

INFORMATION SOCIETY TECHNOLOGIES
(IST)
PROGRAMME

Contract for:
Shared-cost RTD

Annex 1

“Description of Work”

Project acronym: MAGICSTER

Project title: Embodied Believable Agents

Contract number:

Related to other Contract no:

Date of preparation: April 12, 2002

Proposal number: IST-1999-29078

Operative commencement date: September 1, 2000

Contents

| | |
|---|-----------|
| 1 Preamble | 4 |
| 2 Project summary | 4 |
| 3 Project objectives | 6 |
| 4 Participant list | 9 |
| 5 Contribution to programme/key action objectives | 10 |
| 6 Innovation | 11 |
| 7 Community added value and contribution to EU policies | 15 |
| 8 Contribution to Community social objectives | 16 |
| 9 Economic development; scientific and technological prospects | 17 |
| 10 Workplan | 19 |
| 10.1 General description | 19 |
| 10.2 Workpackage list | 33 |
| 10.3 Workpackage descriptions | 34 |
| 10.4 Deliverables list | 39 |
| 11 Project planning and timetable | 40 |
| 12 Graphical presentation of project components | 41 |
| 13 Project Management | 42 |
| 14 Clustering | 44 |
| 15 Other contractual conditions | 44 |
| Appendix A: Consortium Description | 45 |
| EDIN | 47 |
| ROMA | 49 |
| UPENN | 51 |
| DFKI | 53 |
| SICS | 55 |
| BARI | 57 |

AvatarMe 59

1 Preamble

This description of work is a revised version of the original. The revisions reflect the changed status of some partners and some coincidental minor adjustment in the scientific program. The most significant alteration is that UPENN's involvement in the project is changed from the original role as partner to a consultative role. UPENN will continue to support the project mainly via the ROMA and EDIN groups.

2 Project summary

Objectives

- To design a believable conversational interface agent which makes use of gaze, facial expression, gesture and body posture as well as speech in a synchronised fashion.
- To evaluate the use of the conversational agent in laboratory conditions to determine which aspects of the embodied agent are important for what types of human-computer interactions.
- To develop and document the agent architecture and components to enable other research and development teams to prototype and evaluate new versions of the agent interface in new domains and for novel tasks.

Description of Work

Believable animation: to develop an agent which displays natural or believable movement.

This includes modeling not only voluntary movement (e.g. gesture) and involuntary movement (e.g. blinking), but also movements which correlate with communicative acts (e.g. limb gestures, body posture and facial expressions). Whereas research in computer animation has actively studied these different classes of human movement, it is still difficult to procedurally generate believable natural movements within the latter class. Our goal is to explore this problem in the area of facial and non-facial animation.

Believable conversation: to develop an agent which can perform conversational functions such as repair, feedback or turn-taking in a believable way, using verbal and non-verbal signals. This work will build on empirical and descriptive work on person-to-person dialogue, as well as on computational work on end-to-end spoken dialogue systems, and integrate it with work on simulated agents. Of particular concern here will be the development of a formalism that allows us to synchronise the different information signals, rather than assume that a particular channel is always the dominant one.

Believable information delivery: to develop an agent which can deliver information in a contentful and contextually appropriate way. This work will build on earlier work on the generation of spoken and written messages, integrated with work on animation. To increase the believability of the agent, we want to pay particular attention to the generation of messages in a linguistic style which projects the right kind of "personality" for the animated agent.

Evaluation: to evaluate the use of believable agents in information delivery. Throughout the project various aspects of believability will be developed and instantiated in different combinations in subsequent prototypes. This will allow us to evaluate the effects of different types of believable agents in different contexts. However, because the technology itself is novel, we also want to work on the further development of the evaluation *methodology* for MAGICSTER-style technology.

3 Project objectives

As computational devices become more powerful and ubiquitous, people will spend more and more time interacting with computers. Current network facilities already allow us to use computers and telephones to access a wide range of computational information sources. The network of the future will connect not only our PCs and telephones, but also our fridges, toasters and garage doors, allowing a much wider group of users to communicate with objects we never communicated with before.

This poses new challenges for the way communication interfaces are designed. The use of speech is often described as the most natural means of communication most people have at their disposal, and because of this it is sometimes claimed that speech input and output devices are the best way to overcome communication problems between people and machines.

But speech is only part of the solution. There is evidence that it is not speech as such that allows for natural interaction, but the ability to engage in *conversation*, in particular face-to-face conversation.

Conversation involves more than speech. The reason people find conversation an effective means of communication is because they use a variety of signals in combination with speech, such as gaze (making eye contact, looking down, looking at a particular object,...), gesture (pointing at something, indicating how big something was,...), or body posture (turning away from the other person, leaning forward to emphasise a point,...). Building speech-based interfaces which do not allow for these extra conversational signals will not allow for high-bandwidth of communication and will inevitably lead to interactions that break down easily. Conversation also involves more than the transfer of information from one party to another. It involves dynamic modification of each speaker's knowledge of the other. A proper modeling of this is necessary to achieve great efficiency and non-redundancy in communication devices.

It is the design of properly *conversational* communication interfaces that is the main goal of this project, in particular the design of interfaces that make use of non-linguistic information when delivering information. The best way to incorporate non-linguistic conversational aspects into an interface is by allowing users to interact with an agent-like character which has as many visual and other properties of humans (a body, a face, lip movement, moving limbs, etc) as can be exploited in conversation. We call such representations *embodied agents* or *characters*. We also want these characters to be *believable*. This means that, at a minimum, the character's behaviour should not interfere with its primary goals; if its goal is to deliver an educational message, its behaviour should complement or enhance that educational goal. But beyond that, we want interface characters to display the kind of behaviour that might make users believe they are observing a sentient being, with its own beliefs and "personality".

The high-level goals of the project can be summarised as follows:

- design a believable conversational interface agent which makes use of gaze, facial expression, gesture and body posture as well as speech in a synchronised fashion, in interaction with a dynamic model of the user's knowledge state and goals;
- evaluate the use of the conversational agent in laboratory conditions to determine which aspects of the embodied agent are important for what types of human-computer interactions;

- develop and document the agent architecture and components to enable other research and development teams to prototype and evaluate new versions of the agent interface in new domains and for novel tasks.

The research area concerning embodied, believable agents is wide, encompassing various areas of Information Technology as well as social science, human factors engineering and software architecture development. In MAGICSTER, we want to make progress in several of these areas, concentrating on aspects that contribute to the *believability* of the interface agent:

Believable animation: to develop an agent which displays natural or believable movement.

This includes modeling not only voluntary movement (e.g. gesture) and involuntary movement (e.g. blinking), but also movements which correlate with communicative acts (e.g. limb gestures, body posture and facial expressions). Whereas research in computer animation has actively studied these different classes of human movement, it is still difficult to procedurally generate believable natural movements within the latter class. Our goal is to explore this problem in the area of facial and non-facial animation.

Believable conversation: to develop an agent which can perform conversational functions such as repair, feedback or turn-taking in a believable way, using verbal and non-verbal signals. This work will build on empirical and descriptive work on person-to-person dialogue, as well as on computational work on end-to-end spoken dialogue systems and knowledge representation, and integrate it with work on simulated agents. Of particular concern here will be the development of a formalism that allows us to synchronise the different information signals, rather than assume that one channel is always the dominant one (usually the linguistic channel) and adding other information to it.

Believable information delivery: to develop an agent which can deliver information in a contentful and contextually appropriate way. This work will build on earlier work on the generation of spoken and written messages, integrated with work on animation. To increase the believability of the agent, we want to pay particular attention to the generation of messages in a linguistic style which projects the right kind of “personality” for the animated agent.

Evaluation: to evaluate the use of believable agents in information delivery. Throughout the project various aspects of believability will be developed and instantiated in different combinations in subsequent prototypes. This will allow us to evaluate the effects of different types of believable agents in different contexts. However, because the technology itself is novel, we also want to work on the further development of the evaluation *methodology* for MAGICSTER-style technology.

The MAGICSTER technology has a wide range of potential applications, ranging from entertainment over simple information delivery to education. Because we want to develop the generic, basic research aspects of this embodied conversational interface technology, MAGICSTER will not concentrate on one particular application in a particular domain. Nevertheless, for evaluation purposes it is important we instantiate the technology and its components in simulations that are sufficiently real. We will do this by developing MAGICSTER components for a mixture of applications and domains, building on results from other projects, in particular projects where multimodal communication interfaces have been developed but which

do not make use of embodied agents. This will allow us to evaluate and contrast the emerging MAGICSTER technology without the need for elaborate domain-specific modeling work. Evaluation domains include explanations of radiological images, information delivery about museum exhibits, and narrative agents commenting on a user's interaction with the Web.

4 Participant list

| Participant | Short name | Country | Status | Entry date | Exit date |
|--|------------|---------|--------|------------|-----------|
| University of Edinburgh, Division of Informatics | EDIN | UK | C | m1 | m36 |
| Università di Roma "La Sapienza", Dipartimento di Informatica e Sistemistica | ROMA | I | P | m1 | m36 |
| University of Pennsylvania, Center for Human Modeling and Simulation | UPENN | US | P | m1 | m12 |
| Deutsches Forschungszentrum für Künstliche Intelligenz, Intelligent User Interfaces department | DFKI | DE | P | m1 | m36 |
| Swedish Institute of Computer Science | SICS | SE | P | m1 | m36 |
| Università degli Studi di Bari, Dipartimento di informatica | BARI | I | P | m1 | m36 |
| AvatarMe | AME | UK | P | m1 | m36 |

5 Contribution to programme/key action objectives

MAGICSTER proposes to develop enabling technology which can help realise one of the general goals of the IST programme, namely the delivery of high-quality information services to European individuals in a user-friendly manner. There is evidence that the delivery of information in a personalised, conversational fashion results in higher quality information delivery than is otherwise possible; this increase in quality can be measured in terms of the speed with which the information is taken up and the degree to which the information is retained by individuals. There is also evidence that users feel more comfortable interacting with computational devices when the interactions are mediated through embodied agents. This illustrates that embodied believable agents can really contribute to the IST goal of developing high-quality, user-friendly information services. Developing the generic aspects of this technology to enable wide applicability and reuse in a range of applications and domains will also make this an economically viable technology.

We demonstrate and evaluate the viability of the MAGICSTER ideas in a small range of applications and domains. But MAGICSTER will be exploring generic research and development issues which in the longer-term will feed into a number of more specific IST activities. For example, MAGICSTER shares with key action III.4.2 (*Natural Interactivity*) the goal of enhancing human-computer interaction through the development of intuitive interfaces and knowledge representation for interaction. However, whereas III.4.2 focuses on the development of robust and scalable *language* technologies and their integration into highly interactive systems, MAGICSTER is concerned with technology that can be integrated and *synchronised* with these linguistic technologies and enhance them; at the same time, it would be premature to attempt this integration without further basic research and development work. In a similar vein, key action III.3.3 (*Advanced Training Systems*) is concerned with the development of simulation and animation technology which can be shown to improve individual and collaborative learning processes. MAGICSTER will develop technology which, in the longer term, will be of use in these processes; but without this further research and development work, embodied conversational agent technology is not at a stage yet where its educational impact can be properly evaluated by educational professionals.

This also illustrates why MAGICSTER is being pursued as a research activity under *Future and Emerging Technologies* (FET): FET promotes research that is of a longer-term nature or involves high risks for high pay-offs. There is evidence that embodied agents can improve human-computer interaction, and promise high pay-off. We can now begin to see a way in which such agents can be built; as explained in the section on innovation (Section 5), building these agents is not just a matter of incremental engineering, but requires some fundamental rethinking of basic technologies and architectures.

The use of embodied conversational agents carries a risk, in that building the agents the wrong way or using them for the wrong applications may result in a costly lack of improved interaction. Proper evaluation of this type of interface technology has been hindered by the absence of believable embodied agents. As embodied conversational agents begin to emerge, it becomes possible to evaluate them properly, to evaluate their component technology, and to develop an evaluation methodology for this type of technology. The kind of work carried out in MAGICSTER will thus help reduce the risk for future application-oriented projects in this area.

6 Innovation

The basic claim developed in MAGICSTER is not that for human-computer interaction to be effective it should resemble human-human interaction in all respects. However, conversational behaviour which combines linguistic and non-linguistic information is increasingly considered to be an important aspect of human-computer interaction; and embodied, animated agents are an obvious way of achieving this combination of different types of information delivery. At the same time, there is a growing realisation that users assign human-like properties and personalities to the computational devices they interact with; because of this, it is important to pay more attention to the way conversational behaviour is implemented in these computational devices and ensure they achieve a high degree of believability, and perhaps even come across as sentient beings. And whereas there is a lot of historical evidence that anthropomorphic interface designs have failed, it is only recently that we have reached a stage where agents can be implemented with sufficient believable properties to be useful, and to allow for rigorous evaluation.

The work in MAGICSTER is focused on the generation of multimodal, synchronised messages. The prototypes will allow for limited user input in different modalities, but multimodal input is itself not a topic for research in this project.

The type of interface we want to develop goes far beyond the abilities of current user interfaces, focusing on the personalisation of the user interface by developing innovations in the following key technologies:

Animation

A lot of progress has been made in the past few years in the animation of two-dimensional and three-dimensional characters. For the development of embodied conversational agents, considerably advanced animation is needed to provide the agents with at least those human-like characteristics that can be exploited in natural conversation. This includes facial movement (in particular lips, gaze, eyebrows) and body movement (in particular hand movement, direction of the body with respect to the human interlocutor, and movement of the body across the scene the agent is describing) in synchrony with the generated speech output or the user input. Where possible, we will use standards such as the MPEG-4 SNHC standard in the definition of the parameters driving the facial and body models. This will ensure greater portability of the project results.

Synchronisation

There are already agents that display verbal and non-verbal conversational behaviours—agents which move in relation to objects in the interface or which use pointing gestures in combination with spoken or written output. However, the relation between these different types of behaviour tends to be realised as additive, autonomous channels rather than synchronised channels.

