Supporting Verification of the Comprehensive Nuclear Test Ban Treaty Through a Persistent Conversation Interface

Stephanie Guerlain[1], Donald Brown[1], John Green[1], Brian Hashemi[1], Robert Rauschenberg[1], Christopher Stacy[1], Roger Mason[2], and Jane Bohlin[2],
[1]University of Virginia, [2]Pacific-Sierra Research Corporation

Abstract

Workflow management systems have the potential to make explicit an organization's implicit need to initiate, track, and execute artifacts and communications about those artifacts. In this sense, these systems can be described as persistent conversation systems, because they are supporting the communication and coordination of distributed actors through a computer-based medium. Furthermore, such systems have the ability to archive and report on current and past activities and communications. We describe the work that we conducted on building a prototype system that has these characteristics. The system is intended to provide an efficient, effective, auditable means of communication between the various groups that make up the Comprehensive Nuclear Test Ban Treaty, thereby enabling the verification of this very important worldwide legislation.

1. Introduction

Much research in computer-supported cooperative work and persistent conversations focuses on supporting virtual meetings, virtual conversations, digital chatrooms, discussion groups, and the like (e.g., see [7], [8], [10] and [12]). However, structured, distributed discussions are sometimes an organizational requirement, for tracking and auditing purposes. Thus, an important research question to ask is whether it possible to design a system that helps structure and organize communications according to the organization's auditing and tracking needs, but does not inhibit effective communications or activities. What this implies is that the structure of the conversational system should implicitly or explicitly inform the distributed, diverse team members of the both the required activities and information bits that constitute a "good" communication. Further, since the intent of such a structured system is to be able to track progress of activities, there needs to be an effective means to browse and monitor the current and past activities.

Why would an organization want to manage and track distributed activities? In any distributed organization or system, where information, tools, samples, or other artifacts are passed around, it is necessary to be able to track the status of those artifacts. For example, customer packages that are shipped throughout the world need to be tracked in case a package is lost or needs to be recalled. As another example, blood and other biological samples need to be tracked in a hospital organization. In order to audit and track activities and to collect statistics on the organization's effectiveness (number of artifacts delivered on time, for example), it is necessary to be able to track the activities of artifacts, and the people's activities involved with working with those artifacts. Thus, computer-based systems have the potential to aid in structuring activities and communications about these activities in a persistent, auditable way. In this paper, we describe such a system, one that is designed to support international communication and tracking of radionuclide testing samples in support of the Comprehensive Nuclear Test Ban Treaty. Before describing this system in detail, we bring in some theoretical constructs on conversational and workflow management systems.

2. Related work on Conversational and Workflow Management Systems

Tatar, Foster and Bobrow [10] discuss computer-based design for conversation and draw on work that was done on conversational theory by Clark and Wilkes-Gibbs [3] and Clark and Schaefer [2]. They describe how a basic unit of conversation consists of two parts – a *presentation phase* and an *acceptance* phase – where the two together constitute a *contribution*. How people establish that they are talking about the same thing is important. This draws on work on establishing a *common ground* for communication – knowing what your listener knows, and ensuring that each party in a conversation has the necessary knowledge and context to effectively understand the conversation. Thus, a computer-based system that supports conversational activities should help participants to know what item is being talked about, and in what context (i.e., where in the sequence of activities and the

0-7695-0981-9/01 $10.00 (c) 2001 IEEE

sequence of the conversation that this particular discussion item refers to). This is particularly important for asynchronous conversations. In multi-participant, distributed conversations, it is important to know where a message came from, what item is being referenced and where and how this message relates to the sequence of activities and conversation elements that have occurred in the past.

Turoff, Hilz, Bieber, Fjermstad and Rana [12] also discuss basic discourse structures and how they manifest themselves in computer-mediated group communications. They describe general discourse structures such as temporal occurrence, comment/reply hierarchies, and key word association of comments, with people, conversations, and objects of conversation related in a discourse network. Further, they discuss how collaborative discourse structures can be *tailored* to a specific area of knowledge, which limits the available courses of action and structures the conversation to a large degree.

In their study of organizations and computer-mediated communication systems, Winogred and Flores [13] describe how an organization can be modeled as a network of recurrent conversations, where conversations are initiated by a request for information, products, etc., and secondary conversations spawn off that are concerned with conditions of satisfaction for the initial ones. Conversations can be between different organizations, between different groups within an organization or between particular individuals within those groups. These conversations have recurrent themes, and each type of conversation (request for activity, reporting on status of activity, etc.) has its own structure of completion and states of incompletion with associated time constraints.

Winograd and Flores describe the possibility of analyzing these types of conversations and making the structure within them explicit. This is in fact the case with many organizations' requirements to fill out forms when requesting certain information. The form-filling acts as a means to tell the form filler what information is required in order for the form-receiver to successfully process the request.

In order for the organization to successfully meet requests, it is important to be able to manage requests coming in, track whether requests have been fulfilled, and be able to determine the current status of a request in case of inquiry. Winograd and Flores state: "A person working within an organization is always concerned with questions such as 'What is missing?', 'What needs to be done?', and 'Where do I stand in terms of my obligations and opportunities?' In situations where many people must act together, the problem of coordination becomes a crucial one." [13, p. 158]. Many tools, such as word processors, spreadsheets, and the like, are primarily designed to support single-person activities. A coordination system is explicitly designed to support communication, coordination, and conversation among groups.

