

Intelligent C2 Information Technology

Craig S. Anken

Air Force Research Laboratory

Information Directorate

525 Brooks Road

Rome, NY 13441

(315) 330-4833

ankenc@rl.af.mil

Abstract

Present day and future joint warfare will increasingly demand rapid processing and transfer of data for Command and Control (C2). Improved capabilities to take this data and effectively utilize it for decisions is critical. Commanders need to obtain and use timely battle space information to make informed decisions. While capabilities for Air Force C2 has seen some technology improvements in recent years, there still remains a considerable number of legacy systems designed towards independent commands and missions. The C2 system can still be characterized as data-rich, but information poor. A better way of integrating operations, and providing consistent, tactically relevant, accurate, and timely information across commands and services is needed. The real challenge is not just having access to diverse data sources, but in being able to effectively process, analyze and interpret this data to form information, and ultimately knowledge for use in the C2 decision process. Intelligent C2 information technology will enable the Air Force to manage and control its information flows in a manner that is faster, more efficient, and more precise than that of our adversaries. The Research and Development community is currently working a number of technology areas which will have an impact on C2 as we know it today. This paper will present an overview of some of these initiatives. It will look at some of the technology products coming out of these programs and identify areas where current and future intelligent information technology could be utilized within the C2 decision making process.

1.0 Introduction

The process of going from raw data to decisions requires a number of important and increasingly difficult steps. Raw data is continually being collected via a wide variety of sources (e.g., antenna, eyes, radar, satellites etc.). This data consists of bits and bytes that have not been processed, correlated or interpreted in any way. Once processed this raw data ultimately ends up in a structured or unstructured form that people can understand. This paper is not concerned with the processes and technologies necessary to get the raw data to this form. More intelligent ways of doing this are also being investigated and will only help enhance the capabilities discussed here. Thus, for this paper, the term data will refer to the products resulting after this initial transformation has taken place. Figure 1 below helps illustrate the primary steps required to take data and turn it into useful information and knowledge for decision making. The diagram is not intended to represent all the products and technologies needed for this transformation, but to give the reader a notional idea of the concepts involved.

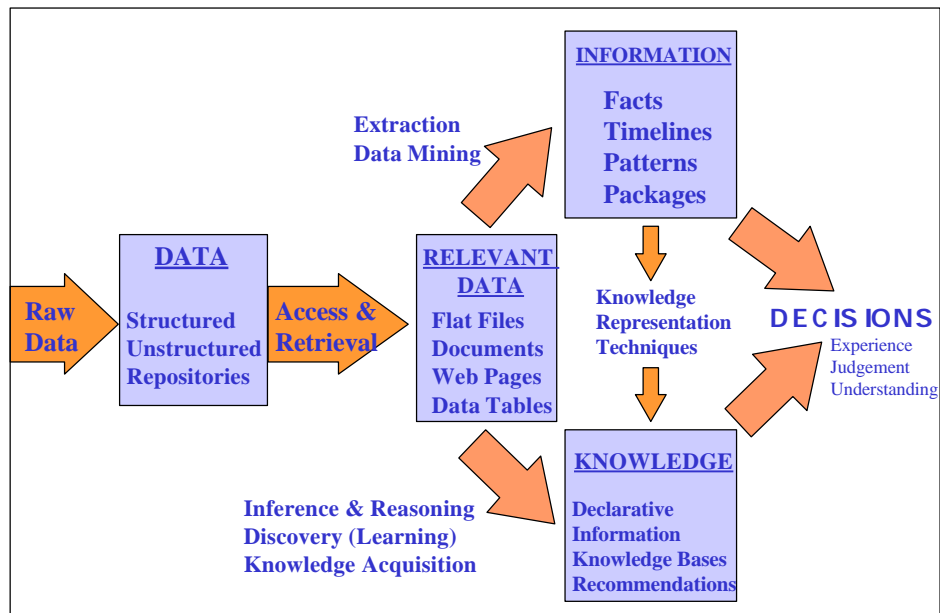


Figure 1: Data to Decisions

2.0 Data to Relevant Data

Anyone who has ever used a search engine on the internet (e.g., Yahoo!, Lycos, Excite) knows the frustrations that can be encountered with trying to find data relevant to their needs. Add to this additional data sources such as imagery products, real audio/video, relational databases, metadata, free text, e-mail, and open source information and the number of data items is virtually infinite. Intelligent access and retrieval is needed to reduce the “number of hits” and help filter the data items relevant to the task at hand. Both the commercial world and DoD are working this problem and a number of capabilities are available for this task. While the objective of this paper is not to discuss this area of research, it is worth mentioning one example of a tool in use within the Air Force today. This tool, being extensively used in the Intelligence community and beginning to gain acceptance in the C2 community is called Broadsword/C2link.

