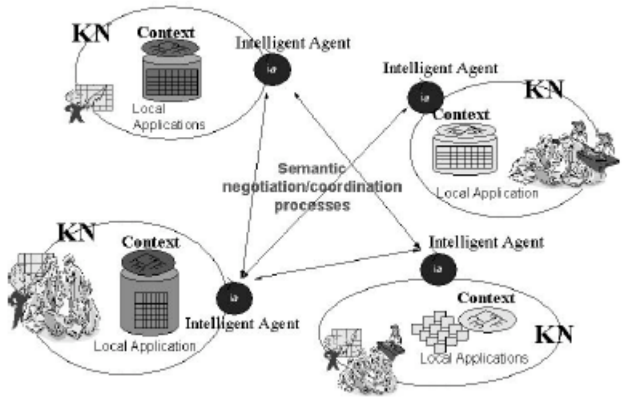



**Figure 11. DKM approach as used in EDAMOK**

This approach was realized as a framework, based on P2P communication infrastructure and agents communication protocols, by XML-based representation language and by meaning negotiation algorithms. The precondition for its working is that when an organization, school or some other institution wants to transform knowledge into a valuable organizational asset, they must use formalization, distribution, sharing, using and re-using of knowledge.

### 4.1.6 Case-based reasoning

Case-based reasoning[30] (e.g. Watson 2002) is a methodology for supporting knowledge management. The set of CBR principles are more fully defined as a cycle comprising six activities or processes, called the CBR-cycle (see Figure 12). The six activities (called the six-REs by the CBR Community) are as follows:

1. Retrieve knowledge that matches the knowledge requirement.
2. Reuse a selection of the knowledge retrieved.
3. Revise or adapt that knowledge in light of its use if necessary.
4. Review the new knowledge to see if it is worth retaining.
5. Retain the new knowledge if indicated by step 4.
6. Refine the knowledge in the knowledge memory as necessary.

[30] http://en.wikipedia.org/wiki/Case-based_reasoning

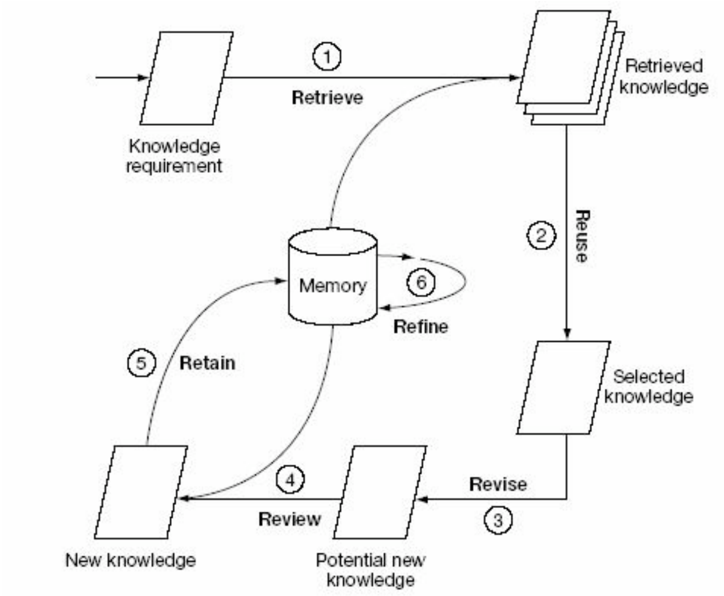



**Figure 12. CBR cycle**

### 4.1.7 Further examples of KM applications and projects

Some further examples of KM applications and projects are:

1. In Know-Net[31], the system is built around a knowledge-networking approach that is a unique fusion of the knowledge-as-a-product (content) and knowledge-as-a-process (context) perspectives to knowledge management (cf. also the conceptual level of the three-layers concept introduced in section 2.2). Two basic KM approaches are integrated:
    a. process-centred: KM as a social communication process
    b. product-centred knowledge processes such as creation, storage and reuse
2. The METOKIS[32] project investigated the use of semantic web technologies for electronic publishing in knowledge-intensive subject fields such as News Services, Education, and Clinical Studies. The project results are a Semantic Content Model and software to handle these semantics-based content objects, as well as a methodology for assessing the cost-benefit and for planning the introduction of semantics-based content applications in knowledge intensive organizations.
3. SAIC's approach[33] to capturing and reusing knowledge has been refined through successful application in many Fortune 500 companies and government organizations over the last several years.
4. The goal of KDE (Hemetsberger and Reinhardt, 2004) is identification of online processes for knowledge creation and sharing.

[31] A European research project (Esprit EP28928): see http://www.knownet.hhs.gov/default.htm

[32] http://metokis.salzburgresearch.at/

[33] http://www.saic.com/km/knowledge.html

5. PCPACK[34] is a powerful business solution that supports the retention, sharing, management and re-use of knowledge. Major global corporations are benefiting from the use of PCPACK. These include Airbus, BAE Systems, British Airways, Corus, GKN, Logica, Motorola, Prudential, QinetiQ, Rolls-Royce and Unilever.

There is a definite link between knowledge management and Intellectual Capital, as stated by a project manager (www.si.is/nhki): "The KM is simply the daily management of your Intellectual Assets for value creation".

Case studies which have gained attention in respect to KM solutions have been reported in (Gamble and Blackwell, 2001). Here, a few examples are listed:

- Swedish company Scandia was probably the world's first to value intellectual assets on its balance sheets. It has created a formalized procedure to capture experiences. While starting new financial services products reduced the time from start to profitability from 2 years to 6 months. The navigator approach takes into account the same set of financial, operational, and customer concerns as the scorecard. But, it makes the need to consider the organization, its structure and processes for nurturing its employees more explicit (Bredahl and Rydén, 2002).
- IC research projects sponsored by the Nordic innovation centre[35]: Nordika stands for "Nordic project for measurement of Intellectual Capital". This project has been initiated by the Nordic Industrial Fund (Nordisk Industrifond, part of the Nordic Council of Ministers) and concerns Denmark, Finland, Iceland, Norway and Sweden. It has started in September 1999. The aim of the project was to create a network of stakeholders of Intellectual Capital accounting and reporting in the Information Technology sector of the Nordic countries. Once established, the network evaluated the possibility and feasibility of launching a Nordic project to produce harmonized indicators for reporting IC in the IT sector of the Nordic countries, with a special focus on small and medium enterprises
- Intranets and groupware technologies have been key in enabling KM practices for AstraZeneca's projects that span the globe, time differences, and functionality. Like all major pharmaceutical companies, AstraZeneca invests heavily in its own research and development. Increasingly its portfolio of drugs contains products that are under license from other companies. The route in for a licensed product can be varied and complicate. Approaches to AstraZeneca can run to thousands a year. Then internal knowledge management consultants initiated a KM approach to the issue. There were two bodies of knowledge that needed management: the scientific and commercial knowledge, and the knowledge scattered all around the company about the status of any particular product under consideration
- Benetton's dynamic production lines keep modifying designs based on knowledge gained from customers. Benetton group have embraced KM in an effort to better leverage knowledge and information, including the process of how data and information is accessed, collected, and stored on a firm's networked computer system. The Group has a program called "Fabrica" for communication and commercial networking (SU-JEONG, 2002).

[34] http://www.epistemics.co.uk/Notes/55-0-0.htm

[35] See: http://www.nordicinnovation.net, Nordika: http://www.nordika.net/, Frame: http://www.icframe.net/, PIP: www.si.is/nhki

## 4.2 Structuring Information and Knowledge

### 4.2.1 Describing knowledge processes through ontology

An approach to describe knowledge is through ontology. Since sub-deliverable 5.2.3 extensively deals with ontologies, here no further details are being provided. The only mention is that this approach is implemented e.g. in the projects Pellucid[36] and Raport[37]. The main ontological model in Raport consists of three ontological models that are based on a generic ontology:

- A workflow ontology: control part, management of full process
- A data ontology: specifications of input and output data
- An user ontology: information about users

In the Raport project, an ontology is created through the CommonKADS methodology (see section 3.3.2).

### 4.2.2 Knowledge Maps

Some authors point to "knowledge mapping" as a feasible KM method to coordinate, simplify, highlight and navigate through complex webs of knowledge possessed by institutions (see Plumley, 2003). Knowledge maps, or k-maps, point to knowledge but they do not contain it. They are guides, not repositories. One of the main purposes of k-maps is to locate important knowledge in an organisation and show users where to find it. Effective k-maps should point not only to people but to documents and databases as well. K-maps should also locate actionable information, identify domain experts and facilitate organisation-wide learning. They should also trace the acquisition and loss of knowledge, as well as map knowledge flows throughout the organisation. Knowledge mapping can offer many benefits including economic, cultural, structural and knowledge returns. In this respect, k-maps are very close to the implementation of the structuring and cartography of knowledge in knowledge repositories (resp. organisational memories, see section 4.1).

K-map examples provided by Eppler (cited in Glassey, 2004) include knowledge application, knowledge structure, knowledge source, knowledge asset and knowledge development maps. Wexler (2001) identifies concept, competency, strategy, causal and cognitive maps. Plumley (2003) suggests that knowledge maps can be procedural, concept, competency and social network maps. A more abstract set of categories focusing primarily on cognitive maps is used by Huff (1990). The analysis of similarities and differences among these various types demonstrates that some classifications are simply different ways of referring to essentially the same maps.

---

[36] http://www.sadiel.es/europa/pellucid/partners.htm

[37] http://raport.ui.sav.sk/

## 4.3 Information Retrieval

Information Retrieval (IR) can be defined as the application of computer technology to the acquisition, organization, storage, retrieval, and distribution of information (Baeza-Yates and Ribeiro-Neto, 1999). The associated research discipline is concerned with both the theoretical underpinnings and the practical improvement of search engine technology, including the construction and maintenance of large information repositories.

IR (see e.g. Jackson and Moulinier, 2002) is an activity, and like most activities it has a purpose. A user of a search engine begins with an information need, which he or she realizes as a query in order to find relevant documents. It is easy to forget that document retrieval starts not with a query but with the indexing of documents.

Automated IR systems were originally used to manage information explosion in scientific literature in the last few decades. Many universities and public libraries use IR systems to provide access to books, journals, and other documents. IR systems are often related to object and query. Queries are formal statements of information needs that are put to an IR system by the user. An object is an entity, which keeps or stores information in a database. User queries are matched to documents stored in a database. A document is, therefore, a data object. Often the documents themselves are not kept or stored directly in the IR system, but are instead represented in the system by document surrogates

An index for the full-text search of electronic documents is generally more exhaustive than the index of any book. One would like to be able to query a collection of documents by matching terms in the query with terms actually occurring in the text of those documents. This ability requires that a document be indexed with all of the words that occur in it, instead of being indexed only by keywords or subject headings provided by an editor or a librarian.

According to Jackson and Moulinier (2002), an index consisting of a list of all the words occurring in all the documents in the collection is called an inverted file, or dictionary. Words are typically stemmed before being stored. Thus, we attempt to conflate all the variants of a word, reducing words like ‘anticipate’, ‘anticipating’, ‘anticipated’, and ‘anticipation’ to a common root, ‘anticipat’, for indexing purposes.

For each token we store the following information:

- Document Count. How many documents the token occurs in. This allows us to compute a useful statistic, called ‘inverse document frequency’ (IDF), for ranking purposes.
- Total Frequency Count. How many times the token occurs across all the documents. This is a basic ‘popularity’ measure that tells you how common the token is.

In addition, for each token, we store the following indexing information on a per document basis (Jackson and Moulinier, 2002):

- Frequency: how often the token occurs in that document. This number is a very rough indicator of whether or not the document is really ‘about’ the concept encoded in the token, or whether it simply mentions the concept in passing.
- Position: the offsets10 at which these occurrences are found in the document. Offsets can be retained for different reasons. Some search engines allow users to search for a query term within n words, say 3, of another term. Other search

engines, like Google, use offsets to generate word-incontext snippets for display, which can be quite effective abstracts for retrieved documents, because they are query dependent. Finally, offsets are sometimes used to highlight query terms in retrieved documents.

### 4.3.1 Tools to support information retrieval

**Search engines** (Jackson and Moulinier, 2002):

- A Boolean search is one in which the user searches a database with a query that connects words with operators, such as AND, OR, and NOT. Such a search is often called a 'terms and connectors' search, since there is a clear distinction made in the query between content-bearing terms and content-free operators based on logical connectives.
- Most web search engines such Google and Lycos are based on a different technology that ranks search results based upon the frequency distribution of query terms in the document collection. As the example suggests, ranked retrieval is usually employed in search interfaces where users are allowed to enter unrestricted 'natural language' queries, without Boolean or other operators. Such a query is then processed by removing stop words, like 'where' and 'do', and performing various manipulations on the remaining words, the most common being stemming. In modern search engines, words are stemmed at index time, and stemming algorithms attempt to identify the root forms of query terms automatically, so that the user does not have to resort to wild cards.
- The probability ranking suggests ranking a document according to its odds of being in the class of relevant documents, rather than the class of non-relevant documents. Probabilistic IR is based on a theory that incorporates a number of underlying assumptions. The most common form of the theory frames the document retrieval problem as one of computing the probability that a document is relevant to a query, given that it possesses certain attributes or features. These features are typically words or phrases occurring in the document, as in the ranked retrieval model.
- Language modelling is a framework that, until recently, had been more commonly associated with speech recognition and generation. The primary difference between what is now being called 'classical' probabilistic IR and language modelling is that the latter seeks to model the query generation process, rather than the pool of relevant documents. Query generation is viewed as a process of sampling randomly from a document, or rather from a document model consisting of terms and their frequencies of occurrence in the document. In other words, we consider the probability that a given document model could have produced the query, and rank the corresponding document accordingly. Documents with a relatively high probability of generating the query are ranked high in the results list.

In recent years, researchers have expanded their concerns from the bibliographic and full-text search of document repositories to Web search, with its associated hypertext and multimedia databases. There is a common confusion, however, between data retrieval, document retrieval, information retrieval, and text retrieval, and each of these has its own bodies of literature, theory, praxis and technologies.

Traditional search engines were never intended to deal with a vast, distributed, heterogeneous collection of documents such as the WWW. The almost complete absence of editorial control over web documents poses special problems, such as coverage, correctness, spamming, dead links, and the manipulation of rankings for commercial advantage.

The Web is indexed by “crawling” it. Jackson and Moulinier (2002) state that a Web crawler is a program that visits remote sites over the Internet and automatically downloads their pages for indexing. Today this is typically done in a distributed fashion, using more than one program. Search engine indexes have grown significantly since 1997. By the end of 2001, Google was indexing an estimated 1.5 billion pages, with runners-up Fast, Altavista, and Inktomi indexing half a billion or more. Indexing the Web is a non-trivial business. A crawler may connect to half a million servers and download millions of pages. Downloaded documents need to be compressed and stored, parsed to extract index terms, and then sorted to generate an inverted index.

Frakes and Baeza-Yates (1992) state that it is hard to classify IR algorithms, and to draw a line between each type of application. However, we can identify three main types of algorithms. There are other algorithms used in IR that do not fall within our description, for example, user interface algorithms. The reason that they cannot be considered as IR algorithms is because they are inherent to any computer application.

The main class of algorithms in IR is retrieval algorithms, that is, to extract information from a textual database. We can distinguish two types of retrieval algorithms, according to how much extra memory we need:

- Sequential scanning of the text: extra memory is in the worst case a function of the query size, and not of the database size. On the other hand, the running time is at least proportional to the size of the text, for example, string searching
- Indexed text: an "index" of the text is available, and can be used to speed up the search. The index size is usually proportional to the database size, and the search time is sublinear on the size of the text, for example, inverted files and signature files

Filtering algorithms are such that the text is the input and a processed or filtered version of the text is the output. This is a typical transformation in IR, for example to reduce the size of a text, and/or standardize it to simplify searching. The most common filtering/processing operations are:

- Common words removed using a list of stopwords.
- Uppercase letters transformed to lowercase letters.
- Special symbols removed and sequences of multiple spaces reduced to one space.
- Numbers and dates transformed to a standard format.
- Spelling variants transformed using Soundex-like methods.
- Word stemming (removing suffixes and/or prefixes).
- Automatic keyword extraction.
- Word ranking.

Indexing algorithms build a data structure that will allow quick searching of the text. There are many classes of indices, based on different retrieval approaches, e.g. inverted files, signature files, tries etc. Almost all types of indices are based on some kind of tree or hashing. Usually, before indexing, the text is filtered.

### 4.3.2 Examples of IR software tools

- ASPseek, http://aspseek.org/
- Egothor, http://www.egothor.org/
- Lemur, http://www.lemurproject.org/
- Lucene, http://en.wikipedia.org/wiki/Lucene
- mnoGoSearch, http://mnogosearch.org/
- Sphinx, http://www.sphinxsearch.com/
- Terrier, http://ir.dcs.gla.ac.uk/terrier/
- Wumpus, http://www.wumpus-search.org/

### 4.3.3 Examples of IR applications

- Search Engines
  - Addresses and Phone Numbers, http://www.switchboard.com/
  - Google scholar, http://scholar.google.com\
  - Google Government, http://www.google.com/unclesam
  - Louisiana State Library, http://www2.state.lib.la.us/databases/
  - ERIC, http://www.eric.ed.gov/
  - High Wire, http://highwire.stanford.edu
  - Virtual Reference Desk, http://www.vrd.org
  - National Research Council of Canada, http://iit-iti.nrc-cnrc.gc.ca
- Digital_Libraries
  - University of California, Berkeley Digital Library, http://elib.cs.berkeley.edu
  - New Zealand Digital Library, http://www.nzdl.org/
  - IEE Digital Library, http://www.ieedl.org/
  - Online Archive of California, http://www.oac.cdlib.org/
  - Bibliotheque Nationale de France, http://www.bnf.fr/
- Enterprise_Search_Engines
  - Google Enterprise, http://www.google.com/enterprise
  - Northern Light, http://www.northernlight.com
  - LexisNexis, http://lexisnexis.com
  - Convera, http://www.convera.com
- Desktop_Search
  - Google Desktop Search, http://desktop.google.com/
  - MSN Desktop Search, http://beta.toolbar.msn.com
  - Yahoo Desktop Search (X1), http://desktop.yahoo.com
  - Copernic Desktop Search Engine, http://www.copernic.com/

## 4.4 Knowledge Analysis

### 4.4.1 Knowledge discovery in databases (KDD)

Knowledge discovery in databases (KDD) can be defined as nontrivial processes of identifying valid, novel, potentially useful, and ultimately understandable patterns in data. According to Fayyad et al (1996), it is an interactive and iterative process with several steps. It means that at any stage the user should have possibility to make changes (for instance to choose different algorithm settings, different data mining tasks or pre-processing data in another way) and repeat the following steps to achieve better results. Data mining is a part of this process.

**Data Mining** (DM) is the most common term used to name the field of knowledge discovery. This confusing use of the terms KDD and DM is due to historical reasons and the fact that most of the work is focused on refinement and applicability experiments of machine learning algorithms from artificial intelligence for the data-mining step. Pre-processing is often included in this step as a part of the mining algorithm.

Han and Kamber (2001) recognized the following steps within the KDD process (compare also with Figure 13) following:

1. *Data cleaning* to remove noise and inconsistent data
2. *Data integration*, where multiple data sources may be combined
3. *Data selection,* where data relevant to the analysis task are retrieved from database (or data warehouse, where data is already cleaned and integrated)
4. *Data transformation* - data are transformed or consolidated into forms appropriate for mining
5. *Data mining* as core of the KDD process, where intelligent methods are applied in order to extract data patterns
6. *Pattern evaluation* – to identify interesting patterns
7. *Knowledge representation* - visualization of mined knowledge.

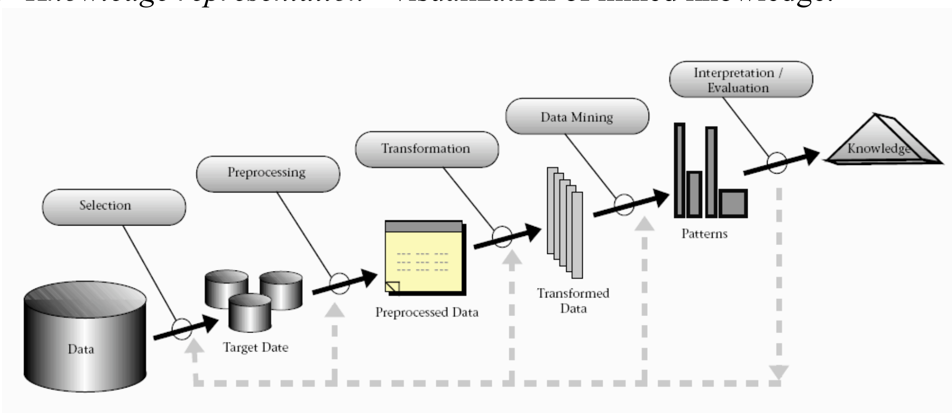



**Figure 13. The KDD process according to Frawley et al (1991)**

KDD is a growing field: There are many knowledge discovery methodologies in use and under development. Some of these techniques are generic, while others are domain-specific:

- Learning techniques may be supervised or unsupervised. In general, supervised learning techniques enjoy a better success rate as defined in terms of usefulness of discovered knowledge. According to Brachman and Anand (1996), learning algorithms are complex and generally considered the hardest part of any KDD technique.

