PNWD-SA-6555

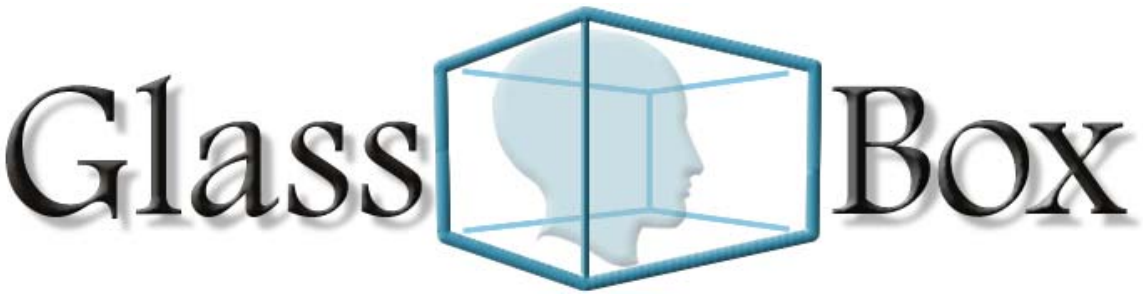
Glass Box: An Instrumented Infrastructure for Supporting Human Interaction with Information

Paula Cowley, Battelle
Lucy Nowell, ARDA
Jean Scholtz, NIST

For Presentation at:

Hawaii International Conference on System Sciences

January 3-6, 2005



Battelle Memorial Institute

Glass Box: An Instrumented Infrastructure for Supporting Human Interaction with Information

Paula Cowley
Battelle Memorial Institute, Pacific Northwest Division
paula.cowley@pnl.gov

Lucy Nowell
Advanced Research and Development Activity
ltnowel@nsa.gov

Jean Scholtz
National Institute of Standards and Technology
jean.scholtz@nist.gov

Abstract

In this paper, we discuss the challenges involved in developing an infrastructure to support a new generation of analytic tools for information analysts. The infrastructure provides data for establishing context about what the analyst is doing with the analytic tools, supports an integration environment to allow suites of tools to work together, and supports evaluation of the analytic tools. We discuss the functionality of the Glass Box, the challenges of evaluating adaptive systems including the capture of data for evaluation metrics, and lessons learned from our experiences to date.

1. Introduction/background

The Advanced Research and Development Activity (ARDA) program on Novel Intelligence from Massive Data (NIMD) is aimed at creating a new generation of analytic tools to support human interaction with information. In short, we are trying to build smart software assistants and devil's advocates that help analysts deal with information overload and avoid analytic errors.

"Interaction" implies mutual exchange. When humans interact, our behavior changes over time as we learn about one another, develop a shared context for dialogue, and establish expectations about one another's beliefs, responses, and behavior. Today's analysis tools leave the entire burden of learning and adapting behavior to human users. The tools themselves do not learn from or about the analyst, the analytic process, or the problem at hand. What is required to make tools function as more equal partners, capable of transforming their performance and their presentation of information over time based on growing knowledge about users?

NIMD research is focused on five areas: modeling analysts and analytic processes, capturing and reusing prior and tacit knowledge, generating and managing hypotheses, organizing/structuring massive data (mostly unstructured text), and human information interaction. At the heart of NIMD is a piece of software called the Glass Box that resides on an analyst's workstation and captures the parts of the analytic process that happen online — queries and the sources to which they are sent, search results, documents viewed, draft reports, etc. — recording what an analyst actually does online during analysis. Features of the Glass Box also encourage the analyst to enter notes about off-line activities, including collaboration and use of hard-copy materials. Much of the NIMD research attempts to automate "learning" from Glass Box data about the analyst and the task at hand to achieve a higher level of synergy between the analyst and the analytic tools. The Glass Box serves as the sensory input mechanism for these tools – the eyes, ears and hands that provide data about analytic activity.

