charles river analytics

+ Charles River Analytics Inc.
625 Mount Auburn St.
Cambridge, MA 02138

www.cra.com

+ p: 617.491.3474
f: 617.868.0780

20 February 2014

Dr. Harold Hawkins
ONR Code 341
Office of Naval Research
875 North Randolph SL
Arlington. VA 22203-1995

Reference: US Navy Contract N00014-12-C-0653: “The Model Analyst's Toolkit: Scientific Model Development, Analysis, and Validation”
Charles River Analytics Contract No. C12186

Subject: Contractor’s Quarterly Status Report #6
Reporting Period: 20-November-2013 to 19-February-2014

Dear Dr. Hawkins,

Please find enclosed 1 copy of the Quarterly Status Report for the referenced contract. Please feel free to contact me with any questions regarding this report or the status of the “The Model Analyst's Toolkit: Scientific Model Development, Analysis, and Validation” effort.

Sincerely,

W. Scott Neal Reilly

W. Scott Neal Reilly
Principal Investigator

cc: Cheryl Gonzales, DCMA
Annetta Burger, ONR
Whitney McCoy, Charles River Analytics

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
APR 2014		**20-11-2013 to 19-02-2014**

4. TITLE AND SUBTITLE	
The Model Analyst's Toolkit:Scientific Model Development,Analysis, and Validation	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Charles River Analytics,625 Mount Auburn Street,Cambridge,MA,02138**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**16**	

Charles River Analytics

Monthly Technical Progress Report No. R12186-06
Reporting Period: November 20, 2013 to February 19, 2014
Government Contract No. N00014-12-C-0653
Charles River Analytics Contract No. C12186

The Model Analyst's Toolkit: Scientific Model Development, Analysis, and Validation

Quarterly Status Report

Principal Investigator: Scott Neal Reilly

Charles River Analytics
625 Mount Auburn Street
Cambridge, MA 02138
617-491-3474

February 20, 2014

1. Executive Summary

The proposed research effort builds on and extends the work of the previous ONR-funded "Validation Coverage Toolkit for HSCB Models" project. The overall objectives of the on-going research program are:

- Help scientists create, analyze, refine, and validate rich scientific models
- Help computational scientists verify the correctness of their implementations of those models
- Help users of scientific models, including decision makers within the US Navy, to use those models correctly and with confidence
- Use a combination of human-driven data visualization and analysis, automated data analysis, and machine learning to leverage human expertise in model building with automated analyses of complex models against large datasets

Specific objectives for the current effort include:

- **Fluid temporal correlation analysis.** Our objective is to design a new method for performing temporally fluid correlation analysis for temporal sets of data and implement the method as a new prototype component within the Model Analyst's Toolkit (MAT) software application.
- **Automated suggestions for model construction and refinement.** Our objective is to design and implement a prototype mechanism that learns from data how factors interact in non-trivial ways in scientific models.
- **Data validation and repair.** Our objective is to design and implement a prototype capability to identify likely errors in data based on anomalies relative to historic data and to use models of historic data to offer suggested repairs.
- **System prototyping.** Our objective is to incorporate all improvements into the MAT software application and make the resulting application available to the government and academic research community for use in scientific modeling projects.
- **Evaluation of applicability to multiple scientific domains.** Our objective is to ensure (and demonstrate) that MAT can be applied to a wide range of scientific domains by identifying and building at least one neurological and/or physiological model and analyze the associated data with MAT, making any extensions to the MAT tool that are needed to support the analysis of such a model.

2. Overview of Problem and Technical Approach

2.1. Summary of the Problem

One of the most powerful things scientists can do is to create models that describe the world around us. Models help scientists organize their theories and suggest additional experiments to run. Validated models also help others in more practical applications. For instance, in the hands of military decision makers, human social cultural behavior (HSCB) models can help predict instability and the socio-political effects of missions, whereas models of the human brain and mind can help educators and trainers create curricula that more effectively improve the knowledge, skills, and abilities of their pupils.

