Service and Agent Oriented Framework for Information Fusion and Knowledge Management

Dr. G. Emami
Global InfoTek, Inc.
1920 Association Drive
Reston, VA 20191 USA

1. INTRODUCTION

In today's military, mountains of data are generated, absorbed, and filed from different sources and presented in diverse ways such as text, audio, video, and images. Network Centric (NC) capabilities turbo-charge data collection and information exchanges result into a "flood" of battlefield data. This data can't be processed fast enough by warfighters to find critical threats and events. As a result, today's network-enabled operations must be augmented with tools to quickly parse vast amounts of data and deliver alerts and critical mission relevant data to key personnel in a clear, concise fashion.

A solution to mining these mountains of data can be achieved by augmenting C4ISR systems and warfighters with technology appliqués called Software Agents. Software Agents (*SA*) work like, and for, humans 24/7, but much faster than is humanly possible. Groups of these Software Agents are the key to unearthing enemy threats with unprecedented timeliness. This timely identification of the enemy is paramount to swift threat neutralization.

Global InfoTek, Inc. has developed MIATA™**,** a service and agent-based system (see Figure 1), and a force multiplier, that enables real-time threat and anomaly detection**.** MIATA's uses "agent-aided" analysis of rapidly changing information that would normally overwhelm users, to provide decision makers with situational awareness. By reducing the workload associated with monitoring massive amounts of rapidly changing data and augmenting the Common Operational Picture (COP), MIATA enables warfighters to focus their attention on critical mission elements.

2. MIATA OVERVIEW

MIATA provides Duty Officers, Chief of Operations, Battle Captains, Commanders, and decision-makers with enhanced situational awareness within their area of interest. MIATA provides automatic threat notification via a variety of visual and messaging mechanisms.

Built upon Software Agent-based capabilities, MIATA provides powerful data ingestion, business rules, data fusion, and anomaly detection in carrying out its multi-faceted and evolving mission. MIATA also supports easy, rapid Battlefield User integration of net-centric systems installed on weapons-system platforms, COTS, Government off the Shelf (GOTS), legacy, and open source applications. This integration enables warfighters to monitor diverse information sources regardless of their origin and the type of data.

MIATA is an Agent-based architecture that is built on the Intelligent Service Layer (ISL™) SOA fabric, Composable

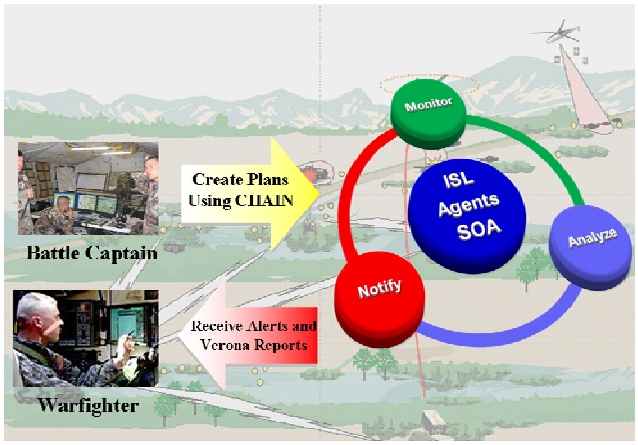

Figure 1 - MIATA Software Agents, on behalf of Users, continuously Monitor and Analyze mountains of battlefield data, and only Notify when a Critical Event is detected.

Heterogeneous Agents for Intelligent Notification (CHAIN) "Pluggable" Framework, and Verona Collaborative Multi-media Knowledge Management Tool.

MIATA allows diverse applications, fusion and reasoning components, including third party software, to interoperate and work together. MIATA also includes a powerful Collaborative Multi-media Knowledge Management Tool, called Verona, which can be used to rapidly organize, present, share, and distribute knowledge in a timely manner.

3. MIATA USERS

MIATA increases the task efficiency of Battle Captain, related to monitoring and detecting threat-related or critical events, which provides an effective Force Multiplier.

