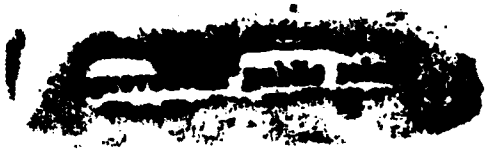
RAND

Database Technology Activities and Assessment for Defense Modeling and Simulation Office (DMSO) (August 1991–November 1992)

A Documented Briefing

Iris M. Kameny

National Defense Research Institute



The research described in this report was sponsored by the Under Secretary of Defense for Acquisition. The research was conducted in the National Defense Research Institute, RAND's federally funded research and development center supported by the Office of the Secretary of Defense and the Joint Staff, Contract No. MDA903-90-C-0004.

Published 1994 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
To obtain information about RAND studies or to order documents,
call Distribution Services, (310) 451-7002

RAND

Database Technology Activities and Assessment for Defense Modeling and Simulation Office (DMSO) (August 1991–November 1992)

A Documented Briefing

Iris M. Kameny

Prepared for the
Under Secretary of Defense for Acquisition

National Defense Research Institute

PREFACE

This report consolidates the findings of the Defense Modeling and Simulation Office (DMSO) Data Base Technology Working Group (DBTWG) and the Information/Data Base (I/DB) Task Group over the period August 1991 to November 1992 in supporting DMSO in promoting the interoperability, sharing, and reuse of databases and models throughout the Department of Defense (DoD) Modeling and Simulation (M&S) community. It is based on a draft written in August 1992 to accompany a briefing on Database Technology Assessment for Modeling and Simulation given to the Defense Science Board Summer Study on 11 August 1992. Appendices containing the agenda and notes of four I/DB Task Group meetings are included.

The work described here was performed for the Defense Modeling and Simulation Office as part of its initiative to strengthen the use of simulation and modeling throughout DoD. RAND's participation in this effort was accomplished for the Under Secretary of Defense for Acquisition, within the Applied Science and Technology Program of RAND's National Defense Research Institute (NDRI), a federally funded research and development center sponsored by the Office of the Secretary of Defense and the Joint Staff.

This work should be of interest to those working in the areas of interoperability of information systems, information resource management (IRM), data dictionary systems, resource directories, data modeling methodologies and tools, data administration, and assessment of data management technology.

CONTENTS

Preface iii
Summary vii

Introduction 1
Goals and Scope 8
Perspectives on the Current State of Database Technology 12
Initiatives Identified by the DMSO Database TWG 28
Status and Discussion of Current CIM Efforts 43
Summary 58

Appendix

A. List of Acronyms 61
B. Document References of I/DB Task Group 65
C. Information/Data Base Task Group Members 72
D. Notes From the 1st I/DB Workshop, February 20, 1992 75
E. Notes From the 2nd I/DB Workshop, March 30–31, 1992 81
F. Notes From the 3rd I/DB Workshop, June 4–5, 1992 100
G. Notes From the 4th I/DB Workshop, October 7–9, 1992 124

SUMMARY

This report presents work that was performed for the Defense Modeling and Simulation Office (DMSO) from August 1991 to November 1992. The first part, on database technology assessment, was performed by the DMSO Data Base Technology Working Group (DBTWG), composed largely of representatives from a number of federally funded research and development centers (FFRDCs) and representatives from the Navy, Army, Defense Information Systems Agency (DISA), National Institute of Standards and Technology (NIST), and George Mason University. The purpose of the initial effort was to provide FFRDC support for the development of the DMSO Master Plan. An earlier document, "Report of the Data Base Technology Working Group (DBTWG)," August-November 1991, summarizes the work resulting from a series of meetings and activities that took place August through November of 1991.

As a result of the assessment, appropriate members of the Information Working Group and the DBTWG came together to form the Information/Data Base (I/DB) Task Group, which since January 1992 has been focused on designing a prototype DMSO Information System and addressing issues affecting the interoperability, sharing, and reuse of databases and models throughout the Modeling and Simulation (M&S) community. At the same time, the DoD Corporate Information Management (CIM) Initiative is addressing the data needs of the DoD community. One of the objectives of the I/DB Task Group is to recognize data needs of the M&S community not likely to be addressed by CIM, other DoD, or commercial endeavors in the near term. These needs are candidate areas for DMSO R&D investment that in turn would contribute to long-term CIM objectives. It is also critical for the I/DB Task Group to continue to monitor CIM activities and help DMSO develop compatible M&S guidelines and procedures whenever possible while pointing out possible incompatibilities with CIM. At the same time, the I/DB will continue to support the development of the DMSO Information System so that the M&S community can share information about M&S happenings, projects, databases, models and simulations, organizations, and so forth.