While there are various software tools that are used by the scientific community to help them develop and analyze their models (e.g., Excel, R, Simulink, Matlab), they are largely so general in purpose (e.g., Excel, R) or so focused on computational models in particular (e.g., Simulink, Matlab), that they are not ideal for rapid model exploration or for use by non-computational scientists. They also largely ignore the problem of validating the models, especially when the models are positing causal claims as most interesting scientific models do. To address this gap, Charles River Analytics undertook the "Validation Coverage Toolkit for HSCB Models" project with ONR. Under this effort, we successfully designed, implemented, informally evaluated, and deployed a tool called the Model Analyst's Toolkit (MAT), which focused on supporting social scientists to visualize and explore data, develop causal models, and validate those models against available data (Neal Reilly, 2010; Neal Reilly, Pfeffer, Barnett et al., 2011, 2010).

As part of the development of the MAT tool, we identified four important extensions to that research program that would further support the scientific modeling process:

- Correlation analyses are still the standard way of identifying relationships between factors in a model, but correlations are fundamentally flawed as a tool for analyzing potentially causal or predictive relationships as they assume instantaneous effects. Even performing correlation analyses with a temporal offsets between streams of data is insufficient as the temporal gap between the causal or predictive event and the following event may not be the same every time (either because of variability in the system being modeled or because of variability introduced by a fixed sampling rate). What we need is a novel way of evaluating the true predictive power across streams of data that can deal with fluid offsets between changes in one stream of data and follow events in the other stream of data.
- Modeling complex phenomena is a fundamentally difficult task. Human intuition and analysis is by far the most effective way of performing this task, but even humans can be overwhelmed by the complexity of modeling the systems they are studying (e.g., socio-political system, human neurophysiology). Automated tools, while not especially good at generating reasonable scientific hypotheses, *are* extremely good at processing large amounts of data. We believe there is an opportunity for computational systems to enhance human scientific inquiry. Under the "Validation Coverage Toolkit for HSCB Models" project, we demonstrated how automated tools could help human scientists to analyze and validate their models against data. We believe a similar approach can be used to help suggest modifications to the human-built models to make them better match the available data. To be useful, however, such automated analyses will need to be rich enough to suggest subtle data interactions that are most likely to be missed by the human scientist. For instance, correlations (especially correlations that take into account fluid temporal displacements) could be used to identify likely relationships between streams of data, but such an approach would miss complex, non-linear relationships between interrelated factors that cannot be effectively analyzed with

simple two-way correlations. For instance, if crime waves are associated with increases in unemployment *or* drops in the police presence, that would be hard to identify with a correlation analysis. We need richer automated data analysis techniques that can extract complex, non-linear, multi-variable relationships between data if we are to effectively suggest model improvements to human scientists.

- Even if a scientific model is sound, if the data sets provided as inputs to the model are unreliable, the results of the model are still suspect. And, unfortunately, data will often be wrong. For instance, HSCB surveys are notoriously unreliable and biased for a variety of reasons, and neurological and physiological data can be corrupted by broken or improperly used sensors. If it were possible to identify when data was unreliable and, ideally, even repair the data, then the models that are using the data could once again be effectively used.
- The MAT tool we developed under the "Validation Coverage Toolkit for HSCB Models" project was focused primarily on assisting social scientists in the analysis, refinement, and validation of HSCB models. In parallel with that effort, however, we also took an opportunity to apply MAT to evaluating neurological and physiological data under the DARPA-funded CRANIUM (Cognitive Readiness Agents for Neural Imaging and Understanding Models) program. We discovered the generality of the MAT tool makes it potentially applicable to a great number of different scientific domains. MAT proved to be a useful, but peripheral tool, in CRANIUM. We believe MAT could be applied to a broader suite of scientific modeling problems than it has been so far.

2.2. Summary of our Approach

To address these identified gaps and opportunities, we are extending MAT's support for model development, analysis, refinement, and validation; enhancing MAT to analyze and repair data; and demonstrating MATs usefulness in additional scientific modeling domains. Our approach encompasses the following four areas, which correspond to the four gaps/opportunities identified in the previous section:

- **Temporally Fluid Correlation Analysis.** We are designing a new method to perform Temporally Fluid Correlational Analysis on temporal sets of data, and we are implementing the method as a new component within the MAT software application. The version of MAT at the beginning of the new effort supported correlation analysis for temporally offset data; it shifts the two data streams being compared by a fixed offset that is based on the sampling rate of the data (i.e., data that is sampled annually will be shifted by one year at a time), performs a standard correlation on the shifted data, plots the correlation value against the amount of the offset, and then repeats the process for the next offset amount. If two data streams are shifted by a fixed offset (e.g., changes in one stream are always followed by a comparable value in the other stream after a fixed time), then this method will find that offset. Under the current effort, we are expanding on this capability to support fluid temporal shifts within the data streams. That is, we are making it possible to identify when the temporal offset between the change in the first data stream and its effect in the second stream is not a static amount of time.