To be able to exploit the affordances of embodiment, a much tighter coupling of the verbal and non-verbal signals is needed. This will require the development of a theory of discourse

which not only distinguishes between conversational goals that are propositional in nature (i.e. involve content) and those that are interactional in nature (i.e. relate directly to the current state of the discourse), but also distinguishes between form and function: a signal like the raising of the eyebrows on the animated face or a falling intonation contour in the synthesised speech can have many different functions; a function like confirming that the agent is still attending to the user's request can be performed by nodding or saying "U-hu". In the agent's actions, it should not be the case that language is primary and the other signals are in addition to the linguistic signal; such an approach will not allow one to fully exploit the embodiment of the agents. Instead, the different types of signal should be treated as equivalent, potentially equally valued alternatives defined by a single core meaning representation, and for any given function the best primary form, whether verbal or non-verbal, should be chosen and if necessary backed up by a complementary signal.

This requires a much richer meaning representation and a much tighter synchronisation of the different information signals than has so far been achieved, not just to achieve the right timing between signals, but also to ensure that the appropriate information signals play their proper role.

User modeling

Any form of user-adapted and context-sensitive system requires the use of an explicit model of the user, their changing knowledge as the conversation unfolds, and the context which the system can exploit. With the amount of information available on the Web, there is a very active area of research on user-adapted information retrieval and information delivery. The innovative aspect of the MAGICSTER approach is that it will attempt to use user models not only to select what information should be delivered, but also to determine *how* the various non-linguistic aspects of information delivery can be brought to bear on delivering the information in the most appropriate way.

Agent modeling

When interfaces look or behave more like people, there is a greater tendency to view them as having a particular personality, and this often leads to a lack of acceptability: for example, the provision of simple explanations can be interpreted as talking down to the user, helpfulness can easily be misinterpreted as meddling. Because of this, there is a need to ensure that agents project the right kind of "personality". This can be achieved by a range of characteristics, such as the choice of linguistic material, facial expression, or body language. We want to examine how these features combine to create individualised agents.

Concept to Speech

An important challenge in natural language generation is the integration of language generators and text-to-speech components into fully integrated concept-to-speech systems. The reason for this is that to ensure that synthesised speech sounds natural, it is important to control rhythm, accent, intonation, stress and other prosodic elements. Text-to-speech systems try to achieve this by analysing the input text. In an integrated system like the MAGICSTER

system, this analysis is not necessary, since the generator already has information such as the intended meaning, the purpose of the utterance, and the history and context of the exchange. Better control over prosody and related factors can therefore be achieved by tightly coupling the language generation and speech synthesis components, so that the generator can pass this information to the synthesiser in a principled fashion. But in addition to this, we want the spoken message to be delivered in synchrony with the movements of the lips of the conversational agent, its gaze and its movements. Whereas standard work in concept to speech generation works from information sources specifically designed for the generation of a *linguistic* message, in MAGICSTER we want to be able to generate from a representation which can also control these other types of behaviour. One of the goals of the early integration of the MAGICSTER speech and animation components is to be able to experiment with this tight coupling of different modules.

Natural speech synthesis

Standard diphone-based concatenative synthesis systems produce intelligible speech, but the result does not sound completely natural, even with plenty of prosodic and other information. An alternative that has been explored recently is unit selection, where an acoustically appropriate unit with the correct prosodic characteristics is selected from a database of units and no further prosodic modification is carried out. This makes it possible to produce sentences that sound much more natural. Especially in the relatively restricted domains of discourse we will be exploring in MAGICSTER it becomes possible to experiment further with a unit selection approach to speech synthesis and to achieve speech output which is superior to that of more general text-to-speech systems. But switching from diphone-based speech synthesis to unit selection has consequences for how synchronisation with lip movement is achieved, and this poses an additional challenge to be tackled in MAGICSTER.

Architectures for embodied agents

The study of architectures for agent-based information delivery is an active area of research. It covers issues such as the use of planning vs schema-based approaches in generating text and movement, the impact of pragmatic factors on deciding what to say and how to deliver the message, and serial versus incremental generation of the message.

Partners in MAGICSTER have independently developed architectures for the synchronised delivery of messages focusing on subsets of the information signals that we will be studying in MAGICSTER, and this will form the basis for the development of the MAGICSTER architecture. But the overriding goal is to integrate these components into a more general purpose architecture, so as to enable exploitation of MAGICSTER results in different applications and with different component technology. There has recently been work on the abstract specification of linguistic generation systems, and one of the strands of work in MAGICSTER will be to try and extend this abstract architecture for MAGICSTER-style information delivery. Advances in this area are expected to contribute to the wider deployment of conversationally interactive systems and the increased reuse of their component technology.

Annotation

After the development and rapid take-up of XML the World-Wide Web Consortium W3C decided that all future W3C standards should be based on and compatible with XML. SMIL, the Synchronised Multimedia Integration Language is just one example of an ever-expanding number of standards that play a role in on-line information delivery. In MAGICSTER we want to ensure that we remain compatible with these standardisation efforts. We will design and document an XML annotation scheme and Document Type Definition which we will use to annotate information with the meta-data necessary to deliver messages in different modalities in a synchronised form. To ensure that the annotation scheme develops as an open standard, results will be made available widely, and the project's Technology Take-up Club will be asked for feedback.

Evaluation

Usability analysis of a system is often performed by measuring whether users can use the system to perform a particular task efficiently, correctly or fast. The kind of interface we are developing in MAGICSTER requires a different type of evaluation, however, since its goal is not necessarily to help users perform a particular task or to save time. Other indicators of success, depending on the application domain, could include whether users acquire and can subsequently use the domain knowledge, whether they use the application for longer periods of time, whether they are willing to use it more often, or whether they feel more relaxed or amused. As part of MAGICSTER we want to contribute to the development of a new evaluation methodology for MAGICSTER-style interfaces. Building on current work in software engineering, information systems and psychology, the evaluation methodology will be developed in tandem with the work on the interface; to ensure wide applicability, the project's Technology Take-up Club will also play a substantial role in this.

7 Community added value and contribution to EU policies

It is still the case that the research and product development work in the domain of information delivery is concentrated on the major European languages. As a result, European citizens from communities of lesser-used languages may become excluded from the digital economy and the worldwide information society. One of Europe's social objectives is to ensure that these communities remain integrated. The mode of information delivery studied in MAGICSTER should help overcome some of these barriers, thus helping to achieve the goal of integration and of providing equal access to basic social and economic infrastructure for all language communities.

In addition to this, there are two other reasons for pursuing this project at the European level:

Complementarity of expertise. MAGICSTER is tackling a complex range of technologies, requiring the integration of animation, verbal and non-verbal behaviour, information extraction and delivery, text generation and speech synthesis. The partners were chosen because they are all specialists in complementary subareas of the required expertise, creating the overall skill set needed for this type of project.

Intercultural expertise. The development of embodied conversational agents and its use of non-linguistic information touches on a lot of issues which may be culturally or linguistically specific. The multilingual and multicultural nature of the MAGICSTER consortium will ensure that no development work is too closely tied to a single cultural practice.

8 Contribution to Community social objectives

The objective of MAGICSTER is to develop technology which will enable the construction of user-friendly information delivery services. This type of technology contributes to the EU's objectives of providing European citizens with better access to more relevant information, removing IT, linguistic and other barriers.

The chosen application domains also contribute to Europe's social objectives. For example, the domain of information delivery in museums contributes to the development of a multicultural society: developments in the cultural heritage domain are moving in the direction of 24-hour museums accessible via the web, and MAGICSTER technology will add to that the ability for cultural information to be accessible to a wider range of people from different cultures. Delivering the information in a manner contextualised to the type of access, the history of the interaction and the interests of the user will also contribute to a deeper understanding of the cultural information and generally lower access barriers to this information.

In other application domains, MAGICSTER technology will also help in achieving social objectives such as improving the quality of public services or improving access to learning.

9 Economic development; scientific and technological prospects

General

MAGICSTER contributes to the development of systems which combine linguistic message delivery with animation in a tightly synchronised fashion. There is a wide range of applications where this type of improved personalised information delivery can be exploited, including education (e.g. tutoring systems, information delivery in museums), e-commerce (e.g. personalised shopping assistants for on-line retail) and entertainment.

MAGICSTER is not trying to develop a product. But it does take seriously the development of concrete instantiations of its research and evaluation ideas, hence the range of prototypes and evaluations it will attempt. This will ensure that MAGICSTER ideas can be presented to a wide range of third parties, which in turn will ensure further scientific, technological and product development. The MAGICSTER dissemination strategy is the crucial building block in this respect.

Dissemination

In the course of the MAGICSTER project we intend to disseminate information about project developments mainly to potential users and vendors in the information world. We intend to use different forums for this:

Conferences, Workshops and Journals. There are several conferences, workshops and journals dedicated to IT issues in agent technology, animation, and language and speech generation. They provide important forums for disseminating scientific and technological results to R&D colleagues as well as to vendors and users in different application areas. But because of the wider applicability of the MAGICSTER technology, we will also report in more general forums, such as *Interacting with Computers*, the *New Review of Hypermedia and Multimedia*, the *Workshop on Intelligence and Multimodality in Multimedia Interfaces*, CHI, and general conferences such as IJCAI and CHI.

Specialised forums. The MAGICSTER work that is being carried out in particular application domains will be presented to conferences of importance for those domains. For example, the MAGICSTER partners active working on the museum application are members of MEDICI (Multimedia for EDucation and employment through Integrated Cultural Initiative). This is a Framework of co-operation established and supervised by DG XIII in conjunction with DG X of the European Commission. The goal of MEDICI is to promote the innovative use of multimedia technology for access to and exploitation of Europe's cultural heritage. MEDICI promotes partnerships between museums, cultural institutions and industry. It encourages the sharing of experiences among its participants, by providing information and other services to facilitate the use of advanced communications and information technologies for access to Europe's cultural heritage. MAGICSTER will make use of this European initiative to access a wider base of potential users and exploiters of one particular aspect of the MAGICSTER technology.

Exploitation

MAGICSTER already contains one partner (AME) with a commercial interest in exploiting MAGICSTER results, representing a wide customer base. Their presence in the consortium provides a direct exploitation opportunity for the MAGICSTER results. Several of the other partners (DFKI, EDIN, UPENN) have proven track records in technology transfer from research projects to commercial products. They will continue to develop these channels to ensure the proper exploitation of the MAGICSTER results.

To stimulate wider exploitation of MAGICSTER results in a European context, MAGICSTER will also establish a MAGICSTER Technology Take-up Club (TTC). The TTC will be a technology dissemination and feedback mechanism for MAGICSTER allowing us to establish a closer and longer-term relationship with users and vendors than is possible through the dissemination mechanisms.

Once MAGICSTER work is underway, the consortium will meet with members of the TTC at regular intervals to seek comment and advice on recent work as well as planned developments, and to give them early insight in MAGICSTER developments. Members of the TTC will be potential users and exploiters of the MAGICSTER technology, amongst them current museum clients of MAGICSTER partners. Their close but long-term involvement in MAGICSTER will considerably shorten the distance between the R&D work and the exploitation of the MAGICSTER results.

10 Workplan

10.1 General description

The workplan consists of 4 major work packages, centered around the major objectives of the MAGICSTER system, and dealing with:

WP1: Animated agents

WP2: Linguistic agents

WP3: Believable agents

WP4: Prototype development and evaluation

We will first briefly discuss the nature of the work to be tackled in each work package.

Work Package 1: Animated agents

This Work Package is concerned with the animation of agents, 3D talking heads as well as fully articulated agents. There is a wide body of work on agent animation, and several of the project partners have already worked on many different aspects of animation. The simulation of muscle contraction (Platt 1985, Waters 1987) and skin elasticity (Lee *et al* 1995) gives 3D human head models a high degree of realism, allowing for heads which show realistic lip movement (Cole *et al* 1998, Cohen *et al* 1996, Pelachaud *et al* 1996). In addition, these agents can move naturally and avoid obstacles in the 3D environment, they can use gaze and head movements to emulate turn taking and other conversational protocols, they can gesticulate while talking and can even show “emotional” reactions to simulated situations and/or interactions with a user (Cassell *et al* 1994, Elliott 1994, Noma *et al* 1999, Rickel and Johnson 1998, Poggi and Pelachaud 1998, Thorisson 1997). The current state of virtual human modeling and innovative work needed in MAGICSTER can be characterized at five different levels, each with a wide range of possible realisations (Badler *et al* 1999):

Appearance. Depending on the application, appearance can go from relatively simple 2-dimensional drawings, to 3-dimensional models with accurate surfaces, accurately modelled clothing and the ability to exhibit physiological effects such as perspiration or injury. For the application domains covered in MAGICSTER, some of these subtleties need not be modelled. But it will be useful to be able to display “emotions” (like irritation), and facial movement (eyebrows, gaze) should be under sufficient control to ensure it can be generated in synchronicity with other information channels.

Function. The way an agent functions can vary from functioning as a simple cartoon character to functioning as a jointed skeleton which correctly reflects the effects of fatigue or load-bearing activities. For the MAGICSTER applications, little is needed beyond realistic looking agents, without the need for deeper modeling of its functions.

Time. The timing of agent behaviour can vary from off-line animation to on-line, parameterized motion synthesis. MAGICSTER agents need to be developed at least up to that level. Further development (developing multiple agents, possibly in crowds or coordinated teams) is beyond the scope of MAGICSTER.

Autonomy. For some applications, agents require little autonomy, with their behaviour scripted beforehand or restricted to reacting to particular events. For the MAGICSTER applications, we need to go beyond that, developing agents which can communicate and which can take initiative in these communications.

Individuality. To be able to study the impact of embodied conversational agents, it is important they are more than just a generic character. Different agents should be able to display personality, gender or age differences.

Different applications require different types of capabilities along these five dimensions. The challenge in MAGICSTER is to build an animated agent with enough parameters to be usable across different application areas. In addition, we want to be able to generate agents displaying visual properties that can be made use of in face-to-face conversation. This requires the development of a representation language for agent simulation which can be interfaced with

the representation language used for the generation of the linguistic message (studied in Work Package 2), and which can also be used to integrate other aspects of agent behaviour which create *believable* agents.