- Probabilistic approach covers techniques that utilize graphical representation models to compare different knowledge representations. These models are based on probabilities and data independencies. Probabilistic techniques may be used in diagnostic systems and in planning and control systems (cf. Buntine, 1996). Automated probabilistic tools are available both commercially and in the public domain.
- The statistical approach uses rule discovery and is based on data relationships. Online analytical processing (OLAP) is an example of a statistically-oriented approach. Automated statistical tools are available both commercially and in the public domain.
- Classification is probably the oldest and most widely-used of all the KDD approaches (Quinlan, 1993). This approach groups data according to similarities or classes. There are many types of classification techniques and numerous automated tools available.
- Pattern detection by filtering important trends is the basis for this KDD approach. Deviation and trend analysis techniques are normally applied to temporal databases. A good application for this type of KDD is the analysis of traffic on large telecommunications networks. AT&T uses such a system to locate and identify circuits that exhibit deviation (faulty behaviour) (Sasiekharan et al, 1996). The sheer volume of data requiring analysis makes an automated technique imperative. Trend-type analysis might also prove useful for astronomical and oceanographic data, as they are time-based and voluminous. Public domain tools are available for this approach.
- Neural networks are particularly useful for pattern recognition, and are sometimes grouped with the classification approaches.
- A hybrid approach to KDD combines more than one approach and is also called a multi-paradigmatic approach. Some of the commonly used methods combine visualization techniques, induction, neural networks, and rule-based systems to achieve the desired knowledge discovery. Deductive databases and genetic algorithms have also been used in hybrid approaches.

CRISP-DM[38] is main methodology that describes full process of KDD. The CRISP-DM project developed an industry- and tool-neutral KDD process model. It contains the corresponding phases of a project, their respective tasks, and relationships between these tasks. The life cycle of a KDD project consists of six phases, see Figure 14. The sequence of the phases is not strict. Moving back and forth between different phases is usually necessary. It depends on the outcome of given phase, which phase, or which particular task of a phase has to be performed as next. The arrows indicate the most important and frequent dependencies between phases.

[38] http://www.crisp-dm.org/

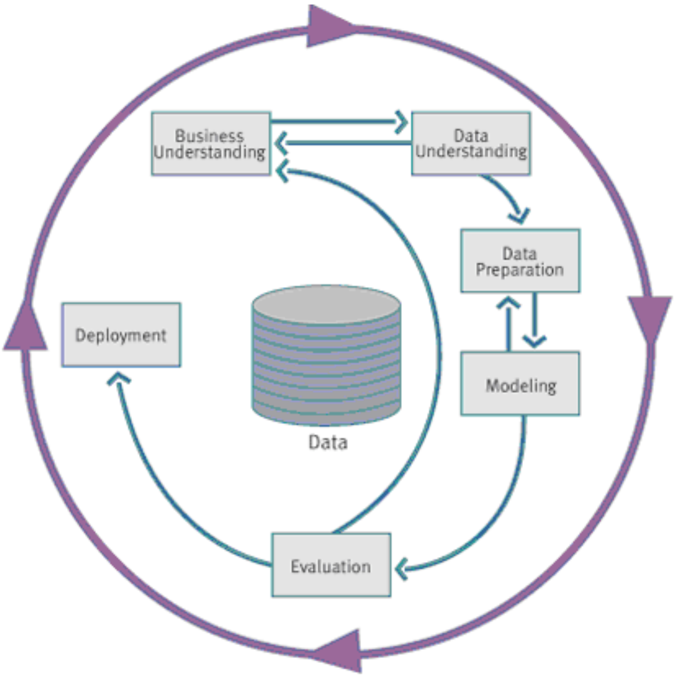



**Figure 14: Methodology CRISP-DM39**

### 4.4.2 Software systems and tools supporting KDD process

The following software systems and tools are examples of support for KDD processes[40]:

- Clementine, http://www.spss.com/clementine/
- Enterprise miner, http://www.sas.com/technologies/analytics/datamining/miner/
- GhostMiner, http://www.fqs.pl/?a=product_view&id=2&lang=en&x=
- DB2 Data Warehouse Edition, http://www-306.ibm.com/software/data/db2/dwe/
- Oracle Data Mining, http://www.oracle.com/technology/products/bi/odm/index.html
- Statistica, http://www.statsoft.com/
- AlphaMiner, http://www.eti.hku.hk/alphaminer/
- Weka 3, http://www.cs.waikato.ac.nz/ml/weka/index.html
- KDD Package (Paralic and Bednar, 2003), http://www.tuke.sk/paralicj/KDD/

### 4.4.3 Selected KDD applications and cases

Some examples of KDD applications are (cf. Edelstein, 1999)[41]:

- Blockbuster Entertainment mines its video rental history database to recommend rentals to individual customers.

---

[39] http://www.crisp-dm.org/

[40] http://www.kdnuggets.com/

[41] http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm

- American Express can suggest products to its cardholders based on analysis of their monthly expenditures.
- WalMart captures point-of-sale transactions from over 2,900 stores in 6 countries and continuously transmits this data to its massive 7.5 terabyte Teradata data warehouse. WalMart allows more than 3,500 suppliers, to access data on their products and perform data analyses. These suppliers use this data to identify customer buying patterns at the store display level. They use this information to manage local store inventory and identify new merchandising opportunities[42].
- The National Basketball Association (NBA) is exploring a data mining application that can be used in conjunction with image recordings of basketball games. The Advanced Scout software analyzes the movements of players to help coaches orchestrate plays and strategies (Bhandari et al 1997).
- Clinical KDD in hospital information systems (Tsumoto, 2000)
- Ireland's Office of the Revenue Commissioners: Integrating Knowledge for Better Customer Service and Improved Business Efficiency[43]
- Web-Based International Trade Knowledge Discovery System (CS Solutions, 2005)
- IBM research group[44], [45]

### 4.4.4 Data mining

Data mining is a key step in the KDD process and covers application of suitable intelligent methods and techniques in order to derive new interesting patterns in data (potentially new knowledge). There are various types of DM, e.g. descriptive data mining, predictive data mining, discovery of association rules or clustering.

The goal of the descriptive data mining is to describe relevant dataset in a concise form, in order to provide general characteristics of analyzed dataset. This procedure is called generalization. Generalization has two basic forms: a characterization and a comparison. The process of a characterization can be automatic or manual. Manual characterization is implemented by OLAP (Online Analytical Processing) operations and process of automatic characterization can be implemented e.g. as attribute-oriented induction (AOI) (cf. Han and Kamber, 2001).

Predictive data mining consists of two main DM tasks: classification and prediction. For this purpose are used:

- Decision trees: Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square

---

42 http://www.teradata.com/t/page/128640/index.html

43 http://www.accenture.com/Global/Services/Accenture_Technology_Labs/Client_Successes/CaseCommissioners.htm#solution

44 http://www.sas.com/technologies/analytics/datamining/miner/semma.html

45 http://domino.research.ibm.com/comm/research.nsf/pages/r.kdd.html

Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a dataset. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CART segments a dataset by creating 2-way splits while CHAID segments using chi square tests to create multi-way splits. CART typically requires less data preparation than CHAID. Typical representative is algorithm C4.5 (Quinlan, 1993).

- Nearest neighbour method: A technique that classifies each record in a dataset based on a combination of the classes of the k most similar record(s) to it in a historical dataset. Sometimes called the k-nearest neighbour technique.
- Bayes classifies predict probability of classification of given example into one of the predefined classes.[46]

Agrawal et al (1993) state that association rule mining finds interesting associations and/or correlation relationships among large sets of data items. Association rules show attribute-value combinations that occur frequently together in a given dataset. A typical and widely used example of association rule mining is Market Basket Analysis: data are collected using bar-code scanners in supermarkets. Such 'market basket' databases consist of a large number of transaction records. Each record lists all items bought by a customer on a single purchase transaction. Managers would be interested to know if certain groups of items are consistently purchased together. They could use this data for adjusting store layouts (placing items optimally with respect to each other), for cross-selling, for promotions, for catalogue design and to identify customer segments based on buying patterns. According to Agrawal and Srikant (1994), Apriori[47] is one of commonly used algorithms.

Clustering is identification of groups of similar objects, or more precisely, the partitioning of a data set into subsets (clusters), so that the data within a subset (ideally) share some common features - often proximity according to some defined distance measure – i.e. are similar to each other and dissimilar to objects from different subsets. Clustering algorithms can be hierarchical or partitional. Hierarchical algorithms find successive clusters using previously established clusters, whereas partitional algorithms determine all clusters at once. Hierarchical algorithms can be agglomerative (bottom-up) or divisive (top-down). Agglomerative algorithms begin with each element as a separate cluster and merge them in successively larger clusters. Divisive algorithms begin with the whole set and proceed to divide it into successively smaller clusters.

### 4.4.5 Specialized DM software

- "Yet Another Learning Environment", http://en.wikipedia.org/wiki/YALE
- Tanagra, http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html
- CART, http://www.salfordsystems.com/cart.php
- The Lumenaut Decision Tree, http://www.lumenaut.com/decisiontree.htm
- XLMiner, http://www.resample.com/xlminer/help/k-NN/knn_intro.htm

[46] http://en.wikipedia.org/wiki/Bayes'_theorem

[47] http://fuzzy.cs.uni-magdeburg.de/~borgelt/doc/apriori/apriori.html

- Auton, http://www.autonlab.org/autonweb/10522
- ARMiner, http://www.cs.umb.edu/~laur/ARMiner/
- Clustan, http://www.clustan.com/

### 4.4.6 OLAP

OLAP (On-Line Analytical processing)[48] is an approach to quickly provide the answer to analytical queries that are dimensional in nature. The term OLAP was created as a slight modification of the traditional database term OLTP (On Line Transaction Processing)[49]. OLAP takes a snapshot of a set of source data and restructures it into a multidimensional OLAP cube. Queries can then be run against this cube, which is created from a star or snowflake schema of tables. At the centre, there is the fact table, which lists the core facts that make up the query. Numerous dimension tables are linked to the fact tables. These tables indicate how the aggregations of relational data can be analyzed. The number of possible aggregations is determined by every possible manner in which the original data can be hierarchically linked. The final result of OLAP techniques can be very simple (e.g., frequency tables, descriptive statistics, simple cross-tabulations) or more complex (e.g., they may involve seasonal adjustments, removal of outliers, and other forms of cleaning the data).

There are three types of OLAP architectures:

- Multidimensional (MOLAP) is the 'classic' form of OLAP and is sometimes referred to as just OLAP. MOLAP uses database structures that are generally optimal for attributes such as time period, location, product or account code. The way that each dimension will be aggregated is defined in advance by one or more hierarchies.
- Relational (ROLAP) works directly with relational databases. The base data and the dimension tables are stored as relational tables and new tables are created to hold the aggregated information. Depends on a specialized schema design.
- There is no clear agreement across the industry as to what constitutes "Hybrid OLAP", except that a database will divide data between relational and specialized storage. For example, for some vendors, a HOLAP database will use relational tables to hold the larger quantities of detailed data, and use specialized storage for at least some aspects of the smaller quantities of more-aggregate or less-detailed data.

The OLAP market continued to grow faster than most other enterprise software sectors. The larger generalist vendors - Microsoft, Oracle, SAP, Business Objects - cannot even measure their OLAP business themselves, because their OLAP capabilities are often delivered as part of larger, bundled products and account for a minority of their revenues. For example, Microsoft Analysis Services, a component of the Microsoft SQL Server, is typically chosen by smaller organizations, while SAP BW and MicroStrategy are much more likely to be found in the largest organizations (cf. Pendse, 2006). Similarly, Business Objects and SAP are relatively stronger in Europe, while the MicroStrategy and Hyperion customer bases have a North American bias. The large non-specialist vendors,

[48] See: http://en.wikipedia.org/wiki/OLAP, http://www.statsoft.com/textbook/stdatmin.html#olap

[49] See: http://en.wikipedia.org/wiki/OLTP

such as Microsoft and Oracle, are stronger in the rest of the world than the smaller BI specialists, who tend to be under-represented outside the major markets. By vertical market, MicroStrategy is particularly strong in retail, Applix in finance and insurance and Microsoft and Business Objects in the IT industry.

### 4.4.7 Examples of OLAP software systems

- Open source OLAP, http://en.wikipedia.org/wiki/Open_source_olap
- Microsoft Analysis Services, http://en.wikipedia.org/wiki/Microsoft_Analysis_Services
- Cognos, http://www.cognos.com/
- Business objects, http://www.businessobjects.com/
- MicroStrategy, http://en.wikipedia.org/wiki/MicroStrategy
- Hyperion, http://en.wikipedia.org/wiki/Hyperion_Solutions_Corporation
- Applix, http://en.wikipedia.org/wiki/Applix
- SAP BW, http://www.thespot4sap.com/Articles/SAP_BW_Introduction.asp

## 4.5 Agent technologies and tools

### 4.5.1 Agents

It is helpful to distinguish three fundamental aspects of agents (Singh and Huhns, 2005): (1) individual agents, (2) systems of agents and their interactions and (3) the environment in which agents operate. For agents exist plethora of labels (e.g. autonomous, intelligent, mobile...), which makes the term “agent” almost meaningless because it can be used too frequently to characterise anything. In this short survey we will use this notion according to the essential properties defined by AgentLink3 Agent Oriented Software Engineering Technical Forum Group (AOSE TFG) (Bernon et al, 2005): an agent is able to act, is autonomous, proactive, communicates with others, and perceives its environment.

Regarding the relation between objects and agents – besides sharing some aspects, they also differ, mainly on notions such as autonomy and interaction. Both agents and objects encapsulate their state, which in objects is determined by the values of a set of variables, whereby in agents this can be defined in terms of goals, beliefs, facts, etc., what determines a mental state. Objects may have control over their state by using private attributes or methods but any public method of an object can be invoked by another object forcing the former one to perform the action described by the method. An agent can determine which behaviour to follow (depending on its goals, its internal state and its knowledge from the environment) and not because someone else forces to do something. Another important difference is the social dimension of agents. Communication between objects is defined in terms of messages that activate methods, but in the agent domain, this communication is richer both in diversity of mechanisms and in the language. The language is namely defined at a more abstract level, in terms of ontologies and speech acts, for instance. This social perspective is reflected also in the definition of organizations with social rules and relationships among agents. Therefore, the use of object-oriented software engineering techniques can be applied for the development of

Multi-Agent Systems (MAS), but some extensions are required to deal with social issues (organization, interaction, coordination, negotiation, cooperation), more complex behaviour (autonomy, mental state, goals, tasks), and a greater degree of concurrency and distribution (Bernon et al, 2005).

#### 4.5.1.1 Agent Environment

Agents, as well services, do not exist and operate in isolation, but rather in some physical or computational environment. There are an unlimited number of environments, but they can be described in terms of the following six characteristics (Russell and Norvig, 2003):

- Observability – an environment is fully observable by an agent if its sensors can detect all aspects that are relevant to its choice of action; it is partially observable otherwise.
- Determinism – an environment is deterministic, from the point of view of an agent, if its next state is completely determined by the current state and the agent's action; otherwise it is stochastic.
- History freedom – an episode is a single cycle of an agent perceiving its environment and taking an action. If the choice of action depends only on the episode itself and not previous episodes, then the environment is episodic. If the current decision affects future decisions, as in deciding on a move in chess, then the environment is sequential.
- Dynamism – an environment is dynamic if it can change while an agent is deciding on the action it should take; otherwise it is static.
- Continuity – from the point of view of an agent, an environment is discrete if the agent perceives it as being in one of a finite number of distinct states, if the agent has a finite number of possible actions, and if there is a distinct set of time points at which it is perceived or actions are taken. If the perceived variables can have a continuous range of values, then the environment is continuous.
- Multiagent – from the point of view of an agent, if there are other agents that can affect its environment and of which the agent is aware, then the environment is considered to be multiagent.

From an implementation standpoint, environments for agents consist of: a communication infrastructure and protocols for interaction, security services for authentication and authorization, remittance services for billing and accounting, and operations support for logging, recovery and validation.

For an agent to act properly in an environment, some combination of its data structures and program must reflect the information it has about its environment. Because this information would reflect the state of the environment according to the agent, it can be termed its knowledge or set of its beliefs. An agent's desires correspond to the state of the environment the agent prefers. An agent's intentions correspond to the state of the environment the agent is trying to achieve, which should be a consistent subset of the agent's desires and directly connected to the agent's actions.

#### 4.5.1.2 Agent Descriptions

Agents can be described at three different levels of abstraction:

1. Knowledge level (or Epistemological Level) – the agent is described by saying what it knows.
2. Logical level – the level at which the knowledge is encoded into sentences.
3. Implementation level – the level that runs on the agent architecture; this level is important only for efficiency.

Research in MAS has recently led to the development of practical programming languages and tools that are appropriate for the implementation of such systems (Bordini et al, 2006). Most research in agent-oriented programming languages is based on declarative approaches. There are many declarative solutions, most of them logic based (e.g. FLUX, Minerva, Dali, Respect), some are based on other formalisms (e.g. CLAIM (Seghrouchni and Suna, 2004) is inspired by ambient calculus). Purely imperative languages are unusual, as in essence they are inappropriate for expressing high-level abstractions associated with agent system design. On the other hand, agent oriented programming languages tend to allow for easy integration with (legacy) code written in imperative languages (e.g. 3APL (Dastani et al, 2004), Jason, Go!).

#### 4.5.1.3 Rules

Rules are needed for expressing the individual decision making of interacting parties, as well as the contracts that bind them to each other. In particular, rules are desirable because they:

- Can be created in a modular manner. It is possible to add rules incrementally. The set of rules being developed becomes in this way more complete and potentially yields behaviour that is closer to what is desired.
- Can be inspected. Unlike imperative programming languages, rules are declarative. Thus, they can be read and understood in terms of their explicit content.
- Are executable. Unlike textual descriptions or even some formal specifications, rules can be directly executed. Thus behaviour specified using rules is attained by executing those rules.

In the Semantic Web vision, rules are higher-level abstraction then ontologies. Rules have natural match with services, especially with regard to their composition. Some potential uses of rules for service computing are:

- Expressing derived concepts in an ontology.
- Expressing the private policies of the different participants in a service composition that would capture how a given participant decides to compose services.
- Expressing how exception conditions of various types are detected and handled. The rules can be used to encode, in essence, flexible transaction models that would otherwise not be possible to encode.
- Expressing the business protocols under which various services interact. Rules can help make the protocols more flexible than if they were hard coded as traditional step-by-step process.

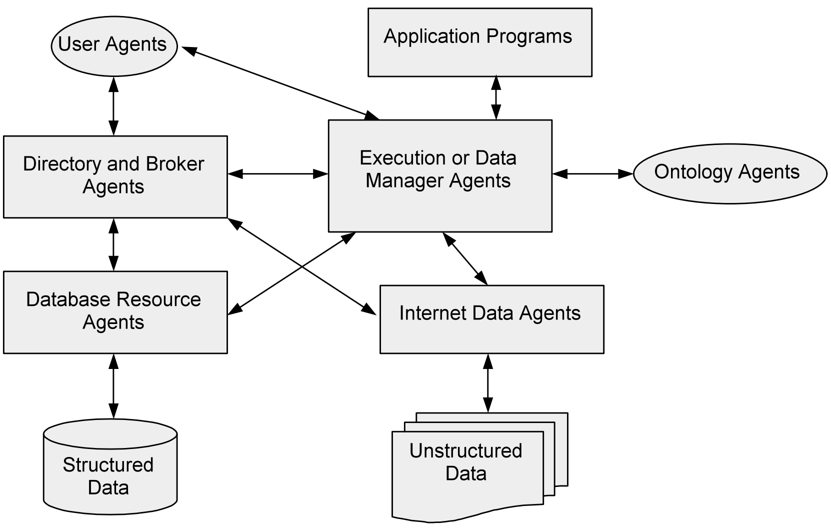
There are four major families of rules: Prolog, production rules such as Jess, event-condition-action (ECA) rules and SQL. Jess[50] is a fast and lightweight rule engine written in Java. It provides to services and agents an ability to reason declaratively using knowledge expressed in rules and facts. To process rules, Jess uses Rete algorithm, which is an efficient mechanism for solving the many-to-many matching problem between facts. Jess supports both forward and backward chaining, working memory queries and the ability to manipulate and reason about java objects directly. For communicating rules among different services or agents – the Semantic Web Rule Language (SWRL) can be used. SWRL is a markup language for expressing rules in a standardized manner and melding them with ontologies. SWRL can be mapped to the proprietary formats of various rule engines that different services might internally employ. By incorporating inference and reaction rules, SWRL covers the major rule families of interest. SWRL extends the set of OWL axioms to include a kind of rules (Horn clauses), thus enabling rules to be combined with an OWL knowledge base.

#### 4.5.1.4 Multiagent Systems (MAS)

Most of the authors agree on viewing a MAS as a system composed of agents that communicate and collaborate to achieve specific personal or collective tasks (Bernon et al, 2005). MAS are appropriate to deal with complex and open problems. The organization facilitates managing complexity by determining structures, norms and dependencies. In some cases, the organization is explicitly a subject of analysis and design (e.g. Zambonelli et al, 2003)), but in certain approaches, the organization emerges at run time (e.g. (Di et al, 2006).

#### 4.5.1.5 Agent Types

To support an architecture in which heterogeneous components can interoperate, negotiate and achieve periodic consistency, a variety of agent roles are needed. Figure 15 shows a multi-agent system architecture in which each agent has a specialized function. The agents communicate using an agent communication language such as FIPA ACL (see section 4.5.1.9), whose sentences wrap a content language such as SQL.




50 Java Expert System Shell, available at http://herzberg.ca.sandia.gov/jess/

**Figure 15 Agent-based system architecture showing the de facto standard agent types**

*User Agents* act as an intermediary between users and information systems, providing access to such information resources as data analysis tools, workflows and concept-learning tools. They support a variety of interchangeable user interfaces (e.g. query forms, graphical query tools, menu-driven query builders and query languages), result browsers and visualization tools.