- **Automated suggestions for model construction and refinement.** We are designing and implementing a mechanism to learn how factors interact in non-trivial ways in scientific models. In particular, we are developing a method for learning disjuncts, conjuncts, and negations. This mechanism starts with the model developed by the scientist user and make recommendations for possible adjustments to make it more complete by performing statistical data mining and machine learning.
- **Data validation and repair.** Recognizing that data contains errors is plausible once we understand the relationships between data sets. That is, if we are able to develop models of the correlations between sets of data, then we can build systems that notice when these correlations do not hold in new data, indicating possible errors in data. For instance, if we know that public sentiment tends to vary similarly between nearby towns, then when one town shows anomalous behavior, we can reasonably suspect problems with the data. There might be local issues that cause the anomaly, but it is, at least, worth noting and bringing to the attention of the user of the data and model. As MAT is designed to help analyze models and recognize inter-data relationships, it is primed to perform exactly this analysis. Existing methods perform similar types of analysis for environmental data (Dereszynski & Dietterich, 2007, 2011). For instance, a broken thermometer can be identified and the data from it even estimated by looking at the temperature readings of nearby thermometers, which will generally be highly correlated.
- **Application to multiple scientific modeling domains.** To ensure (and demonstrate) that MAT can be applied to a wide range of scientific domains, we are identifying and building at least one neurological and/or physiological model and analyzing the associated data with MAT, making any extensions to the MAT tool that are needed to support the analysis of such a model. The initial MAT effort focused on HSCB models; by focusing this effort on harder-science models at much shorter time durations, we believe we can effectively evaluate an interesting range of applications of the MAT tool.

3. Current Activities and Status

During the current reporting period, we made progress on a number of new features. We also received notice that our abstract submission to the American Political Science Association was accepted for presentation at the annual meeting in August.

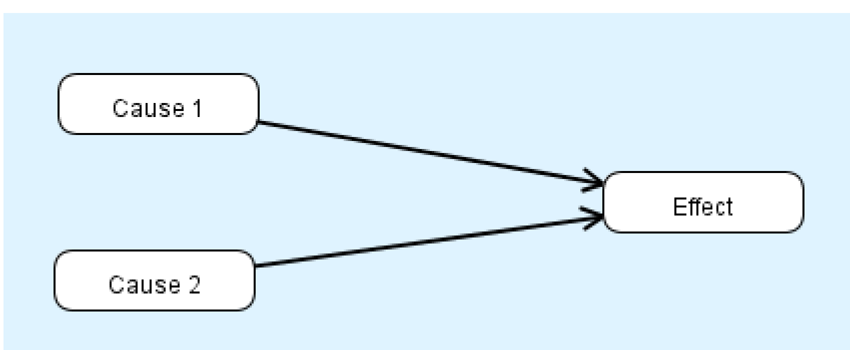
3.1. Comparison of Competing Models

The causal relationship between two concepts is not always clear. Therefore, MAT now allows the user to create multiple causal models for comparison. The following screenshot shows a simple example where the user is unsure whether Concept A causes B or vice versa. Both causal models can be created and validated using the same empirical data. In this example, it is clear that the causal model where Concept B causes A is the better of the two models.