Our original plan for development took as its starting point parameterized action representations first developed by Badler *et al* (1999). A parameterized action representation (PAR) gives a complete description of an action. It is called *parameterized* because an action depends on its participants—e.g. the agents who execute the action, their “personality”, or the objects involved in the action. A PAR includes applicability conditions and preparatory actions that have to be satisfied before the action is executed—a method similar to the use of *subgoals* in planning. For example, if the PAR specifies that an agent should discuss a particular object, but the agent is not close enough to the object or there are several objects nearby, then the agent would have to move towards the object or point to it, either before or during its delivery of the message. If the agent is expected to talk about a particular object but that object has not been talked about before, then the agent may use a pointing gesture as well as a particular intonation contour. If the PAR specifies that a contrast needs to be expressed to what has been said before, a particular intonation contour may have to be produced in combination with a particular facial expression.

A PAR can also describe the purpose of the action, its duration, and the manner in which it should be executed (carefully, quickly,...). The manner in which an action is executed is one of the parameters which can be used to reflect an agent’s “personality”, skill-level, or other individuating characteristics.

A PAR appears in two different forms. Uninstantiated PARs are stored hierarchically in a database called the Actionary. During execution, uninstantiated PARs get instantiated with specific information about the agent, the objects involved, the manner in which the action should be carried out, etc. As part of earlier development work, a module has been developed which converts natural language instructions to instantiated PARs. This was to be used to enable the close interaction between the animations and the natural language work (studied in Work Package 2). Deliverable D1.1 describes this representation.

Several of the partners in MAGICSTER have been involved in animation and human simulation for many years (especially ROMA), and in planning agent behaviour (especially DFKI), and they will be the main partners responsible for this Work Package.

Although capturing the facial expressions, body language and other gestures of users is also of importance in conversational interfaces, it is not a topic that will be pursued in MAGICSTER. Where necessitated by the application domain, this kind of input will be simulated or restricted to simple screen-based or mouse-controlled pointing.

During the first year it became apparent that PAR was not the best representation for our task. In particular it became clear that the walking behaviours and other non-directly purely communicative behaviours (such as grasping and reaching) for which PAR is best adapted were not relevant to the conversational behaviours that formed the central focus of MAGICSTER, and that PARs were both over- and under-engineered for the purpose. Instead we decided to use other inverse kinematics software already available from the UPENN website to describe reaching and deictic gestures and we are defining our own language (APML, developed by ROMA and BARI) to represent communicative behaviours involving gaze, facial expressions and hand gestures. Deliverables D1.2 and D4.2 describe this representation. Simultaneously,

UPENN indicated a desire to move from partner status to consultant status, while continuing to play an active advisory role in the project, albeit at a lower level of person-months than was originally projected for PAR development.

Funding for UPENN advisory activities will be absorbed in the existing budgets for ROMA and EDIN, which now include small items for UPENN consultancy, to be spent on travel for Badler and students.

Work Package 2: Linguistic agents

There are two different strands of work in this Work Package.

Firstly, MAGICSTER can build on substantial expertise in natural language generation, some of it developed by consortium partners, with a particular focus on technologies which allow the combination, within a single system, of deep generation techniques and of template based ones. For example, in ILEX (Oberlander *et al* 1999), generation is centered around operations on a content potential, which is a collection of facts that can be expressed, together with links between them. Facts are simple predicate-argument structures and can be partially or completely canned. Facts are annotated with numerical scores for assimilation (the extent to which they are assumed to be known by the reader), interest (to the reader) and importance (to the information provider). The first two of these are set according to the user model, developed in Work Package 3. Facts are linked by having entities in common (“elaboration”) and by having rhetorical relations to one another. Content is selected in an opportunistic way, starting from the known topic of the text and working out how to include those unassimilated facts that are most interesting or important.

We want to build on the resources that have come out of projects like these, and extend this type of work in three novel directions:

1. Most generation systems have so far looked at modeling messages to the context of delivery. In MAGICSTER it will be important also to tailor messages to the agent “personality”, to ensure that the embodied agent has the kind of personality that the user wants to interact with. This will also involve strong links with the work on identifying user types and the creation of believable agents (Work Package 3).
2. Systems like ILEX already tailor messages closely to the changing knowledge state of the user, both by selecting appropriate content and through the use of sentence-level attributes such as information aggregation, the generation of appropriate referring expressions, and the use of contrastive and other connectives. But the sentence-level adaptation sometimes still results in a stylistically awkward message. We want to improve on this situation by forging a closer link between the changing knowledge state and the emerging message, modeling in particular how the processing of the message itself will change the user’s knowledge state. For this, we can build on recent work on generation (e.g. the SPUD system, Stone and Doran 1997), which uses modal inference to determine the knowledge state of the reader as the message is being generated, thereby increasing the naturalness of the resulting messages. Following this proposal we have started work with Johanna Moore on this topic (see Zinn *et al* 2002 for relevant background). This aspect of the work is closely related to the work on Planning carried out in Work Package 3.
3. Because the tight coupling of different information channels is important, Work Package 3 will look at synchronisation issues which ensure that no single channel always dominates all other channels for information delivery. This will have repercussions for the representation that forms the input to the language generator: whereas most generation systems start from representations designed specifically for the generation of natural language text, the synchronisation requirement will create additional challenges in the design of the representation formalism (developed in Work Package 1), and the

generators developed in Work Package 2 will have to be compatible with the new representation formalism.

The second strand of work in this Work Package is concerned with the generation of speech, and synchronising speech with the other communication channels.

By tightly integrating a natural language generation system with a speech synthesis system, one should be able to get better speech output: whereas a text to speech system needs a detailed analysis of its input text to be able to produce naturally sounding prosody, a speech synthesis system integrated with a language generation system could get that information direct from the generation system, as in the Information Based Intonation Synthesis system of Steedman (1991) and Prevost & Steedman (1994).

Work in this Work Package will be concerned with determining what high-level linguistic information the generator should pass on to the speech synthesiser. The starting point will be recently reported work on discourse-level annotation (e.g. Hitzeman *et al* 1999). This annotation scheme is a set of XML tags (Extensible Markup Language) which serves as the interface between a generator and a speech synthesiser. The generation component generates lexical, structural, semantic and discourse-level information concerning the text it generates, and automatically annotates the text with this linguistic information using the set of XML tags. The speech synthesis component makes use of these tags when determining what the intonation contour, accent placement and other prosodic features of an utterance should be.

We want to build on this work in the development of an annotation scheme for the text generated in the three languages, and compatible with the XML standard for information annotation and exchange (Goldfarb *et al*, 1998) and more specifically with the emerging W3C standards for dialogue annotation and voice markup (Sproat *et al* 1998, Wouters 1999, Dahl 1999, McGlashan 1999). But we also want to improve the quality of the synthesised speech itself. Standard diphone-based concatenative synthesis systems produce intelligible speech, but the result does not sound completely natural. The main reason for this is the limited number of phone variants present in a typical system, and the potential artefacts introduced by concatenating at diphone boundaries. An alternative that has been explored recently is unit selection (Campbell and Black 1996, Black and Taylor 1997). By “unit selection” we mean the selection of some unit of speech which may be anything from a whole phrase down to a diphone or even smaller. Technically, diphone selection is a simple case of unit selection. However, unit selection in our sense is unlike diphone selection in that in unit selection there is more than one example of the unit and some mechanism is used to select acoustically appropriate units with the correct prosodic characteristics. Typically no further prosodic modification is done (see Donovan & Woodland 1995, Sagisaka *et al* 1992). An example of such a system is the ATR CHATR system (Hunt and Black 1996). With unit selection synthesis, it is possible to produce sentences that sound much more natural, although qualitatively speaking the output of such systems is currently still somewhat inconsistent. For purposes of achieving believability, it will be important both to achieve this higher quality speech output in a consistent manner, and that is still a substantial challenge.

For this work, we will use the freely available FESTIVAL speech synthesis system, which already comes with two forms of unit selection, and experiment further with different forms of unit selection. The starting point will be the construction of the unit database. It has been noticed that synthesis using unit selection techniques seem to retain aspects of the original database;

for example, if the database is broadcast news stories, the synthesis from it will typically sound like read news stories (or, more importantly, will sound best when it is reading news stories). Because of this, it is best for purposes of the MAGICSTER prototypes to collect databases of the kinds of descriptions one would typically want to hear.

In addition to this qualitative improvement in speech output, we also want the speech output to be closely synchronised with the animations developed in Work Package 1 (in particular lip movement, facial expressions and gestures). The output of the unit selection process is an accurately time-stamped sequence of phones, which drives the synthesiser. But with this information it should also be possible in theory to have fine-grained synchronisation with the moving lips, and this will be explored further in this Work Package. Other synchronisation issues will be developed as part of Work Package 3.

The work in Work Package 2 will be carried out mainly by EDIN, in close collaboration with the partners responsible for Work Packages 1 and 3. For the prototypes we will allow limited speech input, including intonationally sensitive speech input. But we will use results from other projects and products for this, rather than carry out development work on speech input as part of MAGICSTER.

Work Package 3: Believable agents

One of the important features of the MAGICSTER agent is *believability*. Part of creating a believable agent has to do with the ability to combine verbal and nonverbal behaviour in an appropriate and tightly synchronised way. In addition, the agent should also have the ability to adapt to the user and the context and to show a “personality” and, if needed, a type of “emotional” behaviour. Adaptation requires two types of modeling (of the Agent and of the User or Context) and representation of criteria to match the Agent’s with the User’s characteristics. In MAGICSTER work on believability will focus on synchronisation of verbal and nonverbal behaviour, and on adaptation to affective factors.

Synchronisation of verbal and nonverbal behaviour. The production of agents that engage in multimodal activity is not new. For example, Cassell *et al* (1994) describes conversing agents with automatically generated context-appropriate gestures, facial movements and intonational patterns. The work emphasises the production of non-verbal propositional behaviours that reinforce the speech content. But unlike the MAGICSTER agent, the agents were not designed to run in real-time. They also were not designed to interact with users. Lester *et al* (1997), Rickel & Johnson (1998), and Andre *et al* (1998) all model agents that display verbal and nonverbal conversational behaviours, that can move in relation to other objects in the interface, and that can use pointing gestures in combination with speech or text output. But in these agents the relation between verbal and non-verbal behaviour is additive, and the affordances of the body are not exploited for the kinds of tasks that it performs better than speech. To achieve this, we need to establish a much tighter coupling of the verbal and non-verbal signals. This will require the development of a theory of discourse which not only distinguishes between conversational goals that are propositional in nature (i.e. involve content) and those that are interactional in nature (i.e. relate directly to the current state of the discourse), but also between form and function: a signal like the raising of the eyebrows on the animated face or a falling intonation contour in the synthesised speech can have many different functions; a function like confirming that the agent is still attending to the user’s request can be performed in different ways. This theory of discourse will need to be integrated in the linguistic components emerging from Work Package 2, and in the animation technology developed in Work Package 1.

In the agent’s actions, the different types of signal should be treated as equally valued alternatives defined by a single core meaning representation, and for any given function the best primary form, whether verbal or non-verbal, should be chosen and if necessary backed up by a complementary signal. This requires a much richer meaning representation, which again will need to feed into the work carried out in Work Packages 1 and 2.

Gandalf (Thorisson 1996) comes much closer to MAGICSTER: it is an embodied interface agent inspired by studies of human psychosocial competencies which recognises and displays interactional information such as gaze, simple gestures, canned speech. But it shows limited ability to recognise and generate propositional information, has limited capacity for generating appropriate intonation, and limited ability for synchronising gestures with speech. The architecture proposed for Gandalf and developed further in Cassell *et al* (2000) will be an important starting point for the work in Work Package 4, but MAGICSTER intends to go well beyond Gandalf’s functionality through the development work in Work Packages 1, 2 and 3.

Adaptation to affective factors. Most work on adaptive human-computer interaction

has tended to focus on “rational” aspects of human behaviour, paying particular attention to the user’s knowledge, goals and interests. Recently, the use of other factors in adaptive information delivery has been attracting attention. Nass and colleagues were amongst the first to show that humans tend to attach a “personality” to information technology artifacts and that this personality may manifest itself in any graphical interface (Nass *et al* 1995), Moon and Nass 1996). Since then, much more research has been carried out in this area have begun (see, for instance, the contributions to the UM’99 Workshop on *Attitude, Personality and Emotions in User-Adapted Interaction* and the i3 Workshop on *Affect in Interactions*, both in 1999).

Adapting an animated agent’s personality to the context and enabling the agent to show ‘empathy’ with the presumed emotional state of the User requires solutions to at least three problems: (i) formalising (in the user and the agent model) the characteristics that are relevant for adaptation; (ii) attaching personality and an affective state to the agent in a manner that the user will find believable; and (iii) dynamically updating the agent and user models.

To address the first problem, we want to build on the Five Factor Model. This has been employed in many systems that try to attach a personality to their animated agents or to their graphical interface (see, for instance, Isbister and Nass 1998; Ball and Breese 1998; Andre *et al* 1999). Dominance (extraversion) and agreeableness are the two factors that have been considered in the majority of cases. In MAGICSTER we will refine the computational modeling of these factors and extend them by considering other aspects, such as cognitive overload, doubt, fatigue, boredom, and perhaps even propensity to irony. We will define how these aspects may be formalised in the user model; we will only marginally consider the problem of how the personality and the emotional state of the user may be recognised, by adopting a ‘stereotypical modeling’ approach, and will study how this model can be activated through a fast, non intrusive, initial model building phase and how the agent model can be activated accordingly.

In dealing with the second problem (how to attach a personality to the agent) we will build on the work of Loyall and Bates (1997), Walker *et al* (1997) and Arafa *et al* (1998) to achieve adaptation of text phrasing (Work Package 2) and changes in the external appearance of the agent (Work Package 1). But we also wish to vary more strategical aspects of the agent’s behaviour; to this aim, we will define a structure for the agent model and will study how the formalism of Parameterized Action Representations (PARS, see Work Package 1) may be enriched to enable the representation of affective aspects in the way actions are activated and performed by the agent. The library of PARS will be extended so as to make it possible to represent a variety of planning strategies, allowing different discourse plans to be generated in different conditions.

As for the third problem area (dynamic adaptation of the user and the agent model), we will investigate how these models may be updated after a verbal or nonverbal communicative act. This, again, will enable us to formalise into PARS the effects of these acts on the mental state of the user and the feedback this change will produce on the agent’s mental state. In particular, we will examine problems and advantages of representing uncertainty in this updating process.

The main partners responsible for the work in this Work Package will be DFKI, SICS and BARI, in close collaboration with the partner responsible for Work Packages 1 and 2.

