Collaboration in the Semantic Grid: a Basis for e-Learning

Michelle Bachler[1], Simon Buckingham Shum[1], Jessica Chen-Burger[2], Jeff Dalton[2], David De Roure[3], Marc Eisenstadt[1], Jiri Komzak[1], Danius Michaelides[3], Kevin Page[3], Stephen Potter[2], Nigel Shadbolt[3], and Austin Tate[2]

[1] KMI, The Open University, UK
[2] AIAI, University of Edinburgh, UK
[3] ECS, University of Southampton, UK

Abstract. The CoAKTinG project aims to advance the state of the art in collaborative mediated spaces for the Semantic Grid. This paper presents an overview of the hypertext and knowledge based tools which have been deployed to augment existing collaborative environments, and the ontology which is used to exchange structure, promote enhanced process tracking, and aid navigation of resources before, after, and while a collaboration occurs. While the primary focus of the project has been supporting e-science, this paper also explores the similarities and application of CoAKTinG technologies as part of a human centred design approach to e-Learning.

1 Introduction

The CoAKTinG project[1] aims to advance the state of the art in collaborative mediated spaces for distributed e-Science through the novel application of advanced knowledge technologies. It comprises four tools: instant messaging and presence notification (BuddySpace), graphical meeting and group memory capture (Compendium), intelligent 'to- do' lists (Process Panels) and meeting capture and replay. These are integrated into existing collaborative environments (such as the Access Grid [2]), and through use of a shared ontology to exchange structure, promotes enhanced process tracking and navigation of resources before, after, and while a meeting occurs.

Section 2 gives a context to the work in the Semantic Grid and e-Learning, Section 3 provides an overview of the tools, Section 4 describes the ontology that interconnects them, Section 5 gives a glimpse of current work using the tools, and Section 6 discusses their application to the Learning Grid.

2 The Semantic Grid, Collaboration, and Learning

While the Grid is often thought of in terms of providing a distributed system of high-performance compute resources, this is only one aspect required when supporting successful e-Science. In addition, the Grid must provide structured access to the wealth of data produced and held within it, and an environment within which the collaborative processes of investigation can occur - be this meetings between researchers, or shared access to experiments.

It is the latter facet of collaboration, in particular, which CoAKTinG addresses. Collaboration as an activity can be seen as a resource in itself, which with the right tools can be used to enhance and aid future collaboration and work. Each of the CoAKTinG tools can be thought of as extracting *structure* from the collaboration process. The full record of any collaboration (e.g. a video recording of a meeting) is *rich in detail*, but to be useful we must extract resources which are *rich in structure*.

In essence, this is a process of creating structured knowledge from information, and we must be able to share and re-use the knowledge amongst tools and agents in the Grid - the Semantic Grid.

The issues involved in collaboration are not unique to science; we believe that the techniques and tools employed by CoAKTinG have useful applications in the experimental-based collaborative learning paradigm adopted by the ELeGI project [3].

3 CoAKTinG Tools

3.1 BuddySpace

BuddySpace is an Instant Message client (based on the Jabber protocol) with features that enhance presence awareness. Specifically, it introduces the graphical visualisation of people and the presence on a image or map, as can bee seen in the figure. This allows for multiple views of collaborative workgroups and the immediacy or "at a glance" nature gives users a snapshot of a virtual organisation. In a meeting, the instant message capabilities of BuddySpace naturally provide a "back-channel" to the meeting, for example, conveying URLs of documents discussed or as a non-disrupting communication. For distributed meetings, such Access Grid meetings, the presence of individuals gives an extra indication of co-location (especially if the videoconferencing technology is failing). The back-channel can also be used for meeting control tasks, such as queueing of speakers and voting on issues.