Basic building blocks of conversations include request/promise, offer/acceptance, and report/ acknowledgement. "The development of a conversation requires selection among a certain finite set of possibilities that is defined by the opening directive and the subsequent responses." [13, p. 159]. "In using a coordination system, the individual is faced with a restricted set of possibilities. It is not the same as a face-to-face conversation, a telephone call, or even an electronic message. Because the illocutionary forces and temporality are specified explicitly, it is necessary to be conscious of them and to have a mutually visible manifestation of them." [13, p. 161]. "We find computers embodying possibilities for action within a set of recurrent conversations… the hardware and software are a medium in which requests and promises are made and monitored." [13, p. 173].

Dangelmaier, Kress, and Wenski [4] describe, in a very thorough review, the coordination of distributed worktasks in a company using a computerized workflow management system. Workflow management is a sub-field of computer-supported-cooperative work (CSCW) which deals with the support of time-separated work at different working places (as opposed to real-time coordinated work which is the focus of much of the CSCW literature). A workflow management system is responsible for the management of data flow from one activity to the next activity and also coordinates the actors who are responsible for the execution of the activities. For less structured tasks, there are enhanced requirements for supporting communication and coordination of actors who are working either at different activities of the same workflow or at the same activities but at different workflow instances. A workflow management system should also support monitoring the activities of distributed actors.

Thus, the research just described points to the fact that computer-based conversational systems or a workflow management systems have the potential to make explicit an organization's implicit need to initiate, track, and execute requests. In this sense, these systems can be described as persistent conversation systems, because they are supporting the communication and coordination of distributed actors through a computer-based medium that has the ability to archive and report on current and past activities and communications. We describe next the work that we conducted on building a prototype system that has these characteristics. The system is intended to provide an efficient, effective, auditable means of communication between organizations during radionuclide sample testing, thereby enabling the Comprehensive Nuclear Test Ban Treaty verification process.

3. The Comprehensive Nuclear Test Ban Treaty

Efforts began in the late 1950's to limit nuclear weapons testing. 1963 saw the first in a series of test ban treaties, with the Limited Test Ban Treaty, which constricted airborne radioactivity releases within a country's borders. This essentially required testing to be done underground. Then, in 1971, the Threshold Test Ban Treaty limited the yield of tests. In 1994 arms-control negotiations accelerated, and in 1996 the Comprehensive Nuclear Test Ban Treaty (CTBT) was signed, with individual country ratification to follow. The CTBT is the culmination of years of nuclear testing restriction negotiations [1]. Ratification of the CTBT is thus an important step in preserving humanity and maintaining peaceful relations among nations.

In order for nations to respect the conditions of the treaty, the treaty must be verifiable. As of 8 September, 2000, the treaty has been signed by 160 countries and ratified by 62.

4. Radionuclide testing and messaging with the RAME

In order to detect nuclear tests and verify the treaty, a network of sensors must accurately detect possible violations and reliably relay information. One way to detect possible violations is by monitoring radionuclide levels in the atmosphere. This method involves analyzing air samples taken on filter paper at field stations throughout the world. Stations will perform routine analysis of samples and conduct additional testing if high levels of radiation are detected

The proposed system to accomplish radionuclide testing consists of the following entities: 80 worldwide field stations, 16 certified labs, National Data Centers (NDC's), and the International Data Center (IDC). The field stations and certified labs comprise the International Monitoring System (IMS). The Provisional Technical Secretariat (PTS) oversees the IMS, IDC, and other organizations. The Global Communications Infrastructure (GCI) is a data network that connects these entities.

Sampling starts at the field stations, where air samples are collected on filter paper for 24 hours at a time. Then the sample is allowed to decay for 24 hours to eliminate common cluttering radioactive materials. If initial analysis indicates possible recent nuclear weapons testing, then the IDC, field stations, certified labs, and NDC's must coordinate a series of follow-on activities. The labs partition the sample and send parts to multiple certified labs to conduct further analysis, and to the PTS to archive the sample. Depending on the results, even further analyses may need to be conducted.

4.1 Problems with current sample transfer

Initial field experiments clearly demonstrated the need for a means to track samples more effectively. In these field experiments, nobody knew the status of the filters after they had been mailed. Many filters never made it to the labs. The labs did not know how many filters they would be getting to analyze and when they would arrive. No communications occurred between the stations and the labs. Everything was "controlled" by the CTBT.

Need to enforce communications and communications standards. These early experiments led to the requirement that we develop a communications protocol and messaging system for tracking samples and communications about those samples. The formats, protocols, and communications methods have all been standardized and approved. However, the implementation of these standards is still in progress. Up until now, countries have been voluntarily providing data via email or over the phone, but we now need to shift to a more formal arrangement as the stations are certified into the network.

The messages need to be standardized with a common interface, message generation, and receipt platform so that all stakeholders provide the needed information to the needed members of the network in a consistent, auditable way. The message system needs the form of a Graphical User Interface at all labs and data centers, in which users can generate and receive standard messages. Message production and access should be restricted from different organizations. For example, the field stations and certified labs should have the ability to send and receive messages with the IDC while having other messages blocked. The IDC, however, as the centralized monitoring center, should have access to all messages sent from all labs and field stations, and have the capability to send messages to every node on the Global Communications Infrastructure (GCI).