Work Package 4: Prototype development and evaluation

This Work Package will bring together the different strands of work from the previous Work Packages.

To ensure that an early prototype can be built and developed throughout the project's lifetime, an important aspect of the early work in this Work Package will be the definition of the system architecture. Many of the components to be developed in the other Work Packages have existing prototypes (as detailed above) and descriptions of these will feed into the early architecture work as will earlier architecture work by Prevost *et al* (1999) and Cassell *et al* (2000). A detailed specification of the components that make up a runtime system as well as their interfaces will make it possible to build and test an integrated system early on, and continually improve individual components, without ever jeopardising the integrity of the overall system. AME will be the main partner responsible for this work.

In the course of the project, the prototypes will be instantiated in different applications—an educational application developed by ROMA, a museum application developed by EDIN, and a web-based application developed by SICS, building on earlier work by these partners on interactions involving limited modalities or not involving embodiment (Oberlander *et al* 1999, Höök *et al* 1999). To ensure that each of these partners can integrate the components needed for that application, early agreement on the architecture is needed.

Throughout the integration work, we will conduct evaluations of the emerging technology. There is evidence that the use of animated characters in interfaces makes a difference. For example, Koda & Maes (1996) and Takeuchi & Naito (1995) studied user responses to interfaces with static and animated faces, and users found both of these more engaging and entertaining than faceless interfaces. Similarly, Keisler & Sproull (1997) showed that users cooperate better with interface agents when they have a human face. Andre *et al* (1998) demonstrated that a presentation agent is considered more entertaining and helpful than an interface which does not use an agent. And a study of Gandalf, the agent developed by Thorisson (1996) showed that users consider Gandalf's language skills to be superior when Gandalf uses other conversational behaviour, such as gaze, turn-taking, limited gesture. In other words, without changing the language skills, they are perceived as better by adding additional information signals (Cassell 1999). This is consistent with Oviatt's (1999) work on mutual disambiguation, which shows that information from one channel can be used to disambiguate information or cope with information loss in a different channel.

This work shows that the presence of an animated character can make a difference to the effectiveness of the interface. In addition to this, we also want to evaluate to what extent *embodiment* of agents contributes to the efficiency of the interactions. In particular, we expect well-designed embodied conversational interface agents to address particular needs that are not met in current interfaces. e.g. ways of making dialogue systems robust in the face of imperfect speech recognition. But we also expect improved collaboration between humans and machines, and between humans mediated by machines, requiring a more advanced evaluation methodology: despite the tight coupling of various information signals, we want to be able to detect which signals contribute to which success factors, and which parts of the embodied agents do not contribute much or indeed make interactions worse. Determining this becomes even more crucial when evaluating the believability of the agent, where we need to establish which of the agent's features contribute to the creation of an agent that can be believed and

that users rely on. This requires us to develop a different evaluation methodology. We can build on the work by Reeves & Nass (1996) and Cassell (2000), but also on the evaluation methodologies developed by SICS for the evaluation of their adaptive and embodied interface agents (Höök 1998, 1999, 2000). This makes use of interviews and questionnaires, but also of metaphor analysis to measure the effect and affect of their agents.

References

- E. Andre, T. Rist, J. Mueller (1998). Integrating reactive and scripted behaviours in a life-like presentation agent. In *Proceedings of the Second International Conference on Autonomous Agents (Agents 98)*, 261-268. ACM Press.
- E. Andre, T. Rist, J. Mueller (1999). Employing AI methods to control the behaviour of animated interface agents. In *Applied Artificial Intelligence* 13, 415-448.
- J. Arafa *et al* (1998). Designing and building Personal Service Assistants with personality. In *WECC'98. The First Workshop on Embodied Conversational Characters*. Tahoe City, 1998.
- N. Badler *et al* (1999). Real time virtual humans. *International Conference on Digital Media Futures*. Bradford, UK, April 1999.
- N. Badler *et al* (2000). Parameterized Action Representation for virtual human agents. In J. Cassell, *et al* (eds.), *Embodied Conversational Agents*. MIT Press.
- N. Badler *et al* (2000). To gesture or not to gesture: what is the question? To be presented at *Computer Graphics International 2000*.
- G. Ball and J. Breese (1998). Emotion and Personality in a conversational character. In *WECC'98. The First Workshop on Embodied Conversational Characters*. Tahoe City, 1998.
- K. Binsted (1998). Designing Portable Characters. In *WECC'98. The First Workshop on Embodied Conversational Characters*. Tahoe City, 1998.
- N. Campbell and A. Black (1996). Prosody and the selection of source units for concatenative synthesis. In J. van Santen *et al* (eds), *Progress in speech synthesis*. Springer Verlag, pp279-282.
- J. Cassell, C. Pelachaud, N. Badler, M. Steedman *et al* (1994). Animated conversation: rule-based generation of facial display, gesture and spoken intonation for multiple conversational agents. *Computer Graphics* 28 (4): 413-420.
- J. Cassell (1999). Not Just Another Pretty Face: Embodied Conversational Interface Agents. In *Communications of the ACM*.
- J. Cassell and H. Vilhjálmsón (1999). Fully embodied conversational avatars: making communicative behaviours autonomous. In *Autonomous Agents and Multi-Agent Systems* 2, 45-64.
- J. Cassell (2000). Nudge Nudge Wink Wink: Elements of Face-to-Face Conversation for Embodied Conversational Agents. In J. Cassell, *et al* (eds.), *Embodied Conversational Agents*. MIT Press.
- J. Cassell *et al* (2000). Conversation as a System Framework: Designing Embodied Conversational Agents. In J. Cassell, *et al* (eds.), *Embodied Conversational Agents*. MIT Press.
- C. Castelfranchi, R. Falcone, F. de Rosis and S. Pizzutilo (1998). Personality traits and social attitudes in multiagent cooperation. In *Applied Artificial Intelligence*, 12, 1998.
- M.M. Cohen, R.L. Walker, and D.W. Massaro (1996). Perception of synthetic visual speech. In D.G. Stork and M.E. Hennecke (eds), *Speechreading by Humans and Machines: Models, Systems, and Applications*. NATO ASI Series.
- R. Cole *et al* (1998). Intelligent animated agents for interactive language training. Unpublished manuscript.
- D. Dahl (ed) (1999). Natural Language Processing Requirements for Voice Markup Languages. W3C Working Draft, 23 December 1999. <http://www.w3.org/TR/voice-nlu-reqs>
- F. de Rosis and F. Grasso (1999). Affective text generation. In *Workshop on Affect in Interaction*. October, 1999.
- F. de Rosis, B De Carolis and S. Pizzutilo (1999). Software Documentation with Animated Agents. In *5th ERCIM Workshop on User Interfaces for All*. December, 1999.
- R. Donovan and P. Woodland (1995). Improvements in an HMM-based speech synthesiser. In *Eurospeech95*, Madrid, Spain, pp573-576.
- C. Elliott (1994). Research problems in the use of a shallow artificial intelligence model of personality and emotion. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pp 9-15. Seattle, WA.
- C. Goldfarb and P. Prescod, P (1998). The XML Handbook. Prentice Hall.
- I.Haimovitz (1991). Modeling all dialogue system participants to generate empathetic responses. In *Computer Methods and Programs in Biomedicine*,

- 35, 1991.
- J. Hitzeman *et al* (1999). An annotation scheme for concept-to-speech synthesis system. In *Proceedings of the 7th European Workshop on Natural Language Generation*, pp59–66. Toulouse, France, May 1999.
- K. Höök (1998). Evaluating the Utility and Usability of an Adaptive Hypermedia System. In *Journal of Knowledge-Based Systems* **10** (5).
- K. Höök *et al* (1999). Light-weight measurements of affective responses. CHI short paper.
- K. Höök (2000): Steps to take before IUIs become real. In *Journal of Interaction with Computers*, in press.
- A. Hunt and A. Black (1996). Unit selection in a concatenative speech synthesis system using a large speech database. In *ICASSP96*, vol 1, pp373–376. Atlanta, Georgia.
- C. Isbister and C. Nass (1998). Personality in conversational characters: building better digital interaction partners using knowledge about human personality preferences and perceptions. In *WECC'98. The First Workshop on Embodied Conversational Characters*. Tahoe City, 1998.
- S. Keisler and L. Sproull (1997). Social” Human-Computer Interaction. In B. Friedman (ed), *Human Values and the Design of Computer Technology: CIKM'94*. ACM Press.
- T. Koda & P. Maes (1996). Agents with faces: the effect of personification. In *Proceedings of the Fifth IEEE International Workshop on Robot and Human Communication (RO-MAN'96)*, 189–194.
- D. Laurillard *et al* (1998). Affordances for learning in a non-linear narrative medium. In *American Educational Research Association Conference*. San Diego, April 1998.
- Y. Lee *et al* (1995). Realistic Modeling for Facial Animation. In *Computer Graphics Annual Conference Series*, pp55–62.
- J. Lester *et al* (1996). Focusing problem solving in design-centered learning environments. In *Proceedings of the Third International Conference on Intelligent Tutoring Systems*, pp475–483.
- J. Lester *et al* (1997). The Persona effect: affective impact of animated pedagogical agents. In S. Pemberton (ed), *Human Factors in Computing Systems: CHI'97 Conference Proceedings*, 359–366. New York: ACM Press.
- J. Lester and B. Stone (1997). Increasing believability in animated pedagogical agents. In *Agents'97. Proceedings of the First International Conference on Autonomous Agents*, pp16–21. Marina del Rey, CA. February, 1997.
- J. Lester *et al* (2000). Deictic believability: Coordinated gesture, locomotion, and speech in lifelike pedagogical agents. To appear in *Applied Artificial Intelligence*.
- A.B. Loyall and J.Bates (1997). Personality-rich believable agents that use language. In *Agents'97. Proceedings of the First International Conference on Autonomous Agents*, pp106–113. Marina del Rey, CA. February, 1997.
- S. McGlashan (ed) (1999). Dialog requirements for Voice Markup Languages. W3C Working Draft, 23 December 1999. <http://www.w3.org/TR/voice-dialog-reqs>
- P. Maes *et al* (1995). The ALIVE system: full-body interaction with autonomous agents. In *Proceedings of the Computer Animation '95 Conference*.
- C. Nass *et al* (1995). Can computer personalities be human personalities? *Int Journal Human-Computer Studies*, 43, 1995.
- T. Noma, L. Zhao and N Badler (2000). Design of a Virtual Human Presenter. Submitted to *IEEE Computer Graphics and Applications*.
- J. Oberlander *et al* (1999). Conversation in the museum: experiments in dynamic hypermedia with the intelligent labelling explorer. To appear in *New Review of Hypermedia and Multimedia*.
- S. Oviatt (1999). Mutual disambiguation of recognition errors in a multimodal architecture. In *CHI'99: Proceedings of the Conference on Human Factors in Computing Systems*, pp576-583.. New York, N.Y.: ACM Press.
- C. Pelachaud, N.I. Badler, and M. Steedman (1996) Generating Facial Expressions for Speech. In *Cognitive Science* **20** (1), pp1–46.
- R.W. Picard (1996). Does HAL cry digital tears? Emotion and computers. In D.Stork (ed), *HAL's Legacy – 2001's Computer as Dream and Reality*. MIT Press, 1996.
- S. Platt (1985). *A Structural Model of the Human Face*. PhD thesis, University of Pennsylvania.

- I. Poggi (1996). Mind Markers. In *5th International Pragmatics Conference*. Mexico City, June 1996.
- I. Poggi and C. Pelachaud (1998). Performative faces. *Speech Communication* **26**, pp5–21.
- S. Prevost and M. Steedman (1994). Specifying Intonation from Context for Speech Synthesis. *Speech Communication* **15**, pp139–153.
- S. Prevost *et al* (1999). Face-to-face interfaces. In M. Altom and M. Williams (eds): *Human Factors in Computing Systems: CHI'99*, 244-245. New York: ACM Press.
- B. Reeves and C. Nass (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press.
- J. Rickel and L. Johnson (1998). Dialogs with animated agents in virtual reality. In *Proceedings of the Workshop on Embodied Conversational Characters*, 39–46. Tahoe City, CA. October 1998.
- Y. Sagisaka *et al* (1992). ATR ν -TALK Speech Synthesis System. In *Proceedings of ICSLP92*, pp483-486.
- B. Schneiderman (1998). *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley.
- R. Sproat *et al* (1998). SABLE: A Standard for TTS Markup. In *Proceedings of the 5th International Conference on Spoken Language Processing*, pp1719-1724. Sydney, Australia, December 1998.
- M. Steedman (1991). Structure and Intonation. In *Language* **67** (2), pp260–296.
- M. Stone and C. Doran (1997). Sentence Planning as Description Using Tree Adjoining Grammar. In *Proceedings of ACL35 and EACL8*, Madrid, July, pp198-205.
- A. Takeuchi and T. Naito (1995). Situated facial displays: towards social interaction. In I. Katz *et al* (eds.), *Human Factors in Computing Systems: CHI'95 Conference Proceedings*, 450–455. New York: ACM Press.
- K. R. Thórisson (1996). Communicative Humanoids. A Computational Model of Psychosocial Dialogue Skills. PhD Thesis, MIT Media Laboratory, 1996.
- M. A. Walker, J. Cahn and S. Whittaker (1997). Improvising linguistic style: social and affective bases for Agent personality. In *Agents'97. Proceedings of the First International Conference on Autonomous Agents*, pp96–105. Marina del Rey, CA. February, 1997.
- K. Waters (1987). A Muscle Model for Animating Three-Dimensional Facial Expressions. *Computer Graphics*, **21** (4), pp17–24.
- J. Wouters *et al* (1999). Authoring Tools for Speech Synthesis using the Sable Markup Standard. In *Eurospeech'99*.
- C. Zinn, J. D. Moore and M. G. Core (2002), A 3-tier Planning Architecture for Managing Tutorial Dialogue, in *Intelligent Tutoring Systems, Sixth International Conference*, (ITS 2002), Biarritz, France, June 2002.