Like many software users, analysts already suffer from a plethora of tools that operate in isolation and exchange data poorly. To make the work of analysis easier and to promote adoption of new methods, the NIMD program aims to produce an integrated analytic environment in which the new tools are highly interoperable. In addition to its data capture role, the Glass Box serves as an integration platform for the suite of NIMD research systems. The resulting NIMD Research Environment should allow analysts to focus on what humans do best—the analysis itself—while the assistants aid in doing what computers do best – processing massive amounts of data.

2. NIMD research systems

For NIMD's smart software assistants and devil's advocates to be truly useful, they should have knowledge of the analysts' situation, have some understanding of the context in which the work is being performed, and be able to adapt their actions appropriately. Specifically, they should be able to:

- Take on the mundane tasks computers are good at and analysts are poor at:
  - Finding resources that are relevant, particularly ones that aren't obvious from standard query processing.
  - Processing and prioritizing massive information resources.

- Connecting the dots: Building multiple hypotheses and inference nets to look for potential correlations between pieces of previously unrelated or poorly connected dots.

- Continually monitor information sources for updates or significant changes.
- Discern the type of information the analyst finds relevant and provide dynamic, adaptive filtering to provide the most useful information.
- Help the analyst by monitoring the information the analyst is looking at and noting information that the analyst may be missing.
- Provide reminders of what the analyst already knows.

What do they need to know to accomplish this? Some possibilities include:

- What is the analyst doing, including what kind of task is being worked on and when the task needs to be completed.
- What the analyst has done in the past and, perhaps, what others in the same organization have done.
- How this particular task compares to what the analyst has previously done.
- Whether or not the analyst needs help; whether progress is being made on the task.
- Whether the analyst can be interrupted at any given point for assistance to be provided. This includes the ability of the system to prioritize the messages to be delivered to the analyst.

To be useful to an analyst, the system must learn about that analyst and adapt to the analyst and the situation. Information about the analyst and the situation needs to be captured in such a way that it is useful for the system. Likewise, the analyst needs to understand the capabilities and limitations of the system.

The Glass Box provides an instrumented infrastructure to help the NIMD research systems achieve these ambitious goals. Below we describe progress to date in developing this infrastructure and describe how the NIMD research systems can use it to support the systems' interactions with the human analyst.

3. The Glass Box

The Glass Box provides three major functions (as shown in Figure 1) within the NIMD Program:

- It supplies the "instrumentation" that captures workstation activity as analysts work in the Glass Box. Our instrumentation captures data about the context within which the analysis is taking place and reveals much about the analysis process being used.
- It serves as an integration platform for NIMD research systems so the systems can interact with the Glass Box and with other peer research systems that are handling additional functions. Integration functions and data include a record of current and past activities, logging activities important to the NIMD research systems, live event notification of analyst activity, and enabling communication between collaborating NIMD research systems.
- It serves as a tool for evaluating the effectiveness of research products by recording what really happened during their use. Essentially, the Glass Box serves as a surrogate observer, freeing the human observers from having to record detailed workstation activity so they can focus on higher-level cognitive activities.

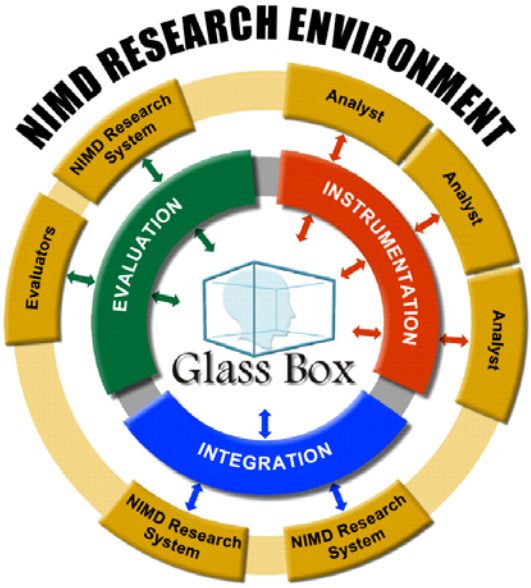


Figure 1: NIMD Research Environment

4. Glass Box instrumentation

The Glass Box serves as a sensory mechanism for the NIMD research system. By capturing user interactions and system activity, the Glass Box allows the NIMD research system to find out what the analyst is doing in near real time. Then it is up to the system to determine what that activity really means and respond appropriately. By supporting the system's ability to retrieve instrumentation data stored from previous activities as needed, the Glass Box also provides access to what the analyst has done in the past. This allows the system to develop and work from a current model of what the analyst is trying to do.