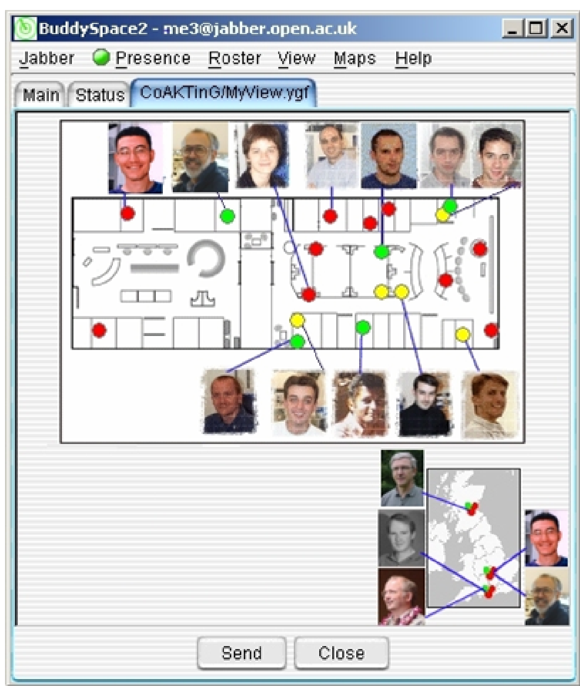

Fig. 1. BuddySpace showing a virtual organisation and presence indicators

For meeting capture purposes, logs of the channel conversations are made. Individual messages are timestamped and possibly examined to see if they control meeting specific messages.

3.2 Compendium

Compendium, first developed in 1993 as an approach to aid cross-functional business process redesign (BPR) teams, has been applied in several dozen projects in both industry and academic settings [4]. Its origins lie in the problem of creating shared understanding between the team members, typical of those attending teams working over weeks or months to design business processes: keeping track of the plethora of ideas, issues, and conceptual interrelationships without needing to sift through piles of easel sheets, surfacing and tracking design rationale, and staying on track and "bought-in" to the project's overall structure and goals [5]. The key feature of the early approach was the combination of an Issue-Based Information System (IBIS) concept-mapping tool [6], which supported informal and exploratory conversation and facilitation, with a structured modelling approach [7]. This allowed teams to move along the spectra of *formal to informal representation*, and *prescribed to spontaneous approaches*, as their needs dictated. It also let them incrementally formalise data [8] over the life of the project. As the approach was tested and refined over the course of several years, additional modelling methods were added, plus tools to transform Compendium's hypertext models into established organisational document forms, and vice-versa [9].