Project Broadsword (and C2Link) provides a method for access to disparate data sources. It allows a user to gather intelligence and C2 information from distributed data sources in support of a particular mission. The system is designed to assist the decision making process by allowing users to query their data-space and returning the relevant data needed to do their job. The main piece of the software is a middleware component called the gatekeeper. The gatekeeper provides brokering services, security auditing, data source communications, and integration of result sets. Data sources are accessed through Data Interface Agents called “plug-ins,” which translate Backus-Naur Form queries into the native query language of the data source. While there are other capabilities comparable to Broadsword/C2link, there is currently nothing that is secure, has a delivery system, deals with free text, and can reach into both the intelligence and C2 world at the same time. [AFRL, 1998]

3.0 Relevant Data to Information

Information is relevant data which has been interpreted and integrated and is now useful in the decision making process. In many AF operations today, much of this task is done manually. We are, however, starting to see some improvement as a number of new capabilities are beginning to mature and be introduced to the operational community. Once such program that is addressing this problem is the Intelligent Integration of Information (I3) program.¹ The goal of the I3 program is to provide more intelligent methods and tools to combine information from autonomous and heterogeneous information resources. As a result, I3 technology will significantly reduce the time and complexity involved in sharing and integrating heterogeneous information.

There are two driving assumptions behind the I3 program. The first is that while the web is very large and contains a lot of relevant data, it is not very useful in its current form. The second is that data sources are many times semantically inter-related. Unfortunately, this relation is not always obvious and done in planned ways. Because of this, the process of integrating multiple information sources can be very complex. Specific problems include the inherent differences in data source construction including schemas, vocabularies and ontologies. While many times there is considerable overlap in the underlying semantics, these differences make integration very difficult to take place. Also, organization of the data can be in completely different formats (e.g., relational, files) and may also be encoded differently such as images or graphics. Methods of accessing and interfacing to the data as well as data granularity often also vary greatly resulting in it being difficult to deduce the intended semantics.

The primary I3 “services” being developed to address these problems include Wrapping, Semantic Integration and Transformation, Coordination and Management, and Functional Extensions. These services are necessary to make it easier to do a number of tasks such as, “wrap” legacy data sources; choose from a large number of available data sources to satisfy an information need; select, coordinate, and manage applications and information; and transform the retrieved data into information that is useful to the decision maker.

Wrappers are used to make data sources comply with standards dealing with their communication, behavior, or to restructure data. To do this they transform the interface or output of the data source or modify the semantics by creating new internal interfaces. **Semantic Integration and Transformation** services support the semantic manipulations needed when integrating and transforming data such as schema integration, data integration, and support for process and physical integration. **Coordination and Management Services** provide support for tasks such as locating the data sources potentially relevant to the task, or helping describe the services, information sources, and steps that must be performed to complete the task. Primary coordination services include resource discovery, brokering, and facilitation. **Functional Extension** services augment functionalities of other I3 Services. They provide the bridge to begin thinking about the transformation from information to knowledge by enriching the semantic

¹The I3 program was initiated by DARPA in the mid 1990's. AFRL participated as both a co-funder and technical agent for many of the efforts. Efforts to transition this technology into Air Force C2 systems is currently ongoing.

expressiveness of information. This is accomplished through capabilities such as activeness, inference, temporal, object-orientation, and persistence services. [Arens, Hull and King 1995]

3.1 Applications

A number of applications have utilized I3 technology since the beginning of the program. Not all of them will be discussed here, but it is worth mentioning a few representative examples. Specific names of companies and individuals involved have purposely been left out.

Research funded under I*3 led to a company that developed commercial implementations of their algorithms for learning which documents users are interested in. The system would learn from observing the text documents read and liked or disliked. This capability was combined with a filtering capability to produce a site that recommends web pages, news articles, etc.