The R&D data areas critical to the M&S community that are not being addressed by other organizations or agencies include:

- Developing an understanding, methodology, and standardization for complex data elements (e.g., rules, objects, networks, images, voice, matrices, etc.);

- Developing data Verification, Validation, and Certification (VV&C) definitions, methodology, management procedures, and guidelines in coordination with the M&S community Verification, Validation and Accreditation (VV&A) needs;

- Developing advanced methods for classifying, locating, and accessing information in distributed Directories, Dictionaries, & Repositories (DD&Rs);

- Developing methods for standardizing domain values, icons, and graphical representations, and addressing the representation and manipulation of domains;

- Addressing the security threat resulting from the use of aggregation and inference techniques applied to the large DD&R data collections.

The data engineering areas critical to the M&S community that are not being addressed elsewhere include:

- Developing database, model, and organization directories as part of the overall development of the DMSO Information System (currently being done);

- Addressing the issue of distributed architecture and access to heterogeneous DD&Rs;

- Addressing the issue of managing and accessing multi-level information in and across DD&Rs.

The I/DB Task Group currently consists of people from FFRDCs, the Services, Joint Staff, DoD agencies, DDI/CIM, DISA/CIM, DARPA, NIST, and some contractors working for government organizations on M&S projects. The I/DB has met four times over the past nine months, and plans to hold meetings every four months in the future to share information and developments, continue to be informed about relevant activities, and discuss and agree on data administration and standardization methodologies and practices.

INTRODUCTION

PURPOSE

The purpose of this document is to provide information to people who are members of the Defense Modeling and Simulation Office (DMSO) Information/Data Base (I/DB) Task Group, those who wish to join the I/DB Task Group, briefers to the I/DB Task Group, and those with an interest in data activities related to modeling and simulation. It consolidates papers describing the DMSO activities in the Information and Data Base technology areas from August 1991 to November 1992.

BACKGROUND

In 1991 the Deputy Secretary of Defense instituted a major new initiative to strengthen the application of modeling and simulation (M&S) in the DoD. Its purpose is to promcte the effective and efficient use of M&S in joint education, training, and military operations, research and development, test and evaluation, analysis, and production and logistics by: (1) establishing Office of the Secretary of Defense (OSD) cognizance and facilitating coordination among DoD M&S activities; (2) promoting the use of interoperability standards and protocols where appropriate; and (3) stimulating joint use, high return on M&S investment. To achieve these goals requires the development and implementation of a DoD M&S policy, establishment of a DoD-wide management structure to coordinate joint M&S activities and requirements, and the formulation and implementation of a long-range M&S joint investment strategy.

A DoD Executive Council for Models and Simulations (EXCIMS) consisting of DoD Component representatives was established as a board to advise the Under Secretary of Defense (Acquisition) (USD(A)) on M&S policy, initiatives, M&S standards, and investments for improving current M&S capability and promising M&S advanced technologies. The Defense Modeling and Simulation Office was established to serve as an executive secretariat for the EXCIMS and to provide a full-time focal point for information concerning DoD M&S activities. The DMSO promulgates USD(A) directed M&S policy, initiatives, and guidance to promote cooperation among DoD Components to maximize M&S efficiency and effectiveness.

To carry out its functions and develop a master plan, the DMSO enlisted the help of several federally funded research and development centers (FFRDCs). A number of functional and technology working groups were established to determine the M&S needs and to evaluate the state-of-the-art with respect to those needs. The functional groups are: education, training and military operations; research and development; test and evaluation; analysis; and production and logistics. The technical working groups are: experiments; architecture, standards, and

interoperability; methodology/applications; information; networking; computers; software; graphics; databases; instrumentation; behavior; and environment.