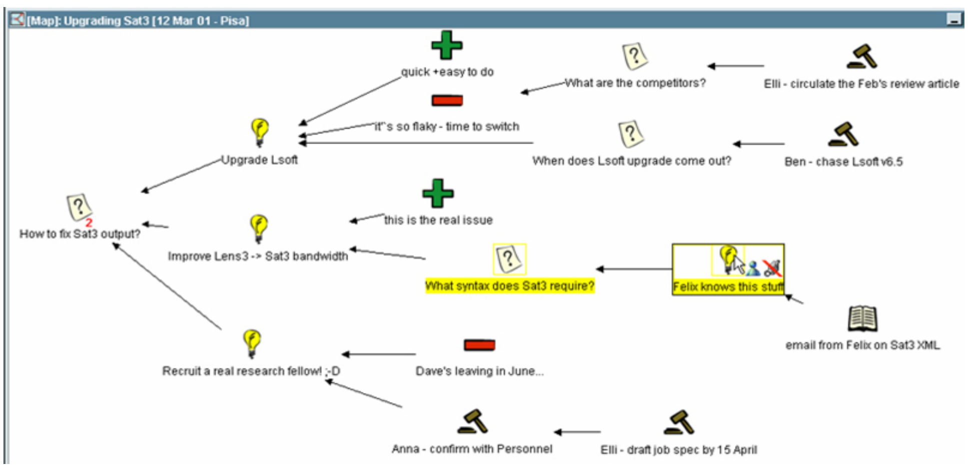

Fig. 2. A Compendium map showing various node types and links

In our experience, Compendium introduces a distinctive element to the design space of knowledge technologies, namely, making meetings into true events for group knowledge creation which leave a *trace* - a structured, collectively owned, searchable group memory that is generated in real time as a product of a meeting. Effective, on-the-fly construction of knowledge resources does not come "for free" - the lower the effort invested at the capture stage (e.g. simply video recording all meetings, or taking conventional minutes), the more work is required for collective reuse and computational support. Naturally, we want quality knowledge resources for minimal effort, and while smart analysis technologies will continue to push the boundaries, there are pragmatic factors to consider: what is possible *now*? Compendium tackles the capture bottleneck that any knowledge construction effort must confront, by investing effort in real time quality capture by a facilitator, mediated and validated by those at the meeting.

3.3 I-X Process Panels

I-X is a suite of tools[10] whose function is to aid in processes which create or modify one or more "product" (such as a document, a physical entity or even some desired changes in the world state). The main interface is the I-X Process Panel (I-P2) which, in its simplest form, acts like an intelligent "to do" list. The panel shows users their current issues and activities, on which Standard Operating Procedures can be applied to manage complex and long-running processes. I-X also has a collaborative element to it, in that issues and activities can be passed between different process panels to enact a workflow across an organisation. Web services can be called to automatically enact steps of the processes involved. Progress and completion reporting between panels and external services is possible. The underlying model on which I-X is based is the <I-N-C-A> Constraints Model[11]. In a meeting scenario, actions raised in a meeting have a direct mapping to <I-N-C-A> activities. Actions created in a meeting specific I-X panel are passed onto the relevant user panel's for individuals, which, on completion report back.

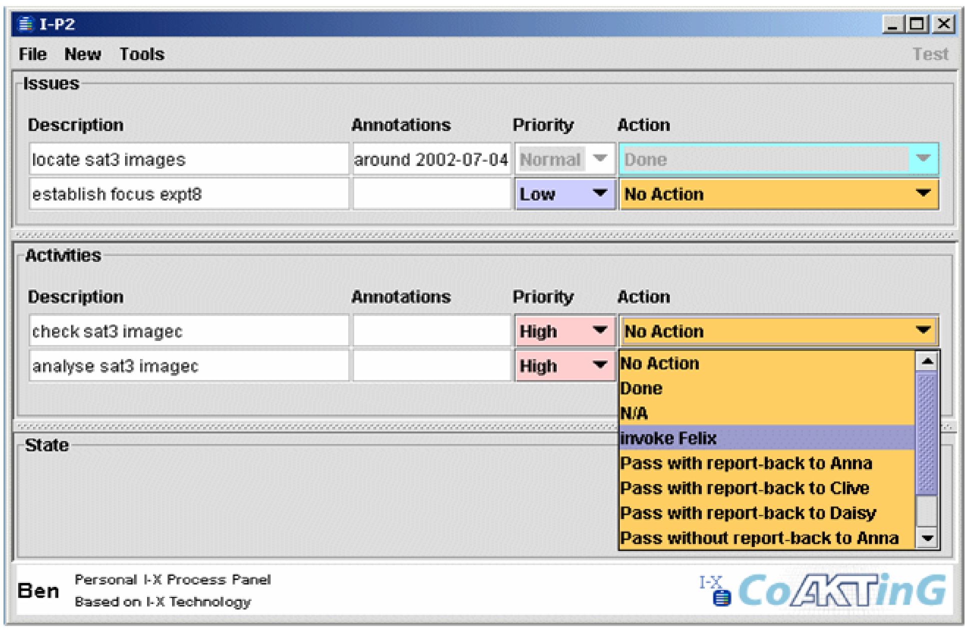

Fig. 3. A I-X Process Panel showing pending issues and activities

4 Meeting Replay

Once a meeting has taken place it can be useful to revisit the ideas and topics discussed. Traditionally, formal minutes are taken to record the salient points, but often these are too brief to be more than a simple aide memoire; in the typical CoAKTinG scenario (such as an Access Grid node) full audio and video logs are available, but conversely these are too verbose to be of practical use. We require the ability to select high-level points of reference from the meeting, then "zoom in" to view detailed records. e.g. a user sees from Compendium notes that a decision was made, but to understand the subtle reasoning behind that outcome wishes to view the video of discussion between participants. Each meeting is described using