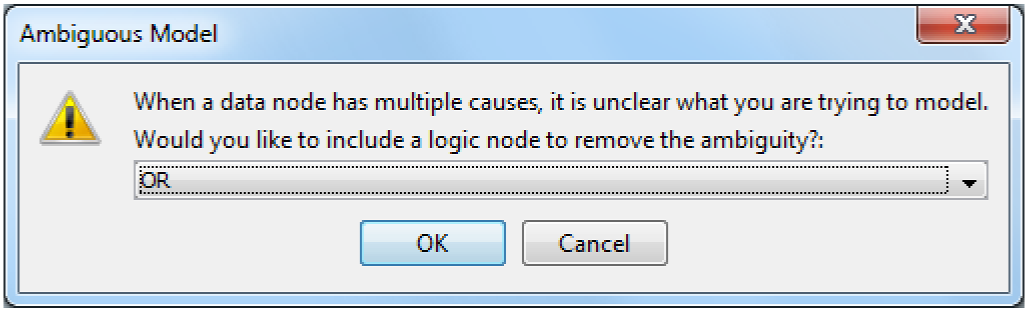
The causal relationship can be further explored by selecting the nodes in the graph to show the valid causal relationships (see following screenshot). An important challenge when including multiple causal models is insuring that it is clear to the user when a node in different models represent the same concept. Therefore, when a concept is selected in one causal model, it also appears as selected in all other causal models that use the concept. The management of concepts in an intuitive way where the user can have both concepts common across models and unique to a single model is still work in progress.

3.2. Handling Ambiguous Causal Models

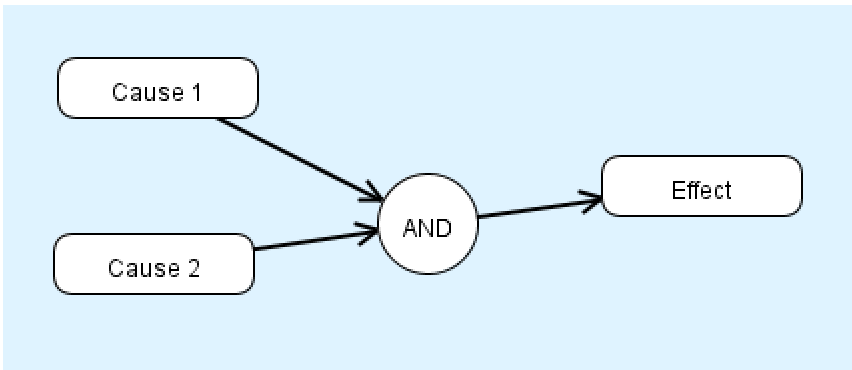
When an effect in the causal model has two possible causes, the user's intentions can be ambiguous. For example, in the following causal model, do both causes need to be present for the effect to occur or is one of the causes good enough?



Therefore, when the user includes a second cause to an effect, the following warning message appears letting the user know of the ambiguity.



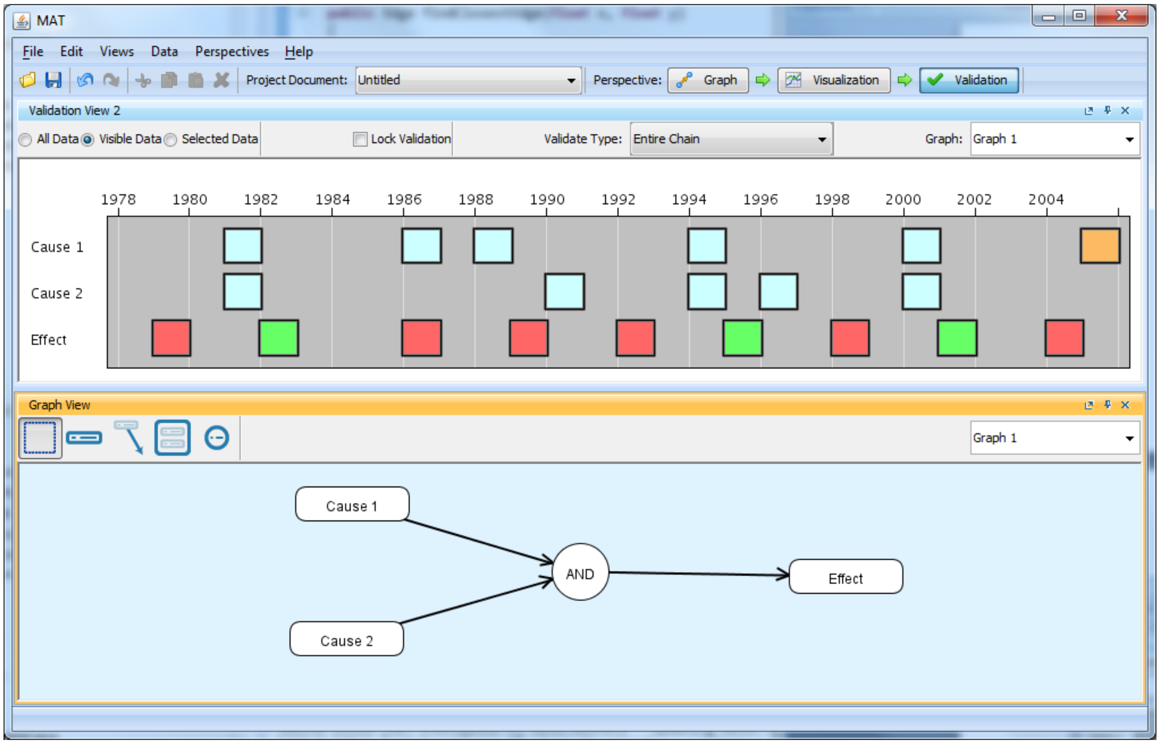
The user can then specify what relationship they meant and the necessary logic will be put in place.