As a result of initial activities, the Information Technical Working Group (ITWG) began to develop plans and design for a DMSO Information System to facilitate coordination among DoD M&S activities. The Data Base Technology Working Group (DBTWG) identified three efforts found critical to M&S needs: need for directories, dictionaries, encyclopedias, and repositories to support timely and cost effective access to, acquisition of, and validation of external and derived databases; interoperability, data integrity and consistency across distributed databases and simulations; and M&S community objective assessment of data management products such as relational DBMSs. COL Jim Shiflett of DMSO asked that a special task group be formed from the ITWG and the DBTWG to address the DMSO Information System in coordination with the first DBTWG identified need for directories, dictionaries, etc. Thus the I/DB Task Group was created.

OBJECTIVES OF THE I/DB TASK GROUP

DMSO is the cognizant office of the Data Administrator for Modeling and Simulation. The core I/DB Task Group people responsible for helping DMSO carry out its tasks include: Cy Ardoin (Institute for Defense Analyses (IDA)), Twyla Courtot (MITRE), Roberta Schoen and Bob Bishop (Defense Technical Information Center (DTIC)), and Iris Kameny (RAND). The broad objective of the DMSO I/DB Task Group is to support DMSO in promoting the interoperability, sharing, and reuse of databases and models throughout the defense M&S community. To accomplish this goal requires data and model administration policies and procedures compatible with those of Corporate Information Management (CIM) and the Services as well as the design and development of a DMSO Information System and appropriate tools. The DMSO Information System will be responsive to problems expressed by the M&S community in knowing who is in the community, what data and models are available, where they are, and who is responsible for them. Not only are there few directories or catalogs of databases and models, but there is no community consensus on definitions of concepts and data elements used in databases and models. The I/DB Task Group recognizes that current DoD CIM, Service, defense agencies, and Joint Staff efforts are addressing similar problems and would like to develop compatible policies and procedures where possible. These would guide M&S organizations as well as individual M&S developers.

As an example, what kind of guidance should be given to a developer of a new modeling system? Should he/she be expected to develop a process model of the system? From that develop a data model? From that develop the data elements and database design for input to the model and for output from the model (which may become input to other models)? How would he/she go about finding out if an appropriate database is already available? If one is not available, then do standard data elements exist that correspond to the data elements required for the database? Where does he/she look for them? In the DoD Dictionary Repository System? In his/her Component's data dictionary? In the functional area data dictionaries

corresponding to functional areas of the model? In the DMSO data dictionary? How does he/she propose new standard data elements?

More general questions about the architecture of the DMSO Information System could reach out beyond just the DMSO community. Should the DMSO Information System be a repository system that includes a DMSO data dictionary and data models? Should DMSO store and maintain sharable databases and simulation models after projects are completed and there is no other place to maintain them? Should DMSO support the maintenance of repositories by Services and other organizations rather than at DMSO? How should different repositories exchange information? Do we need a directory system of repositories and their wares? Of server systems and their services? Should an information system act as a server front end to users to handle their requests by searching other servers and repositories?

So far, the I/DB Task Group has addressed the services, tools, and resources required by the DMSO Information System. The DMSO Information System prototype is being designed and implemented by IDA and will be maintained at DTIC.

The DMSO Information System will provide Service support for: (1) M&S special interest groups including bulletin board, email groups, and automatic forwarding of messages to members at their request; (2) M&S related general announcements and event calendar; (3) M&S common definitions, acronyms, and library references; (4) directories/catalogs of M&S organizations, databases, and simulation models; (5) electronic versions of M&S policy and procedures documents, and other documents; and if required, (6) a repository of simulation models, databases, data models, and a DMSO data dictionary. The DMSO Information System tools will eventually need to include manipulation of flat files, relational databases, data objects, and multimedia objects, and a federated interface to heterogeneous data collections. Resources include support for communications, and possibly extensive storage for models, databases, directories, and the DMSO data dictionary.

Current tasks (as of summer 1992) are described below. The ordering of the tasks is not meant to imply a priority. Since the tasks are meant to complement efforts being addressed by CIM and the Services, their priority will be determined by the relative value and need for these solutions/products across the M&S community.

1. Model and Data Administration Policy and Procedures: develop DoD M&S model and data administration policy and procedures that, if possible, are compatible with CIM, Component, and Joint Staff data and software repository policy and procedures. In particular:

   — The use of process and data modeling tools by M&S organizations and M&S developers (e.g., Integrated Computer Aided Definition Language (IDEF)) as a common base for understanding models and developing

well-designed databases and identifying new data entities, data elements, and relationships.