*Broker Agents* implement directory services for locating appropriate agents with appropriate capabilities. They manage namespace service and may store and forward messages and locate message recipients. Brokers simplify the configuration of multiagent system. An agent requests the broker to recruit one ore more agents who can provide a service.

*Resource Agents* provide access to information stored in legacy systems. The three common types are classified by the resource they represent. Wrappers implement common communication protocols and translate commands and results into and from local access language. Database agents manage specific information sources and data-analysis agents apply machine learning techniques to form logical concepts from data or use statistical techniques to perform data mining.

*Workflow Agents* are a kind of resource agent that applies to different workflows. They can coordinate the workflows they manage and thereby provide for larger, possibly enterprise-wide, workflows.

*Execution Agents* supervise query execution, operate as script-based agents to support scenario-based analyses or monitor and execute workflows. This third functionality can extend over the Web and be expressed in a format such as the one specified by the Workflow Management Coalition (WfMC). Such agents might be implemented as rule-based knowledge systems.

*Ontology Agents* manage the distributed evolution and growth of ontologies. They provide a common context as a semantic grounding, which agents can use to relate their individual terminologies. A third function of ontology agents is providing remote access to multiple ontologies.

#### 4.5.1.6 FIPA Standards

The Foundation for Intelligent Physical Agents (FIPA) is an IEEE Computer Society standards organization that promotes agent-based technology and the interoperability of its standards with other technologies. In the production of these standards, FIPA requires input and collaboration from its members and from the agent's field in general to build specifications that can be used to achieve interoperability between agent-based systems developed by different companies and organisations[51].

The core message of FIPA is that through a combination of speech acts, predicate logic and public ontologies, we can offer standard ways of interpreting communication between agents in a way that respects the intended meaning of the communication. This is much more ambitious than, for example, XML, which only aims to standardize the syntactic structure of documents. To support this, FIPA has adopted and is working on specifications that range from architectures to support agents' communicating with each

[51] Foundation for Intelligent Physical Agents (FIPA): http://www.fipa.org/

other, communications languages and content languages for expressing those messages and interaction protocols, which expand the scope from single messages to complete transactions. In the future, there are plans to extend this even further to cope with longer term relationships between agents.

#### 4.5.1.7 Agent Management Specification

An agent management system, as shown in Figure 16[52], handles agent creation, registration, location, communication, migration and retirement. It provides the following services:

- White pages – include support for agent location, naming, name resolution services and access control services. Agent names are represented by a flexible and extensible structure called an agent identifier, which can support a human-friendly name and transport address, among other things.
- Yellow pages – offer support for service location and registration services, which are provided by the Directory Facilitator (DF).
- Agent message transport service – originally these were standardised based on CORBA's IIOP, but HTTP bindings are now more popular.
- Mobility – agents can also be mobile, wherein their code, their data and the state of their execution may move from one platform or execution environment to another.

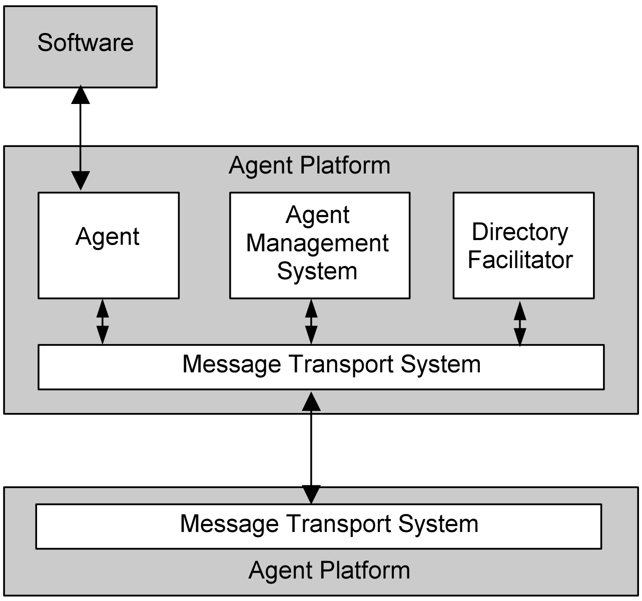



**Figure 16 Agent Management Reference Model**

[52] Foundation for Intelligent Physical Agents (FIPA): http://www.fipa.org/

#### 4.5.1.8 Agent Communication Language

An agent communication language (ACL) provides a domain-independent layer between an application-specific language and underlying message transport protocols. An ACL encodes the most common patterns of communications. It includes elements for specifying requests, commands, statements of belief, commitments and agent management. The most widely used ACL is the one standardized by FIPA.

FIPA ACL is based on speech act theory, because it views communication as action. Speech act theory considers three aspects of a message:

- Locution – how it is phrased.
- Illocution – how it is meant by the sender or understood by the receiver.
- Perlocution – how it influences the recipient.

Communicative acts (CAs) are described in both a narrative form and a formal semantics based on modal logic. FIPA ACL has a library of 22 CAs that includes e.g. the following: Accept-Proposal, Agree, Cancel, Call-for-Proposal, Confirm, Disconfirm, Failure, Inform, Inform-If, Not-Understood, Request, Subscribe. The semantics of the four primitive CAs (inform, request, confirm, disconfirm) are defined in SL and the other CAs are defined using these four basic CAs.

#### 4.5.1.9 FIPA Protocols

Major FIPA protocols could be spited into four groups:

1. Basic protocols: Request, Request-when, Query
2. Cooperation protocols: Propose, Contract Net, Iterated contract net
3. Market mechanisms: English auction, Dutch auction,
4. Middle agent protocols: Brokering, Recruiting, Subscribe

Every ACL message is tagged with the protocol it assumes. Protocols were originally specified using simple interaction diagrams, now are they specified by AUML (Agent Unified Modelling Language - http://www.auml.org/).

### 4.5.2 Agent platforms and tools

Agent platforms support developers by providing a set of reusable components and services for the implementation and deployment of agents. Most of them are compliant with standards. In Europe, the most widely used platform is JADE and it can be considered as the reference FIPA compliant platform. Other platforms are more focused to support agent coordination, such as TuCSoN (Ricci and Omicini, 2003).

JADE (Java Agent DEvelopment Framework)[53] (see also Bellifemine et al, 2001) originates as a collaboration between the research labs of Telecom Italia (TILAB) and Univ. Parma, and currently is distributed as open source software under the terms of LGPL (Lesser General Public License Version 2). JADE illustrates well the implementation of FIPA management architecture components: the Agent

[53] Java Agent DEvelopment Framework (JADE): http://jade.cselt.it/

Communication Channel, the Agent Management System and the Directory Facilitator. Agent communication is performed through message passing, where FIPA ACL is the language to represent messages, and with libraries that implement FIPA protocols, which can be used as reusable components when building agent-based applications. This facilitates the task of developers who can rely on agent lifecycle management by JADE and have some guarantee of interoperability with other FIPA compliant agent systems. JADE supports both reactive and deliberative agents by defining a structure for agent behaviours, which can be Java classes implementing state machines or rule systems, by an integration of JESS in the platform. Furthermore, JADE provides basic set of tools for agent debugging and monitoring, and other common services such as yellow pages, logging and naming[54].

#### 4.5.2.1 Remote Monitoring Agent

The Remote Monitoring Agent (RMA) allows controlling the life cycle of the agent platform and of all the registered agents (see Figure ). The distributed architecture of JADE allows also remote controlling, where the GUI is used to control the execution of agents and their life cycle from a remote host.

#### 4.5.2.2 DummyAgent tool

The DummyAgent tool allows user to interact with JADE agents in a custom way. The GUI allows composing and sending ACL messages and maintains a list of all ACL messages sent and received. This list can be examined by the user and each message can be viewed in detail or even edited. Furthermore, the message list can be saved to disk and retrieved later. Many instances of the DummyAgent tool can be started as and where required.

#### 4.5.2.3 Sniffer Agent

As the name itself points out, the Sniffer Agent is basically a FIPA-compliant Agent with sniffer features. When the user decides to sniff an agent or a group of agents, every message directed to/from that agent/group of agents is tracked and displayed in the Sniffer Agent GUI. The user can view every message and save it to disk. The user can also save all the tracked messages and reload it from a single file for later analysis.

[54] See: JADE Administrator's Guide, 2005 http://jade.cselt.it/doc/administratorsguide.pdf

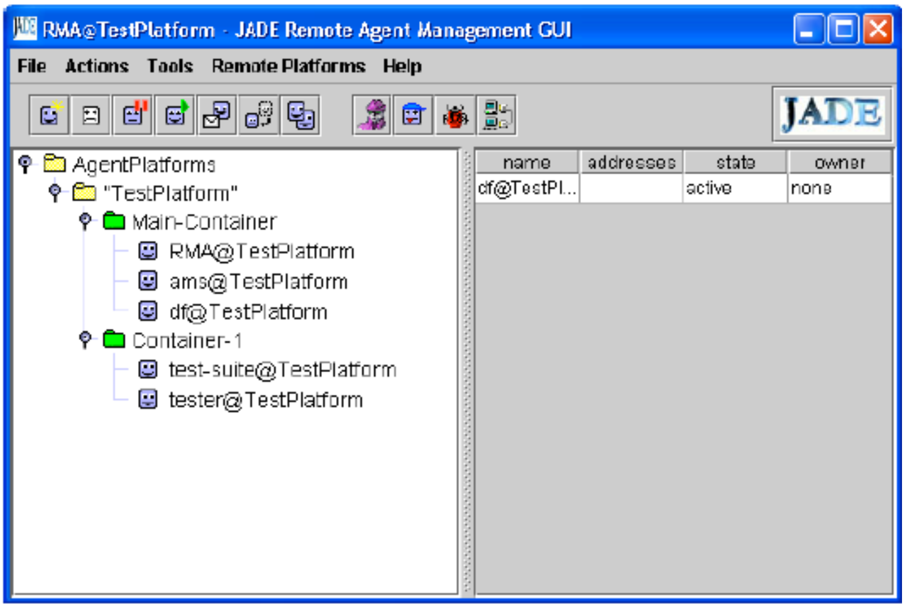



**Figure 17 Snapshot of the RMA GUI55**

#### 4.5.2.4 Introspector Agent

The Introspector Agent tool allows monitoring and controlling the life-cycle of a running agent and its exchanged messages, both the queue of sent and received messages. It allows also monitoring the queue of behaviours, including executing them step-by-step.

#### 4.5.2.5 Directory Facilitator Agent

JADE offers a graphical interface to the Directory Facilitator (DF) agent. This GUI allows to federate with other DFs and to control (i.e. register, deregister, modify and search for agent descriptions) all the network of federated DFs. The GUI and the code of the DF itself can be reused to implement user-defined DF.

The DF provides "yellow pages" services to other agents. Agents may register their services with the DF or query the DF to find out what services are offered by which agents. At least one DF must be resident on each Agent Platform (the default DF). However an AP may support any number of DF's. DF's can register with each other building a federation of DF's.

### 4.5.3 Methodologies for creating agent-based systems

Despite many of relevant results, multi-agent systems have not become widespread as industrial and commercial applications. In order to bridging the gap between agent technology and methodologies or technologies accepted for real world applications some efforts have been done. A survey of agent-oriented methodologies can be found in Shehory and Surm, (2001). Some of the most interesting results are Gaia (Zambonelli et al, 2003) (methodology for agent-oriented analysis and design supporting macro/societal level as well as micro/agent level aspects), MaSE (Deloach et al, 2001) (object-oriented

[55] JADE Administrator's Guide, 2005 http://jade.cselt.it/doc/administratorsguide.pdf

approach for support the complete software lifecycle from problem description to realization) and Prometheus (Padgham and Winikoff, 2002) (iterative methodology covering the complete software engineering process and aiming at the development of intelligent agents using goals, beliefs, plans and events).

As an example, overview of the methodology closely related to the JADE platform proposed in Nikraz et al (2006) is depicted at Figure 18.

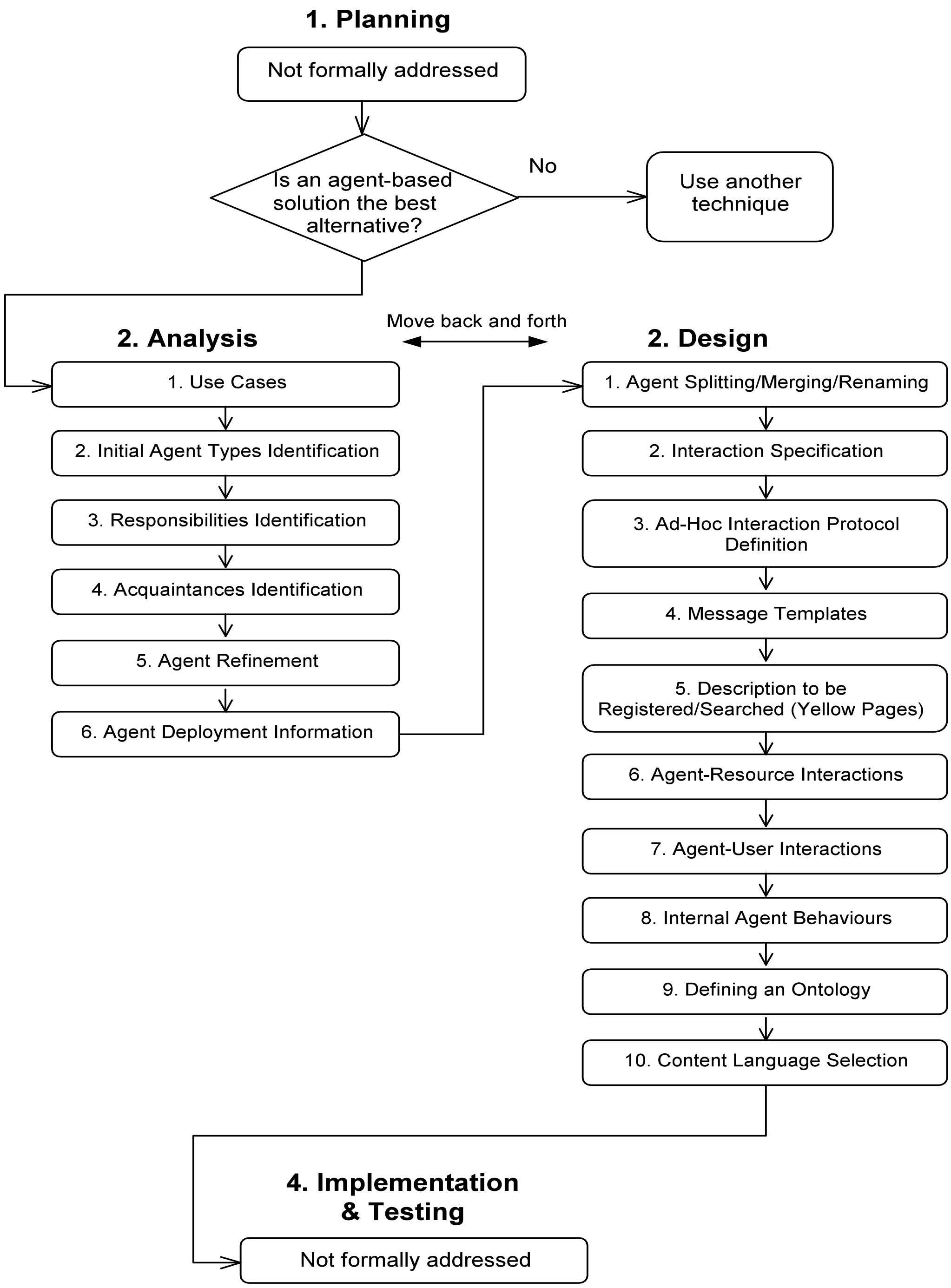



**Figure 18 Overview of a methodology using JADE by Nikraz et al (2006)**

This methodology does not attempt to extend object-oriented techniques, instead focusing on agents specifically and the abstractions provided by the agent paradigm. It combines a top-down and bottom-up approach so that both existing system capabilities (including those provided by legacy software and people) and the applications overall needs (based on the requirements) can be accounted for.

The methodology attempts to formalize the analysis and design phases of the agent-based software development life cycle. The formalization of the planning and implementation phases of the software development life cycle is currently outside the scope of the methodology.

## 4.6 Content syndication in knowledge management

Content syndication is a blanket term used to refer to accessing and publishing web content (text, images, etc.). Content publishers can make their content available through syndication by using available technologies to produce what is known as 'feeds' (e.g. 'blog feeds' or 'news feeds'). These feeds can either show headlines only, headlines and summary, or full content. The focus is mainly on dynamic content that allows people to share information and to interact. In general, content syndication refers to making feeds available from a site in order to provide other people an updated list of content from it (for example one's latest forum postings, etc.). This originated with news and blog sites but is increasingly used to syndicate any information. Anything that can be broken down into discrete items can be syndicated: the "recent changes" of a wiki, a changelog of CVS checkins, even the revision history of a book, etc.

Readers and/or fellow web publishers can access the content (e.g. latest updates) of particular sites with content syndication when they use aggregators and/or feeds generators. Once information about some item(s) is in appropriate format, a feed-aware program can check the feed for changes and react to the changes in an appropriate way.

Content syndication really only covers delivery of content items; it doesn't deal with storage of stuff or keeping track of relationships or anything like that. RSS/Atom can represent a solution to many problems both on the input side as well as on the output side. On the input side, it enables the integration of many (hundreds) of feeds which are skimmed all at once, with much less effort than having to visit all of the content resources individually. It is good on the output side, too - it gives a nice smooth way of getting information (hints, alerts, required information, etc.) out to people who want to read it. Therefore, many solutions (both commercial and open sourced) for knowledge management try to integrate syndication technology seamlessly into its interfaces.

Currently, two leading technologies heavily used are RSS and Atom. Both of them are XML based formats (files must conform to the XML 1.0 specification, as published on the World Wide Web Consortium website) with elements that enable the description of channels and items within these channels. To describe channels and items in more detail, various tags for different bits are used. Although enormously successful, the currently used technologies has one disadvantage from the point of knowledge communication – they are not able to communicate knowledge in a targeted way since all users subscribed to feeds receive the same chunks of content. In order to overcome this drawback, new extensions to basic technologies have emerged.

### 4.6.1 Individualised feeds

IRSS (Individualised RSS) is an evolution of the RSS standard which will allow content to be published in a way that can be targeted, measured and individualised[56]. This means that every subscriber to an RSS feed can receive unique content meant only to him or her or a specific group. It allows for fully individualised communications such as alerts, notices and targeted promotions.

Individualised RSS feeds allow content providers to target, segment and individualise communications much the way they do email messages today. Individualised RSS recipients receive text, images and other bits of content uniquely matched to their expressed interests and desires. The individualised feeds enable providers to communicate with subscribers based on demographics, past behaviour, or any other segmenting attributes.

With these solutions, each recipient gets his or her own unique feed, enabling providers to understand exactly how many and which recipients are picking up their messages. And because each feed is unique to the individual recipient, providers can track and measure subscriber actions all the way down to an individual, facilitating the same behavioural targeting and testing possible in other individualised media. Moreover, providers can actually create a unique message for each user based upon demographic or behavioural data.

But best of all, these individualised RSS solutions do not require any changes on the part of recipients – they can use the same reader they use today.

Three types of IRSS can be identified:

- Metric Enabled, using unique URLs to identify unique users, but their content and structure are always the same. The solution to this problem is adding some additional metadata to the RSS specification, which would allow the aggregators to cache the feed, but still enable the metrics.
- Customised Feeds carry different content items for different users. The content items themselves are the same, but different users will get different items. The solution to this problem is adding additional meta-data to the content item itself, to let the aggregators identify individual content items, regardless of what feed from a certain publisher they appear in.
- Personalised Feeds deliver different content items, for example by including the name of the recipient and data unique to that recipient. This can be solved by metadata, which would tell the aggregators that this content item in fact is unique.

The idea of individualised feeds is being mostly connected with direct marketing activities. An example of such application is where customers must explicitly provide their interest profile data[57]. But it is not the only possible application field. For example, an idea of using IRSS can be used in e-book content distribution[58]. Users can subscribe serially to any of the e-books. Via RSS, they can read a new book - obtaining a few pages

[56] http://www.razorshine.com/archive/2005/10/26/irss-the-evolution-of-rss/

[57] http://www.coravue.com/exec/gt/tpl.h,content=346&

[58] Cf. http://www.russellbeattie.com/notebook/1008220.html

at a time. It is possible to set how often pages are updated or to return to a page already visited.

Another example is CustomScoop[59] which is offering its service on the market of on-line clipping service providers - commercial organisations that filter traditional and on-line media according to customised query profiles. Its customers can now integrate results from other search engines, blogs and individual publishers through RSS. The new integration allows customers to merge news clips from the daily email delivery with items from selected RSS feeds. The results from this merge may then be exported again as individualised RSS feeds.

A solution based on building a social filtering network is provided by Findory[60], which launched RSS feeds for individualised search for news and blogs. Any article customers read through the RSS feed is included in their reading history, and at the same time it teaches more about their interests. The individualised RSS feeds are one-to-one. The feeds are built just for the particular customers with recommended news and blog articles. The individualisation algorithm analyses individual articles, what others have read, and customer's reading history to build individualised pages for the customer. It is as if the entire community of readers shared what they found and quietly recommended interesting articles to each other. Everyone is helping everyone discover articles they might enjoy, all anonymously, all with no effort.

### 4.6.2 Attention.XML

There is one problem with using content syndication technologies in practice - feed readers collect updates, but with too many unread items, how do you know which to read first? Attention.XML[61] is designed to solve these problems and enable a whole new class of blog and feed related applications. It is an open standard, built on open source that helps keep track of what people have read, what they are spending time on, and what they should be paying attention to.