MIATA supports two classes of users:

1. Operation Officers, Battle Captains, or Intel Analysts who can easily use the MIATA GUI to rapidly insert new sets of logic, or rules, to satisfy new operational requirements or perform what-if analysis without any assistance from the technical staff. These users can quickly create plans that specify which data is to be accessed, how a diverse set of algorithms is to be invoked to support threat evaluation, and who needs to receive the results or alerts. CHAIN does not require Battle Captains to become experts on the underlying algorithms or implementation. Powerful algorithms are provided as services, or software agents for the Battle Captain to orchestrate or choreograph plans to support their immediate specialized needs.

2. Duty Officers or warfighters who simply receive the alerts, justification report, threat classifications, and scoring results. These users do not interact with the underlying MIATA's GUI. These users utilize their existing system where MIATA notifications appear.

4. MIATA SOFTWARE SUITE

MIATA, an integrated suite of applications that operates as a single tool, is built on ISL, CHAIN and Verona (See Figure 2). One of its principal components CHAIN (see section 4.2) empowers end-users through its easy-to-use graphical user interface, without programming to rapidly craft event and information monitoring, anomaly detection, attention focusing, and other mechanisms, to process and monitor

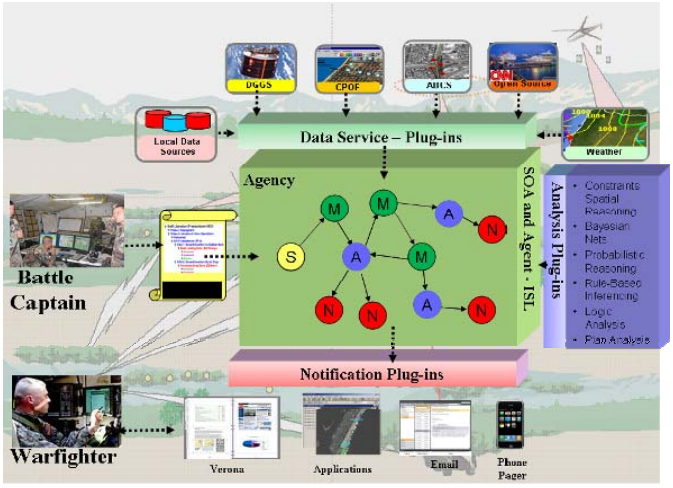

Figure 2 - MIATA is a Plug-in Software Framework for Graphically Building Intelligent Analysis and Notification Plans

massive amounts of data and events in real time. CHAIN agents use ISL's framework (see section 4.1) for agent-to-agent communications, to access diverse information systems, and to interoperate with legacy applications.

MIATA also includes a powerful collaborative multi-media knowledge management tool called Verona that can be used to rapidly organize, present, share, and distribute knowledge in timely manner (see section 4.3). CHAIN agents use Verona to give users consolidated interactive reports that include justification and supporting information for alert responses. This ability saves users critical time that's usually needed for collecting information before responding to an alert.

4.1 Intelligent Service Layer (ISL)

ISL enables agents to communicate and exchange information, access diverse information systems, and interoperate with legacy application using a Service Oriented Architecture (SOA). CHAIN applications as discussed in section 4.2 are built using the ISL framework. While typical SOAs and Web Services have benefits in the enterprise environment, they also have limitations, especially in highly dynamic environments. As a result, ISL was designed to be a lightweight SOA framework to:

1. Extend current enterprise SOA implementations into highly dynamic environments such as battlefields
2. Enable development of intelligent SOA applications that better meet the user's information needs
3. Provide a powerful, light weight, and easy to use integration framework to rapidly incorporate legacy applications into SOA

Used with Web Services, ISL creates an enhanced SOA solution that extends the benefits of enterprise services all the way to users at the battlefield's edge (see Figure 3).

ISL was also designed to address some of the shortcomings of the Web Services in the battlefield. Some drawbacks inherent in the Web Service approach include:

- Reliance on XML; which creates overhead
- Discovery of services is static; services are not automatically registered, discovered, and de-registered
- Web Services are reactive; they run when a message is received, not autonomously
- Web Services are hard for developers to create; developers use vendor tools
- Frequent network disruption reduces the viability of Web Services for mobile computing

Adaptive and Robust SOA

ISL was designed to accommodate challenges associated with highly dynamic communities of cooperating software modules that participants and services would join and leave; with communications services that would be various, intermittent, unreliable or unavailable; and with communities that would form, reorganize and dissipate. ISL meets these challenges with:

- Automatic resource reclamation

- Automatic activation and deactivation of services
- Support for coordinated redundant services
- Store and forward message passing

Intelligent Services for SOA

The ISL contains a software agent framework at its core. Originally developed for DARPA, this agent platform provides increased and unique capabilities to services. Specifically ISL agents provide:

- Autonomous operations for services so they are no longer just reactive
- The ability to work with other agents and services
- Learning capabilities to decide how and when to interact based on network conditions
- Dynamically Deployable Agent Behaviors
  - Automatically distributes service updates
  - Executable behaviors controlled by Java Security Interface
  - Policy-based control allows administration of behavior constraints

Lightweight Framework for SOA

ISL is a lightweight framework that is a perfect compliment to Web Services. Some of ISL's features include:

- Extremely lightweight and does not require the use of XML
- Fully supports XML, including Web Service SOAP messages and WSDL contracts
- Messages can be simple text messages without the overhead found in SOAP message
- Can send binary messages, including Java objects, photos, video, and audio
- Light weight advertisements in ISL provide an efficient alternative to WSDL

The use of ISL results in increased flexibility with reduced computational resources.

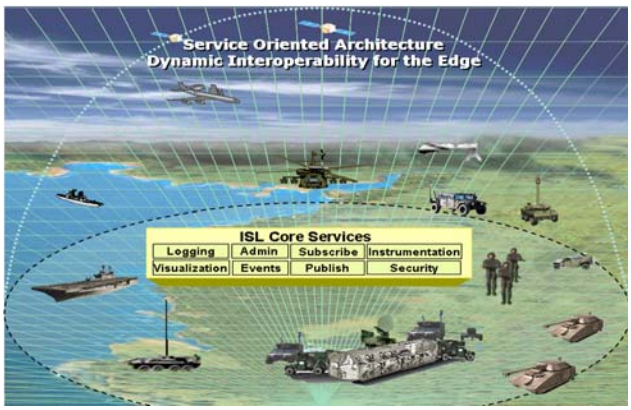

Figure 3 - Intelligent Service Layer and its Supporting Core Services, Dynamic Discovery and Registrations, and Support for Multiple Protocols Extend Current Enterprise SOA implementations in Highly Dynamic Environments, Enable Development of Intelligent SOA Applications that Better Meet the User's Information Needs, and Provide a Powerful, Light weight, and Easy to Use Integration Framework.

Ease of Use

ISL is easy to learn and use. Key attributes include:

- Services are seamlessly integrated with Web Services
- Services automatically provide a WSDL and understand SOAP messages
- Code generation wizards speed creation of services
- Automated tools significantly reduce the time and effort to wrap legacy applications

Secure Interoperability

ISL provides critical capabilities to deal effectively with the demands of distributed security, including:

- Authentication and Authorization
  - Integrates with Java Authentication and Authorization Service (JAAS)
  - Public Key Infrastructure (PKI) compatibility with support of X.509 Certificates
- Confidentiality and Integrity
  - Configurable message queue encryption via SSL; pluggable encryption algorithms
  - Code signing and signature checking integrated with PKI
- Accountability
  - Distributed logging of security relevant events
  - XML encoded event logs suitable for analysis by automated tools

As the foundation for CHAIN applications and associated agents, ISL provides the infrastructure for reliable monitoring and notification of events and automatic notification of the Commander when monitored events (units, equipment, shipments) deviate from their expected norm or projected track.

4.2 Composable Heterogeneous Agents for Intelligent Notifications

CHAIN is a framework for graphically building intelligent analysis or notification plans by end-users and not software developers that monitor data-sources, perform pre-defined analysis in response to events, and report critical events across disparate information sources. CHAIN enables analysts to rapidly assemble and task teams of agents to detect and respond to opportunities and anomalies based on a set of user-definable, agent-executable conditions.

CHAIN is more than a clever acronym. Each word in the name describes a powerful and important concept.

- **Composable:** CHAIN's user-composable software agents assist users to plan, assemble, and execute a task or set of tasks. Users identify which data sources to monitor, the event triggers and the analytical conditions that drive them, and, what notification actions to take based on resultant events.
- **Heterogeneous:** CHAIN's SOA integration framework has been tested extensively by the US Navy and enables run-time integration of distributed, heterogeneous information sources, sensors, web services, agents, and legacy systems.