RDF conforming to the OWL meeting ontology; this represents resources such as: the meeting time, location, attendees, audio/video recordings, any presentations given (and associated web viewable versions), and argumentation annotation from Compendium. The Event / has-sub-event structure held within the RDF is mapped onto a more conventional time-line, which is automatically published using HTML and Javascript on a web site (figure 4). The user can navigate the meeting using the video timeline, or jump to a different point in the meeting by selecting a particular event, such as a slide being presented, or a Compendium node being created. By using the shared AKT reference ontology, we can also link to further information about resources held in other knowledge bases, e.g. when a person is referenced we link to information about them in the populated AKT triple store. We populate the timeline with any temporally annotated information about the meeting that would aid the user in navigation.

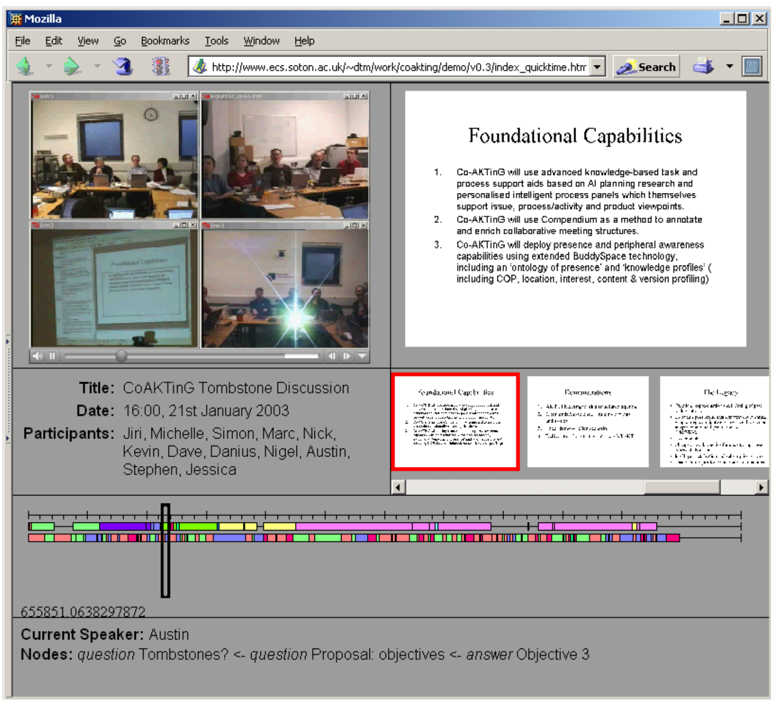

Fig. 4. The meeting replay tool

In CoAKTinG we have experimented with:

- Agenda item
- Slide exhibits
- Compendium node
- Speaker identification
- I-X activity(action item) creation
- BuddySpace chat

By providing all available information we hope to cater for the many activities and contexts of the user, in a seamful[12] manner.

We can categorise the information presented in the entire meeting replay in terms of the dimensions "structured" and "detailed", as shown in figure 5. Video, for example, is high in detail, in that it captures the entire audio and visuals of the meeting. Structurally, it is relatively low, since although there is implicit structure (image frames and audio samples) these do not directly contribute to navigating the structure of the meeting. Video processing could applied to segment the video into scenes but structurally this would not provide much more than Speaker Identification. The Agenda, conversely, is high in meeting structure, but relatively low in the details. Compendium captures a moderate level of detail in a highly structured representation.