Attention.XML is an XML file that contains an outline of feeds/blogs, where each feed itself is an outline, and each post is also an outline under the feed. This hierarchical outline structure is then annotated with per-feed and per-post information which captures such information as, the last time the feed/post was accessed, the duration of time spent on the feed/post, recent times of feed/post access, user set (dis)approval of posts, etc.

Basically it is metadata that records and shares information on the "attention" users give to their RSS feeds and blogs. Attention.xml basically provides a way of describing aspects of a user’s visits to a blog/feed/page/post/item/entry in a machine-readable fashion. This is information that could be extremely useful if captured, to both clients and servers of feeds. Attention.xml could tell us who looks at a blog or feed, how often they look at it, where those viewers come from.

According to its blog, Yahoo! experiments with Attention.XML within 'My Web', Yahoo's new personal search engine integrated with Yahoo! Search[62].

---

59 http://www.masternewmedia.org/news/2005/02/21/aggregate_rss_content_feeds_into.htm

60 http://findory.com/help/personalization

61 http://developers.technorati.com/wiki/attentionxml

62 http://www.ysearchblog.com/archives/000104.html

## 4.7 Emerging Collaborative KM Tools and Technologies

### 4.7.1 Making recommendations

One of commonly performed tasks is selecting one ore more items from a collection of items. This selection can have various forms, e.g. selection of an item from a department store catalogue, selection of a book to buy, or selection of a document to read. The last is relevant for the field of knowledge management – given a knowledge repository where knowledge is represented in the form of text (various document formats, web pages, etc.), an individual faces the problem which text to read in order to find knowledge of some importance for him/her. The original recommendation scenario was filtering through available documents to decide which ones were worth reading.

Recommender systems use the opinions of a community of users to help individuals in that community more effectively identify content of interest from a potentially overwhelming set of choices (cf. Resnick and Varian, 1997). The aim of such systems is to suggest specific items to their users, providing users with a ranked list of the recommended items, along with predictions for how much the users would like them. This is the core recommendation task and it recurs in a wide variety of research and commercial systems. In many commercial systems, only the "best bet" recommendations are shown. There are some modified alternatives as finding all good items.

Many recommendation services allow users to subscribe to syndication feeds (RSS). This allows subscribers to become aware of new resources (for some topic) they could be interested in. In this way, such systems play the role of knowledge repositories with active distribution of knowledge.

### 4.7.2 Collaborative filtering

One of the most successful technologies for recommender systems is called collaborative filtering. The first system to use collaborative filtering was the Information Tapestry project at Xerox PARC. This system allowed users to find documents based on previous comments by other users. Currently, some of the highest profile web sites like Amazon.com, CDNow.com, MovieFinder.com, and Launch.com have made successful use of the technology.

Automated collaborative filtering (ACF) systems predict a user's affinity for items or information/knowledge. Unlike traditional content-based information filtering system, such as those developed using information retrieval or artificial intelligence technology, filtering decisions in ACF are based on human and not machine analysis of content (Herlocker et al, 2000). Each user of an ACF system rates items that they have experienced, in order to establish a profile of interests. The ACF system then matches together that user with people of similar interests or tastes. Then ratings from those similar people are used to generate recommendations for the user.

The underlying assumption of this technology approach is that: those who agreed in the past tend to agree again in the future. Based on this assumption, collaborative filtering systems usually take the following steps:

- a large group of people's preferences are registered;

- using a similarity metric, a subgroup of people is selected whose preferences are similar to the preferences of the person who seeks advice;
- a (possibly weighted) average of the preferences for that subgroup is calculated;
- the resulting preference function is used to recommend options on which the advice-seeker has expressed no personal opinion as yet.

If the similarity metric has indeed selected people with similar tastes, the chances are great that the options that are highly evaluated by that group will also be appreciated by the advice-seeker.

Borghoff and Pareschi (1998) present a knowledge pump[63] system which uses a prediction algorithm employing several components (average population-wide rating, average advisors' rating, and correlation-weighted sum of ratings) in order to function under different circumstances – including the phase of the first deployment of the system.

The above given scheme is valid for so called active user-based filtering. Precondition for active filtering is the fact that the people want to and ultimately do provide information regarding the matter at hand.

ACF has many significant advantages over traditional content-based filtering, primarily because it does not depend on error-prone machine analysis of content (the recommender system has no idea what the content is). The advantages include the ability to filter any type of content, e.g. text, art work, music, mutual funds; the ability to filter based on complex and hard to represent concepts, such as taste and quality; and the ability to make serendipitous recommendations.

It is important to note that ACF technologies do not necessarily compete with content-based filtering. In most cases, they can be integrated to provide a powerful hybrid filtering solution.

Using active filtering, evaluations of recommender systems focus on the explicitly given recommendations; however if users don’t rate items, then collaborative filtering recommender systems can’t provide recommendations. Thus, voluntariness of users to contribute ratings is an essential condition for a recommender system employing active filtering to be successful.

An alternative method of collaborative filtering is called passive filtering. It is based on implicit observations of normal user behaviour. This method does not rely on user’s ratings but it collects information implicitly. The implicit filters are used to determine what else the user will like and recommend potential items of interest. Implicit filtering relies on the actions of users to determine a value rating for specific content, such as repeatedly using, saving, printing an item, etc. An important feature of passive collaborative filtering is using the time aspect to determine whether a user is scanning a document or fully reading the material.

The advantage of passive filtering is broader potential than of active approach. The reason is that only certain types of people will take the time to rate an item/document, while in passive collaborative filtering anyone accessing the repository has automatically given data. Implicit collection of user preferences does not involve the direct input of opinion from the evaluator user, but rather they input their opinion through their actions while working with the knowledge repository. This reduces the demand on the user (avoids cognitive aspects) and it reduces variables amongst users.

[63] See also section 3.4

Item based filtering (Sarwar et al, 2001) is another method of collaborative filtering. This type of filtering was popularized by Amazon.com (users who bought x also bought y). It proceeds in an item-centric manner:

- Build an item-item matrix determining relationships between pairs of items
- Using the matrix, and the data on the current user, infer his taste

ACF systems are stochastic processes that compute predictions based on models that are heuristic approximations of human processes. They base their computations on extremely sparse and incomplete data. These two conditions lead to recommendations that are often correct, but also occasionally very wrong. To be reliable, the system needs a very large number of people (typically thousands) to express their preferences about a relatively large number of options (typically dozens). Therefore, the system only becomes useful only after a "critical mass" of opinions has been collected.

### 4.7.3 Social bookmarking

Social bookmarking tries to organize content items using tags. Social bookmarking sites are an increasingly popular way to locate, classify, rank, and share Internet resources through the practice of tagging. The concept of shared on-line bookmarking dates back to 1996 with the launch of itList.com. The contemporary concepts of social bookmarking and tagging took root with the launch of the web site del.icio.us in 2003. The idea is also exemplified by CiteULike[64] – a social bookmarking site for academic papers, and Yahoo's MyWeb2[65].

In a social bookmarking system, users create, store, and manage lists of resources, which they find useful. Bookmarking systems rank the resources by the number of users which have bookmarked them. The technology of social bookmarking is based on an idea that as people bookmark resources that they find useful, resources that are of more use are bookmarked by more users. Thus, such a system will "rank" a resource based on its perceived utility. This is arguably a more useful metric for end users than other systems e.g. traditional search engines which rank resources based on the number of external links pointing to it.

Users also categorize the resources by the use of informally assigned, user-defined keywords or tags. Multiple tags allow bookmarks to belong to more than one category. This results in a user-directed amateur method classifying information. Most social bookmarking services allow users to search for bookmarks. Although bookmark collections are personally created and maintained, they are typically visible to others which are associated with given "tags," and order them according to their rank. Activities like social bookmarking give users the opportunity to express different perspectives on information and resources through informal organisation structures. In this way, the community of users over time will develop a unique structure of keywords to define resources - something that has come to be known as a “folksonomy” (Guy and Tonkin, 2006).

All tag-based classification of resources (such as web sites, web pages, documents, etc.) is done by human beings, who understand the content of the resource, as opposed to software which algorithmically attempts to determine the meaning of a resource. This

64 http://www.citeulike.org/

65 http://myweb2.search.yahoo.com/

provides for semantically classified tags, which is similar to annotating resources using ontologies as embodied in the idea of the Semantic web. There are drawbacks to such tag-based systems as well: no standard set of keywords (controlled vocabulary), mis-tagging due to spelling errors, tags that can have more than one meaning, and no mechanism for users to indicate hierarchical relationships between tags.

Although bookmark collections are personally created and maintained, they are typically visible to others. Visitors to social bookmarking sites can search for resources by keywords (search for resources that have been assigned that tag), person (to get a sense of the topics of interest), or popularity and see the public bookmarks, tags and classification schemes users have created and saved.

Many social bookmarking services also have implemented algorithms to draw inferences from the tag keywords that are assigned to resources by examining the clustering of particular keywords, and the relation of keywords to one another. It is possible to employ social network analytical methods to understand the information affinities among users.

The apparent success of Internet-based social bookmarking applications begs the question of whether large enterprises or organizations would also benefit from social bookmarking systems. To investigate this question, IBM has designed an enterprise-scale social bookmarking system called “dogear” (Millen et al, 2005), which plays the role of an information resource which can be easily integrated with other corporate applications.

# 5 Potential Uses of Existing Technology for e-Participation

## 5.1 Applications in e-participation

Examples of large KM approaches in the context of e-participation are still very rare. Some examples to mention are:

- Knowledge portals (see e.g. Webocrat, section 4.1.1.1) are an example being partially used in e-government and e-participation settings to support knowledge and information sharing in e-participation contexts.
- OLAP (see section 4.4.6) has been successfully used in e-participation tools, based on MS SQL server, where the technique is used to extract, amongst others demographic data to select target groups for participation based attributes, to monitor usage of respective participation tools and to evaluate the participation for validation of results. All known demographic data is deposited in a data warehouse where it can be retrieved by analysis services to form the multidimensional data cubes. The information can then be retrieved by the use of specialized viewer - interfaces such as Cognos and Knosys or a simple Excel-based plug-in, to extract information from the warehouse cubes. The use of specialised OLAP interfaces allows access to the data warehouse to be given to non-technical staff and reduced the training requirement for more skilled users as no programming language skills are required to extract data from the warehouse. One major benefit is the promotion of participation activities as the solution provides online access to data and can therefore be used market the activity to a particular group through selected channels.

## 5.2 Future Emerging Scenarios of KM in e-Participation Contexts

It seems that KM processes and tools offer a large potential to support e-participation. Yet, this potential has not yet been explored and exploited.

In the future, e-participation processes and application areas could be enriched and supported with knowledge management processes. Assuming that citizens and interest groups, as well as elected representatives engage more intensely via ICT, discussions may take place in a virtual space. In order to participate, the stakeholders need the right information and knowledge to support their arguments. For these purposes, tools and technologies for KM, knowledge engineering and overall knowledge management processes such as supporting the distribution of knowledge, the creation of knowledge and the usage of knowledge are important. Technical support is required in many aspects:

- Structuring knowledge,
- Information and knowledge retrieval,
- (Semi)automatic processing of knowledge via reasoning, intelligent agents, analytical processing, etc.
- Dissemination of knowledge,

- Visualization of knowledge,
- Assessment and evaluation of information and knowledge,
- And possibly more

It is important to note that KM and knowledge technologies are not specific to e-participation. However, these are crucial means to support the various stakeholders in the different areas and types of e-participation as introduced in deliverable D 5.1.

Participation works via a complex network of cooperation of different actors and stakeholders. Several ICT tools exist, which support collaborative work. Knowledge management has to be carefully integrated into these tools to make information available and accessible, and to make it automatically computable. The demands for knowledge enriched collaboration in e-participation can be categorized (and especially assigned to) three levels of participation as settled in D 5.1 (the first – informing – being addressed in dissemination of knowledge):

The lowest level of collaboration in e-participation is consulting. It takes effect when official initiatives by public or private agencies encourage stakeholders to contribute their opinion on specific issues. The task of knowledge management in this context is to present opinions in a well structured way so that participants are able to share and perceive knowledge efficiently and accurately. Besides presenting and visualizing added content, knowledge management should help the contributor to understand the intention of the questioner and to add arguments properly.

Real collaboration consists of collective elaboration of issues like it is often found to be implemented with groupware which includes several collaborative functions. In addition to informing other parties of their own opinions and providing means of commenting a particular subject, as it is done in consulting, an enhanced discussion of contents has to be accomplished. The aim is to produce joint decisions (e.g. proposing and shaping policy) in an environment of partnership. The multiple processes of negotiation and acquiring knowledge have to be managed properly to achieve reproducible results and to show the traces of argumentation and evolution of decisions. Unfortunately, the intensity of negotiations necessary, their length, their course of argument, and the number of parties that are involved are often not foreseeable. Therefore support systems should have a set of highly modular components and services in terms of information and knowledge management. These also need to be flexible and open for ad-hoc needs of collaboration.

Key functionalities that support collaboration include:

- appropriate and convenient management of electronic documents and shared workspaces for providing common views on a particular subject;
- various forms of conferencing on the desktop (bulletin boards, simultaneous conferences, video conferences);
- collaborative writing and white-boarding enabling revisions, comments, and annotations in a shared document;
- idea processing and argumentation focusing on the material content of negotiation and decision making.

These functionalities can be seen as part of a knowledge creation process (see the core processes of knowledge management) that aims at the development of previously non-existent knowledge (especially in groups). The distinct process of knowledge creation depends on the e-participation area it is adapted to.

Some of the participation areas (cf. D 5.1), in which knowledge management could come into consideration, are:

- Community Building / Collaborative Environments
- Consultation
- Deliberation
- Discourse
- Mediation
- Spatial Planning

Based on real collaboration, one may argue that empowering goes an extra mile within the bounds of the decision-making process. As defined in D 5.1, empowering refers to the placement of the final decision in the hands of the public. This would require knowledge enhanced collaborative e-participation tools to provide means for final casting of votes as well.

Empowering expands the list of participation areas to be considered for knowledge management presented above with the following:

- Electioneering
- Voting

# 6 Research Needed to Advance KM for e-Participation

In a recent Demo-net workshop[66] dedicated to KM in e-participation, the following needs to research KM in e-participation contexts have been identified:

- Proper understanding of the needs of KM in e-participation contexts
  - Does e-participation have some unique characteristics that require specific KM tools?
  - Are there differences in knowledge representation, knowledge repositories and knowledge communities between corporate KM and KM for e-participation?
  - Do we need different representation of knowledge for different stakeholders (e.g. citizens and politicians)?
  - How much information should be presented to the citizens? Different amounts, level, granularity of information is required for different stakeholders and situations.
- What kind / types of knowledge need to be managed in e-participation?
  - Rational knowledge, factual, policy-related knowledge
  - Community knowledge, opinions
  - Visceral, emotional knowledge, instinct
  - Rationale for previous decisions
- What aspects do impact the use of knowledge / knowledge processes in e-participation contexts?
  - Trust of knowledge
  - Source of knowledge
  - Level of granularity of knowledge
  - Traceability of contributed (community) knowledge
  - Evolution of contributed knowledge
  - Analysis, structuring, clustering and synthesis of text-based knowledge
  - Scalability and personalization
- Which KM techniques and tools are available and are suitable for which e-participation context?
  - Which KM processes to support (see section 3.2.1) – first we need to identify and analyze the e-participation processes
  - Which current tools need to be exploited, and how
  - Some specific processes that require KM support: conflict resolution, deliberation, consensus building, documentation and traceability of decision making
- What are the knowledge objectives in e-participation contexts?
  - taking informed decisions, better quality decisions, failure in not getting policies right, problem solving, etc.
  - Other, more general eParticipation objectives: conflict resolution, consensus building, making decisions without social break-down
- What are the challenges related to KM processes, technologies applied to e-participation:

[66] The workshop took place on 5th of December, 2006, at ICCS, Athens. The minutes were recorded by Dimitris Apostolou (ICCS).

- Integration of KM processes into e-participation contexts
- Knowledge for managing innovation
- KM for recording decisions
- Mapping KM processes to the policy lifecycle (see D 5.1)
- Ownership of knowledge and KM processes
- Different stakeholders have parts of ownership
- Role of facilitator in the e-participation context
- Scope of different types of knowledge
- Instrument for evaluating different types of knowledge
- Subsidiarity of knowledge – relevant to the appropriate government level (local, regional, etc.)

The concepts, tools and technologies introduced in this sub-deliverable indicate that there is a large potential for applying proper KM processes and support tools in e-participation contexts. However, focused research is needed to gather a better understanding and to develop recommendations and guidelines on how to implement and integrate KM in e-participation tools and applications.

# References


Abecker A., "*Business process oriented Knowledge Management – Concepts, Methods and Tools*", PhD Thesis. University of Karlsruhe, Faculty of Economics. 2004.

Abecker, A.; Decker, S.; Kuehn, O.: Das *aktuelle Schlagwort: "Organizational Memory"*. In: Informatik Spektrum 21 (4) 1998, pp. 213 – 214

Adams, M., D.Edmond, and A.H.M. ter Hofstede. *The Application of Activity Theory to Dynamic Workflow Adaptation Issues*. Proceedings of the Pacific Asia Conference on Information Systems , Adelaide, Australia, June 2003.

Aden, J.; Gora, W. (Eds.): *Informationsverbund Berlin-Bonn*. Fossil, Koeln 1999

Agrawal, R., T. Imielinski, and A. Swami, *"Mining Association Rules Between Sets of Items in Large Databases,"* Proc. 1993 ACM-SIGMOD Int'l Conf. Management of Data, pp. 207-216, May 1993.

Agrawal, R. and R. Srikant, *"Fast Algorithms for Mining Association Rules,"* Proc. 1994 Int'l Conf. Very Large Data Bases, pp. 487-499, Sept. 1994.

Baeza-Yates R., Ribeiro-Neto B.: *Modern Information Retrieval*. New York, NY: ACM Press, 1999

Bardram, J.E.: *Plans as Situated Action: An Activity Theory Approach to Workflow Systems*. In Proc. of the ESCW'97 Conference, Lancaster, UK, September 1997

Bellifemine, F., A. Poggi, G. Rimassa: Developing *multi-agent systems with FIPA-compliant agent framework*. Software Practice and Experience 31(2), 2001, pp. 103-128

Bernon, C., M. Cossentino, J. Pavón: *An Overview of Current Trends in European AOSE Research*. In Informatica Vol. 29, No. 4, November 2005, pp. 379-390

Bhandari I., Colet E., Parker J., Pines Z., Pratap R., Ramanujam K.: *Advanced Scout: Data Mining and Knowledge Discovery in NBA Data*. Data Mining and Knowledge Discovery, Volume 1, 1997, p. 121 – 125

Bonifacio M., Bouquet P., Cuel R. *Knowledge Nodes: the Building Blocks of a Distributed Approach to Knowledge Management*, Journal of Universal Computer Science, vol. 8, 2002

Bordini, R. H., L. Braubach, M. Dastani, A. F. Seghrouchni, J. J. Gomez-Sanz, J. Leite, G. O'Hare, A. Pokahr, A. Ricci: *A Survey of Programming Languages and Platforms for Multi-Agent Systems*. In: Informatica Vol. 30, No. 1, January 2006, pp. 33-44

Borghoff, U. M.; Pareschi, R.: *Information Technology for Knowledge Management*. Springer Verlag,1998.

Bots, P.; Sol, H.; Traunmueller, R. (Eds.): *Proceedings of the IFIP TC8/WG8.3 Working Conference on Decision Support in Public Administration*. North Holland, Amsterdam 1993

Brachman, R.J., and Anand, T. *The Process Of Knowledge Discovery In Databases: A Human-Centered Approach*. In Advances In Knowledge Discovery And Data Mining , eds. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, AAAI Press/The MIT Press, Menlo Park, CA., 1996, pp. 37-57.

Bredahl ACH., Rydén M.: *The Scandia Navigator*. 2002

Breiter, C., Scardamalia, M.: *Surpassing ourselves: An inquiry into the nature and implications of expertise*. La Salle, IL: Open Court, 1993.

Buntine, W. *Graphical Models For Discovering Knowledge*. In Advances In Knowledge Discovery And Data Mining, eds. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, AAAI Press/The MIT Press, Menlo Park, CA., 1996, pp. 59-82.

Cole, Michael; *Cultural Psychology, A Once and Future Discipline*, Belknap Press, Cambridge, 1998

Cole, Michael; Engeström, Yiryi. *A cultural-historical approach to distributed cognition*, in: Salomon (Ed.), Distributed Cognition, Psychological and Educational Considerations, Cambridge University Press, New York, 1993, pp. 1-46.

CS Solutions. Case *Study: Web-Based International Trade Knowledge Discovery System*

Dastani, M.; M. B. van Riemsdijk, F. Difnum, J. C. Meyer: *A programming language for cognitive agents: goal directed 3APL*. In: Proceedings of the 1st int. workshop Programming Multiagent Systems (ProMAS'03), LNCS vol. 3067, Springer Verlag, 2004, pp. 111-130

Deloach, S. A., M. F. Wood, C. H. Sparkman: *Multiagent Systems Engineering*. International Journal of Software Engineering and Knowledge Engineering. Vol. 11, No. 3, 2001, pp. 231-258

Di, G. ,M. Serugendo, M.-P. G. Irit, A. Karageorgos: *Self-Organisation and Emergence in MAS: An Overview*. Informatica Vol. 30, No. 1, January 2006, pp. 45-54

Diaper, D., *Task Analysis for Human-Computer-Interaction*, Ellis Horwood Ltd., 1989

Edelstein H.: *Introduction to Data Mining and Knowledge Discovery, Third Edition.* Two Crows Corporation, 1999

Edwards, E. *Introductory overview,* In Wiener, Nagel (eds.), *Human Factors in Aviation,* Academic Press, San Diego, 1988

Edwards, E. *Man and machine: Systems for safety*. In Proc. of British Airline Pilots Associations Technical Symposium, British Airline Pilots Associations, London, 1972, pp. 21-36

Engeström, Y.: *Innovative Learning in Work Teams: Analyzing Cycles of Knowledge Creation in Practice*. Cambridge, MA: Cambridge University Press, 1999.