Software was developed using I3 developed mediators and wrappers to access C4I information. The prototype software demonstrated the use of mediators to optimize users query in a World Wide Web or DoD's Intelink environment. I*3 mediators were used to access filtering information from diverse information repositories to automatically narrow the search to a specific geographic region of interest.

A capability for automated delivery of diverse C4I information based on operational requirements and user preference profiles was also developed. The prototype was developed in response to DARPA's Battlefield Awareness and Data Dissemination (BADD) project. I3 wrappers were developed for "forces" representing GCCS data and "targets" representing Air Force's CTAPS data. Flexible WWW wrappers were also developed designed to automatically extract weather information from WWW home pages that publish the world major cities' temperatures. A robust CORBA Query Server that receives requests from a client application and uses the appropriate wrapper to automatically query Forces and Targets databases was also implemented.

Technology developed under I*3 was licensed, in conjunction with several other technologies, to three companies who applied the technology to the development of KBs describing biological organisms. It is also being used in a project developed for a pharmaceutical company. The purpose is to assist in the discovery of new pharmaceutical compounds and aid research into competitors' capabilities. It helps provide a workspace where the user can organize tasks to be accomplished, issue queries against diverse data sources, and collect, organize, and display findings.

The Warfighter's Information Packager (WIP) utilizes I3 technology to support heterogeneous data access, customized information delivery and user-defined information packaging. By providing a tool set that establishes an infrastructure of information objects the WIP system provides a mechanism that allow users to express their information needs based on high-level domain modeling. [Arens et al., 1998]

I3 technology is also playing a role in a number of other technology and application programs such as DARPA's Genoa, HPKB, & JFACC programs. It is also being looked at for use by the Air Force C2 community and USTRANSCOM to name a few.

4.0 Information to Knowledge

Knowledge helps us to anticipate events and recognize in advance new or impending developments. It is used to gain a deeper level of awareness of a specific situation through an understanding of what is going on. When combined with a commanders intuition and judgement, it forms the basis for determination of the best possible solutions to the problem at hand. Associated with the creation of knowledge is the need for efficient techniques to reason over declarative information (as opposed to procedural information). Large bodies of declarative information are called knowledge bases (KBs). The combination of these KBs and inference results in knowledge-based systems.

The potential of utilizing KBs to go from information to knowledge is much more realistic today than in the past. In the late 1970s and '80s proponents of AI/KB technology began developing rule-based programs. These systems took a considerable amount of time to put together, didn't work well, were very inefficient, and usually could not answer all the questions. The "brute force" paradigm of problem solving was clearly the prevailing theme. It was around this time that the popularity of "neural nets" began to take hold. But, while useful for some problems, neural nets could not be used to help predict the unforeseen because of insufficient training data. As a result of these limitations, KB technology was seen as a tool with limited use. [Helm, 1996]

It wasn't until the late 80's with the discovery that Bayesian networks² could provide an efficient way to deal with missing information and ambiguity that the promise of KB systems began to be resurrected. Since this time, we've reached a point where the power gained from applying knowledge based technology is now much greater. Techniques for building knowledge bases have advanced significantly. Add to this the increased processing power now available on even desktop PC's as well as advances in how to perform efficient inference and the potential payoff for applying this technology has increased significantly.

The High Performance Knowledge Base (HPKB) program, aimed at enhancing and applying this potential to DoD problems, was initiated in 1996.³ The objective of HPKB is to produce the technology needed to more effectively take the step from information to knowledge. The goal is to enable KB system developers to rapidly construct large knowledge-bases that provide comprehensive coverage of topics of interest, are reusable by multiple applications problem, and are maintainable in rapidly changing environments. The program is centered around three major

²Bayesian networks are complex diagrams that organize the body of knowledge in any given area by mapping out cause-and-effect relationships among key variables and encoding them with numbers that represent the extent to which one variable is likely to affect another. Programmed into computers, these systems can automatically generate

optimal predictions or decisions even when key pieces of information are missing.

³ HPKB is funded by the Defense Advanced Research Projects Agency (DARPA). AFRL is the technical agent for a majority of the program and is actively working to transition results to the Air Force operational commands.

steps associated with the construction and use of large knowledge bases. These steps are building the foundation knowledge, acquiring domain specific knowledge, and applying application-specific problem solvers.