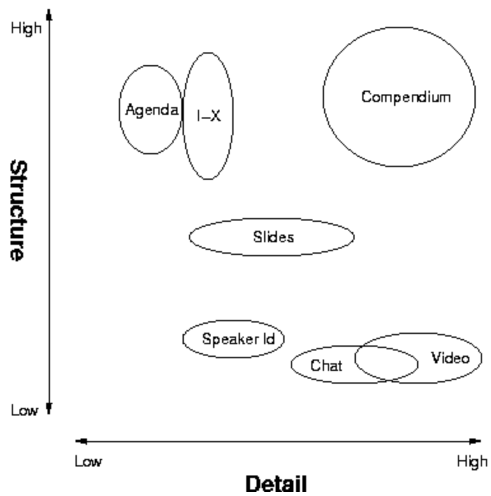

Fig. 5. Meeting Detail and Structure of recorded sources

5 Ontology

The Advanced Knowledge Technologies (AKT) project, with which CoAKTinG is affiliated, has developed a reference ontology [13] to describe the domain of computer science research in the UK, exemplified by the CS AKTive Space semantic web application. Within this domain, its vocabulary is able to express relationships between entities such as individuals, projects, activities, locations, documents and publications. For purposes of capturing meeting specific information, the reference ontology is already suitable for encapsulating:

- the meeting event itself
- meeting attendees
- projects which are the subject matter of the meeting
- documents associated with the meeting, including multimedia

For activities such as meetings, which we wish to index and navigate temporally, the way in which the ontology represents time is of particular relevance. The reference ontology contains the notion of an *Event*, which is a *Temporal-Thing* that can define a duration, start and end times, a location and *agents* involved in the event. More importantly, each Event can express a *has-sub-event* relationship with any number of other Events, and it is with this property that we build up our temporal meeting structure. Within the ontology there are also many Event sub-classes, such as *Giving-a-Talk*, *Sending-an-Email*, *Book-Publishing*, and *Meeting-Taking-Place*.

While the reference ontology provides a foundation for describing meeting related resources, the CoAKTinG meeting ontology (figure 6) extends the OWL version of AKT reference ontology to better encompass concepts needed to represent collaborative spaces and activities, including:

- time properties sufficient for multimedia synchronisation
- distributed gatherings to represent meetings which simultaneously take place in several spaces, both real and virtual
- exhibition of information bearing objects; e.g. showing a slide as part of a presentation
- compound information objects; e.g. to describe a presentation consisting of several multimedia documents
- rendering of information objects; e.g. JPEG image of a slide
- transcription of events; e.g. a video recording of a presentation, minutes of a meeting
- annotation of events; e.g. making a verbal comment, creating a Compendium node