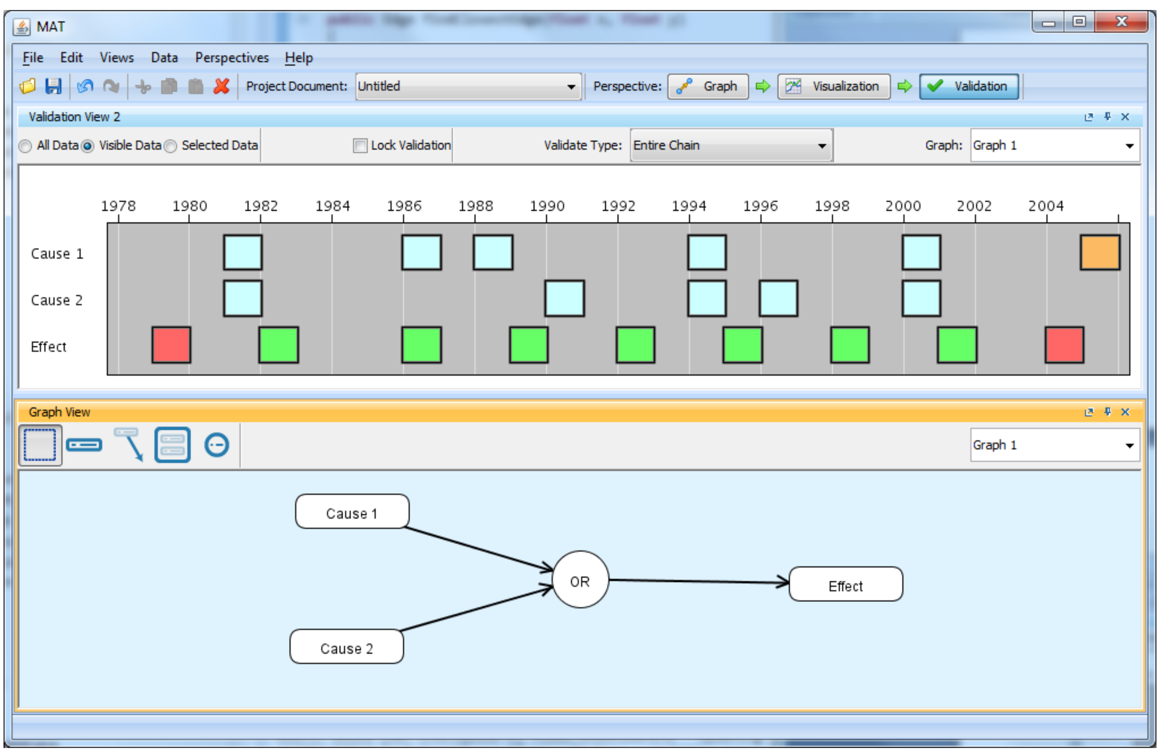
Also, if the model was changed in a way that was not the user's intention, then they can undo the operation and go back to the previous state of the causal model.