Engeström, Y., Miettinen, R. N, Punamäki, R.-L. *Perspectives on Activity Theory: Learning in doing – social, cognitive and computational perspectives*, Cambridge University Press, 1999

Fayyad, U.M., Piatetsky-Shapiro, G., and Smyth, P. *From Data Mining To Knowledge Discovery: An Overview*. In Advances In Knowledge Discovery And Data Mining , eds. U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, AAAI Press/The MIT Press, Menlo Park, CA., 1996, pp. 1-34.

Frakes, W. B. and Baeza-Yates, R. (eds): *Information Retrieval: Data Structures and Algorithms*. Prentice-Hall, 1992.

Frawley, W.J., Piatetsky-Shapiro, G., and Matheus, C. *Knowledge* Discovery In Databases: An Overview. In Knowledge Discovery In Databases, eds. G. Piatetsky-Shapiro, and W. J. Frawley, AAAI Press/MIT Press, Cambridge, MA., 1991, pp. 1-30.

Gamble P., Blackwell J.: *Knowledge Management - A State Of The Art Guide*. Kogan Page Limited, UK/USA, 2001

Glassey, Olivier. "*PETALE: Case Study of a Knowledge Reengineering Project",* In: Wimmer, M. (Ed.). KMGov 2004, LNCS 3035, Springer Berlin/Heidelberg, 2004, pp. 304-309.

Guy M, - Tonkin, E.: *Folksonomies: Tidying up Tags?* D-Lib Magazine, vol. 12, 2006, no1. http://www.dlib.org/dlib/january06/guy/01guy.html

Han J. - Kamber, M.: *Data Mining: Concepts and Techniques*. Morgan Kaufmann. (2001)

Hemetsberger A., Reinhardt Ch.: *Sharing and Creating Knowledge in Open-Source Communities, The case of KDE*. The Fifth European Conference on Organizational Knowledge, Learning, and Capabilities in Innsbruck, Austria, 2004.

Herlocker, J., Konstan, J., Riedl, J.: *Explaining Collaborative Filtering Recommendations*. Proceedings of the ACM 2000 Conference on Computer Supported Cooperative Work, December 2-6, 2000.

Huff A. (ed), "*Mapping Strategic Thought*", Wiley, NY, 1990.

IBM, *Business Process Execution Language for Web Services, 2005, see:* http://www-128.ibm.com/developerworks/library/specification/ws-bpel/

IEEE Standard 610.12-1990. *IEEE standard glossary of software engineering terminology*, 1990.

IEEE Standard 1233-1996. *IEEE guide for developing of system requirements specifications*,1996..

Jackson P., Moulinier I.: *Natural Language Processing for Online Applications: Text Retrieval, Extraction and Categorization*. John Benjamins Publishing Company, 2002.

KBSI: "*The IDEF5 Ontology Description Capture Method Overview*", KBSI Report, Texas, 1994.

Kirwan, B., Ainsworth, L.K. *A Guide o Task Analysis,* Taylor & Francis Ltd., London, 1992

Kraemer, L.; King, J.: *Computer-Based Systems for Cooperative Work and Group Decision Making*. In: ACM Computing Surveys 20, 1988, pp. 115 - 146

Lenk, K.; Brueggemeier, M.; Hehmann, M.; Willms, W.: *Buergerinformationssysteme. Strategien zur Steigerung der Verwaltungstransparenz und der Partizipationschancen der Buerger*. Westdeutscher Verlag, Opladen 1990

Lenk, K.: *Electronic Support of Citizen Participation in Planning Processes*. In: Barry N. Hague and Brian Loader, Digital Democracy. Discourse and Decision Making in the Information Age, Routledge, 1999, pp.87-95

Lenk, K.; Traunmueller, R. (eds.): *Öffentliche Verwaltung und Informationstechnik - Perspektiven einer radikalen Neugestaltung der oeffentlichen Verwaltung mit Informationstechnik*. Decker, Heidelberg 1999

Lenk, K., R. Traunmüller, M. Wimmer, *"The Significance of Law and Knowledge for Electronic Government"*, In A. Grönlund (ed.), *Electronic Government - Design, Applications and Management*, Ideas Group P., 2002, p 61-77

Mach, M., Macej, P., Hreno, J.: Ontology-*based Communication Forum*. Proc. of the Int. Workshop on Intelligent Knowledge Management Techniques IKOMAT'02, Podere d'Ombriano, Italy, 2002, 1544-1548.

Mach, M., Stofanik, V.: *Creating Dedicated Information Collections from the Web*. Proc. Of the 17th Int. Conference on Information and Intelligent Systems, Varazdin, Croatia, 2006, 321-325.

Menne-Haritz, A.: *Prozessgedaechtnis und UEberlieferungsbildung*. In: Metzing, (Ed.): Digitale Archive ein neues Paradigma? Beitraege des 4. Archivwissenschaftlichen Kolloquiums der Archivschule Marburg, Marburg 1999, pp. 283-308

Mentzas, G.N., D. Apostolou, R. Young and A. Abecker. "*Knowledge Asset Networking: a Holistic Approach for Leveraging Corporate Knowledge*", Springer-Verlag, 2002

Millen, D. - Feinberg, J. - Kerr, B.: Social *Bookmarking in the Enterprise*. ACM Queue, vol. 3, 2005, no. 9. http://acmqueue.com/modules.php?name=Content&pa=showpage&pid=344&page=1

Murdoch J., McDermid J.: *Modelling Engineering Design Processes with Role Activity Diagrams*. Society for Design and Process Science, 2000.

Nardi, B. A.: *'Activity Theory and Human-Computer Interaction' in Context and Consciousness: Activity Theory and Human Computer Interaction*. MIT Press, Cambridge, 1996.

NAVSEA, "*Community of Practice Practitioner's Guide, V1.0a*", Office of the NAVSEA Chief Information Officer, U.S. Navy, May 2001

Neches, R., R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, & W. R. Swartout. *Enabling technology for knowledge sharing*. AI Magazine, 1991.

Nonaka, I., Takeuchi, H.: *The Knowledge Creating Company*. Oxford University Press, New York, 1995

Nikraz, M., G. Caire, P. A. Bahri: *A Methodology for the Analysis and Design of Multi-Agent Systems using JADE*. International Journal of Computer Systems Science & Engineering, special issue on "Software Engineering for Multi-Agent Systems", to appear in 2006

Paavola S., Hakkarainen, K.: *"Trialogical" Processes of Mediation through Conceptual Artefacts*. Technical Report for the KP-Lab consortium, University of Helsinky, Finland, 2006

Paralic, J. – Bednar, P.: *A Tool to Support f the KDD Process*. In Journal of Information and Organizational Sciences, Varaždin, Croatia, Vol. 27 (2003), Nr. 1, pp. 15-27

Padgham, L., M. Winikoff: *Prometheus: A Pragmatic Methodology for Engineering Intelligent Agents*. In Proceedings of the workshop on Agent-oriented Methodologies at OOPSLA 2002. November 4, 2002, ACM Press, 2002,, pp. 97-108

Pendse N.: *OLAP report*, Market share analysis, last updated March 2006, http://www.olapreport.com/market.htm

Plumley D., "Process-*based knowledge mapping: A practical approach to prioritising knowledge in terms of its relevance to a business or KM objective*" Knowledge Management Magazine, March 2003, (http://www.destinationkm.com/articles)

Quinlan, J.R. *C4.5: Programs For Machine Learning*. San Mateo, CA: Morgan Kaufmann, 1993.

Raab S., Studer R.: *Knowledge processes and Ontologies*. On-To-Knowledge, 2001

Resnick, P., Varian, H. R.: *Recommender Systems*. Communications of the ACM 40, 1997, 56-58.

Ricci, A., A. Omicini: *Supporting Coordination in Open Computational Systems with TuCSoN*. In: 12th IEEE Int. Workshops on Enabling Technologies (WETICE 2003), Infrastructure for Collaborative Enterprises. IEEE Computer Society, pp. 365-370

Rossen, H: *Vollzug und Verhandlung*. Mohr, Tuebingen 1999

Russell, S. J., P. Norvig: *Artificial Intelligence: A Modern Approach*, Second Edition. Pearson Education, Inc. 2003

Sarwar, B. George Karypis, Joseph Konstan, and John Riedl: *Item-based Collaborative Filtering Recommendation Algorithms*. Proc of the 10th International World Wide Web Conference, 2001, 285–295.

Sasisekharan, R., Seshadri, V., and Weiss, S.M. *Data Mining And Forecasting In Large-Scale Telecommunication Networks*. IEEE Expert: Intelligent Systems & Their Applications 11, 1 (Feb. 1996), pp. 37-43.

Simon, H.: *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. 2nd Edition, Macmillan, New York 1957

Schreiber, G., Hans Akkermans, Anjo Anjewierden, Robert de Hoog, Nigel Shadbolt, Walter Van de Velde, Bob Wielinga. "*Knowledge Engineering and Management*, The CommonKADS Methodology", MIT-Press, 1999

Schwabe, G.: *Die Rolle neuer Informations- und Kommunikationstechnologie für die Bürgerinformation*. In: Information management 2, 1996, pp. 6 - 14

Seghrouchni, A. F., A. Suna: *CLAIM: A computational language for autonomous, intelligent and mobile agents*. In: Proceedings of the 1st int. workshop Programming Multiagent Systems (ProMAS'03), LNCS vol. 3067, Springer Verlag, 2004, pp. 90-110

Shehory, O., A. Sturm: *Evaluation of modeling techniques for agent based systems*. In Proceedings of th 5th International Conference on Autonomous Agents, Montreal, Canada, June 2001, pp. 624-631

Singh, M. P.; Huhns M. N.: *Service-Oriented Computing, Semantic, Processes, Agents*. John Wiley & Sons Inc., 2005

Staab S. , Studer R. , Schnurr HP., Sure Y.: *Knowledge Processes and Ontologies*, IEEE Intelligent Systems, v.16 n.1, p.26-34, 2001

SU-JEONG HWANG: *Knowledge Management (KM) for mass customization in the apparel business.* Raleigh, 2002

Suchman L.: *Response to Vera and Simon's Situated Action: A Symbolic Interpretation*. Cognitive Science, vol. 17, 1993, pp. 71-75.

Tiwana A., "*The Knowledge Management Toolkit*", Prentice Hall, 2000

Tsumoto S.: Clinical *Knowledge Discovery in Hospital Information Systems: Two Case Studies, Principles of Data Mining and Knowledge Discovery*: In proceedings of 4th European Conference, PKDD 2000, Lyon, France, September 2000

United Nations, "*Knowledge Management Methodology: An Empirical Approach in Core Sectors in ESCWA Member Countries*", E/ESCWA/ICTD/2003/9

Van der Veer, G.C., Lenting, B., Bergevoet, B. *GTA: Groupware task analysis - Modeling complexity*, Acta Psychologica 91, 1996

Watson I.: *Applying Knowledge Management: Techniques for Building Organisational Memories*. LNCS #2416, Proceedings of the 6th European Conference on Advances in Case-Based Reasoning, 2002, pp. 6 – 12

Wexler M., "*The who, what, why of knowledge mapping*", Journal of Knowledge Management, 5(3), pp. 249-263, 2001

Wimmer, M. Designing *interactive Systems – Key Issues for a Holistic Approach*. University of Linz, Trauner Verlag, Linz, 2000

Wimmer, M., Traunmüller, R. *Reflecting Distributed Knowledge in IT Systems: A Three Layers Concept*. In Proceedings of the DEXA workshops, IEEE Computer Society Press, Los Alamitos, CA, 2001, pp. 459-466

Wimmer, M., Traunmüller, R. *Knowledge assets in the public sector: some spotlights on administrative knowledge.* In Monica Palmirani, Tom van Engers, Maria A. Wimmer (Eds.) 2003. E-government - Workshop in conjunction with JURIX 2003. Proceedings, Schriftenreihe Informatik # 10, Trauner Verlag, Linz, 2003, pp. 41 - 53

Wimmer, M., Traunmüller, R., Lenk, K. *Electronic Business Invading the Public Sector: Considerations on Change and Design*. In Proceedings of the 34th Hawaii International Conference on System Sciences (HICSS-34), January 3-6, Maui, Hawaii, 2001.

WfMC-XPDL: *Workflow Process Definition Interface - XML Process Definition Language*, WFMC-TC-1025, 1.0 Final Draft, WfMC, October 25, 2002

Zambonelli, F., N. R. Jennings, M. Wooldridge: *Developing Multiagent Systems: The Gaia Methodology*. ACM Transactions on Software Engineering and Methodology. Vol. 12, No. 3, July 2003, pp. 317-370

IST Network of Excellence Project

FP6-2004-IST-4-027219

Thematic Priority 2: Information Society Technologies

**DEMO_net**

**The Democracy Network**

# D5.2.6 – Devices, Channels and Mobile Technologies

| | |
|---|---|
| **Editor :** | **Andreas Rosendahl, Maria Wimmer** |
| **Revision :** | 04 [final] |
| **Dissemination Level :** | [TA p. 63] |
| **Author(s) :** | Andreas Rosendahl, Christian Schneider, Maria Wimmer |
| **Due date of deliverable :** | 30th December 2006 |
| **Actual submission date :** | 6th February 2007 |
| **Start date of project :** | 01 January 2006 |
| **Duration :** | 4 years |
| **WP no.:** | 5 |
| **Organisation name of lead contractor for this deliverable :** | IWVI |

**Abstract:** The sub-deliverable 5.2.6 will separate communication channels, application technology and devices. It will show how these depend on each other and, therefore, will analyze the features, requirements, user preferences, means of service delivery as well as public value of devices and their specific technologies. The importance and impact of limitations, advantages, conditions, business models and the public value of proper channels and devices for the targeted e-participation arena as well as the intended stakeholders have to be taken into account.

**Project funded by the European Community under the FP6 IST Programme**

# Executive Summary

An inclusive European information society and the accessibility of services through a range of communication channels are crucial to enable eParticipation for all (cf. D5.1 section 6.4). To allow access for various user types and social groups, their specific means and channels for communication should be addressed. The terminus "channel" in D5.1 made no distinction between communication channels, application technology or devices.

The sub-deliverable at hand tries to clarify the different understandings of technology and devices on the one hand. On the other hand, their dependencies among each other will be shown. Consequently, features, requirements, user preferences, means of service delivery as well as public value of devices and their specific technologies will be analyzed. The importance and impact of limitations, advantages, conditions, business models and the public value have to be taken into account.

The sub-deliverable is organized as follows:

First, the introduction sets the scope and ground of understanding for devices and channels in e-participation.

Chapter 2 focuses on communication channels for eParticipation. Aspects such as flexibility of users, mode of transmission, infrastructures for bearer services, mobile provider services and data access, internet application and services, and general aspects of availability of channels and cots will be discussed.

In chapter 3, device classes will be analysed along their interaction and usability aspects, the primary channels used for transmission, and the types of applications these devices are used for. PCs, mobile phones and digital TV will be investigated.

Subsequently, the importance of mobile technologies and digital TV channels is discussed in view of eParticipation. Chapter 4 discusses therefore issues of diffusion, personalization and localization for mobile technologies and devices, as well as digital TV.

Chapter 5 investigates the impact of devices and channels on eParticipation

Chapter 6 concludes the report with reflections and an outlook.

# 1 Introduction

An inclusive European information society and the accessibility of services through a range of communication channels are crucial to enable eParticipation for all (cf. D5.1 section 6.4). To allow access for various user types and social groups, their specific means and channels for communication should be addressed. The terminus “channel” in D5.1 made no distinction between communication channels, application technology or devices.

The sub-deliverable 5.2.6 will not only separate these forms of technology but furthermore will show how they depend on each other. Therefore features, requirements, user preferences, means of service delivery as well as public value of devices and their specific technologies will be analyzed. The importance and impact of limitations, advantages, conditions, business models and the public value have to be taken into account.

In deliverable D1.1 (sections 5.6-5.8) important eParticipation challenges and needs to be kept in mind in regard to socio-technical, technological and deployment issues have already been specified. Hereafter these issues are sorted in respect to two important views:

- eInclusion / digital divide for the purpose to increase access to information for all, consequently referring to challenges and needs for considering the following aspects:
    - accessibility (barrier-free tools)
    - usability
    - age, gender, social context
    - public access terminals and mobile phones
    - inclusive deployment[67]
- eParticipation technology for the purpose to sustain eParticipation, i.e. challenges and needs to support:
    - the combination of online and offline participation
    - personalization options
    - identification means
    - mobile technology
    - use the interactive potential of ICT
    - understand the business models of eParticipation tools in respect to channels and devices.

Investigations are needed to properly understand the impact of multiple channels in eParticipation systems. In specific, any eParticipation tool being used in a certain eParticipation context has to have an interaction interface for any of the users of the tool or application. Different channels and devices are available for that purpose. A thorough consideration shall provide a proper understanding of employing the right devices in the targeted eParticipation arenas and for the intended target groups.

[67] Inclusive deployment will mean introducing a range of technologies and approaches simultaneously, which will be a challenge (cf. D1.1 section 5.8)

As stated in D5.1 (section 6.5), eParticipation services and information offers via electronic channels need to be simple, effective, easy-to-use and functional. Also the look-and-feel and the fun-factor should not be underestimated. A similar challenge has been recognized in consumer products: these have to be designed in the right way to achieve broad success. Lina Bonapace [6] introduces a hierarchy of user needs when interacting with products (shown in Figure 1). Besides pleasure (cf. fun-factor in D5.1), which takes the highest level in the consumer hierarchy, usability is a main aspect of user satisfaction. In contrast to that, functionality always stands back.

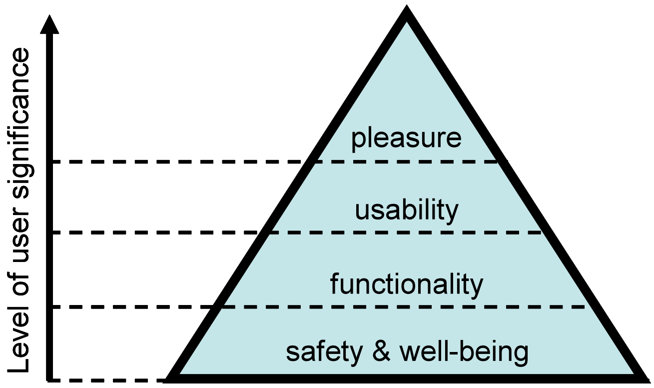



**Figure 1:** Hierarchy of user needs when interacting with products [5]

When considering important matters of channels and devices in terms of eParticipation, it is significant not just to treat the possibilities of certain technologies, but also to evaluate to what extent these technologies provide pleasure and usability. In the following sections considerations will be made by discussing

- communication channels available (chapter 2),
- categories to describe and categorize devices (chapter 3),
- the importance of mobile technologies for eParticipation (chapter 4), and
- the anticipated impact of devices and channels on eParticipation (chapter 5).

# 2 Communication Channels for eParticipation

When discussing communication channels for eParticipation, a distinction among the communication channels can be made on a conceptual level. Even if it is not easy to make a distinct classification between all kinds of technologies, notions of ICT like DSL, internet, telephony or email have to be investigated separately because these are not at the same conceptual level. To start with, we classify different points of view as follows, even though this structure may come under scrutiny:

1. Flexibility of Users
2. Mode of Transmission
3. Infrastructure
4. Mobile Provider Services
5. Mobile Data Access
6. Internet applications and services.

These views will be further described in the following sections of this chapter.

## 2.1 Flexibility of users

Communication networks can either base on **fixed** line technologies (POTS[68], ISDN, DSL) that cord their clients to a cable. Respectively users are not able to leave the house or the small range of their personal (unlicensed, cf. section 2.5) radio link that is attached to their connection. Other networks are inherently **mobile** and enable their user to move freely through provided areas. In terms of flexibility there are two types of communication networks:

- fixed
- mobile

Users are more flexible when using **mobile** networks. Unfortunately they have to deal with high costs and different restraints. Mobile technologies with higher transmission rates are often not available in areas of low population. **Fixed** lines usually have higher transmission rates and lower costs than mobile lines. Again broadband access is often left blank in the rural environments so that communication is confined to circuit switched connections that are almost ineligible for intensive use of modern internet applications.

## 2.2 Mode of Transmission

Two main modes of transmission can be differentiated, which characterize data communication:

- circuit switched
- packet switched

---

[68] Plain Old Telephone Service

**Circuit switched** connections are still standard for voice connections over telephone and used for internet-connections if necessary. Drawback is the time-based billing that naturally comes along with this technology. Actually there are flat rate products available but these are only valid for voice communication. So circuit switching is not appropriate for a permanent (resp. always-on) connection to the internet. Beside that, the transmission rates are quite low (56-128kbit/s). Therefore **packet switching** is the connection-mode that is adapted to the internet. It allocates just the bandwidth that is actually needed so it is possible to charge it based on the transferred data or flat rate. This also enables push-services. In the fixed-line sector, packet switching mostly comes along with broadband connection with transfer rates around several Mbit/s. Although packet switched connections are most preferable, in the majority of cases (especially in mobile solutions) they have one drawback: asymmetry of upload and download transmission rates. While a user is browsing the web or downloading email the transmission may be satisfactory, but uploads of content and information will need a period several times longer with asymmetric connections.