Need to maintain archives. Archiving message communications is critical to verification of the treaty's regulations. Messages should be associated with particular air samples and archived in a database. The archiving will facilitate trend analysis for radionuclide levels over a certain time period. The message archive will also provide support for decisions about retesting and analysis. If a country requests additional tests of a sample, the message archive will provide the history of the sample and be a record of the course of events for testing. A country can make policy decisions based on this course of events and have this archive as documentation of the testing process.

4.2 Developing the Radionuclide Analysis Messaging Environment (RAME)

For the above reasons, it was necessary to create a system that formalizes the messaging activities. The object of this study was to formulate an explicit set of message types, and to develop a computer-based messaging environment to support standardizing and archiving these messages. The major parts of this project included designing the types of messages that should be sent based on envisioned activities and communication requirements, and developing a computer-based messaging system that enables the different organizations to follow this messaging structure. This system, called the Radionuclide Analysis Messaging Environment (RAME) consists of generic message types, a data warehouse structure, audit trail capabilities, user interfaces, and a communications infrastructure. The system is intended provide an efficient, effective, auditable means of communication between organizations during sample testing, thereby facilitating the treaty verification process.

4.3 Objectives

In summary, our goal was to improve the treaty verification process, specifically the radionuclide testing procedures, by accomplishing the following objectives:

1) Reduce message variability by creating standard messages using a common GUI interface provided at all sites, and limit the ability of lab technicians to create their own formats. The purpose of the standard format is to facilitate message archiving and the receiver's understanding of the message, and ensure inclusion of all pertinent information into the message.
2) Reduce the time required to send messages, using standardized message formats. For example, the lab workers can select a message type to send. Then a message template will appear with a standard format including a set order of information.
3) Improve the security and validity of information concerning samples by restricting message access and generation from some organization types and creating the common interface.
4) Archive message traffic for future access and analysis. The standard messages and interface will create information of a consistent format that can be stored in existing Oracle databases.

4.4 Standard message types

Based on an analysis of the required activities by the different member organizations, we generated a set of standard message types (see Figure 1) that cover the majority of communications that are expected to take place regarding sample movement. Message creation follows a consistent pattern. Figure 2 illustrates the sequence of these messages, from the initial request for analysis capabilities to the last analysis response.

Message Type	Description
Request Availability Message (RAM)	The IDC asks a lab if it has the capability to perform analysis
Laboratory Availability Message (LAM)	The lab responds with a yes or no
Send Sample Message (SSM)	The IDC tells the field station where to send the split samples
Dispatch Notification Message (DNM)	The field station tells sample split recipients that the samples are on their way
Notification Receipt Message (NRM)	The sample recipients indicate receipt of the sample
Further Analysis Message (FAM)	The IDC asks the labs to perform further analysis
Further Analysis Response (FAR)	The lab responds with a yes or no

FIGURE 1 – Types of messages generated during additional sample testing.

The chain of messages for a particular sample begins when further analysis is requested on a sample that has been categorized as a preliminary level four or five event (high levels of radioanuclides detected). The initial correspondence is a Request Availability Message (RAM) requesting the use of a particular laboratory's facilities for analysis from the IDC. The station should then respond to this request using a Laboratory Availability Message (LAM), indicating whether the laboratory is capable of performing the necessary analysis. This process iterates until a suitable lab has been identified. Then, using a Send Sample Message (SSM), the IDC instructs the field station where the sample has been collected to partition the sample into a specified number of pieces, and mail it to specified laboratories. The corresponding labs are carbon copied this message. In turn, the field station sends a Dispatch Notification Message (DNM) to the IDC and laboratories acknowledging that the sample partitions have been mailed, as well as the date, time, and courier tracking number of the samples. Upon receipt of the bar-coded sample partition, the lab sends a Notification Receipt Message (NRM), allowing officials to review the efficiency of the sample distribution process.

Based on the results of the desired analyses, the IDC can either instruct to send the tested sample to the IDC for archiving or request further analysis using a Further Analysis Message (FAM). The laboratory then sends a Further Analysis Response (FAR) acknowledging whether it is able to perform the required analysis. This process is iterated as during the initial lab selection until a suitable laboratory is found. The messaging process generally concludes when all of the corresponding partitions for a sample are finally received at the IDC after satisfactorily completing the testing procedures.

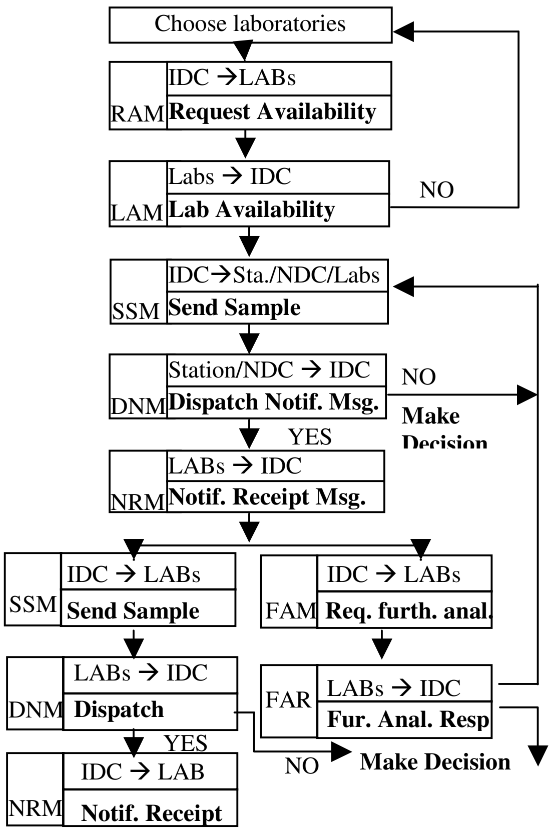


FIGURE 2- Flowchart of messages during additional sample testing.