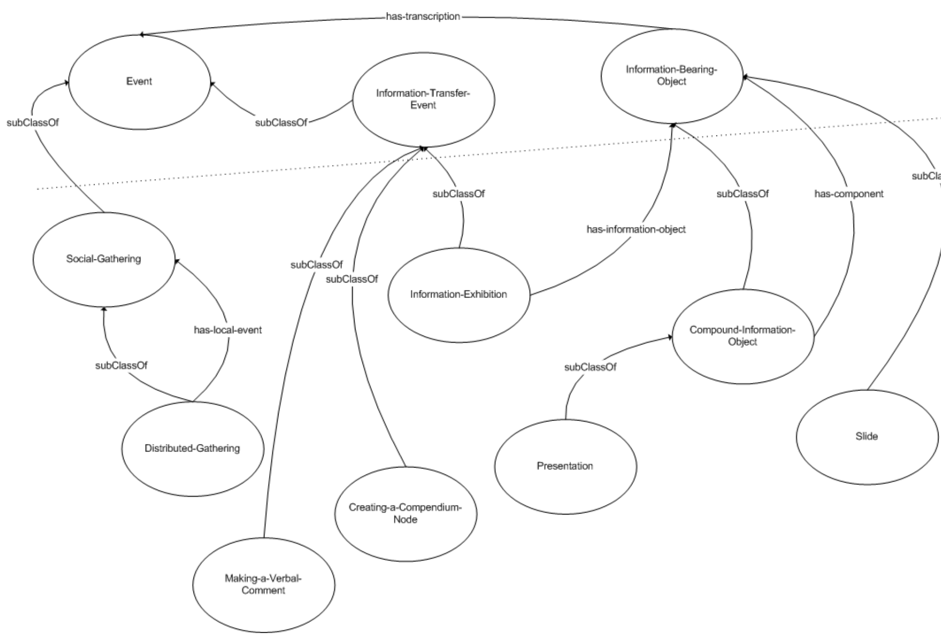

Fig. 6. A simplified representation of the meeting ontology

When a meeting takes place we "mark up" the event with metadata - details such as those listed above - to build a structured description of the activities that occur. Through use of an ontology shared and understood by several different tools, we can lower the workload needed to provide usable and useful structure.

6 Case Studies

6.1 CombeChem

The CombeChem project aims to enhance structure property correlation and prediction by increasing the amount of knowledge about materials via synthesis and analysis of large compound libraries. Automation of the measurement and analysis is required in order to do this efficiently and reliably while ensuring that wide dissemination of the information occurs together with all the necessary associated background (raw) data that is needed to specify the provenance of the material. The project aims for a complete end-to-end connection between the laboratory bench and the intellectual chemical knowledge that is published as a result of the investigation; this necessitates that all steps in the process are enhanced by a suitable digital environment. CombeChem has achieved many parts of this ambitious programme, e.g. the smart laboratory (smarttea.org), grid-enabled instrumentation, data tracking for analysis, methodology for publication@source, process and role based security and high throughput computation.

The CoAKTinG tools provide support for the e-Science process in CombeChem and they also enable the digitisation of 'missing links' in the processing chain which form part of the typical collaborative scientific processes that we are attempting to enhance using the grid infrastructure: support of the experimental process, tracking and awareness of people and machine states, capturing of the discussions about data as well as the traditional metadata, and enriched meta-data regarding these components to support interlinking.

The BuddySpace systems can be adapted to show and track the interactions between the staff and equipment using the National Crystallographic Service (NCS), providing information to their users about the state of the service. Compendium provides the harness to ensure more adequate capture of the discussions in analysis, while Process Panels provide the means to initiate and track key tasks and issues. Additionally the ideas from CoAKTinG provide different techniques to achieve the necessary multi-user interaction in real time over the network and give CombeChem the opportunity to implement the "video interaction" collaboration part of CombeChem using event based ontologies to annotate real time streaming media and content.

These various components are valuable complements to CombeChem individually but jointly are even more powerful. For example, Process Panels can exploit the presence information derived from BuddySpace with respect to instrument status and operator availability to offer more informed task delegation options. This completes the chain of digital support and capture, maximising the potential for re-use of the digital information in support of the scientific process.

The following figure illustrates one particular aspect of the deep integration – the application of the Process Panel tool to the laboratory, building on the process capture work of CombeChem's Smart Tea team.

Figure 7 shows a screen capture of an I-X Process Panel and its Map Tool resulting from our initial experiment. The Map Tool depicts a real Chemistry lab where both fixed and mobile entities are represented. The positions of mobile entities such as movable equipment and technicians are updated automatically through the (World) State sub-panel. By sharing information with BuddySpace, (dynamic) properties of devices are also described in the same panel. At this particular point in time, it shows Technician-2 is in front of the Rotary Evaporator and about to carry out the sub-process "Remove solvent from the-mixture using Vacuo results in Compound", having completed the previous steps in this process. In our investigation, the process decomposition facility of the I-X Activity sub-panel supports views of different levels of abstraction that fits nicely with different chemists' (and labs') practice. Activities, issues, annotations and constraints may be recorded directly or via Compendium where in-depth discussion has taken place. Static and dynamic process editing provide great flexibility as processes are modifiable at run-time in response to unexpected changes. The ability to store, retrieve and refine process models is important in the Chemistry domain where existing processes are constantly reviewed and modified to discover or synthesise new chemical compounds.

This facility alone makes I-X a valuable back-end component for integration with the existing CombeChem Grid.