The current Glass Box instrumentation captures analyst workstation activities including keyboard/mouse data, window events (active window, active application, location on screen, etc.), file save events, copy/paste events, and Web browser activity (URLs, page contents, images, queries, and query results). The Glass Box makes extensive use of a relational database to store time-

stamped events and a hierarchical file store where files and the content of web pages are stored. The Glass Box "snatches" a copy of every file the analyst saves so we have a complete record of the evolution of documents. We also explicitly store the contents of every web page the analyst visits so we can recreate the content as it existed at the time the page was visited. We capture screen images to see what the analyst saw during the session.

The Glass Box's Control Panel allows the analyst to suspend and resume recording to make sure the software does not inadvertently capture sensitive or proprietary information. Analysts also have the ability to delete data inadvertently recorded.

The instrumentation we have described to this point is relatively passive. Except for the responsibility of the analyst to turn on or turn off Glass Box recording, there is virtually no interaction with the analyst and no cooperation is required from the applications we have instrumented. Most of what this instrumentation captures is obtained by observing the application from the operating system level.

Much of the analytic process occurs offline: in the mind of the analyst, in collaborative conversations with other analysts, and in odd moments of reflection while doing other tasks, to mention just a few possibilities. These activities are important to tools that attempt to model analysts and analytic processes. The Glass Box provides means for the analyst to record comments and notes about off-line activities. The analysts use the Glass Box annotation tool to note meetings, casual discussions, items of interest from the media, hypotheses being considered, plans for the day, etc. Analysts can also annotate and mark the relevance of citations, files, and excerpts from files. These annotations often provide the NIMD researchers with important context that supplements the recorded workstation activities, but they are written for processing by humans rather than by the NIMD research systems. Figure 2 shows an annotation window.

The Glass Box software also provides tools for examining the data that have been collected. The Glass Box Review Tool provides analysts and researchers with a quick review and analysis capability, including the capability to observe:

- Progression of activity over time
- Repeated patterns over time
- Information clusters (i.e., areas where analysts repeatedly spend time)
- Breadth of scanning and narrowing of focus.

The Review Tool provides three different and complementary ways of visualizing the Glass Box data. The **Tabular Review** window shown in Figure 3 displays a chronological record of recorded events that can be explored in detail. The display can be filtered in various ways (using specialized queries) to study different types of activities and relationships in the data. The Tabular Review window also allows the events to be displayed in a context and view that is very close to what was seen by the analyst (we call this view "deja view"). **Over-the-Shoulder** and **Timeline Review** windows provide additional ways of visualizing the data that help to develop an understanding of the structure, timing, and content of analyst activity. Over-the-Shoulder Review provides screen images of what the the analyst was seeing on his or her workstation at the selected point in time. The screen capture operates at a pre-selected rate, currently one image per second. The Timeline Review provides a time-based visualization of multiple events in relation to each other. This view shows the flow of information across activities, the specific points in time where analysts performed mouse actions and keyboard inputs, and the resulting effects on applications and windows.