4.5 System architecture

We developed a prototype system that includes the message forms and database schema to produce a messaging environment to accomplish our objectives. The RAME consists of web-based templates that facilitate message creation and viewing. Java applets parse the web forms into generic message forms that are then sent over an existing email system for transfer. A central server stores the message data in a relational database. Figure 3 illustrates the physical location of system components. The IDC contains the server that stores the RAME web forms, code, the Oracle database, the existing email system and parser. The field stations, IDC, labs, and NDC's have web browsers. The web forms are the basis of our messaging environment, as they display information in standardized formats. When a user creates a message through a form and sends it over the network in the existing email system, the parser disassembles the message and stores it in the database. The existing email system also transfers the information to the web browser of the message recipient, who will view the messages through another web form. Users access audit trail information from the database with the web browser. The audit trail includes summary information about messages for a particular sample. All information transfer occurs over the GCI.

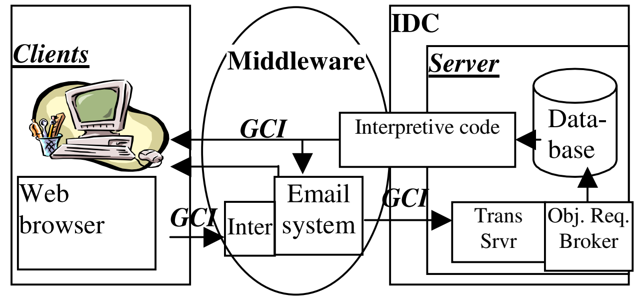


FIGURE 3 – RAME Client/ Server System Architecture

4.6 User Interface

Within the scope of this protocol, the message senders log on to the system, identifying themselves and their organization (see Figure 4). Once a user is logged on, their name and other information is retrieved from the database and the user is transferred to a general inbox environment (see Figures 5 and 6 for Inbox views from different organizations).

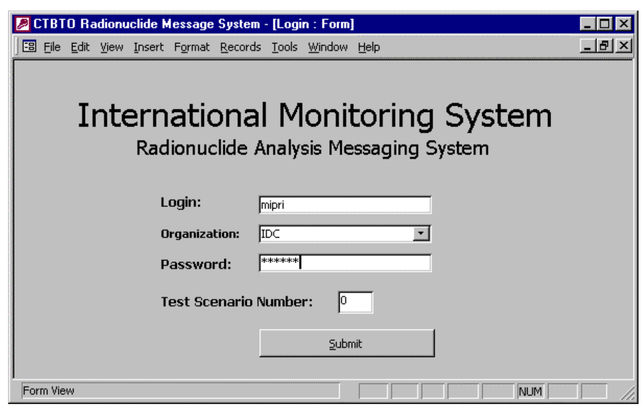


Figure 4. The Logon Screen

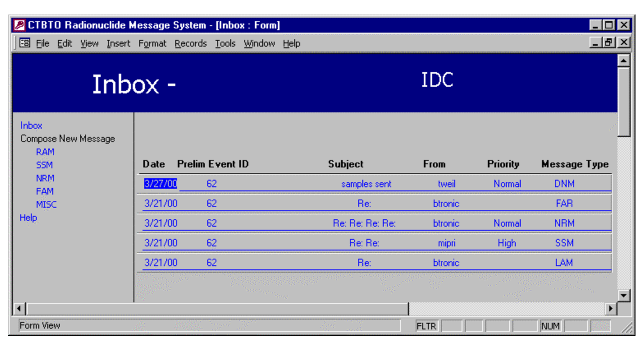


Figure 5. Inbox view for the IDC

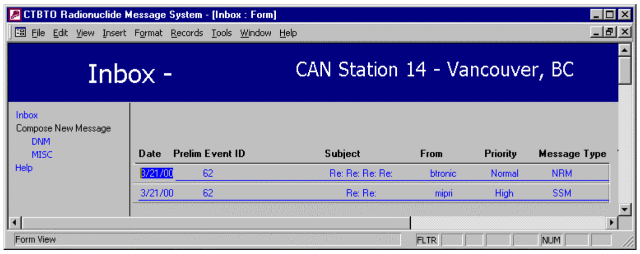


Figure 6. Inbox view for a Field Station (message creation options are more restricted)