## 2.3 Infrastructure for Bearer Services

Several different technologies can be distinguished, which are being used to build up an infrastructure in order to provide clients with voice communication and internet, among them are:

- DSL
- Broadband cable
- POTS/ISDN
- Mobile networks
- Satellite
- Broadcast Networks

Information can be accessed over many different kinds of infrastructure. Some of them are originally intended for circuit switched telephony (**POTS/ISDN**, **mobile networks**) and others where introduced for broadband connections to the internet (**DSL**, **cable**). First, telephone lines where used to provide data-connections and now broadband internet connections are utilized to offer telephony at a favorable price (VoIP). Information over **satellite** connections can be accessed from all over the world but solutions at consumer cost need a conventional line for uplink and are mainly a substitute for broadband access in areas without DSL or cable. Besides that there are **broadcast networks** (one way), like TV or radio, which supply citizens with information by cable or over the air. Partially digital TV is provided with a backward channel, too.

## 2.4 Mobile Provider Services

Following major services are provided by mobile carriers:

- Telephony
- SMS (Short Message Service)
- MMS (Multimedia Messaging Service)
- Video telephony
- Data services (mainly internet)

The key providers of mobile services still make their main revenue by **telephony** and **SMS** – the core services they started with. In Europe these services are provided over networks that are based on, or evolved from the GSM standard. New technologies for data transfer and further services were added over the years. Besides **data services,** other value-added services like **MMS** and **video telephony** have been integrated. Compared to fixed line internet connections, mobile data access tends to be way more expensive.

## 2.5 Mobile Data Access and Bearer Services

Technologies for mobile data access can be divided into two groups:

- Licensed
  - CSD[69] (9,6 or 14,4 kbit/s) ([14], p. 172)
  - GPRS[70] (approx. 40 kbit/s) ([14], p. 178)
  - EDGE[71] (up to 384 kbit/s) ([27], p.40)
  - UMTS (approx. 384 up to 2000 kbit/s) ([14], p. 174)
  - (WiMAX)
- Unlicensed
  - WLAN/WiFi (11, 54, 108 Mbit/s) [29]
  - Bluetooth (1-3 Mbit/s)[72]

Technologies with a **licensed** spectrum of frequencies to provide services are e.g. GSM providers all over Europe. They licensed their part of the available wave bandwidth. Several systems extend GSM to provide and enhance mobile data access (e.g. **CSD**, **GPRS**, **EDGE**). New frequencies are mainly assigned by auction (e.g. 3G **UMTS**). **WiMAX** is a relatively new and uncommon standard which will rather appear as an alternative for DSL than as a real mobile solution [29]. Like consumer-DSL the licensed packed switched technologies are always asymmetrical, so that the maximum uplink speed is only a fraction of the actual transfer rate.

**Unlicensed** technologies can be installed and operated by anyone (**WLAN**, **Bluetooth**). They have a limited output power and range. Yet there are also providers that use WLAN hotspots for commercial broadband connections. Again, the more sophisticated and faster a data service is, the more focused it is around congested areas.

At the moment, GPRS (licensed model) is the only packet switched data service that is available in almost every area that is endued with GSM.

## 2.6 Internet applications and services

The different types of fixed line or mobile internet connections can provide IP access to a multitude of applications and services, such as:

- Email
- Web
- WAP

---

69 Circuit Switched Data

70 General Packet Radio Service

71 Enhanced Data Rates for GSM Evolution

72 http://www.bluetooth.com/Bluetooth/Learn/Technology/Compare/

- Downloads
- Streaming
- Instant messaging
- etc.

First these applications were exclusively PC-based. Main applications still are **email** and the World Wide Web (**web**). Bit by bit they have also been implemented on mobile networks. Some of them stayed basically unchanged (email) and others were customized to the constraints of mobile networks and devices (e.g. **WAP**). As mobile devices get more and more powerful they converge to a greater extend towards PC-based internet applications (cp. WAP 2.0 and web). Besides powerful devices, **streaming** implicitly needs a capable network to work properly. Other applications (e.g. **instant messaging**) work perfectly with relatively low transfer rates as long as it concerns a packet switched connection (e.g. GPRS).

## 2.7 Availability of Channels, and Costs

From a pan-European view on channels, costs and availability of mobile as well as stationary services and internet connections differ largely. Recently, some recommendations were made to reduce the roaming costs of mobile channels among European countries. Focused investigations will need to be made to further evaluate the pricing models of the different channels.

# 3 Analysis of Device Classes

The list of devices used as interaction interfaces in eParticipation tools and applications contains numerous electronic appliances that differ deeply in their principles. Since mobile technologies are of special importance when promoting eParticipation (cf. chapter 4) a classification of different device categories is useful. We use mobility as a point of separation: **Mobile phones**, **PDAs**, **tablet PCs** as well as **notebooks** are mobile in terms of their:

- size (they can be carried around more or less conveniently),
- integrated input and output interfaces (do not need additional periphery),
- independent power supply (in form of a rechargeable battery).

In consequence the user is able to operate these devices nearly location-independent. But this does not mean that they are all equally suitable for mobile applications. Figure 2 indicates that the particular level of mobility can be quite diverging and is often inversely proportional to the resources of the system (ability vs. mobility). **Notebooks** are endued with similar performance and operating systems as desktop PCs. Unfortunately they have to be carried with both hands during operation. To perform any kind of proper input, a notebook has to be placed on a plane surface. With its pointing device (touch pad or track point), a large display and a hardware-based QUERTY-keyboard, the modes of input and output are quite similar to fixed PCs. In opposition to that, a **tablet PC** can be used while standing (e.g. supported by arm and chest) because of its compact design and its possible use with a stylus as a pointing device.


**Figure 2: Mobility vs. performance of mobile devices.**

Although tablet PCs and PDAs have great similarities in their way of use and interaction, the separation of mobile devices into two distinct groups (yet) takes place between these breeds. Unlike tablet PCs, **PDA**s do not have a PC-like performance and connectivity. In addition, completely different operating systems (e.g. Windows Mobile, PalmOS) are used for these somewhat smaller devices. The handheld operating systems meet with the shortened resources and the comparatively small screen diagonals (below 4″ resp. 10cm).

In contrast to tablet PCs, it is possible to wear these reduced types in the pocket of your jacket and this enables a whole new grade of mobility.

Based on this distinction, we hereafter understand PDAs and **mobile phones** as mobile devices since they dimension the borderline of real mobility. Not least they belong to this category because of their number of built-in interfaces to provide wireless network connectivity needed for mobile applications.

Besides the PC-like (laptops, tablets) and mobile devices (PDAs, mobile phones), we determine diverse TV interfaces as a third category of devices. These types of access to eParticipation can break ground to whole new groups of users. Consequently, we further analyze four main categories of devices as follows:

1. Personal computers (section 3.1)
2. Mobile devices (section 3.2)
3. TV-based interaction and delivery (section 3.3)

The characteristics of a device category have vital impact on the used channels and the requirements for their applications. In this respect the attributes of usability will be of special interest, as it primarily affects the success of devices as well as their applications. Rosson and Carroll [21] characterize usability as the quality of a system regarding ease of learning, ease of use and user satisfaction. Terasewich [26] points out that usability "also deals with the potential of a system to accomplish the goals of the user". When it comes to application, Terasewich also underlines that "… application developers must look carefully at potential users, devices, and contexts of use." The interaction of users, devices and applications will be investigated in the next sections. The three device categories will be described along the following arguments:

- In which way does the user interact (input/output/usability)?
- What are the characteristic channels (cf. section 2)?
- Which applications can be used on the designated devices?

The outcomes of the following sections will be vital to the discussion of the impact of devices and channels on certain eParticipation areas in chapter 5.

## 3.1 Personal Computers (PCs)

The term "personal computer" or "PC" is used with different meanings. In this deliverable, it defines usual desktop computers or computers that have comparable performance as well as interaction workflows and are able to run a standard Microsoft Windows operating system or an equivalent like Linux or Mac OS. Barry Brown [7] predicted already in 2000 that multi-function devices like the PC will have broad success in homes rather than those often mentioned single-function "appliances" that are aspired in the ubiquitous computing world (cf. brief outlook on future in chapter 6). Brown points out that "The multi-functional nature of PCs means that they can be adopted and used for a huge range of different activities. Home office, game playing, web browsing, email, digital graphics, personal finances are all activities carried out within the one box." The success of PCs is emphasized by the great variety of its representatives:

- Desktop PCs or Apple equivalent (iMac, Power Mac, Mac Pro, etc.)
    - External display for visual output
    - External keyboard and mouse for input

    - Lots of interfaces and I/O ports for printers, multimedia, etc.
- Notebooks (e.g. running Windows XP, Linux or Mac OS)
    - Integrated display, keyboard and pointing device (e.g. touchpad)
    - Similar external interfaces and ports as a desktop (USB, VGA, etc.)
    - Can often be used like a desktop by attaching external devices
- Tablet PCs e.g. with Windows XP Tablet Edition
    - A notebook reduced to the display
    - Stylus as a pointing device
    - Expandability is reduced but still comparable to a notebook
- UMPC[73] (Universal Mobile PC) aka Urigami
    - A highly mobile tablet PC with Windows XP functionality
    - Demonstrates the new borderline to PDAs

Every category itself has different versions of devices that vary e.g. in size, computing power and connectivity. The fact that an UMPC is almost incomparable with a desktop PC but neighbored versions of different classes show great equality, illustrates the variety of multifunctional devices that have emerged in the last years. Many other multifunctional devices nowadays base on PC-technology. Examples are video game engines or HTPCs[74] for digital TV. Kiosk systems[75] also usually base on simple PCs. Because of the differing modes of interaction during their every-day usage they are not part of the PC section.

In the interaction of users through output devices, usability aspects play an important role. These will be discussed in the following subsection.

### 3.1.1 Interaction and Usability

The great variety of devices that work with the same kind of advanced operating systems is amazing. Originally, the user interfaces for operating systems of personal computers have been designed for full desktop configurations with external display, mouse and keyboard. For a few years now, screen resolutions of tablet PCs or UMPCs match those of external displays (notebook display resolutions were able to catch up). Consequently, output interaction is not a big deal any more, when using standard features of the OS. The only restrictions may come from requests of full-sized and well-stocked web pages or software that assumes high resolution displays at the output device.

Input interaction is more critical (cf. mobile devices in subsection 3.2.1) regarding the portable PC-versions. It definitely is a challenge to design these computers in a way that keeps the OS usable and controllable in all situations and at all times.

[73] http://www.microsoft.com/windowsxp/umpc/default.mspx

[74] Home Theatre PC

[75] Due to time constraints, kiosk systems could not be investigated for this deliverable.

Notebooks integrate a somewhat reduced keyboard and simulate the mouse with a pointing device that stays attached to the computer encasement (a touchpad, a trackpoint[76] or both). Trackballs are barely integrated nowadays. Compared to the reduced keyboard the adopted pointing devices are afflicted with higher usability drawbacks. Pointing is ponderous and less precise. Most times this only results in a slower workflow. When it comes to applications that need a higher level of accuracy and/or speed, also quality suffers. Examples are interactive image processing and gaming. Often notebooks and their smaller relatives integrate additional functions that have to be purchased and attached separately with desktop PCs. Examples are webcams, microphone, loudspeakers and wireless interfaces.

Since tablet PCs and many UMPC (see Figure 3 for some examples) have no built-in keyboard but a quite intuitive pointing device (stylus), text input is the bigger problem regarding this kind of portable computers. The input of text is accomplished mainly with software keyboards that are operated with the stylus. UMPCs have an alternative input mode where the device is held with both hands and characters are typed with two thumbs[77]. Some UMPCs comprise a small hardware keyboards whose size and layout is not comparable to a PC or notebook keyboard and thereby do not allow fast typing with ten fingers. Using the stylus as a pointing device is more intuitive than using touch-pad or trackpoint – as long as the user picks elements or moves the mouse cursor directly. In other applications the pointing device is not used for absolute picking that is limited to the visible screen. In single-person games for example the pointing device (usually a mouse) is used for relative changes, which are added to the displayed position so that a stylus is not very intuitive either.

All PCs, even the portable ones, are mostly used in a settled position (e.g. sitting on a chair): The desktop PC and its periphery need a determined position respectively around a desk. A laptop has to be placed on a plain surface. Tablet PCs can also be operated while standing, but the user is still not very mobile. Because of input procedures optimized for mobility, low weight, and compact size, the UMPC is carried safer in one's hands during use. Thereby it comes close to a PDA. Unfortunately, due to its size and inflexibility, an UMPC is still experienced as a cutback while on the move or when being interrupted by other tasks.

Besides the hardware restrictions of input and output interaction with PC-devices, general usability of applications and operating systems may impose constraints. The PC is not specialized for a single field of use like many consumer devices. It is a multifunctional machine that can be utilized for a broad range of applications. In reverse the PC is a "usability disaster" which limits its range of user types [17]. Barry Brown then brings it to the point: It "frustrates and teases" the user and exhibits a behavior we would never tolerate from a car [7]. Users have to learn how to use and maintain a PC operating system and its applications. Additionally, OS and applications often behave quite unpredictable and quite some experience may be needed to manage such situations. All this makes a PC reasonably user-unfriendly – especially for certain user groups. Furthermore the computer industry has an orientation to the young [7] so that the use of PCs stays a challenge to the elder. One attempt to acquaint older people with PC

[76] The trackpoint is a small analogue joystick in the middle of a Lenovo notebook. Other vendors also integrate this kind of pointing device but might have other names for their systems.

[77] Consider that usual tablet PCs with Windows XP tablet edition are only usable with the stylus (no touch screen).

applications it the Simplico[78] PC that has been introduced by Fujitsu-Siemens in July 2006. It has reduced functionality, fonts easy to read and a simple usability concept.


**Figure 3: Examples for common UCMP devices**79

### 3.1.2 Primary channels

Prior to the prevalence of packet switched broadband connections via DSL and cable, internet had to be accessed via POTS or ISDN lines. Today these technologies are still used a lot, because broadband connections are not yet available in many rural areas. Some users also do not want to spend another monthly fee for broadband.

For the most ambitious PC users, however, broadband connection is a must-have. Besides the higher transfer rate, the most common charge modes (either data based or flat rate) of broadband connections are an important feature for internet applications. They allow an always-on connection without extra expenses, which makes the internet much more valuable. Email can be received right when it arrives on the mail server, and users can be contacted via instant messaging services (e.g. ICQ or Windows Messenger) around the clock. UMTS-technology is not only used for mobile phones and smart phones, but also as an affordable alternative for wired packet switched connections. Even UMTS-routers (see Figure 4 for an example) that provide access to several devices over the same connection are available[80].

---

[78] http://www.fujitsu-siemens.de/home/aktionsangebote/simplico/index.html

[79] UMPC images: http://www.teltarif.de/arch/2007/kw03/s24591.html

[80] http://www.teltarif.de/arch/2007/kw04/s24655.html


**Figure 4:UMTS-router that provides access to several devices (XSBox R4v)**

Another future alternative for DSL is WiMAX, which could also be provided in rural areas or as a competitive product in urban districts. Perhaps WiMAX chips will be built into notebooks in the future like it is already common for WLAN. Until now, native access to licensed mobile networks is not very common to PC devices (only some UMPCs have built-in GSM or UMTS modules). This feature has generally to be added with an auxiliary PC-card or via Bluetooth connection to a mobile phone.

WLAN is generally used to allow mobile access for portable PCs around fixed line connections and/or just to connect the stationary PC with less wiring. Most internet routers for DSL already have an integrated access point to provide WLAN in the user's residential area. WLAN is also offered punctual in some public areas (hot spots) but the range of costs goes from free to extortion.

Main internet applications for personal computers are email and web browsing. Besides simple surfing on html pages, web browsers provide additional functionality. Examples are audio/video players (with streaming), RSS[81]-feed subscription, and downloading. Some of these features have to be integrated with plug-ins of stand-alone applications (e.g. windows media player). The user often has the choice, whether s/he takes a web-based application or a stand-alone program for a certain task. The following subsection will cover these different types of internet applications.

### 3.1.3 Types of internet applications

Different kinds of applications for PC-devices exist. A lot of them use or need an internet connection. In this subsection these applications are divided into two groups:

1. Standalone software
2. Browser based applications (BBAs)

**Standalone applications** (aka desktop applications) directly use the operating system as their platform. In fact they benefit from all advantages and capabilities the OS offers. All the functionalities that do not need data from the internet can also be accessed while the system is offline.

**BBAs** are built on top of web browsers and use the web as a platform. Three critical characteristics can be identified when looking at their capabilities [24]:

- Page orientation
- Statelessness
- Limited local computation.

[81] Really Simple Syndication

Like delays in the connection, discontinuities as well as clumsy interfacing, these characteristics may lead to limited functionality, errors, confusion, and lower usability. However, there are many reasons why BBAs dominate the world of internet applications today [24]:

- User comfort
- Availability (no installation required)
- Platform independency
- Seamlessness

Platform independency and the fact that no installation is necessary, are the main reasons why BBAs are that popular. Users often have the choice whether to use desktop software or a BBA. Instant messaging can be used via a special client that offers great functionality and control, or by launching a web-site provided by the instant messaging service and thereby getting access without any program installation. The BBA also has the advantage that configurations and data are stored on a web-server and therefore do not need to be backed up or transferred to another PC. This is a possibility also standalone internet applications use (e.g. instant messaging or mail clients).

A catchphrase to be mentioned in this respect is **Web 2.0**, since it includes the change from software to services: Google for example is not selling software or systems, only services. The company never sold or packaged any piece of software but made millions by letting the user (directly or indirectly) pay for the offered service. It is “infoware” instead of software. And Web 2.0 is the full realization of the true potential of the web platform [18]. Mark Silver stresses, that the browser has to be improved and replaced at the same time [24]. Today’s web applications try this by refining and optimizing web-pages and – regardless of the statelessness of the used HTTP protocol – by giving them the usability and appearance of “stateful” PC-based software that is not strictly page-oriented any more. A mixture of technologies that is used for this is AJAX (**A**synchronous **J**avaScript **a**nd **X**ML). BBAs are also replaced. A lot of software that is not browser based is still classed among Web 2.0 applications. This is also shown by Tim O’Reilly who, among others, coined the notion of Web 2.0. He gave some examples to illustrate the difference between Web 1.0 and Web 2.0 [18] (see Table ).

By these examples another characteristic of Web 2.0 applications turns up: The new era of Web 2.0 is “harnessing the collective intelligence of web users”, and “hyperlinked documents” change to “overlapping communities” as Koji Zettsu states [30]. This social software includes human communication and collaboration with the aid of self-organizing social networks and virtual communities [2]. All these attributes will make Web 2.0 applications and the growing spirit of Web 2.0 invaluable for eParticipation. Besides wikis and blogging also forums, instant messaging and social bookmarking can be counted to social software (for some further details see Deliverable D 5.1). One defining quality of a Web 2.0 application is the fact that its use adds value on itself and that a broad use of it is desirable and valuable for clients and service providers [1]. Or with the words of Tim O’Reilly [18]: “Users are developers”.

| Web 1.0 | | Web 2.0 |
|---|---|---|
| Britannica Online | → | Wikipedia |
| mp3.com | → | Napster |
| Personal Websites | → | Blogging |
| Publishing | → | Web Services |
| Content Management Systems | → | Wikis |

**Table 1: Differences between Web 1.0 and Web 2.0 according to Tim O'Reilly [18]**

In the end, the vast number of knowledge and information that will be acquired by this collective intelligence of the web, its users and their usage, has to be adequately managed (cf. sub-deliverable D5.2.5 about knowledge management).

State of the art PC applications deal more and more with participation, communities and collaboration. Sometimes their success is shown by raw numbers. User generated online video (UGOV) for example made 47% of the whole online video market in the US in 2006[82]. While working on this deliverable one time, over 8 Million users were online in Skype, which also shows that a BBA is not the only means to enable high participation. But BBAs definitely can make services available to a broader audience.

The principles of Web 2.0 might even come back to the desktop PC itself by forming a new semantic desktop that reflects the personal view and the knowledge of the user [22]. Once more this reveals the convergence of desktop and web, and it shows that both are learning from each other. Yet, standalone applications as well as BBAs are not strictly bound to PC devices. Many applications and interaction technologies can also be ported to, and used on other device categories, like mobile phones, PDAs, TV set-top boxes or kiosk systems. A better understanding of the premises and limitations of these other device classes will be considered in the following two sections.

## 3.2 Mobile Devices

As stated above we understand PDAs and mobile phones both as mobile devices. Nevertheless, their use differs slightly: Utilizing the stylus and/or a built-in keyboard, the user of a PDA can control the device more functionally and often faster. On the contrary one needs both hands during input. Hardware keys, soft keys and a centric control element (cursor key) are used for interaction on common cell phones so they can be controlled with one hand and with less attention in most scenarios. Nowadays the junction from PDA to mobile phone is performed fluently and a lot of devices are available that do not allow a precise naming of relationship to one of these two types. They describe a group of devices that evolved from two different progressions which are moving more and more towards each other: On the one hand the size reduction of portable computers with the objective of getting handier devices and on the other hand the ongoing upgrade of mobile phones with numerous additional functionalities. Similar to both of these evolutionary movements is the extension of their capabilities of data communication.

[82] Screendigest (15.01.07): http://www.screendigest.com/press/releases/FHAN-6XDN28/pressRelease.pdf

Extreme results of this ambition are devices with IrDA, Bluetooth and WLAN interfaces as well as access to mobile networks over GPRS (MDA vario, Nokia 9500). Meanwhile several of these chatty devices are also able to connect to 3G networks (e.g. Nokia N95, MDA Pro). Usually mobile applications want to enable services that have to be independent of daytime and the user's whereabouts. So the connection to a cellular network is a must have. The other interfaces can be a helping support for useful optimizations of services.