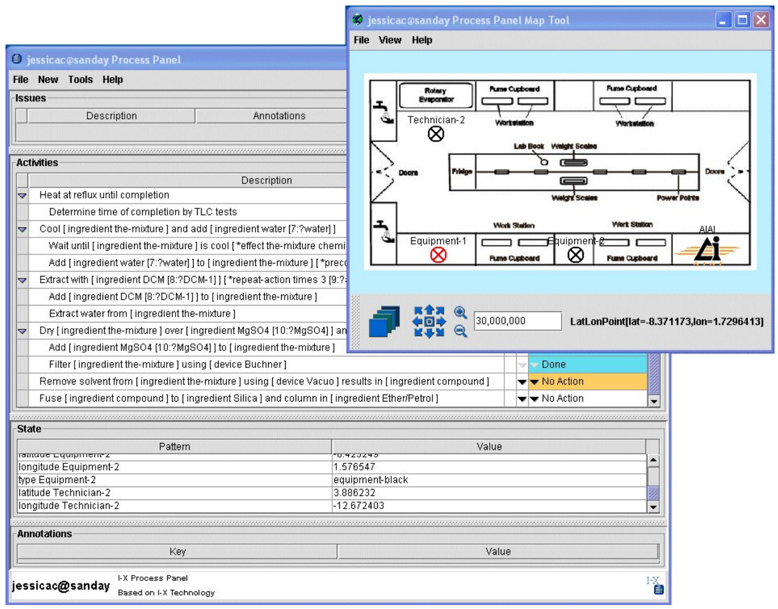

Fig. 7. I-X Process Panel configured for e-Chemists

6.2 Scientific Exploration on Mars

As part of long-term research into manned Mars missions, NASA's Work Systems Design and Evaluation group conducts annual field trials of its agent-based software and robots at the Mars Society's Desert Research Station (MDRS) in Utah, USA. As a part of the most recent trial, several CoAKTinG tools have been used to support the collaboration that occurs between the astronauts on "Mars" and the distributed groups of support scientists on Earth (know as the Remote Science Team, RST, and in this particular case specialists in geology).

The role of the RST is to analyse the data collected by the astronauts during their EVAs on the planet surface, and the subsequent debrief at the Mars base (which is videoed to provide a detail-rich recording). Throughout the EVA semantically annotated data is collected using the NASA agent robots. Communication delays between Earth and Mars mean that the usual means of collaboration of at a distance, such as real-time conversations and the sharing of computer screens, are impractical. This is further complicated by the international composition of the RST, who will be collaborating across many time zones.

During their debrief, the astronauts use Compendium as a dialogue mapping tool to capture the structure of the meeting. This is sent back to Earth, along with the video recording, where the CoAKTinG ontology is used as a mediator to produce a Meeting Replay.

This Replay is then viewed by the distributed members of the RST, in conjunction with the Compendium map of the debrief. When the RST meet virtually, any one member can take navigational control of the Replay so as to highlight relevant sections to the other RST members. The RST meeting itself is also captured using

Compendium, and the map is sent back to Mars with the RST analysis - this is used to plan for the next EVA. Throughout the mission, and especially during their meetings, the virtual community of the RST is supported by BuddySpace.

7 Applicability to e-Learning

It is apparent that there are many similarities in supporting the collaborations involved in e-Science and e-Learning; indeed this is one of the reasons why the (Semantic) Grid is a suitable approach to the human centred design of e-Learning. It is a short step from the remote experiments and collaboration of CombeChem and Mars exploration to virtual teaching laboratories and experimentation; the interactions, conversations, and enhanced presence which are key to Learning Grids and Virtual Communities.