The first step is building the foundation knowledge by selecting, composing, extending, and specializing components from a library of reusable ontologies and common domain theories. Capabilities being developed here include tools to edit, extend, and enable easy browsing, visualization, and understanding of large ontologies. This also includes tools to import data dictionaries, data schemas, lexicons, and other source information to create initial ontologies and skeleton theories for later elaboration. An “upper ontology” has been developed composed of foundation theories for space, time, actions and processes, belief and uncertainty, simple physical behavior, and other commonly used concepts. In order to complete the second step of Acquiring Domain Knowledge, acquisition technology is being developed to help populate a complete knowledge base. Capabilities being developed include information extraction tools to automatically discover and extract new information to be entered in the knowledge base. This extraction is done from a variety of text documents such as manuals, textbooks, dictionaries, reports, web pages, and messages. Machine learning and knowledge discovery tools are also being developed to automatically and intelligently assist humans in finding useful facts from data and information. The final step in building a KB system is to build application specific problem solving modules. These modules need to be tailored to the unique requirements of a particular application. This portion of the program is being accomplished through a challenge problem-based evaluation system that is driving development. Challenge problems are currently aimed at dealing with Crisis Management and Battlespace Understanding functions. [Gunning, 1996]

Overall the first year of HPKB the program was very successful in demonstrating the enormous reasoning power and decision support capabilities that KB technology has to offer. A number of products and publications further describing this work is available.

4.1 Ontologies

One important area of research that deserves further discussion is that of ontology⁴ development. While ontology development tools are usually grouped under the KB technology area, advances in this area can have a significant impact at every step of the information food chain. Because of this, a coordinated effort was needed to "harmonize" the numerous ontology efforts. HPKB has, to some degree accomplished this, but work is still needed. An important outcome of HPKB has been the development of ontologies and ontology tools that can be used to support the entire data transformation process.

At the data level, ontologies will help provide the semantic foundation needed to advance beyond the current keyword-based retrieval and indexing techniques by allowing documents to be referenced by a description embedded in the link. With ontologies to “guide the way”, the path

⁴ An ontology is a specification of a set of domain concepts to be used for expressing information and knowledge in a given domain. It may include types of entities (objects), attributes and properties, relations and functions, and constraints. The purpose of an ontology is to provide thorough definitions for the concepts referenced in a domain model.

from relevant data to information to knowledge will be much less confusing. Extendable ontologies and ontology development tools, available at all levels of data transformation, will provide developers with a foundation towards achieving the desired interoperation of components.

4.2 Applications

As mentioned HPKB has been working problems associated with crisis management and battlespace understanding. Some of these include estimating plausible national/strategic courses of actions for an hypothesized crisis, determining possible enemy work arounds to inflicted battle damage, and identifying functionally significant patterns of movements for vehicles, equipment and military units.

Some other potential applications of HPKB technology being looked at include applying a knowledge base/problem solving capability to enhance air operations planning and targeting against theater ballistic missiles, target identification & selection, intrusion detection and information analysis. All of the above share a common requirement of battlefield knowledge: for hypothesizing the composition and location of enemy forces, for reasoning about possible troop movements, for monitoring the execution of operational and logistics plans, or for identifying the most critical targets to accomplish an objective.

Technology from this program is also supporting a number of other programs. Some of these include Command Post of the Future (CpoF), Control of Agent Based Systems (CoABS), GENOA, and Information Assurance (IA).

5.0 References

[AFRL, 1998] Air Force Research Laboratory. *Project Broadsword Functional Document*. Prepared for: 497 th Intelligence Group / Intelligence Systems Directorate (497IG/IND) Bolling Air Force Base, Washington DC, March 1998.

[Arens, Hull, King 1995] Yigal Arens, Richard Hull, and Roger King. *REFERENCE ARCHITECTURE for the Intelligent Integration of Information*. VERSION 2.0 (Draft) August 22, 1995.

[Arens, Lee, Zev 1998] Yigal Arens, Yongwon Lee, Marc Zev, Weixiong Zhang, Jon Dukes-Schlossberg. Warfighter's Information Packager. American Association for Artificial Intelligence, 1998.

[Gunning, 1996] Dave Gunning. *High Performance Knowledge Bases Program Description and Information Package*. Unpublished, 1996.

[Helm, 1996] Leslie Helm. *Improbable Inspiration*. Los Angeles Times, <http://www.latimes.com>, October 28, 1996.