Besides several limitations (which are described latter on) mobile devices have considerable advantages that make them preferable. The main advantages are:

- High portability
- Less attention necessary
- Usable with one hand (esp. cell phones)
- Mobility during use
- High potential of personalization
- High availability
- Location independence

The more powerful a device is and the more its way of usage heads toward PDA and further (e.g. notebook) the more these advantages of mobility vanish. Consequential they primarily take effect in compact mobile phones.

### 3.2.1 Interaction and Usability

Mobile devices include several limitations, which come along with their reduced design that makes them so flexible. These restraints are quite natural to most consumer products: Mobile devices usually have less memory, are smaller and have lower-resolution displays, provide fewer colors, as well as they offer different I/O mechanisms compared to personal computers [5]. One problem especially concerning powerful devices is the shortened battery power. Power supply will stay a major challenge and a barrier as cell phones receive more and more functionality and wireless interfaces like 3G and WiFi [25]. Due to their small size human interaction stays the focal problem for mobile devices. Real mobile hardware has no room for user-friendly periphery. So developers of user interfaces have to deal with small screens and limited input interfaces so far.

Usability of mobile devices has to be considered from the two modes of interaction: input and output interaction:

**Input interaction**

Input stays the most critical point when using mobile devices. Especially the input of long texts is often performed in a very uncomfortable and tedious way. On mobile phones keys (0-9,*,#) are arranged in the typical 3x4 layout. The digits 2 to 9 are linked to 3 or 4 characters each so that the whole alphabet can be typed. The first character on a key is chosen by pressing it once, the second by pressing it twice, and so on. In conclusion many more keystroke actions have to be performed than on a regular keyboard. The fact that maximum two thumbs can be used for typing decreases the input speed furthermore. Text prediction tools like T9 can lower the problem only marginally. At least for the average consumer real improvements are achieved not before devices are provided with QUERTY keyboards (hardware or software) [26]. Unfortunately this again restrains the flexibility and mobility of the respective device and its use.

Menu navigation is less critical than text input. Modern mobile devices provide several soft keys and a cursor key to navigate through menus and to choose certain options. The cursor key evolved from simple 2-way keys or scroll wheels to 5-way joystick-like buttons which can be pushed in four directions to navigate and pressed to choose an option. This layout makes navigation of any kind faster and more efficient.

Speech recognition was assumed as one possible solution to the problem of text input on mobile devices. After providing some devices with speech dialing and speech control these efforts have not been continued yet. Text input through speech recognition is quite error-prone as well as dependent on ambient noise and the particular user's voice. Additionally, any kind of input in public areas (which is very common for mobile devices) could grant an unwanted insight of ones privacy as well as could assume some autistic features when performed with speech input.

Besides the traditional user interfaces, like keypad and microphone, further types of input means have been developed for mobile devices. They are more used for situational and contextual control of an application than for text input or menu navigation. With their help the user has to do less input himself. This could e.g. be an automatic determination of the user's position, which makes the input of the own location redundant during the search for e.g. the nearest gas station. Any kind of context should be included (not exclusively location information) e.g. the available communications infrastructure, the current physical conditions, the user's social setting, or the user's emotional state [23]. Other alternative ways of input are the reading of barcodes (e.g. UPC or DataMatrix) with a built-in digital camera or the direct transmission of application data over the integrated means for short-range communication [26].

**Output interaction**

Especially for mobile phones the output of information is way less critical than the input. Nowadays color displays with several thousand colors are built-in solely. By now, a large part of devices dispose of screen resolutions, which were characteristic for the superior device class a few years ago. Even quite compact smart-phones without touch-screen have the same resolution as relatively young pocket-PCs (e.g. T-Mobile SDA II 320x240 pixels, Nokia E60 352x416 pixels). However the displays' small dimensions alone result in restrictive capabilities. Pictures and longer texts have to be zoomed and scrolled. An enlargement of displays to present multimedia and text content in a more user friendly way entails a decline of flexibility due to size, weight and higher power consumption.

Apart from a visual output of information, handhelds – especially mobile phones – imply some kind of "natural" audio output. In addition to the output of the interlocutor's voice, modern mobile phones and of course PDAs provide audio entertainment features and are capable of real sound and MP3 output. Besides completing the multimedia capabilities, sound generation is mainly a supportive element of optical output as well as for the input over different interfaces. Besides the output of status signals when the display is out of sight (ring tone as prime example), sound as feedback for input of any kind is of importance. It is used as an easy perceivable feedback for key presses and voice control, because keypad (haptic) and display (visual) often give limited or no feedback themselves ([26], p.2). Cell phones and PDAs (at least those with phone ability) imply an additional kind of haptic feedback in form of a vibrating function. With acceptable discretion it informs the user of a phone call, of new messages or other events. It can also be used as a supporting tool in applications. Especially in games it is an often added sensation to achieve a higher grade of immersion.

### 3.2.2 Primary channels used by mobile devices and applications

Plain voice communication is still the most commonly used channel for mobile devices, respectively cell phones. The success of telephony is based upon its simple usage (which is not very different to fixed line telephony), high availability and the broad network effects. These highly evolved network effects result from several characteristics of mobile phones:

- Every mobile phone is capable of voice telephony,
- everybody who has a mobile phone is able to use it at least for voice communication, and
- any other phone (mobile and fixed) can be reached through telephony
- regardless of its technology or standard.

The importance of consolidated network effects is shown in relation to the quite new technology of mobile video telephony. Only a small number of phones and not all geographical regions are capable of video communication (UMTS standard is needed – see section 2.5). Combined with a small group of users willing to pay higher costs for a video call, these properties result in a limited network and an overall usage of video telephony that falls short of expectations. Compared to asynchronous services like SMS or email, telephony as a synchronous service enables a more effective communication that makes it easier to solve problems for example [25].

The short message service (SMS) is the second cash cow of mobile providers. Similar to telephony the user networks are highly developed (in fact every cell phone is capable of receiving and sending SMS) and especially a young community claimed it as its new kind of communication. Meanwhile many fixed line phones can also be reached with a short message or are even able to send SMS. Although the short message service is asynchronous and quite expensive compared to the amount of information that is delivered (160 characters at approximately € 0.20) it has several advantages that make it preferable in many cases. Often the user is not willing to talk to somebody (maybe s/he is not in the mood or his/her situation does not allow it) so that asynchronous communication is preferred to deliver information. SMS can also be the right choice in a reverse case when the addressee of communication is not available (e.g. the phone is switched off or in an inappropriate time of day). Since SMS is a store and forward service the receiver is even able to read a message that was sent when the phone was switched off.

Until now the multimedia messaging service (MMS) was not a big success; even though it is a highly improved version of the SMS with higher text capacity and media content like pictures, video and sound. The issue might be quite similar to the relation of mobile video telephony and plain voice calls.

Another well developed service in mobile networks is email, which was especially boosted by the blackberry device. The blackberry combines a cell phone with QWERTY-keyboard and an efficient email system that pushes email to the device right when it arrives at the user's mail account instead of requesting (pull) new email periodically. By now, email push services are available on many other vendors' devices, too.

Besides the push services the pull technologies are still widely used. Mail can be fetched and written with built in email clients or over web application (mainly WAP pages). In most cases, packet switched technologies are used to build up a channel to the internet. Usually a GPRS connection will be sufficient to receive text based email.

When demands on transfer rates are higher than those of plain text, bandwidth can become one barrier for widespread mobile support [25]. This could appear, when email contains large attachments or media content. Also downloads and media streams are often too large for regular mobile internet connections but could just as well be a problem for constraint devices. In the bounce of broadband internet over fixed line connections (e.g. DSL) also web pages became exceedingly bulky. Subsequently with CSD or GPRS it is usually no pleasure to browse the regular web utilizing a mobile device, even if its display, input units and general capabilities are acceptable. This problem can be tempered by using a 3G (UMTS) connection, which offers higher quality and transfer rates. Unfortunately these connections are not available in some regions and they require special devices. The use of WLAN or other technologies for unlicensed mobile access are even less widely spread. Fortunately emerging technologies allow relatively transparent shifting between various networks, such as GPRS, UMTS and WLAN [25].

Like voice communication or SMS, GPRS is available in all areas that are provided with GSM. Also the majority of devices support GPRS. Therefore it is still the basis for mobile internet connections and applications that should be available everywhere and for everyone.

In terms of usability for mobile web browsing, the WAP standard was introduced as a basis for mobile information. However, a network's quality of service (QoS) is phrased in a variety of factors (bandwidth, delay, etc.), which cannot be separated by the user: The acceptance of mobile internet services correlates with the quality of service experiences by the user, and this perception is highly application-specific. Planning of service and technology rollout is often network centric. Network providers often underestimate that supply does not create demand on its own, but results in new applications (and the devices), which will not be widely accepted if these have no advantage for the common users or not enough participants to build a network. If users stay with common applications the existing networks will be adequate [12].

A mayor problem for a wide use of mobile internet and applications are the high costs of mobile data connections. Connecting the own device to the residential internet or the home PC can reduce these costs. By synchronizing interesting documents or media with the PC, the user can enjoy information without paying the expensive fees – even though s/he may not have the latest data. The Apple iPod and its pod-cast show quite vividly how appealing it is to benefit from information on a mobile device without being dependent on a mobile connection.

The next subsection will show some important characteristics of successful mobile applications.

### 3.2.3 Characteristics of mobile applications

The constraints of small mobile devices (as stated in section 3.2.1) are the main reasons for differences between PC- and mobile applications. The advantages of mobile devices are bought dearly with a list of restraints, which make it impossible to directly port a PC-based application into a mobile application. Any development has to assign the specific constraints of devices and the characteristics of mobile channels and data communication. Neuhaus ([15], p. 108) defines the following important demands on mobile applications:

- User-friendliness in all respects
- High level of security
- Up-to-date appreciable and reliable information
- Immediate  and simple access
- Low costs.

The relevance of user-friendliness is emphasized by the fact that mobile devices and applications are mainly consumer applications (cf. consumer products in chapter 1), which have to consider clueless and inexperienced users. The level of usability of a mobile application heavily depends on how it follows the specific constraints of the devices it is used on. In order to provide a simple and intuitive interface, it has to be reduced to essential elements and functionality. Additionally the user should be able to trust an application at any time. Since mobile networks often struggle with availability and stability, mobile applications are often a subject to distrust. Therefore a mobile application should always be predictive and avoid safety risks at any time. Due to costly mobile data connections, only relevant and needed information should be transmitted. Since mobile applications are often utilized in essential situations, the use of services and the transmission of information have to be fast, reliable and straightforward.

Beside the demands on mobile applications, Neuhaus ([15], p. 108) describes the advantages of mobile applications as follows:

- Flexible
- Location-independent
- Timesaving
- Relevant
- Beneficial
- Available at any time
- Up-to-date.

PC-based applications are mainly superior to mobile applications. Therefore, mobile applications need to precisely distance themselves by addressing and focusing their specific advantages. Simply porting an application to mobile interfaces will always disappoint users. A simple duplicate of an application or a web page will suffer poor navigation and information overload and, consequently, will never reach the quality of the original implementation [8]. Terasewich [26] points out that “it may not make sense to perform certain tasks through specific wireless devices, or through any wireless devices at all.”

A foundation for reasonable and promising mobile applications is the frequently mentioned domain of user customized services [3] [20]. One appropriate way to customize applications is via location based services (LBS) that are able to deliver information depending on the position of the user. Neither the manual input nor the knowledge of the own location is necessary for that. Positioning takes place by utilizing the cell phone network, GPS information or a synthesis of both technologies. User specific information could also be provided locally by using near-field communication like Bluetooth or RFID[83]. A very simple way to provide customized information is to use information that is directly available from the network providers (e.g. billing information when using value added services). Agent technologies (cf. Deliverable 5.1, section 4.1, and sub-deliverable D 5.2.5) could help to offer self-learning personalization that uses former user inputs. To avoid tedious input as well as expensive data transfer it is helpful to store user information on mobile phones persistently.

Besides personalization the most important feature when designing mobile applications is simplicity. This quality is not only compelled by the constraint of devices, but equally by the character of users. Functionality has to stand back for usability ([5], p. 7-8).

Important matters for users of mobile devices are:

- High grade of personalization (lower complexity, higher value)
- Local storage of as much information as possible (avoid long/expensive transfer)
- As simple as possible (consider user's point of view)
- Only as complex as necessary (reduce functionality as far as possible).

One big challenge in the design of mobile applications is the diversity of mobile devices. Dozens of different screen resolutions, input options and user interfaces need either a development that fits the application to various device properties or an abstract programming interface that lets standardized devices consider how the user interface should behave and be displayed.

### 3.2.4 Relevant projects with mobile devices

Useme.gov: http://www.usemegov.org

Mobile Phones for Youngsters: http://www.evoice-eu.net/, See also

http://www.ccre.org/news_detail_en.htm?ID=602

MobiLife – Life goes mobile! www.ist-mobilife.org

## 3.3 TV based Interaction and Delivery

When speaking about TV-based interaction and delivery, mainly the advantages of digital TV are expressed, such as stated in [19]: Digital TV has a much wider choice of TV and radio channels than a common TV. For digital TV, four or six digital channels are being offered instead of one analog. And the number of channels will increase in the future. As a consequence, a higher flexibility for broadcasters is offered in current and future digital TV as well. Although digital TV has a compression of signals, it has a better picture and sound quality.

Several interactive services are already offered via digital TV [19]:

- Integration of web technology
- Video-on-demand (the world's most popular interactive service for TV)
- Media Integration (banking, shopping, games).

[83] Radio Frequency Identification

In terms of different types of multimedia services, three levels of interactivity can be distinguished [19]:

- Local interactivity, i.e. no return path such as multi-camera angle or replay chosen "locally"
- One-way interactivity, i.e. sending messages to the service provider, e.g. for voting
- Two-way interactivity ("true" interactivity), i.e. a user gets a response on his/her message from the provider via the return path or the broadcast channel.

The two-way interactivity can further be differentiated between two levels:

- low level: view data send back, e.g. pay-per-view
- high level: continuing two-way exchange, which is fundamental, e.g. for chat.

Another model of interactivity is shown in Table 2; it distinguishes five levels of interactivity [19].

| Level | Description | Applications | Forward Channel | Return Channel |
|---|---|---|---|---|
| 0 | Call to the service provider by the viewer | Pay per view, pay TV | Broadcast Network | Usually telephone |
| 1 | Pseudo-interactivity | Teletext, Internet access, games | | Information locally stored on TV |
| 2 | Basic interactivity from remote control | Commerce, Internet | | Wireless (e.g. GSM, DECT, UMTS) is plug free and doesn't require professional installation<br>Integrated return path: suitable for cable the antenna can be used for the back signal |
| 3 | Use of return channel for video reception | Video on demand; internet video streaming, video telephone, two way video | Switched Network | Telephone up-to-date with: ISDN/ADSL/UMTS |
| 4 | Network with full service | Two way video; professional use i.e. telemedicine, videoconferencing | | Modem cable; VSAT (satellite) |

**Table 2: Levels of interactivity**

The change from cathode ray tube to LCD and the general extension to higher functionality in the world of consumer devices might open the users' minds for new applications based on the good old TV.

However, one also has to be aware of the fact that most people just want to soak up / absorb / consume (what's on) the television / TV and relax while doing so (let televison just wash over them). Pagani emphasises that this is a very passive occupation, with some consumer usage characteristics as follows [19]:

- heavy video
- Information is medium because of one-way communication (cf. D5.1)
- Consumption is entertainment-based (more a leisure activity than a learning environment)
- Social/family access
- Centrally generated (by the service provider)
- Passively received, user is unable to influence content flow (linear in form)
- Long form programs.

If we succeed in turning around the passive consumption towards a more active interaction, an innovation path between television and online services may be exploited ([19], p. 49). An important point are above mentioned features, with which digital TV offers more choice of (specialized) programs and therefore a more personalized program set. Also, in future TV, receiving via set top box or decoder (size of VCR, close to TV set) may be integrated.

### 3.3.1 Interaction with TV interfaces

Studying the input devices in more detail unveils that TV has a lower portability than PCs and mobile phones. Apart from that, the size of input devices for TV is small in comparison to the PC. Table 3 [19] reports the differences in devices.

| **Characteristic** | **TV** | **PC** | **Mobile phone** |
|---|---|---|---|
| Size of display device | Large | Large | Small |
| Size of input device | Small | Large | Small (keypad) |
| Portability | Low | Medium | High |

**Table 3: Physical Characteristics of Consumer Devices according to Pagani [19]**

Chorianopoulos [9] states that interaction techniques for interactive television are complex and difficult to use for a wide range of viewers, because there is a higher complexity than for analogue TV (e.g. the number of channel or multi-dimensional navigational metaphors). The author further points out that digital TV is too near to PC interfaces, and remote control for interacting via digital TV is yet too complex for a broad range of users. The traditional user interface is not adequate for the ordinary TV viewer. Another aspect is that TV is a shared medium, while mobile phones and PCs are almost always personal devices.

In a study reported in [11], different interaction devices have been analysed. The results showed that users preferred multimodal systems with combined speech and remote control. Further on, the study pointed out that speech interaction provides a good environment for speech control (privacy), while the limited interaction with the device and the need for convenient user interaction claim for alternative input devices.

In interacting through digital TV, interaction with any device requires a proper user interface (UI) for a wide range of users. Chorianopoulos [9] defines a set of characteristics of a good UI for digital TV interaction devices:

- fast task completion
- consider entertainment, laid-back posture, visual language
- aesthetically pleasing UI over efficiency
- depending on application (entertainment, learning, game-play)
- traditional UI-concepts can not be adopted without regarding the specifics of the context and purpose
- support familiarity
- user is regarded first as a viewer and then as a user.

### 3.3.2 Channels for TV based applications

A number of channels exist for TV-based applications. We distinguish among three types: broadband TV, uplink via IP connection (can also be a dial-in), and broadband internet.

For the transmission, DVB and Internet Protocols can be differentiated. DVB standards are published by the European Telecommunications Standards and they cover large geographic areas. The history of the DVB standards began in 1994 with the DVB-S satellite transmission standard, and is now a de facto world satellite transmission standard for digital TV applications. There is a list of related standards, but our focus rests upon the DVB-C, DVB-T and DVB-S standards.

To give an overview of transmission protocols, some characteristics of DVB and DSL standards are given below:

- Wire (DVB[84]-C) (coaxial, optic fiber cable delivery mechanism)
  - Large number of channels
  - DVB-C is closely related to DVB-S
  - based around 64-QAM, although higher order modulation schemes are also supported
  - no direct return channel available, so the interactivity must be provided, e.g. over telephone uplink
  - Broadband interactive services can be integrated
  - Can also implement internet and telephone services and therefore replace these additional networks (and costs)

[84] DVB = Digital Video Broadcasting, http://www.dvb.org

- Wireless (DVB-S) (satellite)
  - Coverage for large geographic areas
  - Vast quantity of channels, great transmission capacity
  - No direct return channel available (interactivity can be provided e.g. over telephone uplink)
  - Broadband interactive services can be integrated
  - Can also implement internet and telephone services and therefore replace these additional networks (and costs)
- Wireless (DVB-T) (airwaves)
  - the youngest of the three core DVB systems and the most sophisticated
  - allows service providers to match, and even improve on, analogue coverage - at a fraction of the power
  - extends the scope of digital terrestrial television in the mobile field, which was simply not possible before, or with other digital systems
  - Regional
  - Limited number of channels
- Internet Protocol, e.g. Digital Subscriber Line (DSL)
  - At least 6 Mbit
  - Download speed of consumer DSL services ranges from 256 kilobits per second (kbit/s) to 24,000 kbit/s
  - No broadcast (not enough channels available)
  - On-demand mode for digital TV via remote server (experience of switching between different channels locally)
    - Consequently, on-demand mode can provide unlimited number of channels
  - Uplink and therefore interactivity inherently integrated.

### 3.3.3 Range of applications

A lot of applications exist, which can be offered also via digital TV. The Electronic Program Guide (EPG) is such an example. According to Pagani [19] ”EPG is an essential, navigational device allowing the viewer to search for a particular program by theme or other category and order it to be displayed on demand. Ultimately, EPGs will enable the TV set to learn the viewing habits of its user and suggest viewing schedules”. Another example is pay-per-view with video-on-demand or TV shopping, with the presentation of the products. Sometimes it is possible to purchase the products directly. Furthermore, interactive games and convenient business applications are offered via digital TV, e.g. small web games, or TV banking.

# 4 Importance of Mobile Technologies and Digital TV Channels for eParticipation

## 4.1 Mobile Technologies and Devices

The importance of mobile devices and technologies has already been stressed in several reports of Demo-net (e.g. D1.1 or D5.1). Different reasons exist for the special interest in mobile technologies being explored for eParticipation although these have lots of restraints when compared with PCs and fixed line internet connections. In digging deeper into the issue of mobile devices (cf. discussion chapter 3), only PDAs or mobile phones, which are connected to a mobile network provider and which might be used on the run, can be considered real mobile devices. At a first glance, the largest (and often single spotted) benefit of such mobile technologies is that these enable a consumption of services and applications on the move.

In regards to eParticipation, further advantages are being offered by mobile channels, such as:

- High diffusion
- Personalization
- Localization.

The advantages are detailed in the subsequent sections.

### 4.1.1 High diffusion

Before GSM became widely diffused, the analog mobile telephone networks in Europe had not the state of a mass media. Device costs as well as rates of calls and calling plans were relatively high so that only those who severely needed mobile communication could afford it. Due to the fact that the technology was not yet mature and powerful enough, these analog networks even never intended to provide access to the main part of the population.