— Standard data naming conventions, schema, and definition processes. Issue: the M&S community also needs to represent more complex kinds of data in the data dictionary such as structured objects, tables, and matrices of information (e.g., environmental data, emissions tables, satellite data, etc.); aggregate data; composite data; and data elements that are semantically complex such as probability of target acquisition, probability of hit, and probability of kill.

— Issue of need for DMSO data dictionary, repository for models and databases, or repositories distributed across M&S community. Does DMSO need to build and maintain these? Whether yes or no, how should multiple repositories interact? Who determines what should be placed in a repository? Who is responsible for maintaining repository directories and deciding when to archive and destroy products?

— Architecture of DMSO Information System and/or repository with respect to interfacing with DoD Data Dictionary, Component data dictionaries, etc. Are these being conceived of as distributed communicating servers that can each be addressed by "his own user"?

2. DMSO Information System: Currently prototype development is proceeding at IDA and the operational version will be installed at DTIC. Another near-term task is to use the I/DB as an example special interest group to test out the bulletin board and group features of the DMSO Information System.

3. Directories: (1) a directory of M&S organizations is being compiled; (2) a schema design for a directory of database and database directories using IDEF1X methodology is currently under development; and (3) future plans call for development of a schema for a directory/catalog of models and simulations in consort with or taking advantage of efforts being done by other organizations (e.g., Army, J-8 OASIS, J-MASS, J-6). An issue is the search/key word terms/hierarchy for the database and model base directories. We need to find out if other groups have developed such, if they are applicable, and how they will be maintained.

4. Understanding and evaluating IDEF tools and methodology: explore IDEF tools and methodology to better understand and evaluate them and future enhancements in order to inform the M&S community as to recommended usage and shortcomings.

5. Definitions, terms, acronyms, references: effort has been expended on reaching agreement on definitions and terms, collecting acronyms, and library references. Issues: Who will maintain and update these? Where will the library of hardcopy documents be kept? How will these be made available to the M&S community?

6. Legacy models and databases: This task is just starting to be addressed. Experience with "standalone" databases indicates that in many cases, each existing data element will be able to be mapped to one or more standard data elements and that the mapping information could be maintained as metadata with the database and as alias information with the standard data. This is a human intensive activity and an issue is how to pay for the effort since the cost is usually too great for the using project and the benefit is to many projects and users over time. Legacy models present a more difficult problem though CIM is experimenting with tools for reverse engineering. Part of this task will be to examine these tools and if they appear reasonable recommend further evaluation by running one or more test cases. The goal for legacy models would be to try to develop, as automatically as possible, process and data models, and to identify input and output data descriptions. Again, this will probably be a very human intensive activity and the cost will probably be an issue.

I/DB TASK GROUP MEETINGS

The I/DB Task Group meetings began in February 1992 with a small core of members. The purpose was to discuss issues related to the tasks that were being undertaken and to report on progress. Several briefers were invited to address relevant issues. More and more people in the M&S community learned of the meetings and asked to join in. Currently, there is a membership (active and inactive) of around 60 people. The meetings have evolved from being working meetings to being a means of facilitating information exchange among the M&S community by introducing people to others with similar needs and problems, briefing relevant M&S data-related projects, inviting guest briefers on methods, standards, issues, and relevant activities going on in different organizations such as the Director of Defense Information (DDI)/CIM office, Defense Information Systems Agency (DISA)/CIM, Joint Interoperability Engineering Organization (JIEO), the Services, NIST, and ARPA. DMSO has encouraged the continuation of the meetings, which are held approximately every four months, since they appear to be answering an M&S community need.

ORGANIZATION AND STRUCTURE OF THIS DOCUMENT

This document consolidates the findings of the DMSO Data Base Technology Working Group and the I/DB Task Group over the period from August 1991 - November 1992 in supporting the DMSO in promoting the interoperability, sharing, and reuse of databases and simulation models throughout the DoD Modeling and Simulation (M&S) community. It is based on a draft written in August 1992 to accompany a briefing on Database Technology Assessment for Modeling and Simulation given to the Defense Science Board Summer Study on 11 August 1992.

The next section of this document is the annotated Defense Science Board briefing. It is followed by a series of appendices. Appendix A contains a list of

acronyms. Appendix B contains a list of documents relevant to the subject area. Appendix C contains the current list of I/DB Task Group members. Appendices D through G contain the notes from each of the four I/DB Task Group meetings held between February and November 1992.