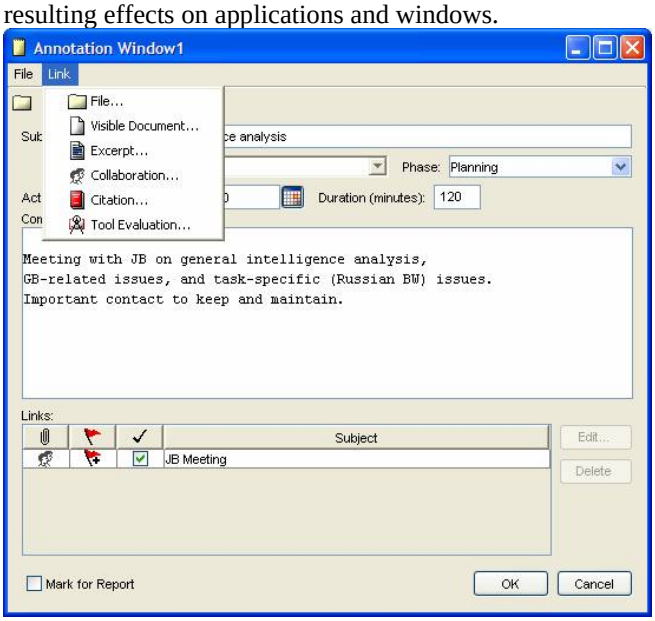


Figure 2: Glass Box Annotation window

The Glass Box provides hooks for applications to instrument themselves from the inside, based on what they need from our instrumentation infrastructure. The next section discusses how this part of the Glass Box is evolving to meet NIMD integration requirements.

5. Glass Box as an instrumentation infrastructure for integration

NIMD's ambitious research goals cannot be met fully by any single member of the NIMD research community. Thus the goal is for NIMD research efforts to be integrated via the Glass Box to create the NIMD Research Environment, as shown in Figure 1. In addition, NIMD researchers are also finding it mutually beneficial to integrate with each other because many have complementary research scopes.

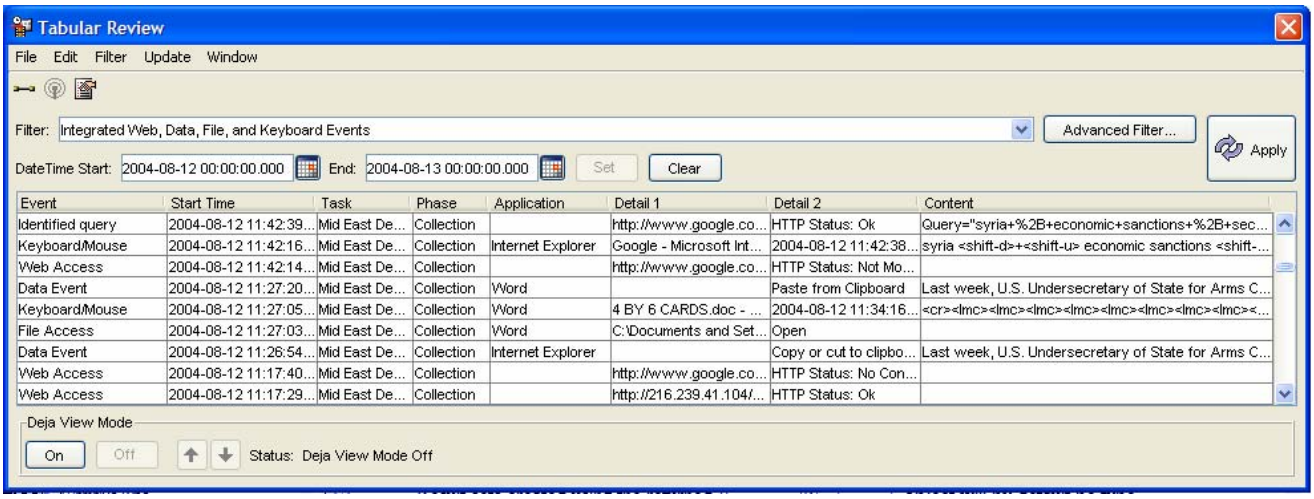


Figure 3: Tabular Review showing integrated web, data, file, and keyboard events

The Glass Box research team has implemented an Application Programming Interface (API) [1] as its major integration product. Through the API, the NIMD research systems can plug into the Glass Box's instrumentation infrastructure to store data in the Glass Box datastore, log activities important to each system, receive notification about analyst activities, retrieve data from the datastore, receive live event notification information about what the analyst is doing, and share information between themselves. The following scenario (see Figure 4) describes how NIMD research systems can use the API to collaborate: A NIMD research system called NRS-A has written a query capability and it uses logging to record the query and its results. Another NIMD research system called NRS-B is developing hypotheses and finds that queries provide valuable input in the formulation of hypotheses. So NRS-B subscribes to be informed of queries from NRS-A. NRS-B uses the information from the API to formulate new hypotheses and informs the analyst that it has some new hypotheses for the analyst to consider.