- **Agents:** CHAIN's framework leverages the distributed nature of agent communication within the ISL framework. CHAIN Plans are able to communicate with agents distributed across multiple systems and platforms to increase system scalability and redundancy.

- **Intelligent:** CHAIN's out-of-the-box agent-driven logic can be easily extended to include COTS, GOTS, legacy, and open source analysis engines that support anomaly detection, rule-based reasoning, inferencing, and heuristics.

- **Notification:** CHAIN alerts and notifies users when reportable events occur. Notification agents help tailor the information to the user by identifying the who, what, when, and how of reporting through a variety of reporting devices such as email, instant message, telephone, pager, PDA, and digital notebooks.

CHAIN gives an organization a 365x24x7 solution to automatically monitor, analyze, and notify decision makers of opportunities and threats. It provides an effective solution for situational awareness of massive amounts of continuously changing data and information. It offers users the means to perform event correlations, and rapid analysis when complex and seemingly unrelated events occur. CHAIN provides an easy graphical interface for rapidly changing and visually documenting these analysis plans.

These alerts are delivered in the user's environment (Map, Email, Messaging, and Portal). CHAIN also provides a human-in-the-loop solution for agile data fusion where operations officers or analysts can invoke and experiment with alternative strategies of fusing data sources using a drag-and-drop composability tool.

CHAIN's open, extensible architecture enables the development of intelligent and automated fusion capabilities that help Watch Standers to manage information by exception and avoid receiving false or repeated alerts in response to events. CHAIN system uses a plug-in architecture and enables incorporation of any third party reasoning, data mining, and rule based engines.

CHAIN's capabilities includes over 50 components consisting of geospatial analysis, data mining, pattern recognition, rule based inference, and neural network that support anomaly and threat detection. CHAIN includes:

- A suite of software agents that access and monitor diverse information sources such as any relational databases, RSS news feeds, emails, web pages, Google Search, automatic identification systems (AIS) data, and C2 mission specific such as GCCS track database.

- A collection of agents for Pattern Recognition, Behavior Prediction, and Threat Evaluation. These agents provide data mining, machine learning, expert systems, classification, regression, clustering, and association rules based on open source algorithms.

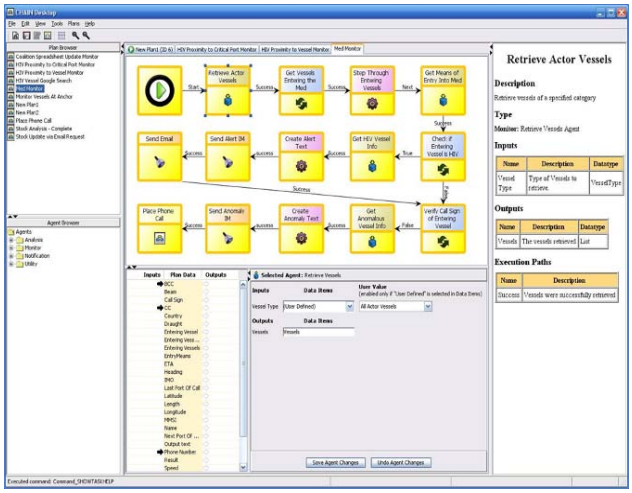

Figure 4 - CHAIN User Interface Allows Users to Graphically Chain Together Agents and Services to Monitor Diverse Data Sources and Information Systems.

These analytical capabilities are offered as software agents or as services. We provide algorithms such as: IBk- the k-nearest neighbor learner, C4.5 decision trees, Naive Bayes, Holte's 1R, support vector machines, LVQ, and Immunos-81. Our solution can embed either commercial (e.g. ILog) or open source (e.g. JESS) rule engines.

- Geospatial Reasoners for C2 Domain - Our solution includes a number of Geographic Information System (GIS) software agents that support navigational reasoning such as: closest point of approach, distance to a land mass, distance between vessels separated by land mass, and other GIS analytical functions.

- Knowledge Discovery, Predictive Analysis, Entity Extraction and Entity Resolution Agents - provides the Intelligence Community with technology, services, and solutions for Knowledge Discovery, Predictive Analysis, Entity Extraction and Entity Resolution, and a SOA framework for agile Analytic Flow.