Database Technology Assessment for Modeling and Simulation

Presented to the Defense Science Board
Summer Study on Modeling and Simulation

11 August 1992

Outline

Goals and Scope

Perspectives on the current state of database technology

Initiatives identified by the DMSO Database TWG

Status and discussion of current CIM efforts

Summary

Goals and Scope of DMSO Data Base Technology Working Group

Goals:

- Assessment of state-of-the-art and research areas
- Identification of key issues and potential DMSO initiatives
- Recommendations

Scope:

- For four (4) phases of M&S development and use
  1. Identifying and preparing data for M&S
  2. Executing the simulation
  3. Storing and analyzing the results
  4. Management support for designing experiments and maintaining record of experimental runs.

GOALS

The DBTWG overall goal is to develop technical guidance for the DMSO as to what is available from the database community, what the future looks like, what the shortfalls are, and which of these could be critical to DMSO and need support. The first step was to assess the state-of-the-art and research areas with respect to broad M&S needs. From this assessment we identified key issues critical to M&S use of databases and database technology and transformed these into four potential DMSO initiatives.

SCOPE

To be complete in handling diverse needs of the M&S community, the assessment would need to cover: database management systems; knowledge base management systems; systems for handling very large specialized data collections, such as an image management system; and systems that handle very large collections of eclectic data, information, and programs such as repositories or archives. The latter type of system may be tightly integrated (e.g., a single repository either centrally managed or distributed) or loosely integrated (e.g., a front end to autonomous heterogeneous data, knowledge, and file systems). During this short period the DBTWG concentrated mainly on DBMS technology including directories, dictionaries, and repositories of data, models and information to support interoperability, sharing, and reuse across the M&S community.

We tried to scope the effort in terms of what database technology has to offer to M&S development in the following phases: (1) identifying and preparing data inputs for M&S; (2) executing the simulation; (3) storing and analyzing the results; and (4) management support for designing experiments and runs and maintaining records of such. Because the DMSO effort is directly concerned with interoperability among distributed models, we addressed distributed processing and security concerns appropriate to each phase.

Phase 1: Identifying and preparing data inputs for M&S requires the use of: information resource management tools and techniques such as directories, dictionaries, and repositories; standards/conventions for representing and manipulating data elements, objects, higher level concepts, database schemas, etc.; representation of data constraints (e.g., range of values, enumerated list of values); and techniques for version control across databases. Technology support in this area is basic to providing interoperability across models since it provides the tools for developing definitions and agreements on the meaning of data concepts and their names. Directories, dictionaries, and repositories can enable M&S builders to locate data of interest. Data element/objects/schema, standards/conventions/definitions can reduce ambiguity and redundancy. Data constraints can be used in error checking the data.

Phase 2: Executing the simulation: simulation unique databases derived from Service, DoD agency, etc. databases can be managed by a database management system (DBMS) accessible to the simulation. This could be a relational, extended-relational, object-oriented, or intelligent data or knowledge-based management system. The use of such a tool could offer: (1) structured support for representation of objects and their relationships to each other (and their relationships to related multimedia objects such as an engineering drawing of the tank object or a satellite image that contains the installation object); (2) generic methods for manipulating/reasoning about the objects making use of their relationships such as command structure, support structure, and aggregation/disaggregation in support of variable resolution; (3) triggering of other data changes; and (4) persistent storage of simulation objects. A database system using history-preserving techniques could serve as a persistent store for data entities during simulation execution as well as for post-analysis of the simulation results.

Phase 3: Storing and analyzing the results of M&S experiments/runs: storage of database results in a DBMS can support many purposes such as improving the model, use of the results by other models, analysis and evaluation of the results for a single exercise or run, comparison of results across runs or models, and replaying the simulation. DBMS tools could include simple statistical tools and graphics or more sophisticated object-oriented or knowledge-based management tools to support, for example, spatial and temporal reasoning about results, and abstractions, or aggregations of results--aggregating detailed actions of objects into more interesting events that are more amenable to human understanding and analysis than individual actions. If a sophisticated data/knowledge management system was used in execution, it may serve this purpose also.

Phase 4: Management support for designing model runs and experiments and maintaining records of experiments: requires a management tool in the M&S environment that will support the exercise developer or model runner in designing and keeping track of related information for various exercises or runs. This includes defining the schedule runs for sensitivity analysis, tying each uniquely defined run with information such as date, analyst, version number of model, version number of environment, inputs, and outputs. This will support the comparison of results across runs and models in making clear what the differences were between the objects of comparison.