In closing, we summarise where the CoAKTinG tools (and if not specific tools, the concepts underlying them) can be transposed into the Learning Grid:

- BuddySpace, with its notions of enhanced presence and communication, can be used to create a Virtual Community consisting the individual students and teachers.
- Compendium can be used to capture collective thinking within a learning group who are physically distributed, and used to plan, structure, and access, other learning resources.
- I-X Process Panels can be used to plan and structure learning tasks, goals, and experiments, and provides a mechanism for tracking issues and tasks when part of a collaboration.
- The use of a shared semantic ontology amongst the tools provides a sum greater than the parts. Structured metadata from the various tools can be combined with new material to create further services such as the Replay Tool which can be used to review results from collaborative experiments and tasks.

8 Conclusions

This paper has introduced the tools that have been developed by the CoAKTinG project and identified how they are typically used in meetings, and in support of collaborative science in the Semantic Grid. It has also shown how they are being being explored in scenarios such as CombeChem and future Mars exploration, and how this experience can be applied to the construction of a Learning Grid.

Acknowledgements

This work is funded by the UK Engineering and Physical Sciences Research Council under grant number GR/R85143/01, in association with the Advanced Knowledge Technologies Interdisciplinary Research Collaboration. The AKT IRC research partners and sponsors are authorised to reproduce and distribute reprints and online copies for their purposes notwithstanding any copyright annotation hereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of other parties. We are grateful to members of the CombeChem team for their assistance with the case study: Jeremy Frey, Gareth Hughes, Hugo Mills, monica schraefel and Graham Smith. CombeChem is also funded by the EPSRC under grant number GR/R67729/01. ELeGI funding.

References

1. a: The CoAKTinG project. http://www.aktors.org/coakting/ (1)
2. a: The Access Grid. http://www.accessgrid.org/ (1)
3. Gaeta, M., Ritrovato, P., Salerno, S.: ELeGI: The European Learning Grid Infrastructure. In: 3rd International LeGE-WG Workshop: GRID Infrastructure to Support Future Technology Enhanced Learning, Berlin, Germany (2003)
4. Conklin, J., Selvin, A., Shum, S.B., Sierhuis, M.: Facilitated Hypertext for Collective Sensemaking: 15 Years on from gIBIS. In: Proceedings The Twelfth ACM Conference on Hypertext and Hypermedia (Hypertext '01). (2001) 123–124
5. Selvin, A., Sierhuis, M.: Towards a Framework for Collaborative Modelling and Simulation. In: Workshop on Strategies for Collaborative Modelling and Simulation, CSCW'96: ACM Conference on Computer-Supported Cooperative Work, Boston, MA, USA (1996)
6. Conklin, J., Yakemovic, K.C.B.: A Process-Oriented Approach to Design Rationale. Human-Computer Interaction **6** (1991) 357–391
7. Selvin, A.: Supporting Collaborative Analysis and Design with Hypertext Functionality. Journal of Digital Information **1** (1999)
8. Shipman, F., McCall, R.: Supporting Knowledge-Base Evolution with Incremental Formalization. In: Proc. ACM CHI'94: Human Factors in Computing Systems, Boston, Mass., ACM Press: New York (1994) 285–291
9. Selvin, A., Shum, S.B.: Repurposing Requirements: Improving Collaborative Sensemaking over the Lifecycle. In: Profess'99: International Conference on Product Focused Software Process Improvement, Oulu, Finland. (1999) 539–559
10. Tate, A., Dalton, J., J. Stader, J.: I-P2 - Intelligent Process Panels to Support Coalition Operations. In: Proceedings of the Second International Conference on Knowledge Systems for Coalition Operations (KSCO-2002), Toulouse, France (2002)
11. Tate, A.: <I-N-C-A>: an Ontology for Mixed-initiative Synthesis Tasks. In: Proceedings of the Workshop on Mixed-Initiative Intelligent Systems (MIIS) at the International Joint Conference on Artificial Intelligence (IJCAI-03), Acapulco, Mexico (2003)
12. Chalmers, M., MacColl, I., Bell, M.: Seamful design: Showing the seams in wearable computing. In: Proceedings of IEE Eurowearable 2003, Birmingham, UK (2003) 11–172
13. a: The AKT Reference Ontology. http://www.aktors.org/ontology/ (1)