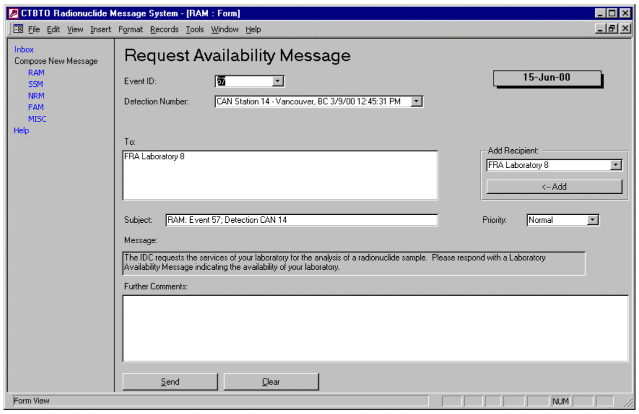


Figure 7. Request Availability Message Creation Form

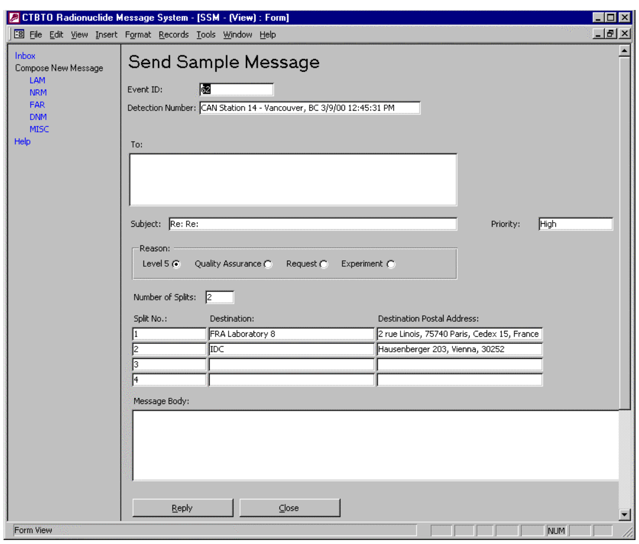


Figure 8. Send Sample Message Creation Form

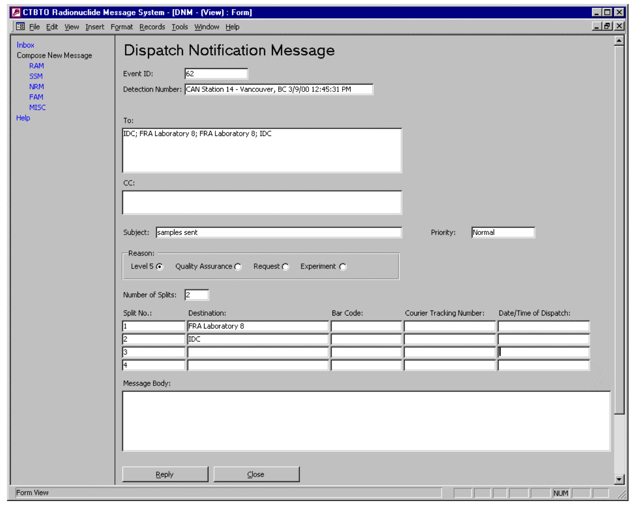


Figure 9. Dispatch Notification Message Creation Form

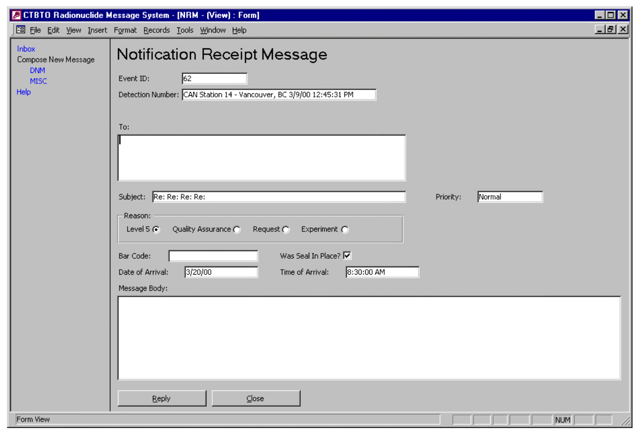


Figure 10. Notification Receipt Message Creation Form

Message Creation Forms. A menu of messaging actions is available on the left frame within the inbox, and allows for selection of message type, after which a standardized message template is displayed (see Figures 7, 8, 9 and 10 for sample message creation forms). The date, time, and user generating the message automatically appear. The message standardization includes aspects such as drop down menus, pre-populated fields within messages, automatic tables, and user reminders/ help functions. The sample identification is selected from a drop down menu when the initial RAM is generated, but is pre-populated during subsequent exchanges. All messages include a free text option. This is so additional comments can be added by the user of the system if desired. Once the user completes a template, the user may send the message. If users want to send each other messages that do not follow the messaging structure, they can fill out a "MISC" (miscellaneous) message, which is just a free-form email message (not shown).

Once a message is sent, the information is parsed and archived into the database. After this message transfer from the sender (client) to the database (server) occurs, other clients can reference the message. In order to receive the message, the receiving organization checks for new messages, selects the message, reads the message, and responds if necessary. Responding to a message will automatically bring up the right type of message (e.g., responding to a Request Availability Message will automatically bring up a Laboratory Availability Message, with the message pre-populated to the extent possible, and form-filling capabilities for replying). Each type of message has a free-text option in addition to the form-filling of required entries.

4.7 Audit Trail

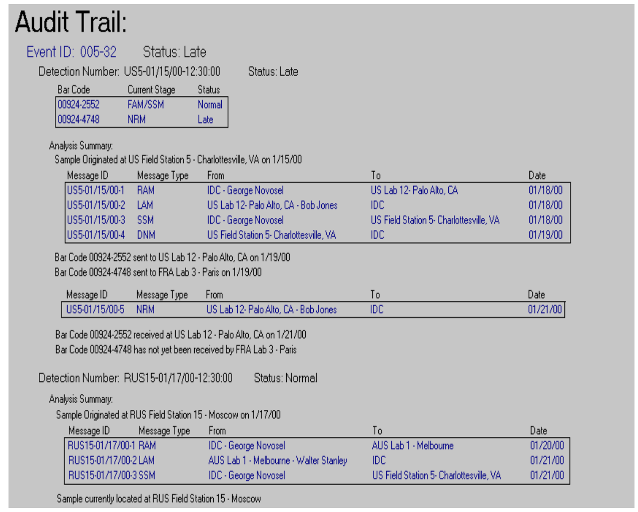
The audit trail is a summary report of the analysis process built from the information contained in the central database. It displays the current status of the analysis, a summary of the messages pertaining to a particular sample, and a record of the handling of each sample partition.