3.3. Exploration of Causal Models Using Undo and Redo

To encourage the exploration of causal models, actions performed on the models are undoable. So, continuing with the previous example, the following screenshot shows the results after validating the model with data. In the bottom row of the timeline chart, the green features are supported by evidence, while the red ones are not.



If the user wanted to see how the model changed using an OR logic node instead of an AND logic node, then the user can change the type of node and the validation will automatically update to the following.



The undo and redo operators in the top left of the window allow the user to switch back and forth between the previous and current version of the causal model. The user could continue making adjustments to the model to see how it changes the data validation and then use the undo operator to return to the original version of the model. This provides another way for the user to explore their causal models and make model adjustments to see the influence on validation results.

4. Planned Activities

During the upcoming reporting period, we plan to focus on the following tasks:

- Extension and enhancement of our feature recognition system. One of our novel methods for causal analysis is to identify meaningful events in data streams and then looking for repeated, temporal patterns of such events. This requires being able to identify meaningful events in some reliable way. In our previous status report, we discussed our developing mechanism for identifying events in data; we are now attempting to identify *meaningful* events. We currently use a mechanism from document analysis that is used to identify the meaningful words in a document called TF-IDF (for *text frequency-inverse document frequency*), which points to words that are frequent in a document and infrequent elsewhere. In MAT, we call out events that are similarly representative as meaningful. Unfortunately, this means that uncommon events are missed even when they are extreme, so we are enhancing TF-IDF to also account for the magnitude of events as being indicative of meaningfulness. We anticipate this enhanced functionality will be implemented during the next reporting period.
- Model learning. We already have a simple model learning component in MAT that finds plausible feature-based causal relationships in the available data. During the next reporting period, we will extend the capability to also support learning simple combination models using logical combination operators (e.g., and, or, not). We will also use other causal analysis metrics (e.g., Granger causality) to help constrain/prioritize the search as we are currently using brute-force search methods.
- Releasable demo. We plan to have a new version of MAT that can be delivered to the research community in time for the annual review in June.

5. Evaluation and Transition

We continue to focus on making MAT available to the government and academic research communities and to look for opportunities to use MAT on a variety of ongoing research efforts. Table 1 summarizes our progress in this regard to date. We will continue to update this table as we make additional progress and will include it as a regular part of future status reports.

Program	Customer	Comments
On-going efforts		
Tourniquet Master Trainer (TMT) (Phase I SBIR)	US Army's Telemedicine & Advanced Technology Research Center (TATRC)	MAT is being used to visualize and analyze data from sensors on a medical manikin that indicate whether a number of novel medical devices used to combat junctional and inguinal hemorrhaging are being applied properly. This program is about to begin a Phase II where MAT will continue to be used both by Charles River Analytics and our partners at the University of Wisconsin.
Laparoscopic Surgery Training System (LASTS) (Phase II SBIR)	US Navy's Office of Naval Research (ONR)	Under lasts, Charles River and Caroline Cao at Wright State University are using MAT to analyze data collected from the location of the laproscopic surgery tools during an experiment. Surgical tools are instrumented with markers and 3D data is collected on their location as the person performs the task. This is an ongoing Phase II SBIR program.
Cognitive Readiness Agents for Neural Imaging and Understanding Models (CRANIUM) (Phase I SBIR)	US Navy's Office of Naval Research (ONR)	MAT was used to visualize and extract patterns of stress and workload from neuro-physiological data for training systems. This was a Phase I SBIR program that did not progress to Phase II.

Business Intelligence Visualization for Organizational Understanding, Analysis, and Collaboration (BIVOUAC) Phase II SBIR	US Navy's Space and Naval Warfare Systems Command (SPAWAR)	MAT is being evaluated as part of the BIVOUAC SBIR program, which provides data analysis and visualization for Enterprise Resource Planning (ERP) systems for the Navy. This is an ongoing Phase II SBIR program.
Adaptive toolkit for the Assessment and augmentation of Performance by Teams in Real time (ADAPTER) (Phase I SBIR)	US Air Force Research Lab Human Effectiveness Directorate (AFRL/RH)	MAT is being used to analyze neuro-physiological data from cyber operators to evaluate cognitive workload during team-based cyber operations. This is an ongoing Phase I SBIR program. A Phase II proposal has been submitted and is being reviewed by the Government.
Anticipated Efforts		
Enhancing Intuitive Decision Making Through Implicit Learning (I2BRC) (ONR Basic Research Challenge BAA)	US Navy's Office of Naval Research (ONR) Charles River is a subcontractor to DSCI MESH Solutions, LLC	The intention is to use MAT to help analyze neuro-physiological data to help better understand how implicit learning and intuitive decision making work. This is an ongoing BAA program, though no data has yet been collected to analyze.
A system for augmenting training by Monitoring, Extracting, and Decoding Indicators of Cognitive Load (MEDIC)	US Army's Telemedicine & Advanced Technology Research Center (TATRC)	We are evaluating the practicability of using MAT to analyze and visualize neuro-physiological data from combat medic trainees to identify periods of stress and cognitive overload. This is a recently started SBIR Phase I program in its initial design and exploration phase.