10.2 Workpackage list

| WP | WP Title | Lead | Pmonths | Start | End | DelNo |
|-------|---------------------------|------|--------------|-------|-----|-------|
| WP1 | Animated agents | ROMA | 58 | 1 | 33 | D1.x |
| WP2 | Linguistic agents | EDIN | 36 | 1 | 33 | D2.x |
| WP3 | Believable agents | BARI | 48 | 3 | 33 | D3.x |
| WP4 | Prototypes and evaluation | AME | 57.6 | 3 | 36 | D4.x |
| WP5 | Management | EDIN | 20.9 | 1 | 36 | D5.x |
| TOTAL | | | 220.5 | | | |

| | EDIN | ROMA | UPENN | DFKI | SICS | BARI | AME | TOTAL |
|--------|------|------|-------|------|------|------|-----|--------------|
| WP1 | 3 | 18 | 5 | 9 | 6 | 8 | 9 | 58 |
| WP2 | 18 | 6 | 0 | 3 | 4 | 2 | 3 | 36 |
| WP3 | 3 | 3 | 0 | 12 | 12 | 12 | 6 | 48 |
| WP4 | 8 | 7 | 0 | 6 | 7.6 | 8 | 21 | 57.6 |
| WP5 | 10 | 2 | 0.9 | 2 | 2 | 2 | 2 | 20.9 |
| TOTAL: | 42 | 36 | 5.9 | 32 | 31.6 | 32 | 41 | 220.5 |

10.3 Workpackage descriptions

| | | | | | | | |
|---|------|------|-------|------|------|----------------------------|-----|
| Workpackage number: WP1: Animated agents | | | | | | Start date: month 1 | |
| Participant: | EDIN | ROMA | UPENN | DFKI | SICS | BARI | AME |
| Personmonths: | 3 | 18 | 5 | 9 | 6 | 8 | 9 |

Objectives:

Development of a formalism for representing the conversational agent's planned behaviour and for driving the animation.

Development of improved animation facilities that allow fine-grained control of subtle parts of the animation, such as gaze, lip movement, facial expression and body movement.

Incorporation of advanced planning techniques to plan the synchronised delivery of information using different information channels.

Description of work:

The starting point for this work are the Parameterized Action Representations described in section 9.1. Early work will be concerned with the development of a common representation platform for use in the other workpackages. After that, work will concentrate on improving the rendering component, and integrating advanced discourse planning facilities in the action representations.

Prototypes in Work Package 1 will be concerned mainly with the display of particular types of animation. Integration of these animations with other types of behaviour (e.g. linguistic behaviour) is the subject of Work Package 3.

Other work from Work Package 1 feeding into Work Package 3 will be the development of Actionaries for the chosen application domains.

Deliverables:

D1.1: Parameterized Action Representations for embodied conversational agents – initial specification (month 6).

D1.2: MAGICSTER embodied conversational agent – first specification (month 12).

D1.3: Prototype of the MAGICSTER agent for application 1 (month 21).

D1.4: Prototype of the MAGICSTER agent for application 2 (month 28).

D1.5: Prototype of the MAGICSTER agent for application 3 (month 33).

Milestones and expected results:

month 12: first specification of the animation part of the MAGICSTER agent

months 21, 28, 33: subsequent prototypes of the animated agent, embodied in the chosen application domain and integrated in Work Package 4 with other information delivery components.

| | | | | | | |
|---|------|------|------|------|----------------------------|-----|
| Workpackage number: WP2: Linguistic agents | | | | | Start date: month 1 | |
| Participant: | EDIN | ROMA | DFKI | SICS | BARI | AME |
| Personmonths: | 18 | 6 | 3 | 4 | 2 | 3 |

Objectives:

The development of language generation and speech synthesis technology which can use representations compatible with the representations needed for animating agents.

Description of work:

We will be developing novel language generation technology which allows messages to be generated tailored to the changing discourse context and reflecting the “personality” of the animated agent. Of particular importance will be the development of generation components which can make use of a formalism compatible with the one developed in Work Package 1 for the representation of animated agents.

We will also be developing speech synthesis technology, which is closely integrated with the language generation technology, and ultimately with the animations. Where application domains allow, the speech synthesiser will take into account domain constraints and thus deliver higher quality speech.

Deliverables:

D2.1: Initial description of language and speech generation architecture for MAGICSTER (month 6).

D2.2: MAGICSTER language and speech generation - first spec (month 12).

D2.3: Language and speech generation for prototype 1 (month 21).

D2.4: Language and speech generation for prototype 2 (month 28).

D2.5: Language and speech generation for prototype 3 (month 33).

Milestones and expected results:

month 12: initial specification of the language and speech components;

months 21, 28, 33: instantiations of the language and speech components for the chosen applications.

| | | | | | | |
|---|------|------|------|------|----------------------------|-----|
| Workpackage number: WP3: Believable agents | | | | | Start date: month 3 | |
| Participant: | EDIN | ROMA | DFKI | SICS | BARI | AME |
| Personmonths: | 3 | 3 | 12 | 12 | 12 | 6 |

Objectives:

To create a believable agent which can combine verbal and nonverbal behaviour in an appropriate and tightly synchronised way.

To deliver a theory of multimodal discourse and a representation formalism which feeds into the work in Work Packages 1 and 2.

Description of work:

We will be developing a representation formalism and architecture which allow us to treat different types of information signal as equal alternatives, so that for any given discourse function the best primary form (verbal or non-verbal) can be chosen and if necessary backed up by a complementary signal.

We will also be developing the theoretical groundwork for agents that display “personality” and related individualising features. Of particular importance will be the ability to (dynamically) update the agent model, without losing the coherence of the individualising feature.

This work will feed into Work Packages 1 and 2. Specifically, we want to examine whether it is possible to extend the Parameterized Action Representations developed in Work Package 1 to take account of affective aspects of agent performance.

Deliverables:

D3.1: Initial requirements for believability (month 10)

D3.2: Believable information delivery for prototype 1 (month 21)

D3.3: Believable information delivery for prototype 2 (month 28)

D3.4: Believable information delivery for prototype 3 (month 33)

Milestones and expected results:

month 10: initial requirements on believable agents; months 21, 28, 33: different aspects of believability integrated and evaluated in subsequent prototypes.

Workpackage number: WP4: Prototype and evaluation **Start date:** month 1

| | | | | | | |
|----------------------|------|------|------|------|------|-----|
| Participant: | EDIN | ROMA | DFKI | SICS | BARI | AME |
| Personmonths: | 8 | 7 | 6 | 7.6 | 8 | 21 |

Objectives:

Development of the MAGICSTER architecture.

Delivery of integrated prototypes and evaluation of the MAGICSTER technology.

Description of work:

Early work in this work package will be concerned with the specification of the overall MAGICSTER architecture. This will allow all partners to deliver component technology according to carefully defined specs, and will allow the project to deliver gradually improving prototypes.

The application domains for the prototypes will be chosen from the partners' research portfolio, so as not to create the need for additional resources in creating ontologies, termbases or other resources needed for the domain. In the course of the first year, the application domains for prototypes 1, 2 and 3 will be decided, taking into account the functionality needed for these applications and the availability of non-MAGICSTER resources.

Prototypes will be evaluated by users and/or in laboratory conditions, as appropriate to the application domain. These evaluations will automatically feed into the next prototype. At the end of the project, an overall evaluation report will be produced.

In a separate deliverable, the MAGICSTER technology will be documented for future exploitation and development work.

Deliverables:

D4.1: Initial architecture specification (month 8).

D4.2: Final architecture specification; definition of application domains and prototypes (month 12).

D4.3: First prototype, integrating D1.3, D2.3 and D3.2 (month 21).

D4.4: Second prototype, integrating D1.4, D2.4 and D3.3 (month 28).

D4.5: Third prototype, integrating D1.5, D2.5 and D3.4 (month 33).

D4.6: Final evaluation report (month 36).

D4.7: Technical documentation (month 36).

Milestones and expected results:

month 12: MAGICSTER architecture defined;

months 21, 28 and 33: prototypes in the chosen application domains;

month 36: final evaluation reports on the prototypes and technical documentation.

| | | | | | | | |
|--|------|------|-------|------|------|----------------------------|-----|
| Workpackage number: WP5: Management | | | | | | Start date: month 1 | |
| Participant: | EDIN | ROMA | UPENN | DFKI | SICS | BARI | AME |
| Personmonths: | 10 | 2 | 0.9 | 2 | 2 | 2 | 2 |

Objectives:

To ensure the proper execution of the project work, and a wide dissemination of the project results.

Description of work:

The management structure for MAGICSTER is described in section 12. In brief, there will be a Management Board consisting of the project manager (from EDIN) and a representative from each of the partners (the site managers). There will also be a Technology Take-up Club, which the project will use to seek directed feedback on the ongoing work.

The Management Board will be responsible for the proper execution of the work, the correct expenditure of project funds, and the delivery of the contractually required reports. It will also be responsible for outreach activities, as detailed in the section on project management and dissemination (section 12).

The consortium will produce:

- semestrial progress reports, following the prescribed IST format;
- annual public reports, made available on the Web and in other suitable formats;
- final report according to the prescribed HLT format;

In addition to this, the consortium will also provide:

Project Description. A possibly revised and more detailed version of the workplan in the technical annex, taking into account changes that may have occurred between the preparation of the Technical Annex and the Operative Commencement date of the project.

Technology Exploitation Plan. This will summarise the MAGICSTER foreground rights, and specify the consortium's plans for the exploitation of project results.

Deliverables:

quarterly and annual progress reports;

public annual reports

month 2: project description (D5.1)

month 36: technology exploitation plan (D5.4)

10.4 Deliverables list

| DelNo | Deliverable Title | Date | Type | Diss |
|-------|--|----------|------|------|
| D1.1 | PARs - initial spec | month 6 | R | CO |
| D1.2 | Embodied conversation agent - first spec | month 12 | R | RE |
| D1.3 | Prototype of animated agent for prototype 1 | month 21 | R,P | PU |
| D1.4 | Prototype of animated agent for prototype 2 | month 28 | R,P | PU |
| D1.5 | Prototype of animated agent for prototype 3 | month 33 | R,P | PU |
| D2.1 | Initial description of lge and speech generation architecture | month 6 | R | RE |
| D2.2 | Linguistic agents - first spec | month 12 | R | CO |
| D2.3 | Lge and speech generation for prototype 1 | month 21 | R,P | PU |
| D2.4 | Lge and speech generation for prototype 2 | month 28 | R,P | PU |
| D2.5 | Lge and speech generation for prototype 3 | month 33 | R,P | PU |
| D3.1 | Initial requirements for believability | month 10 | R | CO |
| D3.2 | Believable information delivery for prototype 1 | month 21 | R,P | PU |
| D3.3 | Believable information delivery for prototype 2 | month 28 | R,P | PU |
| D3.4 | Believable information delivery for prototype 3 | month 33 | R,P | PU |
| D4.1 | Initial architecture specification | month 8 | R | CO |
| D4.2 | Final architecture specification; definition of application domains and prototypes | month 12 | R | RE |
| D4.3 | First prototype, integrating D1.3, D2.3 and D3.2 | month 21 | P | PU |
| D4.4 | Second prototype, integrating D1.4, D2.4 and D3.3 | month 28 | P | PU |
| D4.5 | Third prototype, integrating D1.5, D2.5 and D3.4 | month 33 | P | PU |
| D4.6 | Final evaluation report | month 36 | R | PU |
| D4.7 | Technical documentation | month 36 | R | PU |
| D5.1 | Project description | month 2 | R | RE |
| D5.4 | Technology Exploitation Plan | month 36 | R | RE |