The Glass Box API is designed to support scenarios such as the one above. The API is logically broken into segments based on functionality:

1. The DataAccessAPI allows the system to read data from the Glass Box database and to obtain file locations for items in the file store. This is oriented around the functions our own Glass Box software performs so the system can retrieve the kinds of data retrievable through our review tools.
2. The WriteReadAPI provides functionality for the system to write its own data to the Glass Box database and read that data back. The system can write opaque character large objects (CLOBS) or binary large objects (BLOBS) along with associated metadata to the Glass Box database. The content of the objects is opaque to Glass Box so interpretation of that data is the responsibility of the NIMD research system.

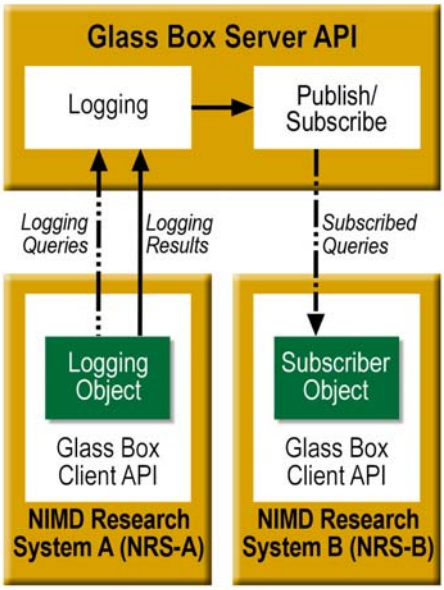


Figure 4: Using the Glass Box API to log and subscribe

3. The ReviewFilterAPI provides NIMD research systems with access to the same filters available in the Glass Box Review Tool. These filters use complex heuristics to filter noise out of the data and are impossible to replicate using straight SQL queries. The system can get the names of the available filters, select and execute a filter for a specified timeframe, and receive an array of data rows with each row having an array list of field names and an array list of values. It can use this

mechanism, for example, to filter data using our comprehensive review filter to see sessions, non-redundant window events, primary web events, file save events, copy/paste events, and logging events.

4. The publish/subscribe function allows the system to subscribe to topics of interest. The system can create and destroy topics; be notified on any newly created topics; determine whether to subscribe to a topic; subscribe to a topic; receive notification of events associated with that topic (e.g., calls to the WriteReadAPI associated with a topic); and retrieve the data associated with that topic.
5. The logging function can be used for testing, tuning, and evaluation. Much interesting information is not captured by the passive instrumentation and is known only to the application. We have designed a flexible logging capability where the application specifies its own tags using XML and stores them as CLOBS in the Glass Box database. This allows NIMD research systems to log events useful for evaluation. Some standard tags may be required to support the evaluation process. We expect log analysis to be done outside the Glass Box using XML-aware tools.

6. Glass Box as an evaluation medium

Evaluating the NIMD research efforts is also a major research project. We have identified three objectives: first, give the research teams feedback on how their software is being used and what needs to be done to improve it; second, provide information to the program management to assess the progress that the program as a whole is making; and third, provide evidence of the impact of the systems on the analytic process and impact of the products on helping analysts decide which NIMD technologies to adopt. To accomplish this, we need to identify metrics and measures and we need to obtain the measures from the data captured by the Glass Box.

6.1. Metrics and measures

The analysis process is a complex, iterative process that includes: understanding the question/task; framing the problem and identifying assumptions; discovery/exploration; synthesis and analysis; collection planning and "what if" analysis and simulation; and the conveyance of knowledge to others.