Outline

Goals and Scope

Perspectives on the current state of database technology

Initiatives identified by the DMSO Database TWG

Status and discussion of current CIM efforts

Summary

Relevance Assessment: Database Technology

M&S Relevance	Status	Type	Representative Products
High	Commercially available	• Relational • GIS (geographic)	Oracle, Teradata, Sybase, Ingres ARC/INFO, Intergraph
	Emerging	• Object-oriented • Distributed—Heterogeneous • Secure	• Object Store, GEMSTONE, ONTOS • Oracle Star*, Sybase* • Teradata, Sybase, Oracle
	R&D	• Scientific • Extended RDBMS	•—— • Postgres, Starburst
Medium	Commercially available	• Test • Fault tolerant	•Topic, GESCAN, FDF, BASIS • Non Stop/SQL, Teradata
	Emerging	• Distributed—Homogeneous	Sybase, Ingres Star, Oracle Star
	R&D	• Historical • Multi-media • Real-time • Main Memory	• —— • ——
Low	Aging	• Network • Hierarchical	• IDMS • IMS, System 2000

*With gateways.

The table above is a high level assessment of current DBMS technology and its relevance to M&S. The first column, "M&S relevance," groups database technologies into three categories (high, medium, and low) with respect to their importance/relevance to *new* M&S activities. The "status" column indicates the DBTWG consensus on the state of the technology. "Commercial" indicates that the technology is available off-the-shelf and has been around for a while. "Emerging" indicates that the technology is available off-the-shelf but is a new product that has not been well tested by consumer use or is a product in beta test. "R&D" indicates that research prototypes have been built but the product is not available off-the-shelf. "Aging" indicates the technology is old and will probably not be supported in the future. The "type" column is an attempt to list the prevalent types of commercial and research DBMS systems in terms of a particular feature. In some cases, the feature is a database model, e.g., relational, network. In other cases, it is a capability in a particular area, e.g., secure, fault-tolerant, or real-time. Note, however, that a particular DBMS could be of more than one type (e.g., Teradata implements a relational model as well as being a fault-tolerant system). For each type of DBMS technology, where possible, we have furnished a short list of example systems.

High Relevance: For the present and near term, relational database management systems and geographic information systems are, and will continue to be, the most stable, reliable, and powerful DBMS products on the market and will play an important role in supporting M&S activities. Object-oriented DBMSs, distributed heterogeneous DBMSs, and secure DBMSs are beginning to emerge as commercial products The M&S community can expect that reliable and powerful implementations of these products will become available and will provide important capabilities in support of M&S. Extended relational DBMSs and scientific DBMSs are now only research areas. However, should they mature into commercial products, they could provide data management capabilities of great importance to the M&S community.

Medium Relevance: Text DBMSs and fault-tolerant DBMSs are currently available as mature commercial products. We expect that there are, and will be, a number of special applications in the M&S community that will require the capabilities provided by such products. Distributed homogeneous DBMSs are emerging as commercial products. While these products do not offer the ability to connect multiple heterogeneous DBMSs, their ability to connect multiple DBMSs all from the same vendor will have a number of uses within the M&S community. Multimedia, real-time, and historical DBMSs are still in the research stage. However, the promise of such systems could ultimately provide such capabilities as direct DBMS to model connections in which the DBMS would provide the model inputs and capture the model outputs.

Low Relevance: Database management systems based on the older technologies such as hierarchical and network DBMSs are currently in place and will likely remain in place for a number of applications within and peripheral to the M&S community. While such systems are not likely to provide any new or advanced capabilities to the M&S community, many of the legacy systems they support are likely to provide data to the M&S community for some time into the future.

The types of current DBMS technology are briefly described below:

Relational: a data model based on the theory of mathematical relations, domains, and ranges. DBMS based on the relational model are characterized by data stored in tables such that each row represents a record of data and each column corresponds to a field of the record.

Object-oriented: a data management methodology based on the concepts of object-oriented programming such that objects are persistent and the data management system maintains not only the data as objects, but also the behaviors of the objects stored as methods.

Extended relational: a data model that extends the scope of the pure relational model by including more semantics of the data in the form of complex data types, rules, constraints, and triggers.