- A collection of alerting agents that notify users using phone, email, user alerts to the portal, multi-media briefing books, and conversion of track information into OTH-Gold format messages for transmission to GCCS, and user in the-loop GUI to manually approve updating of GCCS based on the received OTH-Gold messages.

4.3 Verona

A third principal component of MIATA, Verona™, is a collaborative knowledge management tool that is used to rapidly organize, present, share, and distribute knowledge. Verona allows the users to establish a collaborative environment to rapidly bring together geographically dispersed teams to review the situation and collaborate on potential course of action.

Verona provides a powerful, flexible, scalable, user-friendly digital notebook (see Figure 5). Within a notebook, users can further organize their information into separate chapters. Each chapter can have pages and each page can contain many living interactive objects that are created or discovered by users. It allows users to easily customize and categorize their information for maximum impact. Analyst can quickly access diverse information from multiple distributed sources and manage them all within Verona notebook. Objects dropped into Verona are displayed on a page in content-specific ways and are LIVE and interactive! This means each of the objects can be activated and manipulated using the application that was originally used to create it.

Verona can be used to:

- Collect and assemble digital objects into a common framework
- Annotate objects with text or voice annotations
- Present the contents of those objects in a coherent and simple manner
- Capture and store metadata and links associated with the source material
- Share investigative material in a collaborative environment.

Verona uses commercially-available open-standards technology to create a powerful, scalable, flexible digital container to capture, store, retrieve, and display information and perform collaboration, and dissemination of digital objects.

Verona uses a content-centric model and a familiar notebook-like metaphor. This approach provides a seamless information space that spans the user's desktop, network file systems and repositories, and the Internet. It enables vast amounts of information to be harvested using existing tools, and organized as objects in a notebook which can be manipulated by users using their familiar toolsets. Verona uses the local file systems to manage notebooks that enable users in the field equipped with PDAs and laptops to collect evidence and populate the notebook. Verona can also use local or remote servers for managing its content for sharing notebooks within a workgroup.

Verona is unique in its focus of organizing both static and live information visually in a way that aligns with the way users thinks of information. It is designed to support almost any form of information through an extensible, component based architecture. This design allows it to act as a single container that bridges the users desktop, repositories, and the Internet (or intranet). This seamless integration, combined with a rich

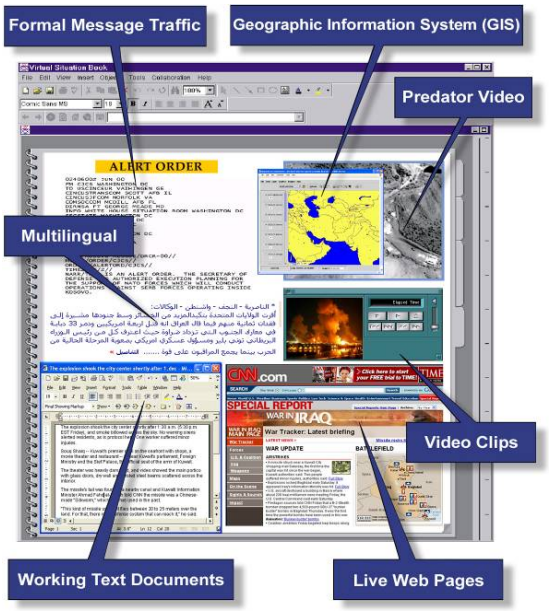

Figure 5 – Verona Provides a Powerful, Flexible, Scalable, User-friendly Digital Notebook. Within a Notebook, Users Can Further Organize Their Information into Separate Chapters.

collaborative metaphor that goes beyond sharing a single view, distinguishes Verona from other tools for information management. Verona provides the ability to easily capture, store, display, search, retrieve, and collaborate on a nearly limitless variety of digital assets. These assets include text, graphics, audio and video files, web pages, Microsoft Office products and web cams. Files are added to the notebook by a simple drag & drop, whereupon they are inserted as either an embedded object or a linked reference to objects in a shared, collaborative data store. Verona works on any Windows-based operating system, including desktops, laptops, and tablet-computers.