Traditionally, analysts have spent more time doing research and producing final analysis products than doing analysis. We are hoping that the NIMD research systems (1) will enable more time to be spent in analysis and less time spent in research and producing final products or (2) will enable more analysis to be performed in the same or less amount of time.

Our process of developing metrics and measures has been to work with the research teams to determine what they envision as the impact of their software on the analysts and on the analysis process itself. While our metrics are currently still in development, we have completed five pilot evaluations and have used the metrics shown in Table 1.

Process metrics such as the rate of information growth, number of documents read, reading time per document, and the number of queries are attempts to compare the research phase of traditional analysis using traditional knowledge worker tools to analysis using the new NIMD research systems, which are intended to improve the analysis process. To produce these metrics, we collect measures from the Glass Box and from data supplied by the research systems.

Table 1: Process and product metrics

Process Metrics	Product Metrics
• Solution time (for problem-solving software) • Number of documents read • Reading time/document • Number of queries • Comparison to expert • Rate of information growth • Number of system-generated hypotheses considered by the analyst • Number of relevant documents / query	• Accuracy • Quality, including confidence in recommendations

6.2. Using Glass Box for evaluation

To provide NIMD researchers with actual analysis data, several analysts have been performing analysis using the Glass Box. The data have been supplied to NIMD researchers and have been one of the sources used for development of user models and for initial development of metrics.

During the first phase of the program, two analysts worked in the Glass Box. Each analyst was working on a separate task. In order to understand what information was captured by the Glass Box and how accurately it represented reality, we evaluated the Glass Box itself [2]. Two people observed the Glass Box analysts for two days. We videotaped the analysts and recorded time stamped observational notes, but we did not intervene. At the end of the day, we conducted debriefing sessions to ask about particular activities that we did not understand. The data from the Glass Box were taken to our evaluation

team, and they analyzed the analysts' activities based solely on the data. Then we compared our observations to the evaluation team's analysis to determine what was missed, if anything. We then produced an annotated version of a portion of the Glass Box data to help the research teams understand and interpret the lower-level data.

Our observations of Glass Box analysts and the corresponding data analysis suggest that:

- The gaps in which there is no analyst activity can be detected and correspond with observed offline activities.
- However, there is currently no way of capturing the activities that occur in these gaps, such as meetings, phone calls, offline reading and thinking, and e-mail activity on the analysts' second computers unless the analyst annotates such activity.
- The time periods of data collection and report generation can be obtained from Glass Box data.
- While we can capture physical interactions with application software, capturing the analysts' intent is only possible through analyst-initiated annotation.

7. Lessons learned

We have learned that instrumentation is a difficult technical problem in itself. Most applications were never intended to be instrumented. With the variety of ways in which user interfaces can present themselves, interpreting the pretty array of pixels on the screen to derive meaning has been more difficult than anticipated. In addition, each application we instrument brings its own unique set of instrumentation challenges. As a result, we have been limited in the number of applications we provide detailed information about. For example, our analysts were initially using two different browsers. At our request, they now are only using Internet Explorer. This limits the amount of software development and maintenance we have to invest in developing and maintaining our software to deal with browsers.

Analysts who work in the Glass Box are fully aware that their workstations activities are being recorded and that the data they generate will be distributed to the NIMD research community. In addition, the analysts are protected under human subjects regulations. Although this provides protection to our analysts, it also places additional requirements on the Glass Box instrumentation. Some of these requirements include not recording passwords and other personal or identifying information and taking measures to preserve anonymity of the analysts in documents, email, and other products. The plethora of ways in which such data can end up in the Glass Box datastore has made protecting our analysts a challenge. We have avoided capturing keystrokes in some cases and have post-processed the data in other cases. Live event notification via the API also requires that data be processed before delivery to the NIMD research systems. Our insight into the difficulty of this problem has increased as we have discovered additional novel ways that systems solicit and record such information. In actual use, personal and identifying data would not be distributed to researchers, but nonetheless, we anticipate that the collection of such data will certainly be a privacy issue that will have to be dealt with by organizations using such systems to generate their own data.