The (technologically and economically) standardized GSM technology made it possible to provide mobile access to a mass-market. Business competition, higher demands and effective production techniques enforced one another. By that, calling plans, calls and devices got permanently less expensive. This led to over 550 million subscribers in 50 European countries[85]. Until today, the spread of mobile phones in western European countries almost reached market saturation. Eastern European states catch up with a growth of over 40% per year[86]. If the economic divide (as one stage of the digital divide) in industrialized countries is a non-issue, then there is definitely no such divide regarding mobile phones [16].

---

[85] Data: http://www.gsmworld.com/gsmeurope/

[86] GSM Subscriber Statistics: http://www.gsmworld.com/news/statistics/

The mobile phone has become the electronic device or maybe even the computer of all social classes and a lot of citizens without PC and Internet access can be reached with mobile service distribution. Table 4 illustrates the significance of mobile connections in terms of telecommunication access in the 25 EU Member States [13].

| | |
|---|---|
| Internet Access | 43,0% (of households) |
| Broadband Internet Connection | 10,6% (of population) |
| Fixed line connections | 49,6 (of 100 citizens) |
| Mobile connections | 89,6 (of 100 citizens) |

**Table 4:** Telecommunication Access in the EU (25 Members, 2004) [13]

Even though the numbers for the mobile connections have to be put into perspective of the fact that many citizens have more than one mobile connection at their disposal, the penetration of nearly 90 % shows the impact of mobile technologies and their advantages over fixed line connections. Also the number of fixed vs. mobile connections has to be considered properly: fixed lines are usually serving several citizens in a common household, while a mobile phone is rather a private device owned by a person.

Mobile technology needs to be investigated also for its use in eParticipation. Since this is the mostly spread device and channel, it has a high potential to reach out widely. However, so far we are lacking proper applications in eParticipation for such channels and devices.

### 4.1.2 Personalization

As shown above mobile telephony expresses a unique and omnipresent form of human communication. The telephone is no longer an interface of a ménage or a firm but has emerged to a personalized anchor of communication and information. A mobile phone is *my phone* and a call number is *my* personal number at which only *I myself* can be reached. By that highly personalized services get feasible and this opens a new range of possibilities for both, providers and customers.

The advantages of personalization and wide penetration entail that mobile communication and devices have a high potential for eParticipation, which needs yet to be explored.

### 4.1.3 Localization

More than with all other types of information access, additional information about the user's position and situation can be explored when using services based on mobile phones. More specifically, context-based and location-based services can be introduced which offer higher functionality and an optimized usability (which are otherwise mutually exclusive attributes). These services may also be explored for eParticipation services.

Future installation of GPS or equal systems into mobile phones ex factory will further advance such possibilities.

## 4.2 Digital TV

In respect to digital TV, the spread is not yet equally broad in different European countries, although initiatives to provide DVB options have been launched widely. Other problems of digital TV are that interaction is rather complex, and users may not have yet the proper TV and interaction devices. Digital TV is seen to become more relevant in the near future, when a broad range of services is available and the population has migrated to new devices. However, when this will have happened, digital TV may be an instrument available per family / household. And for eParticipation actors, it may become a means to reach out to every citizen in collaborative social contexts, where themes of local democracy and political decision-making may be discussed via digital TV.

# 5 Impact of Devices and Channels on eParticipation

Deliverable 5.2 aims at investigating emerging tools and technology for eParticipation. The specific report on devices, channels and mobile technologies aimed at investigating general aspects of channels (wired and wireless) and devices (fixed and mobile) with the purpose to gather a better understand of deploying certain channels and devices for eParticipation. Issues relevant for implementing eParticipation solutions, which can be identified at the side of the interaction devices and the transmission channels, are among others the bandwidth, the reliability of connections, the processing power, the options of the device's interfaces and the user mobility (cf. as well a discussion in deliverable D 5.1 section 6.4). These aspects are further elaborated in respect to two different perspectives:

- User's Perspective: Does a transmission channel / an end device support a user to
  - read and understand information?
  - form an opinion on a subject?
- Provider's perspective (e.g. government): Do these technologies and devices help to
  - provide information?
  - collect information?

The advantages and disadvantages of certain device types and channels will be related to the work in D5.1 (section 2). The questions are: What is useful, what is not useful, and which eParticipation areas can be supported best? If necessary, the user and provider perspectives will be further differentiated in terms of the stakeholders' special foci. These stakeholders could be citizens, politicians, government or others (see also D 5.1 discussion on the stakeholders of eParticipation).

Table 5 shows the differences between the consumer expectations on different devices and channels introduced in this sub-deliverable. As stated in [7], PCs are not considered a shared device, while TV is. Mobile phones are much individual mediums.

| Consumer expectations in TV space | Consumer expectations in PC space | Consumer expectations in the mobile phone space |
|---|---|---|
| Medium, stable pricing of goods | High, unstable pricing of goods | Low unstable pricing of goods |
| Infrequent purchase (one every 7-11 years) | Frequent purchase (every 18 months to 3 years) | Frequent purchase (every 18 months to 3 years) |
| Little requirement for software and peripheral upgrades | High requirements for software and peripheral upgrades | Medium requirement for software and peripheral upgrades |
| Works perfectly first time | Probably will not work perfectly first time | Probably will work first time |
| No boot-up time | Long boot-up time | No boot-up time |
| Low maintenance | High maintenance | Low maintenance |
| Low user intervention | High user intervention | High user intervention |
| Little of no technical support required | Substantial technical support required | Little technical support required |

**Table 5: Differing Consumer Expectations for Different Platforms [19]**

Barriers for the wide use of PCs and the Internet in eParticipation contexts are usability and costs. The assumptions that PCs are foremost for well educated, skilled and rich populations [7], while rural areas and socially and economically weaker populations and individuals cannot afford it (i.e. a high risk of digital divide), are no longer that high of risk, since PCs have become affordable by almost everyone by now [16].

According to Brown [7] "…computers have been seen by many as nerdy, boring, unsociable, unhealthy and – as a killing blow – male". With many joint computer games, at least the young generation is embarking on PCs and network connections. However, still local differences exist with broadband or no-broadband. A chance is WiMAX, because rural areas will be better provided with broadband internet allowing always on connections.

Some social groups adapt to mobile communication faster, which could be an opportunity to advance digital inclusion through mobile computing devices [28].

Members of teams must often be able to make decisions by interacting with a large number of people and be geographically mobile at the same time [25].

To avoid a usability divide, the target eParticipation applications have to fulfill user needs of all target groups. An adjustable application that fits many kinds of user groups would be desirable, but is difficult to implement. Consequently, trade-offs need to be decided. These should not be on the side of weaker groups.

Another aspect to be considered properly relates to digital TV [7]: often, groups often jointly watch TV. The problem here is that the strongest individual / group will dominate, compared to the power over remote control. Yet, TV could also be a powerful approach to involve the uninvolved, those who do not participate actively by searching for information and active participation. They could be involved while just consuming TV offers by giving them an option to take part (e.g. polling or chat).

Citizens with special needs (e.g. blind citizens) have their own devices to interact and stay tuned. The availability of such devices has to be taken into mind when developing eParticipation tools and applications, i.e. offers need to be available on any device suitable for the purpose without excluding certain stakeholder groups. Accessibility can be achieved by focusing once more on the same needs that apply to general usability: Simplicity and unambiguousness.

# 6 Concluding Remarks

When discussing about emerging eParticipation tools, proper understanding of the end devices and the channels employed for specific applications has to be acquired. Devices are the primary means through which stakeholders participate and interact with an application or with other stakeholders in a virtual environment.

Since not every device is suitable for every eParticipation area and application aimed at, decisions and designs for eParticipation applications have to be made thoroughly. Trade-offs need to be compensated as much as possible in order not to delimit the outreach of the application. Also the transmission channels and their limitations need to be taken into account.

The sub-deliverable at hand has investigated the specific characteristics of channels and devices in general, and in respect to eParticipation. However, one has to notice that much of the discussion on devices and channels is not specific to eParticipation areas respectively – any application in business and government has to cope with such restrictions of end devices and channels.

The main arguments of discussion were:

- limitedness of channel / devices
- restricted availability (digital divide)
- outreach of channel
- penetration of devices.

These arguments have an impact on eParticipation tools and applications and, consequently, need to be studied well when developing eParticipation solutions. The discussions in this sub-deliverable should provide some better insights.
As regards the future, trends indicate the following aspects:

- Convergence of devices
- Convergence of channels
- Ubiquitous computing

These may have further impact on the success and development of eParticipation tools and applications. Further research is needed to understand the impact thereof for specific eParticipation environments.

# References


[1] Auer, S., Dietzold, S., Riechert, T. (2006): OntoWiki - A Tool for Social, Semantic Collaboration. In: The Semantic Web - ISWC 2006, LNCS 4273/2006, Springer Verlag, Berlin / Heidelberg, pp. 736-749.

[2] Bächle, M. (2006): Social Software. In: Informatik-Spektrum. April, 2006, vol. 29, iss. 2, pp. 121-124.

[3] Benkenstein, M., Kohrmann, O. (2003): Personalisierung multimedialer mobiler Dienstleistungen – konzeptionelle Grundlagen und empirische Analysen auf Basis der Adoptionstheorie. In: Kruse (ed.), MultiMedia Mobil - Dienste und Inhalte über mobile Plattformen. Verlag Reinhard Fischer, Munich (D), pp. 111–132.

[4] Berglund, 2004

[5] Bloch, C., Wagner, A. (2003): MIDP 2.0 Style Guide for the Java 2 Platform, Micro Edition. The Java Series...from the Source, ed. Friendly, Addison-Wesley, Boston (MA).

[6] Bonapace, L. (2002): Linking Product Properties to Pleasure: The Sensorial Quality Assessment Method - SEQUAM. In: Pleasure With Products: Beyond Usability. Green and Jordan (eds.), Taylor & Francis, London (GB), pp. 189-217.

[7] Brown, B. A. T. (2000): The future of the personal computer in the home: A research note. In: Personal and Ubiquitous Computing. March 2000, vol. 4, iss. 1, pp. 39-44.

[8] Chan, S. S., Fang, X. (2001): Usability issues in mobile commerce. In: Proceedings of the Seventh Americas Conference on Information Systems. Strong and Straub (eds.), Association for Information Systems, Atlanta (GA), pp. 439-442.

[9] Chorianopoulos (2006)

[10] Constantiou, I. D., Polyzos, G. C. (2003): The Impact of Technology Advances on Strategy Formulation in Mobile Communications Networks. In: Mennecke and Strader (eds.): Mobile Commerce - Technology, Theory and Applications. Idea Group Publishing, Hershey, pp. 99-121

[11] Ibrahim, 2003

[12] Koivisto, M., Urbaczewski, A. (2004): The relationship between quality of service perceived and delivered in mobile Internet communications. In: Information Systems and E-Business Management. September 2004, vol. 2, iss. 4, pp. 309-323.

[13] Lumio, M. (2006): Statistik kurz gefasst - Industrie, Handel und Dienstleistungen - Telekommunikation in Europa. 09/2006, Eurostat, European Commission, ISSN 1561-4832, http://www.eds-destatis.de/de/downloads/sif/np_06_09.pdf.

[14] May, P. (2001): Mobile Commerce - Opportunities, Applications and Technologies of Wireless Business. Cambridge University Press, Cambridge.

[15] Neuhaus, D. (2003): Mobile Ticketing - Killerapplikation in der mobilen Welt. In: MultiMedia Mobil - Dienste und Inhalte über mobile Plattformen. Kruse (ed.), Verlag Reinhard Fischer, Munich (D).

[16] Nielsen, J. (2006): Digital Divide: The Three stages, Jakob Nielsen's Alertbox, November 20, 2006 (http://www.useit.com/alertbox/digital-divide.html)

[17] Norman, D. (1998): The invisible computer. MIT Press, Cambridge (MA).

[18] O'Reilly, T. (2005): What Is Web 2.0. [cited 2007 January 8]; Available from: *http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html.*

[19] Pagani, M. (2003): Multimedia and Interactive Digital TV - Managing the Oppertunities Created by Digital Convergence. IRM Press, Hershey (GB).

[20] Pietralla, J. T. (2003): Die neue Welt der Mobiltelefone - Die Schlüssel zum Erfolg. In: Referate des Kongresses "Mobil mit digitalen Diensten", Tagungsband. Sigle and Thielmann (eds.), moderne Industrie Buch, Munich (D), pp. 48-61.

[21] Rosson, M. B., Carroll, J. M. (2002): Usability Engineering: Scenario-Based Development of Human-Computer Interaction. Academic Press, San Francisco (CA).

[22] Sauermann, L., Grimnes, G. A., Kiesel, M., Fluit, C., Maus, H., Heim, D., Nadeem, D., Horak, B., Dengel, A. (2006): Semantic Desktop 2.0: The Gnowsis Experience. In: The Semantic Web - ISWC 2006, LNCS 4273/2006, Springer Verlag Berlin / Heidelberg, pp. 887-900.

[23] Schmidt, A. (1999): Implicit human computer interaction through context. In: Personal Technologies. 4, iss. 2,3, pp. 191-199.

[24] Silver, M. S. (2006): Browser-based applications: popular but flawed? In: Information Systems and E-Business Management. October 2006, vol. 4, iss. 4, pp. 361-393.

[25] Sorensen, C., Gibson, D. (2004): Ubiquitous Visions and Opaque Realities: Professionals Talking about Mobile Technologies. In: The journal of policy, regulation and strategy for telecommunications, information and media. 2004, vol. 6, iss. 3, pp. 188-196.

[26] Terasewich, P. (2003): Wireless Devices for Mobile Commerce: User Interface Design and Usability. In: Mobile Commerce Technology, Theory and Applications. Mennecke and Strader (eds.), Idea Group Publ., Hershey, pp. 26-50.

[27] Turowski, K., Pousttchi, K. (2004): Mobile Commerce. Springer-Verlag, Berlin/Heidelberg.

[28] Wareham, J., Levy, A., Shi, W. (2004): Wireless diffusion and mobile computing: applications for the digital divide. In: Telecommunications policy: the assessment, control and management of developments in telecommunications and information systems, Elsevier Science, Amsterdam (NL), pp. 439-457.

[29] Weiss, M. e. (2006): WiMAX - General information about the standard 802.16. Rhode&Schwarz, Munich (D), http://www.rohde-schwarz.com/appnote/1MA96.

[30] Zettsu, K., Kiyoki, Y. (2006): Towards Knowledge Management Based on Harnessing Collective Intelligence on the Web. In: Managing Knowledge in a World of Networks, LNCS 4248/2006, Springer Verlag, Berlin / Heidelberg, pp. 350-357.

# Conclusions and Reflections on the Emerging Technologies for eParticipation

In this report we have covered a wide range of tools and technologies, at various levels of maturity from established to emerging to conceptual, and each in turn shows strengths and weaknesses for the enhancement of eParticipation.

The report has identified the main problem areas for each tool and technology in turn.

Often the tools discussed suffer from the fact that they are designed and implemented for other or more general uses than eParticipation, resulting in mis-applied focus, inefficiency and poor user adoption. There is a wide-ranging misconception that existing tools can be easily adapted for use in eParticipation (a view often forwarded by stakeholders/providers trying to advance their already existing solutions) and that a diverse selection of disjoined tools each serving their own purpose will prove successful. It is apparent that we need fewer, more comprehensive platforms ensuring interoperability and integration with a wide range of facilities specific to eParticipation.

The fact is that sometimes subtle differences in focus can have disastrous results with respect to public adoption of the tools and technologies and their usefulness in the context of eParticipation. Often a lack of research hampers the advancement of the technology and its successful application for eParticipation.

In this report we have discussed the potential use of some examples of tools and technologies emerging in the eParticipation fast growing scenario. Here are listed some of the important issues raised within the subsections of this deliverable such as the devices, access and information structure for knowledge sharing.

When discussing about emerging eParticipation tools, proper understanding of the end-user devices and the channels employed for specific applications has to be acquired. Devices are the primary means through which stakeholders participate and interact with an application or with other stakeholders in a virtual environment. One must notice that much of the discussion on devices and channels is not specific to eParticipation areas respectively.

The main arguments of discussion were the limited use of channels and the restricted availability. These arguments have an impact on eParticipation tools and applications and, consequently, need to be studied well when developing eParticipation solutions.
As regards of future channel selection, trends indicate the following aspects:

- Convergence of devices
- Convergence of channels
- Ubiquitous computing

These may have further impact on the success and development of eParticipation tools and applications especially those controlled by other means and methods. Further research is needed to understand the impact thereof for specific eParticipation environments.

Argumentation Support Systems can help participation in various kinds of goal-directed dialogues in which arguments are exchanged. Their potential relevance for eParticipation is apparent, since the goal of eParticipation is to engage citizens in dialogues with government about such matters as public policy, plans, or legislation. Surely argumentation plays a central role in this process. In a public consultation, for example, citizens are given an opportunity to not only make suggestions, but also support these suggestions with arguments but this is not done by using specific argumentation systems. This report discussed number of argumentation support systems and associated tools. Some of these focus on the visualization of arguments and here the graphical notation and user interface are important features. Others focus on providing analysis of the situation but typically with a more limited graphical user interface. The systems presented allow users to access various levels of information, to be able to focus on specific information and to have the ability to organize the gathered data to construct an effective argument – all of which are required for eParticipation.

Knowledge management tools and technology adopted from the eBusiness world are proving to be the back-bone of the knowledge industry. Simultaneously a knowledge based industry can not function with out an efficient collaborative system enabling the sharing of information and knowledge through the growing dispersed manners of teleworking and globalism.

In a recent Demo-net workshop[87] dedicated to KM in e-participation, the following needs to research KM in eParticipation where identified-some of which are closely related to collaborative system needs:

- What kind / types of knowledge need to be managed in e-participation?
- Which KM techniques and tools are available and are suitable for which e-participation context?
- What are the challenges related to KM processes, technologies applied to e-participation:
    - Integration of KM processes into e-participation contexts
    - Knowledge for managing innovation
    - KM for recording decisions
    - Mapping KM processes to the policy lifecycle (see D 5.1)
    - Ownership of knowledge and KM processes
    - Different stakeholders have parts of ownership
    - Role of facilitator in the e-participation context
    - Scope of different types of knowledge
    - Instrument for evaluating different types of knowledge
    - Subsidiarity of knowledge – relevant to the appropriate government level (local, regional, etc.)

The concepts, tools and technologies introduced in this sub-deliverable indicate that there is a large potential for applying proper KM processes and support tools in e-participation contexts. However, focused research is needed to gather a better understanding and to develop recommendations and guidelines on how to implement and integrate KM in e-participation tools and applications

In eParticipation there is a clear requirement to better understand how technology can support informed debate on issues but there are a few main obstacles in achieving this. The first is that the public deliberation is typically on complex issues and therefore there are typically a large number of arguments and counter arguments to consider which when presented in linear text can be confusing for the public at large. Secondly, it is not obvious that many people actually have the necessary critical thinking skills to deliberate

[87] The workshop took place on 5th of December, 2006, at ICCS, Athens. The minutes were recorded by Dimitris Apostolou (ICCS).

on issues. In can be seen that the type of argumentation support systems and tools described in this report have the potential to add value to current eParticipation methods. Thirdly the usefulness of the information collected and stored needs to structure in an accessible way for creating new knowledge with advanced KM and made available in user-friendly, easily accessible collaborative environments. And the user must have some freedom of channel emphasising the importance of Natural language processing, ontologies and semantic webs to the enhancement of eParticipation.

All of the above should make it clear that there are no tools already available. That it is a misconception that single tools do already exist that fit the purpose. It should be clear that we need comprehensive platforms with a wide range of facilities and that there are many emerging technologies with good potential emerging or already available but there is lack of research, lack of integrating the many different tools available, the lack of standards that ensure interoperability.

Research is needed, research involving both testing and trying with the end user in well structured pan-European pilots. It becomes painfully obvious when looking at the advancement taking place in the on-line game world where hundreds of thousands of youth and adults share a cyberworld through the ultimate use of available multimedia, 3-D and visualisation technology, as well as extensive knowledge management to collect and distribute all undergoing events and creations how far we are behind in using these technologies for the advancement of eParticipation in the context of eGovernance..

## Future of eParticipation

The reader may ask himself why new technologies important to the scene discussed like GIS based systems, three dimensional representation, visualisation; areas of artificial intelligence and so on were not discussed in this report. There is a good reason. Very little research is available in this fastest growing market of ICT's and very little of it relates to the use of emerging technologies in eParticipation.

The tremendously exciting field of virtual world interaction has become feasible only within the last couple of years with the advent of Massive Multiplayer On-Line Games (MMOG), opening up countless avenues of research. The companies developing these games have found an entirely new set of challenges and problems unique to the melding of real and virtual universes, perhaps indicating both the possibilities and dangers of eParticipation when applied to such environments. But here also the industry and user communities are developing new collaboration systems and techniques, argumentative support systems, arbitration, negotiation and participation, all by direct experimentation. This is the field of ICT fastest to make use of new and emerging technologies putting them to unforeseen uses. The use of semantic web, ontologies and Natural language are already entering the cyber world of games.

Even though the initial attractiveness of such systems tends to be the beautifully crafted gameworlds and colourful graphics, certainly providing unprecedented possibilities in user interaction and user interface tools, the powerfully immersive experience and sense of community should give us reason enough to conclude that in this direction lies.

A word of warning must though be issued: eParticipation is eGovernance and must obey the rules of personal data protection, documentation, traceability and accountability.

However interesting and eye-catching the ICT world becomes eParticipation tools must abide the law and regulations. Also in this area the constrictions issued must be clear and therefore there is the need for multidisciplinary research team efforts when exploring the growing potential of new and emerging technologies to the advancement of eParticipation.