This information will be used by various groups concerned with the radionuclide verification process. For instance, members of the IDC will employ the audit trail in supervising the analysis process by using it to track the current location of each sample and to identify any phases of the analysis that need expediting. In addition, national officials from participating countries will use the audit trail to ensure that the analyses are being completed in a satisfactory manner. The audit trail will also be used to assign responsibility for the well-being of samples should they be lost, stolen, or tainted, thus increasing the security of the system. In this way, the audit trail serves to increase the overall effectiveness of and the level of confidence in the radionuclide analysis process.

Figure 10 shows a sample audit trail. It is organized according to the event hierarchy present in the database design. Under each event are listed the corresponding detection numbers; under each detection number, the corresponding bar codes. The status of each entity is also given, to allow users to determine the current analysis situation.

Following the status report is the analysis summary. This contains a record of the message traffic concerning the particular sample, including the date, message type, sender, and recipient of each message. Interspersed with the message summary is text detailing the handling of the sample partitions, including where each partition was sent, when it was sent, and when it was received. This provides an easy way to track the movement of each sample.

The audit trail will be accessed from the IDC web page over the GCI. Two different methods of access will be allowed. First, users will be able to search for any event or detection number they may be interested in. Second, users will have the ability to browse through all completed or current analyses. Once the audit trail is selected, an initial view appears displaying only events, detection numbers, and their statuses. This view is expandable by clicking on the desired detection numbers. This expanded view is shown in Figure 10.



Audit Trail:

Event ID: 005-32 Status: Late

Detection Number: US5-01/15/00-12:30:00 Status: Late

Bar Code	Current Stage	Status
00924-2552	FAM/SSM	Normal
00924-4748	NRM	Late

Analysis Summary:

Sample Originated at US Field Station 5 - Charlottesville, VA on 1/15/00

Message ID	Message Type	From	To	Date
US5-01/15/00-1	RAM	IDC - George Novosel	US Lab 12- Palo Alto, CA	01/18/00
US5-01/15/00-2	LAM	US Lab 12- Palo Alto, CA - Bob Jones	IDC	01/18/00
US5-01/15/00-3	SSM	IDC - George Novosel	US Field Station 5- Charlottesville, VA	01/18/00
US5-01/15/00-4	DNM	US Field Station 5- Charlottesville, VA	IDC	01/19/00

Bar Code 00924-2552 sent to US Lab 12 - Palo Alto, CA on 1/19/00

Bar Code 00924-4748 sent to FRA Lab 3 - Paris on 1/19/00

Message ID	Message Type	From	To	Date
US5-01/15/00-5	NRM	US Lab 12- Palo Alto, CA - Bob Jones	IDC	01/21/00

Bar Code 00924-2552 received at US Lab 12 - Palo Alto, CA on 1/21/00

Bar Code 00924-4748 has not yet been received by FRA Lab 3 - Paris

Detection Number: RUS15-01/17/00-12:30:00 Status: Normal

Analysis Summary:

Sample Originated at RUS Field Station 15 - Moscow on 1/17/00

Message ID	Message Type	From	To	Date
RUS15-01/17/00-1	RAM	IDC - George Novosel	AUS Lab 1 - Melbourne	01/20/00
RUS15-01/17/00-2	LAM	AUS Lab 1 - Melbourne - Walter Stanley	IDC	01/21/00
RUS15-01/17/00-3	SSM	IDC - George Novosel	US Field Station 5- Charlottesville, VA	01/21/00

Sample currently located at RUS Field Station 15 - Moscow

Figure 10 – Sample Audit Trail Screen Shot

4.8 Simulations

Simulations were conducted to verify that the system design meets the requirements, to detect any errors in the design (e.g. data structure, database population, interface, and sequencing errors), and to gain usability statistics. We developed test scenarios and a test user briefing about the overall system and the users' particular roles in the test. The briefing describes the sample testing process and the flow of messages between organizations (see Appendix 1 for an example test scenario). Three subjects worked at a time, acting as different users of the system (representing the IDC, a lab, and a field station). Subjects were asked to follow through the scenario, sending each other messages in a process that mimics the actual flow of messages during additional sample testing. Subjects were asked to talk aloud as they worked. Their activities were videotaped for analysis of usability problems. After testing, subjects filled out a questionnaire, which asked for demographic information about themselves as well as subjective ratings of the usability of the system.

A total of 9 subjects (3 groups of 3 working at a time) worked through our scenarios. Subjects were students at the University of Virginia, who are unfamiliar with the radionuclide analysis and messaging process. This testing was clearly a first step, which served to give us feedback on our designs without the difficulty and expense of setting up a more realistic field test across organizations in several countries. Clearly, further testing will need to be done with actual users of the system once our design is further along. Having naïve subjects is perhaps not an unrealistic test of performance, as the eventual users of the system will be diverse, from different countries, with a possibly high turnover rate. Thus, we would like a system that is intuitively easy to learn and use and supports teaching the diverse organizations of the expected pattern of behavior through the design of the user interface.