R = Report

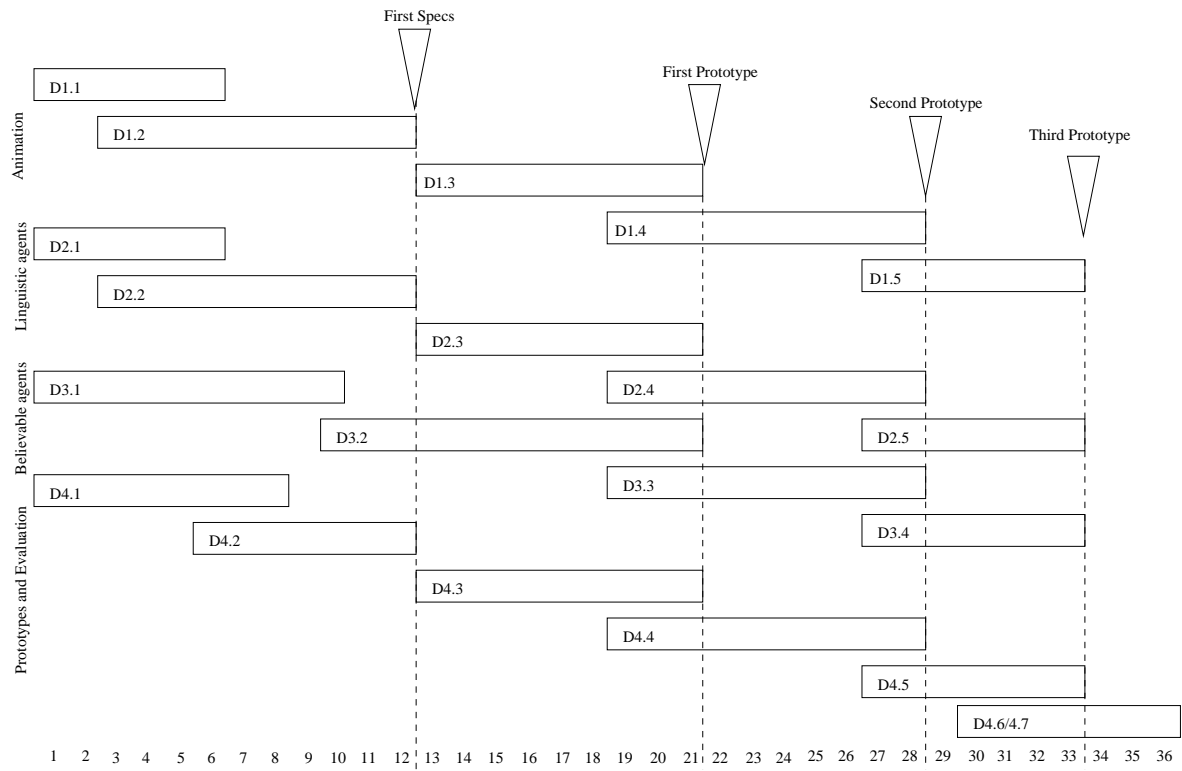
P = Prototype

CO = Confidential, Consortium and Commission services only

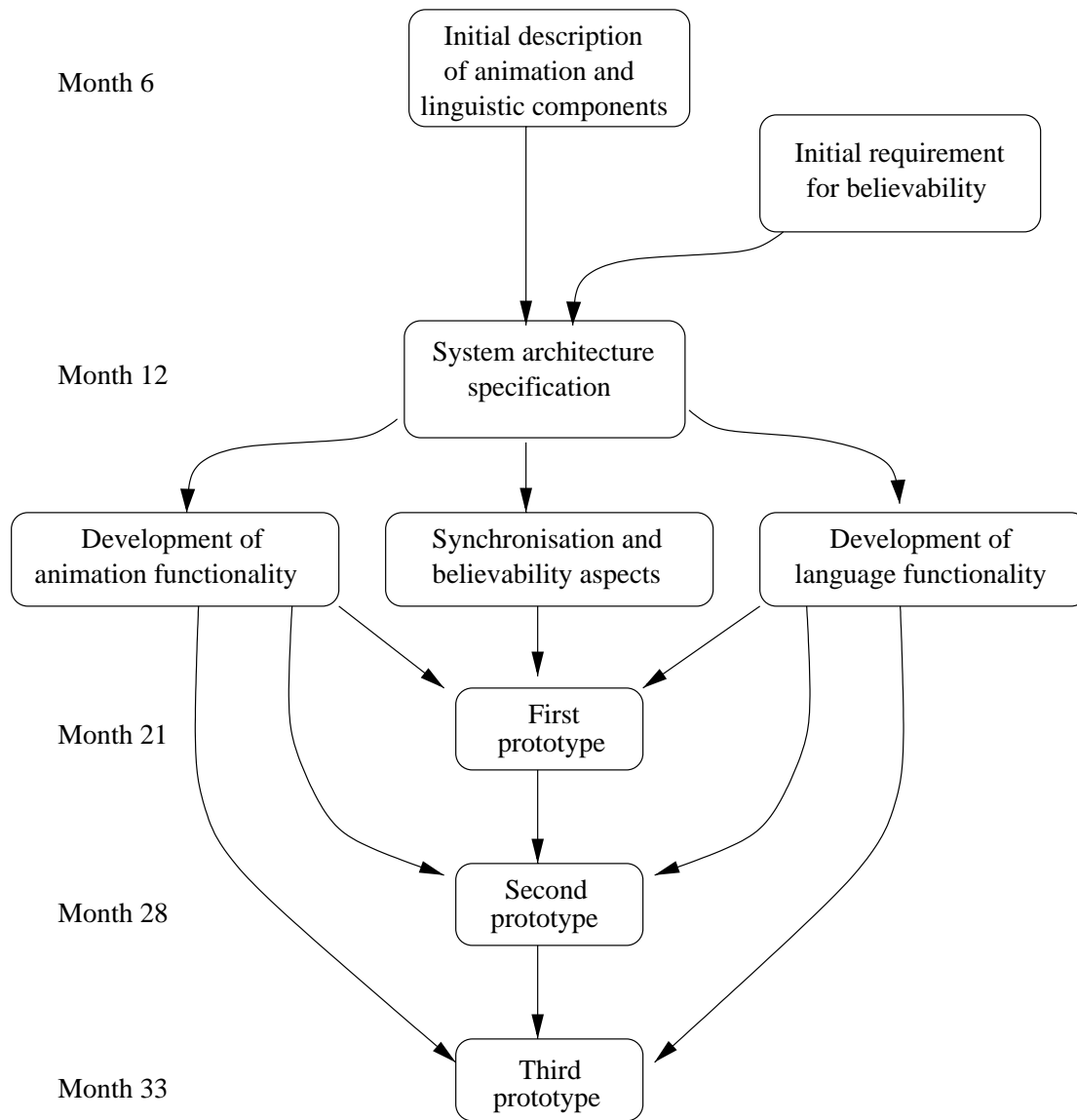
PU = Public

RE = Restricted to Consortium, Commission Services and Technology Take-up Club

11 Project planning and timetable



12 Graphical presentation of project components



13 Project Management

The management of MAGICSTER is organised as follows:

Site Managers: senior people in their organisations, who will be responsible for informing the coordinator of progress and potential problems concerning the work carried out at their site; ensuring the correct and auditable expenditure of project funds at their site; and submitting cost statements to the coordinator at the times stipulated in the contract.

Site Managers will attend meetings of the Management Board, or send a fully authorised representative.

Work Package Leaders: The WP Leaders coordinate the work relevant to each WP, to ensure that the work meets the objectives of those WPs, that the deliverables are finished by the due dates, and that work spanning WPs is coordinated properly. They report to the Coordinator on progress and any potential problems. WP Leaders will attend meetings of the Management Board, or be represented by their Site Manager.

Coordinator: His responsibilities will include: submission of reports and deliverables to the Commission; general liaison and communication with the Commission services; financial management of the project, including receipt and distribution of project payments; chairing the Management Board.

The Management Board will be chaired and convened by the Coordinator and consist of the Site Managers, with the WP Leaders in attendance. The Management Board will deal with resource allocation, monitoring of progress against the work plan, refinement of the plan where needed, and communication with the Technology Take-up Club. The Management Board will meet around the time of the 4-monthly project meetings. Additional meetings of the Management Board may be requested by any of the Site Managers or the Coordinator. It will conduct the rest of its business electronically or via teleconferences. For all problems arising in the project, a consensus decision will be sought. Where such a consensus can't be reached, a majority vote will be sought from the seven Site Managers; if there are absentions and no majority is reached, the Coordinator has a deciding vote.

A Technology Take-up Club (TTC) will be established around the end of the first project year. The composition of the TTC will be based on proposals from members of the Management Board, in particular from the user sites in the MAGICSTER consortium, and will consist of representatives from potential users and relevant IT vendors. Members of the TTC will be asked to comment on the internal and publicly available deliverables and work plans. This will ensure an external control on the quality of the work delivered, and provide them with a considerable competitive advantage should they wish to exploit MAGICSTER technology.

Agreements and Licences

All partners will be entitled to use MAGICSTER technology for internal purposes. However, the partners will identify information and software that is proprietary and brought to the project

as background material. They will identify the terms and conditions under which this material is being made available by the owner to the other contractors. The contractors will identify material that is protected by copyright, and ensure that this protection is maintained during its use within and after the project. The partners will agree and confirm the intellectual and exploitation rights and ownership pertaining to project results in a Consortium Agreement, to be negotiated and signed at the time of signature of the project contract.

Dissemination

Results of the project will be disseminated through the normal channels of presentations at specialised workshops, conferences and journals. But a more important channel for dissemination of emerging project results will be through presentations to the Technology Take-up Club, as well as to occasional targeted dissemination workshops, organised around the delivery of the major prototypes.

Dissemination is considered to be an inherent part of all Work Packages and a responsibility for all partners. There is a separate line in the Coordinator's budget to finance common costs arising out of dissemination activities.

14 Clustering

15 Other contractual conditions

Since one of the MAGICSTER partners is situated outside the European Union, there may be some travel by partners outside the EU Member states.

Three partners will be subcontracting some of their work. We provide summaries of these subcontracts here. More details on this are given in the partners' site descriptions (Appendix A).

ROMA: collaboration with the linguistics department of the University Roma Tre on the use of emotion in agents; collaboration with the Institute of Phonetics and Dialectology of the Italian National Research Council on the synchronisation of speech and movement in Italian;

BARI: evaluation work on the animated agents, to be subcontracted out to Reading university (UK); some corpus collection work for one of the applications, to be subcontracted out to Rome university's radiology department;

AVATARME: some of the implementation work will be carried out by ADA (Bulgaria). AvatarMe closely manages this subcontract work and retains all rights to the results. ADA was incorporated at the start of 1997 and performs contract Research and Development for 3D Scanners and AvatarMe. The company has specialized in the fields of avatars, 3D data capture, vision, image processing and automatic human model generation. Supplementary funding for ADA has been sought.

In the original contract UPENN has an item under Other Significant Project Costs, labelled "Graduate Tuition". This allows a graduate student to participate in the MAGICSTER project. This item has been moved to the ROMA budget to reflect the changed status of UPENN, and will be paid for the same purpose under the heading of consultancy.

Appendix A: Consortium Description

The consortium carrying out the MAGICSTER projects consists of six research and development groups and a SME specialising in multimedia product development:

EDIN: the Institute for Communicating and Collaborative Systems, the Language Technology Group and the Centre for Speech Technology Research, based at the University of Edinburgh, have extensive expertise in natural language generation, information extraction and speech synthesis; they are also the coordinators of MAGICSTER and have coordinated a wide range of European and other projects before.

ROMA: the Dipartimento di Informatica e Sistemistica of the University of Rome “La Sapienza” has a substantial R&D track record in animation and agent representation, some of it in collaboration with UPENN, who now act in a consultant role.

UPENN: the Center for Human Modeling and Simulation at the University of Pennsylvania specialises in computer graphics and animation, especially for representing synthetic humans. They participated as a full partner with no funding from the EU for the first 12 months and made considerable contributions to the project. They continue to be involved as consultants.

DFKI: the Deutsches Forschungszentrum für Künstliche Intelligenz specialises in innovative software technology based on Artificial Intelligence methods. Its Intelligent User Interfaces group is one of the leading research and development groups in the area of animated interface agents

SICS: the Swedish Institute of Computer Science pursues a wide range of research interests, but will particularly contribute expertise in the area of human-computer interaction, intelligent interfaces and social aspects of animated agents.

BARI: the research group on Intelligent Interfaces from the Department of Informatics at the University of Bari has a long track record in user modeling, user-adapted interaction, as well as in the area of agents which can display personality.

AVATARME: UK company developing and marketing photorealistic avatars of people. Its flagship project is the AvatarBooth, which is used for the fast creation of 3D avatars.

These teams have a proven track record and complementary expertise in the areas of crucial concern to the project, viz. animation, language engineering, agent modeling, and software engineering and component delivery. Through subcontracting and through interactions with related projects, three of the partners also bring in application domains and users. The relevant expertise and contributions of each of the MAGICSTER partners can be represented schematically as follows:

| | Animation | Language and Speech | Agent behaviour | Software Engineering | Application and Users |
|-------|-----------|------------------------|--------------------|-------------------------|--------------------------|
| EDIN | ✓ | ✓✓ | | ✓✓ | ✓ |
| ROMA | ✓✓ | | ✓ | | ✓ |
| UPENN | ✓ | | | | |
| DFKI | | ✓✓ | ✓✓ | ✓ | |
| SICS | ✓ | ✓ | ✓✓ | ✓ | |
| BARI | ✓ | | ✓✓ | | ✓ |
| AME | ✓✓ | | | ✓✓ | ✓✓ |

In the next sections we provide more detailed overviews of each of the partners, as well as short CVs and recent publications of key MAGICSTER personnel.

EDIN

The Institute for Communicating and Collaborative Systems at the University of Edinburgh is coordinator of this proposal. The work at Edinburgh will be collaborative between the Institute, the Language Technology Group and the Centre for Speech Technology Research. This brings together one of the largest groups of language and speech engineers, cognitive scientists and interface designers in the UK. They have substantial expertise in dialogue engineering, speech synthesis (its FESTIVAL synthesis system has been licenced to over 3000 industrial and academic users), and natural language generation. They are members of the World Wide Web Consortium, and have been actively involved in the development of XML and related annotation standards. Principal proposers are Jean Carletta, Marc Moens, Jon Oberlander, Mark Steedman, Paul Taylor.

Jean Carletta is a Senior Research Scientist with the Language Technology Group in the Institute for Communicating and Collaborative Systems. She has been involved in projects to develop speech and language mark-up technologies. One such project, MATE, has recently reviewed several of the levels of language mark-up needed for the MAGICSTER project and has developed a workbench of tools supporting their use.

Marc Moens is manager of the Language Technology Group and will act as project coordinator. He has been involved in around 20 European Union projects, either as researcher, site manager or coordinator. His current research focuses on information extraction, text summarisation and other forms of document handling.

Jon Oberlander is lecturer in the Institute for Communicating and Collaborative Systems. He holds two university degrees: BA in Philosophy, Cambridge 1983; PhD in Cognitive Science, Edinburgh 1987. His interests are in discourse theory, natural language generation, and multimodal reasoning and communication. He is one of the designers of the ILEX generation system for information delivery to museum visitors. This will be one of the application and evaluation domains in MAGICSTER.

Paul Taylor holds a BEng in Acoustics and a PhD in Linguistics. He holds a lectureship at the University of Edinburgh and is director of the Centre for Speech Technology Research at the University. He has worked in many areas of speech technology, specifically speech synthesis, speech recognition and intonation. He is an experienced system developer, having contributed to many large scale speech systems. His main achievements in this area have been the CHATR system, developed while at ATR, Japan, and the FESTIVAL speech synthesis system which he has been project leader and co-developer on for the last 4 years. FESTIVAL is a widely used open-source synthesis development system and toolkit, and is used as the basis of many labs' research around the world. To date it has been ported to 20 languages. His main scientific achievements have been in the areas of intonation, speech synthesis interfaces, unit-selection synthesis and intergration of synthesis with natural language generation systems. He is author and co-author of over 40 journal and conference papers.

Mark Steedman is Professor of Cognitive Science and Director of the Institute in Communicating and Collaborative Systems in the Division of Informatics. He is a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI). His research covers issues in computational linguistics, artificial intelligence, computer science and cognitive science, including syntax and semantics of natural language, parsing and comprehension of natural

language discourse by humans and by machine, natural language generation, spoken discourse including dialogue by animated agent, and the place of intonation and prosody in grammar.

Some recent publications:

J. Carletta (1996). Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics* **22**(2), pp249–254.

J. Carletta *et al* (1997). The Reliability of a Dialogue Structure Coding Scheme. *Computational Linguistics* **23**(1), pp13–31.

J. Carletta *et al* (1999). The MATE Annotation Workbench: User Requirements. In *Proceedings of "Towards Standards and Tools for Discourse", ACL-99 Workshop*. University of Maryland, June 20-26.

J. Carletta *et al* (2000) The effects of multimedia communication technology on non-collocated teams: a case study. In *Ergonomics*, to appear.