Verona embodies a knowledge-centric model of information management. It provides a mechanism for a user to collect, organize, and interact with the content of information the way he or she thinks about a problem, rather than having to work with disjoint tools needed to manipulate information. This approach enables the thematic organization of information in a way that aids and supports the work processes and cognitive models associated with particular tasks, processes, and individuals.

CHAIN agents are able to automatically create a Verona Notebook and populate it with all the supporting multimedia information that is needed for providing justification for the

alerts (see Figure 6). This capability enables users to have, in one place, all the supporting documents and information that are needed to respond to an alert.

5.0 MIATA DEPLOYMENT

MIATA has been successfully deployed to variety of applications and have undergone numerous enhancements and formal evaluation by the US Government. For example in a formal evaluation by the Second Fleet, the final experiment report indicates that MIATA provided a highly significant increase in task efficiency for Operations and Duty, and a six-fold increase in the number of "critical events" that a Duty Officer can effectively monitor. The following statements are sample excerpts from the final report:

- The "capability has demonstrated potential to significantly increase knowledge management efficiency (both time and accuracy) across several mission areas in several data-heavy departments. For example, "critical events" reporting task that usually took about 3-4 hours could now be done in approximately 30-45 minutes."

- "I found the user interface to be fairly intuitive and friendly", and "At first some training is needed but is easy to understand. There isn't much to learn before using this system, it's actually pretty straightforward."

- "I liked creating the CHAIN to do what you want without writing code, it's like programming without being a programmer."

- "Starting from day two of the LOE, there was sufficient confidence in the integrity of the [critical event] output on the part of one [duty officer] to warrant employing it. CHAIN system was used for over 2 years by the Sixth Fleet to monitor High Interest Vessels within their Area of Responsibility."

MIATA has been successfully deployed in the Intelligence Community to enable Intel analysts to create analytical workflow. MIATA is being used to process massive amount of unstructured data using third party analytical tools. The initial phase of the project was successfully completed two months ahead of its schedule. Due to its success it is being used in a number of additional applications.

United States Transportation Command (USTC), is developing an application using CHAIN to integrate disparate supply and transportation data sources and to create an event monitoring system for use by USTC operations personnel and customers.

MIATA is being considered for a number of applications in Cyber Security arena. MIATA agents can monitor massive amount of rapidly changing information for intrusion detections, denial of service attacks, or insider threat detection.

6.0 SOME KEY MIATA BENEFITS

MIATA is the first of its kind toolset that empowers users, with limited technical support, to dynamically compose new capabilities for information capture, assessment and rule based alerting in response to changing threats and operational requirements. For example, users can orchestrate any available tools to execute on-the-fly OPSCENTER workflows in a highly scalable fashion, without any programming. MIATA also offers the ability to work with "any data using any tools at any time." MIATA empowers operation officers, commanders, and analyst, with minimal technical support: to rapidly craft analytic processes centered on anomaly detection, attention focusing, and other mechanisms to manage massive all source data. Without any programming, analysts and planner can orchestrate any available tools to execute on-the-fly analytic flows in a highly scalable fashion.

MIATA has already proven its utility in deployment that provided users with improved situational awareness through automate monitoring and notification processes for critical events and anomalies.

7.0 CONCLUSIONS

The Mission Intelligence Analysis and Threat Awareness (MIATA*)* capability, the result of over 10 years of development, enables user-composable, real-time threat and anomaly detection, and, automatic user-friendly threat notification. MIATA accomplishes this by real-time processing mountains of battlefield data using a pluggable framework that includes the ISL, CHAIN, and Verona tools, and "Software Agents".

MIATA manages all battlefield data because it has been designed to extend the power of the SOA enterprise to the edge using the powerful lightweight ISL. As a result, MIATA augments the COP, reduces the workload associated with monitoring massive amounts of rapidly changing data, reduces all sources of risk, and best of all, increases efficiencies related to monitoring and detecting threat-related events.

MIATA is the first of its kind toolset that empowers users to dynamically compose new capabilities in the field, without any programming. This user-friendly approach is built into MIATA with the combination of the ISL, CHAIN and Verona tools.

MIATA is at a high technical readiness level. It has already proven its utility in deployment by providing Navy, Intelligence Community, and United States Transportation Command users with improved situational awareness through automate monitoring and notification process for critical events and anomalies.