4.10 Usability Testing Results

Our testing results showed that although our system supported the sending and receiving of messages as desired, there are several usability issues to overcome. Of primary concern was that users did not take advantage of the form-filling features and automatic reply functions which are intended to greatly facilitate the process. Thus, we will add critiquing functionality [6], to alert the users to the need to fill out the required fields if they do not do so. Once users started to notice the form fields, they began to use them instead of filling in the message details in free text format in the available message body. We also changed the name of the generic "Reply" button to be more specific, such as "Respond with a Laboratory Availability Message." This improved usability to a great degree since the Reply button would automatically bring up the appropriate message type in the sequence of activities. We are also considering ways to make the required form fields more salient and the free-text textbox less salient, perhaps even requiring users to select a button to add in additional comments to their message if desired.

The audit trail design performed well in preliminary usability testing. Testers were able to locate and understand the information they were looking at in the vast majority of cases. Suggestions include providing separate sections for the message and handling summaries, so as not to break up the flow of the message stream and make the audit trail easier to view.

The testing conducted to date points to some usability issues with actually having people fill out the required fields in the interface but does not address deeper issues on tradeoffs between rigidity and the need for freeform conversation. Until we can conduct a more realistic field study, we can not determine whether situations will arise involving unexpected communications which do not fit the forms provided by the system and to see how subjects will handle this.

Clearly, overall usability needs to be improved before our system can be deployed. An alternative approach to the "Inbox/e-mail" metaphor explored so far is the use of workflow management screens as a means to show users where they are in their required activities (i.e., which samples are pending, what messages need to be attended to, who is waiting for a response, etc.). We will begin to explore these types of user interfaces next. One question to answer with such follow-on studies is whether a task management type of interface can support coordination, communication and cooperation without explicitly filling out message templates that look like email messages or whether a smart combination of the two makes the most sense.

5. Conclusion

We have discussed the early design activities for developing an international, coordinated workflow management system. As part of this project, we modeled the envisioned workflow activities, developed a set of message protocols that mapped onto this activity, developed a client-server system architecture to support communications, and developed a prototype user interface and test scenarios to begin to test the feasibility and usability of our envisioned system. This system can be classified as a type of persistent conversation system, as it enables the archiving, tracking, monitoring, and review of distributed communication activities. Ours was only a first step in the development of a worldwide system to support the activities and verification necessary to successfully monitor countries' compliance with the Comprehensive Nuclear Test Ban Treaty.

It is our intention that the architecture that we are designing will enable the growth of the system to accommodate future uses and needs. For example, by separating the message protocols from the user interface through a parser engine, we can improve or even develop different user interfaces for each of the different organizations or countries using the system without greatly affecting the other components of the system. Clearly, once we have confirmed that we have the right kinds of message types included in the system, we will need to develop additional monitoring and workflow tracking displays beyond the archiving reports and message displays that were described here.

This work follows on well from the work done to date on conversational systems, computer-supported cooperative work, and workflow management systems. Today's technology with client-server architectures enables the design of these kinds of systems and they will become more prevalent as time goes on. Much of the work that has been done in groupware, planning and scheduling systems will help inform us of the means to move forward in developing effective, distributed workflow systems. Literature from linguistics, conversation analysis and group coordination help inform us of the cognitive activities that these types of systems need to support, and thus help us to design effective tests of these systems. We now have to develop distributed usability tests to evaluate the coordinated activities of many people, beyond the traditional, one-person, one-computer kinds of tests that are typically done to date. Workflow management is inherently a group activity, and thus it is necessary that each member of the group is able to perform their activities in a flexible way, while still meeting the organizations' monitoring and tracking requirements. Much work remains to be done in designing, understanding and evaluating such systems.

6. References

[1] F. Barnaby and T. Dokos. "Negotiations for a CTBT 1958-1994: Analysis and Evaluation of American Policy" Journal of Strategic Studies v 20 n 2 (1997): 147.

[2] H. Clark and E. Schaefer, "Contributing to Discourse," *Cognitive Science*, vol. 13, pp. 259-294, 1989.

[3] H. Clark and D. Wilkes-Gibbs, "Referring as a Collaborative Process," *Cognition*, vol. 22, pp. 1-39, 1986.

[4] W. Dangelmaier, S. Kress, and R. Wenski "Telework Under the Co-Ordination of a Distributed Workflow Management System" *Workflows 'In the Wild' / GROUP'97: International Conference on Supporting Group Work,* pp.128-137, 1997.

[5] H. Dewar, "Senate Rejects Test Ban Treaty." The Washington Post 14 Oct. 1999, A1, A6, A7.

[6] S. Guerlain, P. Smith, J. Obradovich, S. Rudmann, P. Strohm, J. Smith, J. Svirbely, and L. Sachs, "Interactive critiquing as a form of decision support: An empirical evaluation," *Human Factors*, vol. 41, pp. 72-89, 1999.

[7] G. Mark, "Merging Multiple Perspectives in Groupware Use: Intra- and Intergroup Conventions," *GROUP'97: International Conference on Supporting Group Work,* pp.19-28, 1997.

[8] M. Morris, T. Plant, and P. Hughes, "CoOpLab: Practical Experiences with Evaluating a Multi-User System Cooperative Systems." *Proceedings of the HCI'92 Conference on People and Computers VII* pp.355-368, 1992.

[9] J. Nielsen, Usability Engineering. San Francisco, CA: Academic Press, 1993.