Model Analyst's Toolkit for High-level Fusion (MAT-HF)	US Army Research Laboratory (ARL)	Extend MAT for ARL research objective in high-level information fusion, exploitation, social network analysis and knowledge management research. A BAA white paper submission has been requested and is being written.
Minerva: Decision-making Under Threat: Understanding Factors that Influence Political Preferences in Persistent Conflict	US Navy's Office of Naval Research (ONR) Charles River is a subcontract to Northeastern University	MAT will be used to understand the impact of threat perception on political preferences in societies experiencing persistent conflict, and modeling the way security perceptions are propagated by the media and ultimately influence decision-making.
Intelligent Model Analyst's Toolkit with Machine Learning (iMAT-ML)	NASA Goddard Space Flight Center	We have proposed a data analysis workbench that extends MAT with analysis and machine learning algorithms and an expert system that (1) helps configure the work bench to achieve analysis objectives and (2) learns optimal approaches. A Phase I SBIR proposal has been submitted and is under review.

Table 1. MAT Transition and Use Progress

In addition we have provided copies of MAT to the following institutions based on their requests for the software: the University of Michigan, Arizona State University, Kansas State University, University of California at Los Angeles, the Naval Medical Research Unit at Wright Patterson Air Force Base, Concordia University (Montreal), the University of Wisconsin, and the Air Force Research Laboratory's Human Effectiveness Directorate.

Finally, we submitted a paper abstract on using MAT for data-driven model refinement and validation to the American Political Science Association that has been approved for a presentation at the annual conference in August.

6. Budget and Project Tracking

As of January 31, 2013, we have spent $483,033, or 52% of our total budget of $928,224, in 47% of the scheduled time. Our current funding is $618,457, so we have spent 78% of our available funding.

We believe we are in good shape to complete the project on time and on budget.

7. References

Dereszynski, E. W.& Dietterich, T. G. (2011). Spatiotemporal models for data-anomaly detection in dynamic environmental monitoring campaigns. *ACM Transactions on Sensor Networks (TOSN),* 8.

Dereszynski, E.& Dietterich, T. (2007). Probabilistic models for anomaly detection in remote sensor data streams.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and Control*, 2, 329-352.

Manning, C., Raghavan, R., and Schutze, H. (2008). *Introduction to Information Retrieval.*: Cambridge University Press.

Myers, C.& Rabiner, L. (1981). A level building dynamic time warping algorithm for connected word recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 29, 284-297.

Neal Reilly, S. (2010). *Validation Coverage Toolkit for HSCB Models*. Charles River Analytics Inc.

Neal Reilly, W. S., Pfeffer, A., and Barnett, J. (2010). A Metamodel Description Language for HSCB Modeling. In *Advances in Cross-Cultural Decision Making*.

Neal Reilly, W. S., Pfeffer, A., Barnett, J., Chamberlain, J., and Casstevens, R. (2011). A Computational Toolset for Socio-Cultural Data Exploration, Model Refinement, and Model Validation. In *Human Social Culture Behavior (HSCB) Focus*.

Olszewski, R. (2001). *Generalized feature extraction for structural pattern recognition in time-series data*. Carnegie Mellon University, School of Computer Science.

Salvador, S.& Chan, P. (2007). Toward accurate dynamic time warping in linear time and space. *Intelligent Data Analysis*, 11, 561-580.

Sugihara, G., May, R., Ye, H., Hsieh, C., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338, 496-500.