J. Cassell, C. Pelachaud, N. Badler, M. Steedman *et al* (1994). Animated conversation: rule-based generation of facial display, gesture and spoken intonation for multiple conversational agents. *Computer Graphics* 28 (4): 413–420.

J. Hitzeman, A. Black, C. Mellish, J. Oberlander, M. Poesio and P. Taylor (1999). An annotation scheme for concept-to-speech synthesis system. In *Proceedings of the 7th European Workshop on Natural Language Generation*, pp59–66. Toulouse, France, May 1999.

C. Mellish, M. O'Donnell, J. Oberlander and Knott, A. (1998) An architecture for opportunistic text generation. In *Proceedings of the 9th International Generation Workshop*. Niagara, Canada, August 1998

J. Oberlander and C. Brew (2000): Stochastic text generation. In *Philosophical Transactions of the Royal Society of London*, Series A, **358**, to appear.

M. Steedman (1996). Representing Discourse Information for Spoken Dialogue Generation. In *Proceedings of International Symposium on Spoken Dialogue*, pp89–92. Philadelphia, September 1996.

M. Steedman (1999). Information Structure and the Syntax-Phonology Interface. August, 1999.

M. Steedman (2000). *The Syntactic Process*. MIT Press (to appear).

M. Steedman and S. Prevost (1994). Specifying Intonation from Context for Speech Synthesis. *Speech Communication* **15**, pp139–153.

P. Taylor and A. Black (1999). Speech Synthesis by Phonological Structure Matching. In *Eurospeech99*.

P. Taylor and A. Black (1998). Assigning Phrase Breaks from Part of Speech Sequences. In *Computer Speech and Language* **12**, pp99–117.

S. Teufel, J. Carletta and M. Moens (1999). An annotation scheme for discourse-level argumentation in research articles. In *Proceedings of the Ninth Conference of the European Chapter of the Association for Computational Linguistics (EACL-99)*, pp110-117.

ROMA

The *Dipartimento di Informatica e Sistemistica* of the University of Rome “La Sapienza” is the main research department in Computer Science in Rome and one of the largest in Italy. It has a long tradition of research in the area of Artificial Intelligence and Computer Graphics. The Artificial Intelligence research group is mainly working in knowledge representation, planning, search and Natural Language Processing. A particular interest is on the definition of methods and techniques for reasoning about actions, and for the design and the realization of actual agents (in particular mobile robots) that are able to accomplish complex tasks in real environments. On the graphics side, special interests are on the creation of systems which automatically generate and animate conversations between multiple human-like agents with appropriate and synchronized speech, intonation, facial expressions, and hand gestures. Some of the work ROMA is responsible for will be carried out in collaboration with the Radiology and Image-Based Diagnosis Section, Department of Experimental Medicine and Pathology at the University of Roma. In the Radiology and Image-Based Diagnosis Section of the Department of Experimental Medicine, about 80.000 diagnostic tests of various types are performed every year, with the most advanced technologies. This Section is also a ‘Center of Excellence’ in digital image processing. This will constitute one of the application domains in MAGIC-STER. Principal proposers are Catherine Pelachaud and Marco Schaerf, in collaboration with Emanuela Caldognetto and Isabella Poggi (subcontractors).

Catherine Pelachaud received a PhD in Computer Graphics at the University of Pennsylvania, Philadelphia, USA 1991. In 1993-1994 she was involved, with Norman Badler (UPENN) and Mark Steedman (EDIN) in a Project funded by a National Science Foundation grant, that implemented a system which automatically generates and animates conversations between multiple human-like agents with appropriate and synchronized speech, intonation, facial expressions, and hand gestures. In 1993 she was part of the organisation of a workshop, sponsored by NSF, on standards for facial animation hold at the University of Pennsylvania. In 1993-1996 she was a post-doctorate at the University of Rome at the computer science department with a Human Capital and Mobility grant. In 1998 she worked in the EAGLES group (Expert Advisory Group on Language Engineering Standards), project funded by the European Community to promote standards and distribution of resources in the spoken language field. Her research interests include computer animation, conversational agents, human behaviour simulation and multimedia systems.

Marco Schaerf is currently Associate Professor of Computer Science at the University of Roma “La Sapienza”. He published over 15 journal papers and 40 conference papers on Artificial Intelligence and Computer Graphics. His research interests include: Logics for Knowledge Representation, Nonmonotonic Reasoning, Logic Programming, Theory of Agents, Theory of Databases, Computational Complexity, Geometric Reasoning and Computer Graphics. In 1997 he won the Italian AI*IA Prize for the best young researcher in Artificial Intelligence. He teaches Computer Graphics for Computer Engineering students. He is currently involved in many national and international projects where he focuses on the combination of Artificial Intelligence methodologies and Computer Graphics tools and techniques.

Isabella Poggi is an Assistant Professor at the Department of Linguistics of the University Roma Tre. Her main interests are in linguistic education, pragmatics and theory of communication, multimodal communication and psychology of emotions. Adopting a cognitive

approach she participates in the construction of a goal and belief model of mind and social interaction with Castelfranchi at the Institute of Psychology at the National Council of Research in Rome. She has written books and papers on a cognitive approach to language teaching, interjections, gestures and other nonverbal behaviors, deception, and emotions such as guilt, shame, compassion.

Emanuela Magno Caldognetto is a senior researcher and research manager at the Institute of Phonetics and Dialectology of C.N.R. (Italian National Research Council). She has been supervisor of several national projects. From 1996 to 1998 she was the supervisor of the Executive Division for the “Co-ordination of gestures and words in normal and pathological subjects” within the co-ordinated national project on Multi-modal communication. She also supervises projects on Planning of verbal and co-verbal gestural utterances, and on the production and perception of lips and jaw articulatory movements. Her main areas of interest range from the study of macro-planning and execution processes on the basis of analyses of short-term linguistic malfunctions (slips of the tongue) in normal subjects to the research on the relationships between co-produced verbal and gestural messages.

Some recent publications:

J. Cassell, C. Pelachaud, N. Badler, M. Steedman *et al* (1994). Animated conversation: rule-based generation of facial display, gesture and spoken intonation for multiple conversational agents. *Computer Graphics* 28 (4): 413–420.

E. Magno-Caldognetto and I. Poggi (1997). Micro- and macro-bimodality. In *Proceedings of ESCA, AVSP'97 workshop*. Rhodes, September 1997.

C. Pelachaud, N.I. Badler, and M. Steedman (1996). Generating Facial Expressions for Speech. In *Cognitive Science* 20 (1), pp1–46.

N. Pezzato and I. Poggi (1998): The alphabet and the lexicon of eyes. In *6th International Pragmatics Conference*. Reims, July 1998.

I. Poggi (1996). Mind Markers. In *5th International Pragmatics Conference*. Mexico City, July 1996.

I. Poggi and E. Magno Caldognetto (1996). A score for the analysis of gestures in multimodal communication. In *Proceedings of the Workshop on the Integration of Gesture and Language in Speech, Applied Science and Engineering Laboratories*, pp235–244.

I. Poggi and C. Pelachaud (1998). Performative faces. *Speech Communication* 26, pp5–21.

I. Poggi and C. Pelachaud (1999). Emotional Meaning and Expression in Performative Faces. In *International Workshop on Affect in Interactions: Towards a New Generation of Interfaces*. Siena, October, 1999.

I. Poggi and C. Pelachaud (2000). Facial Performative in a Conversational System. In J. Cassell *et al* (eds), *Embodied Conversational Characters*. MIT Press.

UPENN

The University of Pennsylvania is America's first university and one of its foremost institutions of higher education. It is a national leader in interdisciplinary programs that combine academic theory with professional practice. The faculty and students in the Department of Computer and Information Science engage in a variety of research programs at the forefront of computer science, including artificial intelligence, computer vision, computer graphics, robotics, natural language processing, expert systems, knowledge representation, learning, computer architecture, massively parallel systems, networking, distributed computing, theory of computation, logic, advanced data management systems, biomedical computing, algorithms, computational biology, and medical informatics. The department also has a number of ongoing research collaborations with national and international organizations and laboratories, and has been the recipient of several National Science Foundation (NSF) infrastructure grants, major Army Research Office (ARO) awards, and Defence Advanced Research Project Agency (DARPA) awards and has led five departments at the University of Pennsylvania in a Cognitive Science research program, which has resulted in the award of an NSF Science and Technology Center grant—the first ever to support research in this area.

For the first 12 months of the project the MAGICSTER work will be carried out in the *Center for Human Modeling and Simulation* (HMS). The Center continues to advise the project in a consultative role. The overall goals of the Center are the investigation of computer graphics modeling, animation, and rendering techniques, especially for synthetic humans. Major foci are in behavior-based animation of human movement, modeling through physics-based techniques, biomedical modeling, applications of control theory techniques to dynamic models, and understanding the relationship between human movement, natural language, and communication. The Center originated the Jack software which is used worldwide for human figure animation and human factors analysis. Jack is a software package developed at the Center and is available from Engineering Animation Inc. Jack provides a 3D interactive environment or controlling articulated figures. It features a detailed human model and includes realistic behavioral controls, anthropometric scaling, task animation and evaluation systems, view analysis, automatic reach and grasp, collision detection and avoidance, and many other useful tools for a wide range of applications. Their Parameterized Action Representations will be a unifying factor across several of the workpackages in MAGICSTER. Principal proposer is Norm Badler.

Norman Badler is Professor of Computer and Information Science and Director at the Center for Human Modeling and Simulation. Active in computer graphics since 1968, his research focuses on human figure modeling, manipulation, and animation. He is the originator of the "Jack" software system (now a commercial product from Engineering Animation, Inc.). Badler received the BA degree in Creative Studies Mathematics from the University of California at Santa Barbara in 1970, the MSc in Mathematics in 1971, and the Ph.D. in Computer Science in 1975, both from the University of Toronto. His main interests lie in virtual (graphical) human models that use cognitive science principles to engage in conversation, the use of natural language to control virtual human animations and to create natural user-interfaces, and real-time 3D graphical simulation.

Some recent publications:

N. Badler (2000). Animation 2000++. *IEEE Computer Graphics and Applications* **20** (1), pp28–29.

J. Cassell, C. Pelachaud, N. Badler, M. Steedman *et al* (1994). Animated conversation: rule-based generation of facial display, gesture and spoken intonation for multiple conversational agents. *Computer Graphics* 28 (4): 413–420.

S. Chopra-Khullar and N. Badler (2000). Where to look? Automating attending behaviors of virtual human characters. *International Journal on Agents and Multiagent Systems*, to appear.

N. Badler, R. Bindiganavale, J. Allbeck, W. Schuler, L. Zhao, M. Palmer (2000). Parameterized Action Representation for virtual human agents. In J. Cassell (ed.), *Embodied Conversational Agents*. MIT Press.

N. Badler, M. Palmer, R. Bindiganavale (1999). Animation control for real-time virtual humans. *Communications of the ACM* **42**(8), pp64–73.

N. Badler, D. Chi, S. Chopra (1999). Virtual human animation based on movement observation and cognitive behavior models. *Computer Animation Conference*, Geneva, Switzerland, May 1999.

DFKI

Founded in 1988, DFKI today is one of the largest nonprofit contract research institutes in the field of innovative software technology based on Artificial Intelligence (AI) methods. DFKI is focusing on the complete cycle of innovation, from world-class basic research and technology development through leading-edge demonstrators and prototypes to product functions and commercialization. Based in Kaiserslautern and Saarbrücken, the German Research Center for Artificial Intelligence ranks among the important “Centres of Excellence” worldwide. An important element of DFKI’s mission is to move innovations as quickly as possible from the lab into the marketplace. Only by maintaining research projects at the forefront of science can DFKI have the strength to meet its technology transfer goals. DFKI’s five research departments are directed by internationally recognized research scientists: Information Management and Document Analysis (Director: Prof. A. Dengel), Intelligent Visualization and Simulation Systems (Director: Prof. H. Hagen), Deduction and Multiagent Systems (Director: Prof. J. Siekmann), Language Technology (Director: Prof. H. Uszkoreit), Intelligent User Interfaces (Director: Prof. W. Wahlster). There is an additional subgroup, the “Siemens Tele-Cooperation Center” (Siemens Telekooperationszentrum (STZ)) that is part of the Intelligent User Interfaces Department.

At DFKI, all work is organized in the form of clearly focused research or development projects with planned deliverables, various milestones, and a duration from several months up to three years. DFKI benefits from interaction with the faculty of the Universities of Saarbrücken and Kaiserslautern and in turn provides opportunities for research and PhD thesis supervision to students from these universities, which have an outstanding reputation in Computer Science. DFKI has also successfully conducted a number of projects for the European commission under the programmes ESPRIT, TELEMATICS, and Networks of Excellence. Furthermore, it has been playing a major role in the VERBMOBIL project which is considered to be one of the largest projects on machine translation world-wide.

The Intelligent User Interfaces group has been engaged in the development of animated interface agents for more than 5 years now. Such agents go far beyond conventional interfaces since they promote the emerging paradigm shift from direct manipulation to highly personalized interfaces. The group has been conducted a number of industrial and academic projects on lifelike agents including the BMBF-funded project PPP (Personalized Plan-Based Presenter). With the PPP Persona, the group developed one of the first virtual presentation agents world-wide. This system is based on DFKI’s award-winning presentation planning technology (IT Prize 1995). Another relevant project is the BMBF-funded project AiA (Adaptive Communication Assistant for Effective Infobahn Access). With the Persona-Enabling Toolkit PET, this group provides a software package for the creation of lifelike characters which can easily integrated into web interfaces. Based on this toolkit, the group developed a number of innovative web applications, such as virtual shopping assistants or travel guides. The EU-funded i3 project PUPPET (The Puppet Theatre of Virtual Worlds) explores new forms of learning by the development and evaluation of novel interactive virtual environments. Today, the group is considered to be one of the leading research teams in the area of animated interface agents. Principal proposer is Elisabet André.

Elisabeth André is a principal researcher at DFKI GmbH. She has been leading various industrial and academic projects in the area of intelligent user interfaces, including the BMBF-funded

projects PPP and AiA, DFKI's part of the GIF-funded project IMAP and DFKI's part of the EU-funded project Puppet. She is the Chair of the ACL Special Interest Group on Multimedia Language Processing (SIGMEDIA). Furthermore, she is on the Editorial Board of Artificial Intelligence Communications (AI COM) and the Area Editor for Intelligent User Interfaces of the Electronic Transactions of Artificial Intelligence (ETAI).