[10] D. Tatar, G. Foster, and D. Bobrow, "Design for conversation: Lessons from Cognoter," *International Journal of Man-Machine Studies*, vol. 34, pp. 185-209, 1991.

[11] Preparatory Commission for the Comprehensive Nuclear-Test-Ban Treaty Organization "Procedures for Dispatching Samples and Incorporating Reports from Radionuclide Laboratories Into IDC Products", Vienna, Austria. 10 February 1999.

[12] M. Turoff, S. Hiltz, M. Bieber, J. Fjermestad, and A. Rana, "Collaborative Discourse Structures in Computer Mediated Group Communications," *Journal of Computer-Mediated Communication*, vol. 4, 1999.

[13] T. Winograd and F. Flores, *Understanding Computers and Cognition: A New Foundation For Design*. Norwood, NJ: Ablex Publishing Corporation, 1985.

Appendix 1 – Test Scenario Briefing

The Comprehensive Nuclear Test Ban Treaty (CTBT) is being verified with a network of field stations, laboratories, and an International Data Center (IDC). Field stations collect air samples and send analysis data to the IDC. The IDC interprets this data and determines if a possible recent nuclear event has occurred somewhere near the collection site. If questionable conditions exist (considered a "level 5" event in some circumstances), the IDC can request additional analysis by laboratories.

Throughout this additional analysis process, messages must be sent between the field stations, laboratories, and the IDC. The Radionuclide Analysis Messaging Environment (RAME) facilitates message communication between these organizations. The following message types convey sample information:

Message	**Acro-nym**	**From**	**To**	**Purpose**
Request availability message	RAM	IDC	labs	request availability of services
Laboratory availability messages	LAM	labs	IDC	confirmation or denial of services
Send sample message	SSM	IDC	station labs	notification to send samples to labs
Dispatch notification message	DNM	station labs	IDC	confirmation of sample dispatch
Notification receipt message	NRM	labs IDC	IDC/la bs	confirmation of sample receipt
Further analysis message	FAM	labs	IDC	request for further analysis
Further analysis response	FAR	IDC	labs	confirmation or denial of services
Miscellane ous	MSC	anyone	anyone	miscellaneous

Each of you will act as a member of a field station, a laboratory, or the IDC. Please follow the test scenario to simulate communications among these organizations. Use the help menus as necessary to complete your tasks.

Tester	Username	Organization	Password
1:	mipri	IDC	boxout
2:	btronic	Lab 8	tofutti
3:	tweil	Field Station 14	225

Tester 1: You are a member of the IDC. Your username is "mipri" and your system password is "boxout". Log in to the system with test scenario 2. A level 5 event was detected and you must find out if lab number 8 is available to perform additional analysis. Compose a request availability message to accomplish this. The sample you are dealing with has event number 62 and detection number "VAN – 12:00". Send this message with normal priority.

Tester 2: You are a member of Lab 8. Your username is "btronic" and your system password is "tofutti". Log in to the system with test scenario 2. Check your messages by clicking on the date of the message in the inbox and verbally explain what the message means. Reply to the IDC's message with a LAM, stating that your lab is available to analyze the sample but that the earliest the IDC will receive an analysis report will be two days after receipt.

Tester 1: Check to see if Lab 34 has responded to the availability request by clicking on the date of the message in the inbox. Explain the lab's response. Reply by sending a SSM message to field station number 14, the IDC, and lab 8 asking field station 14 to split the sample with detection number "VAN – 12:00" and to send the split sample to these locations: the IDC, and lab 8. You are sending this high priority message because the level 5 event was detected. Let the field station know that samples must be sent no later than this afternoon.

Tester 3: You are a member of field station 14. Your username is "tweil" and your system password is "225". Check your messages by clicking on the date of the message in the inbox, and explain the message from the IDC. Reply to the IDC with a DNM to indicate that you have sent the sample splits. Carbon copy that message to all the locations that are to receive a split of the sample. Include the following barcodes: for destination IDC: 16725, for destination Lab 8: 38746. This is a high priority message sent because of the level 5 event. State in your message that all the sample splits will be sealed and sent out by Friday at 1700 local time. We are using DHL as the mail deliverer.

Tester 2: Check your messages. Reply by sending a NRM message to the IDC and field station 14 stating that you have received the sample with bar code 6789 for additional testing from the field station. Indicate in the message that you received the sample on 6 March 2000 at 0830. Include that the seal was in place on the sample when it arrived and the reason for receiving the sample was the level 5 event.

Tester 1: Check your messages, and reply to the notification receipt message with a FAM. Send a message to the lab requesting further analysis on a sample. In the message, ask the lab if it has the capability to perform the additional testing tomorrow.

Tester 2: Check your messages and reply with a FAR message to the IDC confirming that you can perform additional testing. Indicate in the message that your lab does have the capability to perform these tests tomorrow.

Tester 1: Compose (you cannot reply to the previous message) an NRM message to lab 8 stating that you've received the sample data and that no further action is required. Include the proper event ID (62), detection number (VAN – 12:00), destination (lab 8), and reason (level 5). Indicate that you received the data on Friday at 1100.

Tester 1: Check the audit trail to research past messages. Select the trail by event number and query for event number 62.

1. Find the date of the Send Sample Message from the IDC to the field stations.
   Date:____________
2. Determine on what day the sample reached the laboratory from the field station.
   Date:____________________
3. Determine the barcode of the sample split sent to the Lab 8.
   Barcode:_____________________
4. Find the day that the further analysis was requested of Lab 8.
   Date:____________________