Pilot evaluations are currently being conducted with the various NIMD research teams. These evaluations not only provide feedback to the researchers about the use of their systems but also provide data about what types of information the system needs to log within the API framework to facilitate refinement and future evaluation efforts. The Glass Box serves well as a "surrogate observer" of workstation activity while evaluating NIMD research systems. The Glass Box logs workstation activities to a level of detail that no human observer could achieve and provides the ability for the human observers to review those observations at many different levels of detail. However, only a portion of the overall analysis process is represented by capturing workstation activity. Much of the analysis process happens off line. Because analysts are usually pressed for time, they will not routinely record these off-line activities. However, the Glass Box analysts have explicitly been asked to record their off-line activity via the Glass Box annotation feature. NIMD researchers have found these annotations to be extremely useful in understanding the analysis process these analysts used and making sense of the workstation activity. The issue is that these annotations can only be processed manually – by humans reading the annotations after the fact. The question is: can NIMD research systems glean this type of knowledge from the raw Glass Box data capture? How much understanding the NIMD research systems will need about off-line activity to be truly useful remains to be seen. We are also waiting to see whether the analysts consider the annotations to be useful enough to keep recording them when no longer required to do so.

8. Future work

In this paper, we have described an infrastructure developed to provide data for establishing context for supporting human interaction with information. The infrastructure also provides a means for evaluating the use of systems that enable such interaction. While we have made good progress in capturing much of what the analyst is doing, major challenges remain:

Analysts have made it clear that they have neither time nor tolerance for a host of new tools that each operates in isolation, with unique input requirements, user interfaces, and operating characteristics. Analysts need one seamless

environment for analysis – an environment that supports interoperable use of specialized tools as they are needed.

One key to achieving such integration is providing shared knowledge bases that use a limited number of knowledge representation formats. The design of these knowledge bases and knowledge representation standards has barely begun, but it is clear that the Glass Box must support or host them in its role as an integration platform.

We are also working to:

- Capture collaborations that occur outside of the electronic environment.
- Ensure that the data provided by the instrumented infrastructure are sufficient to provide the context needed for the software being developed in the NIMD program.
- Provide effective support that allows different NIMD research systems to interoperate and communicate with each other and with the Glass Box.
- Develop and implement policies for privacy in data capture.
- Ensure that the capture environment provides solid benefits to the analysts in return for manual entry of annotations.

We continue to release improved versions of the Glass Box. More in-depth observations of the analysts have been conducted and an annotated set of Glass Box data produced. This will allow us to determine what analyst activities are being captured and what activities are not. Four analysts are now working on multiple tasks and a collaborative task is planned for 2005. While these conditions are much more realistic, they will present even more challenges for capturing the analysis process. For example, this requires that we attempt to match activities to the applicable task when analysts are switching among several tasks even if they forget to inform the Glass Box that they have switched tasks.

As NIMD research systems are integrated into the NIMD Research Environment, we will conduct more observations of the analysts working in the Glass Box and will analyze the captured information to determine how well we have captured data describing the actual context and use of the NIMD research systems within the analysis process. Our expectation is that the infrastructure the Glass Box provides will advance the ability of the NIMD research systems to be informed of what the analyst is doing, interact and collaborate with the analyst, and collaborate with other systems to produce an improved environment for analysis. All of this is aimed at improving the ability of the human to interact with and meaningfully use the massive amount of information awaiting analysis.

Acknowledgements

This work was funded by the Advanced Research and Development Agency's Novel Intelligence from Massive Data Program.

References

[1] Gorton, I., Haack, J. N. 2004. Architecting in the Face of Uncertainty: An Experience Report, International Conference on Software Engineering.

[2] Scholtz, J., Morse, E., and Hewett, T. 2004. In Depth Observational Studies of Professional Intelligence Analysts. Human Performance, Situation Awareness and Automation Technology Conference.