Some recent publications:

E. André and T. Rist (1996). Coping with Temporal Constraints in Multimedia Presentation Planning. In *AAAI-96*, 1996.

E. André, J. Müller and T. Rist (1996). WIP/PPP: Knowledge-Based Methods for Fully Automated Multimedia Authoring. In *EUROMEDIA '96*, pp95–102. London, UK

E. André (1997). WIP and PPP: A Comparison of two Multimedia Presentation Systems in Terms of the Standard Reference Model. In *Computer Standards and Interfaces* **18**, pp555–564.

E. André, G. Herzog and T. Rist (1998). Generating Multimedia Presentations for RoboCup Soccer Games In *RoboCup-97: Robot Soccer World Cup I*, pp200–215. Springer.

E. André, T. Rist and J. Müller (1998). Integrating Reactive and Scripted Behaviors in a Life-Like Presentation Agent. In *Proceedings of the Second International Conference on Autonomous Agents (Agents '98)* pp261–268.

E. André, T. Rist and J. Müller (1998). WebPersona: A Life-Like Presentation Agent for the World-Wide Web. In *Knowledge-Based Systems* **11** 1, 25–36.

SICS

The Swedish Institute of Computer Science, SICS, is an independent non-profit research institute. SICS's mission is to contribute to the competitive strength of Swedish industry by conducting advanced and focused research in strategic areas of computer science, actively promoting the use of new research ideas and results in industry and in society at large. SICS works in close collaboration with industry and the national and international research communities.

SICS's research focus is on distributed and networked interactive real-time multimedia systems and applications, spanning from infrastructural issues to software methodologies to human-computer interaction. SICS actively collaborates with industrial and academic partners. The core of SICS research is financed by the state-owned Ireco Holding AB (Institute for Research and Competence Holding AB) and companies associated with FDF (Foreningen for Datateknisk Forskning). The main national sources of competitive project funding are currently the Swedish Agency for Innovation Systems (VINNOVA) and SITI (Swedish Research Institute for IT). Research contracts assigned by the FDF members and other organizations, Swedish and foreign, speed up dissemination and uptake of research results in industry and public sector. SICS also actively participates in collaborative R&D programs, both national and international, such as ESPRIT and ACTS funded by the European Commission, and Real World Computing funded by the Japanese government. SICS has a well-developed collaboration pattern with high-tech SMEs in Sweden, carrying out joint projects and acting as an external R&D resource for selected SMEs. SICS also has a proven record of disseminating and promoting industrial deployment of its research findings, including establishing of spin-off companies, as well as licensing of its software and patents.

The broad spectrum of SICS capabilities includes Computer and Network Architectures, Formal Design Methods, Intelligent Systems Laboratory, Neural Networks and Real-time Computing, Human Computer Interaction and Language Engineering, and Interactive Collaborative Environments. In MAGICSTER, it will be mainly researchers from the latter groups that will participate. Principal proposers are Kristina Höök and Per Persson.

Kristina Höök (PhD 1996., PhLic 1991, MSc 1987) is a researcher at SICS. She is group leader of the PERSONA group within the Human Computer Interaction and Language Engineering HUMLE lab at SICS. The PERSONA group has about seven members. The research focus around issues of the design and evaluation of computer applications that rely on social navigation, such as chat and MUD systems, recommender systems, trails and readware. Her research interests lie within the field of Human-Computer Interaction (HCI) which constitutes a subfield of computer science. HCI also involves knowledge of psychology (cognitive psychology, social psychology, cultural psychology, etc.), and benefits from work in the design sciences (engineering, industrial design, graphics design, etc.). Her work has therefore always been interdisciplinary. Within HCI she has worked with how to design, implement and evaluate Intelligent interfaces. Intelligent interfaces are such that they try to model some aspect of the user and adapt to his/her perceived needs, knowledge or other capabilities. She has constructed and evaluated two main applications: a route guidance system (reported in her licentiate thesis) and an adaptive on-line manual (reported in her dissertation). Now working with the problem of information overflow, but more directed at the problem of navigation. Also a continued interest in individual differences, studying among other things, the problem of disabled users.

Per Persson is currently concluding his PhD in film studies at the Department of Cinema Studies, Stockholm University. The dissertation will deal with processes of reception and comprehension of editing and cinematic characters. His general interest lies in the interface between psychology and cinema studies and the way moving images in all forms are appropriated, cognitively and emotionally. For the moment he is a visiting scholar at the Department of Communicative Arts, University of Wisconsin, Madison. In addition to this, he has been working for SICS and the PERSONA project over the last eighteen months, in which he has participated with comparative texts about spatiality and navigability in cinematic vs. digital interfaces. During 1998 he was also the responsible for the design of Agneta and Frida, a system incorporating emotion and narratives into web browsing through two characters. The Agneta and Frida system will be one of the application domains used in MAGICSTER for prototypes and evaluation.

Some recent publications:

Kristina Höök (1997). Evaluating the Utility and Usability of an Adaptive Hypermedia System. In *Journal of Knowledge Based Systems*.

Mattias Forsberg, Kristina Höök and Martin Svensson (1998). Design Principles of Social Navigation. In *4th Workshop on User Interfaces for All*. Stockholm, Sweden: ERCIM.

Kristina Höök (1998). Steps to Take Before IUI Becomes Real. In *Interacting with Computers*.

Kristina Höök and Martin Svensson (1999). Evaluating Adaptive Navigation Support. In *Proceedings of the International Conference on Intelligent User Interfaces*, Los Angeles, CA: ACM.

BARI

The University of Bari is one of the largest Universities in Southern Italy. Its Department of Informatics was set up 25 years ago, with the aim of organising research and teaching in the Informatics domain. It includes several Research Groups, whose interests range from Theoretical Computer Science to Software Engineering, Artificial Intelligence, Pattern Recognition and Human-Machine Interaction.

In the Intelligent Interfaces Research Group research is focused on all aspects of Human-Machine Interaction in which some form of “intelligence” is demonstrated: in particular, User Modeling and User-Adapted Interaction, methods of Natural Language Generation, Dialog Management. More recently, the interest of this group has turned towards believable agents and, in particular, towards all aspects of interaction that involve “personality” and “emotional” factors.

Earlier project work of the group include EPIAIM, a European project aimed at prototyping a decision-support system in the area of epidemiology; OPADE, a European project aimed at prototyping a decision-support system in the area of drug prescription; GOLEM, a system to simulate interaction between agents with different “cooperation personalities”; GENET, a dynamic generator of hypermedia from a discourse plan; XDM-Gen, an animated, knowledge-based and “personality-rich” agent that describes how to use an application.

Work done by the BARI in these projects regarded not only the study of methodological aspects, but also the prototyping of working modules, and BARI will continue to play an important evaluation role in MAGICSTER, some of it in collaboration with the Experimental Psychologist Dianne Berry. Principal proposers: Fiorella de Rosis

Fiorella de Rosis is associate Professor of Medical Informatics and Intelligent Interfaces in the 5-year curriculum of Informatics at the University of Bari. Her research interests are focused on the following areas: formal specification and evaluation of user-adapted interfaces; generation of user-adapted NL and multimedia messages; believable agents. She coordinates the Intelligent Interfaces Research Group of the Department of Informatics, where several Projects in this area are being developed. In the last 10 years, she coordinated several National Projects on Medical Informatics, participated in two European Community Projects in Health Telematics, (EPIAIM and OPADE) and coordinated, for several years, an ERASMUS PIC on Medical Technology. She is the author of more than 100 papers, most of which are articles in refereed journals or book chapters.

Dianne Berry is the Director of the Department of Psychology, University of Reading. Her research interests are focused on communication of health related information: she currently investigates what information people should be given about health topics, and how this information should be presented to enable people to assess risk and make effective decisions. Another research topic looks at the effectiveness of different media in multimedia learning systems. She is investigating, at present, whether different media (e.g. text, speech, pictures, video) are differentially effective for conveying different types of information, whether combinations of different media are more effective than the use of single media, and what functions particular media, and their combinations, play in supporting learning processes. Research in some aspects of the two mentioned area was made in cooperation with the Department of Informatics, University of Bari.

Some recent publications:

F. de Rosis, F. Grasso and D. Berry (1999). Refining instructional text generation after evaluation. In *Artificial Intelligence in Medicine* **3**.

F. de Rosis and F. Grasso (1999). Affective Text Generation. In *International Workshop on 'Affect in Interaction'*, Siena.

C. Castelfranchi, F. de Rosis and F. Grasso (1999). Deception and suspicion in medical interactions, towards a simulation of believable dialogues. In D Levy and Y Wilks (Eds). Kluwer.

I. Poggi, C. Pelachaud and F. de Rosis (2000). Eye communication in a conversational 3D synthetic agent. In *AI Communications*, to appear.

B. De Carolis, F. de Rosis, C. Andreoli, M. L. De Cicco and V. Cavallo (1998). The dynamic generation of hypertext presentations of medical guidelines. In *The New Review of Hypermedia and Multimedia* **4**.

C. Castelfranchi, R. Falcone, F. de Rosis and S. Pizzutilo (1998). Personality traits and Social Attitudes in Multiagent Cooperation. In *Applied Artificial Intelligence* **12**.

AvatarMe

AvatarMe Ltd started in 1998 as a spin-out from 3D Scanners. Since 3D Scanners was formed in 1991, it has become the world's leading supplier of laser-stripe scanning systems to the automotive industry. 3D Scanners won the European Grand IT Prize in November 1996, and with it a prize of 200,000 EURO. Many of the prize-winning 3D Scanners development team have moved to AvatarMe Ltd.

AvatarMe was spun out of 3D Scanners to focus on market opportunities for photorealistic avatars of people. AvatarMe develops and markets pioneering solutions for populating the web with photo-realistic 3D avatars. Its flagship product—the AvatarBooth—is a world first in 3D human capture technology, automatically generating 3D avatars via breakthrough software concepts. The largest installation is a string of three AvatarBooths in the London Millennium Dome which will populate the web with around 300,000 avatars of people during the year 2000. The technology stems from an extremely successful i3 EU project: 'Populate' in which commercial orders larger than the project budget were received before the end of the project. AvatarMe envisages a future with more than 20,000 AvatarBooths around the world and is developing a low-cost booth that will be the key to this vision.

AvatarMe is a virtual company; a central team of 5 people manage a multi-national network of over 50 people and 12 companies who together developed the AvatarBooth at breakneck pace through 1998 and 1999. It is a privately funded company that had sales in excess of 1.2 million Euro in 1999. AvatarMe has had close involvement in all stages of the elaboration of the VRML h-anim standard (AvatarMe's CEO was a founder and is a co-chair of the h-anim working group).

AvatarMe subcontracts the majority of its R&D to other companies in Europe including ADA Bulgaria. AvatarMe closely manages this subcontract work and retains all rights to the results. ADA was incorporated at the start of 1997 and performs contract Research and Development for 3D Scanners and AvatarMe. The company has specialized in the fields of avatars, 3D data capture, vision, image processing and automatic human model generation. It has been deeply involved in the development of world-leading software products both as an originator and an implementer of innovative algorithms and approaches, thus accruing wide experience in creating robust, easy to use and productive packages. The company has also been involved as a subcontractor in the EU i3-Net POPULATE project (Jul 97-Sep 99), playing a key role in the research and in the software module integration and interfacing and in the EU 'Home-VR' project where it was involved in animating photorealistic avatars in virtual worlds. The company has 18 employees.

AvatarMe Ltd intend to subcontract some of its software research to ADA Bulgaria. The subcontract will be on a 100% cost basis. This subcontract is very good value for AvatarMe and the European Union because:

- ADA Bulgaria employs extremely high quality, experienced, skilled and productive researchers
- The man-month costs at ADA Bulgaria are about 50
- It fulfills the EU RTD objective of encouraging East European countries to develop their expertise

Key personnel of AvatarMe and ADA Bulgaria are Stephen Crampton, Phil Hand, Juriy Radkov and Nadia Dalaktchieva.

Stephen Crampton - Chief Executive, AvatarMe is the founder of 3D Scanners and AvatarMe Ltd. He is a Mechanical Engineer (Imperial College, London) and has an MBA from INSEAD. He has worked at Vickers Shipbuilding, a builder of nuclear submarines, and at RDP (a robotics company) where he led project teams and performed Marketing and Financial Management roles. He has extensive experience managing European Research Projects. In 1986, he was Project Manager of EP 1217 project, which built CIM components in a human-centred way, drawing heavily on human factors input from industrial psychologists and social scientists. Stephen has built 3D Scanners and AvatarMe from scratch into profitable companies with excellent R&D teams renowned for their research, original thinking and innovation.

Phil Hand - Chief Architect, AvatarMe joined 3D Scanners in 1993 and moved across to AvatarMe in 1999. He studied at Imperial College, London. Phil was the lead software engineer on 3D Scanners' ModelMaker research team that won the European IT Grand Prize in Nov 1996. As Chief Architect at AvatarMe he has laid down the ground-breaking structures with which photorealistic avatars of people are generated and animated.

Juriy Radkov - Managing Director, ADA Bulgaria is a technical manager and researcher with 10 years' experience in software development. He received education at both the University of Sofia and the Royal Holloway And Bedford New College, University of London. In 1994, he joined 3D Scanners as a software engineer in its R&D team, and so has a thorough knowledge of 3D Scanners and AvatarMe's recent involvement in EU projects. He founded ADA in Bulgaria in March 1997, since when he has managed all ADA's subcontracting work with AvatarMe and 3D Scanners, including for the Populate project. Juriy and his team of developers, consequently, are thoroughly grounded in 3D Scanners' work to date on avatars.

Nadia Dalaktchieva - Software Developer with 6 years' research experience. She has an MSc in Bioengineering (Technical University of Sofia). Nadia joined the Laboratory of Biomedical Engineering at the Bulgarian Academy of Sciences, where her work focused on feature analysis and speech recognition. In April 1997, she joined ADA, since when her research work has focused on 3D data and image processing, and virtual humans modelling. She worked extensively on the research subcontracted to ADA from 3D Scanners under both the Populate and HomeVR projects. In particular, she developed the seamless avatar model.