GIS: a geographic information system that has special support for geographic data including spatial indices, spatial queries, and geometric operators to efficiently manage large vector and raster-based geographic datasets.

Distributed: a DBMS that supports distribution of data in a transparent way to users and application programs. In general, a single global schema is supported.

Distributed homogeneous: a distributed DBMS that requires that the same data model and same DBMS be used at all nodes.

Distributed heterogeneous: a distributed DBMS that allows different DBMS and different data models at each node.

Secure: a DBMS that is evaluated to support data security as specified by the Trusted Database Management System Interpretation of the Trusted Computer System Evaluation Criteria issued by the National Computer Security Center, April 1991.

Multimedia/Image: a database and corresponding database management system that maintains repositories of multimedia data such as images, video, voice, text, as well as traditional structured data. These DBMS may support different index, query, and update capabilities for each different medium and enable the combination of media in a single application.

Scientific: a database and corresponding database management system that specializes in supporting technical and scientific datasets such as those required by the human genome project, meteorological studies, and astronomical studies. These databases tend to result from complex data collection activities and require special preprocessing and statistical analysis.

Fault-tolerant: a database management system that uses redundancy to ensure fault-tolerant behavior (e.g., a data item stored redundantly on different storage devices, redundant or standby DBMS processors, redundant or standby networks).

Real-time: a database management system that is optimized for real-time transaction processing to drive time-sensitive applications such as real-time process control and high-risk environmental monitoring.

Main memory: a database management system that resides in main memory and manages data that resides only in main memory (no data resides in secondary storage).

Historical: a database management system that physically appends data changes rather than replacing existing data values with the changed ones. It may also provide a temporal query language.

Network: an outdated data model based on network structures that paved the way for separating the physical storage of data and the logical organization of data

records and fields, thus supporting physical and logical data independence. These DBMS generally supported a procedural data manipulation language but no declarative query language.

Hierarchical: an obsolete data model based on a hierarchical tree organization that was well suited to managing hierarchical data. However, because of the limitations of tree structures (that is, each node has a single parent), the pure hierarchical model was not sufficient for many real world applications.

TECHNOLOGY PERSPECTIVE: DATABASE TECHNOLOGY

AREA	CURRENT STATE-OF-THE-ART			TRENDS
Memory	16 MB chips			64 MB ——> 1 GB
	Type Size	Access Time	Transfer Rate	Higher volumes (e.g., 10 GB mag disk, 100 GB tape cartridge, 5TB tape carousel); faster access (1 ms disk arrays, 5 sec JB/carousel); faster transfer (15 MB/s)
Secondary storage	2.5 GB mag disk	8 ms	8 MB/s	
	1.0 MB solid state	40 ms	3 MB/s	
	Disk arrays	10 ms	20 MB/s	
Archival Storage	500 GB optical JB	45 sec	1 MB/s	
	500 GB carousel tape	45 sec	1 MB/s	
Database machine processors	Teradata (Intel 80486) Transputer-based Connection-machine based			Teradata --> RISC Commercially available Commercially available
Data models	Relational Geographic Information Systems (GIS) Object Oriented (OO) — limited			Extended relational, intelligent and active DB, multimedia, strong commercial OO, integrated OO, relational, and GIS
User interfaces	Command-line Forms Graphical User Interfaces (GUI)			Multimedia, hypermedia, more GUI, customized user DB navigation
Security	System high No evaluated DB security products			C2 relational, multilevel B1 --> B2 relational
Distributed data management	Homogeneous Client-server architectures			Heterogeneous, federated
Open systems: Interoperability	Limited			Continued improvement (e.g., RDA Standard)
Portability	Supported by SQL standard			(?)
Scalability	Primarily influenced by parallelism			(?)

The nine fundamental technology areas that will drive the advancement of database technologies in support of M&S are: memories, secondary and archival storage, database machines, data models, user interfaces, security, distributed data management, and open systems.

Memory: Database management systems are very large, very complex programs. Memory capacity and speed are fundamental to performance for all large-program systems including DBMSs. Greater memory capacity means more data and software can be held in the computer ready for immediate access without the need for paging, thus increasing performance. Ultra-large main memory systems could even eliminate the need for secondary storage of data in that they could support main memory databases.

Secondary and Archival Storage: Database management systems typically manage large volumes of data--far larger than the memory capacity of the host processor. Thus, a major activity of a DBMS is to access data on secondary or archival storage. This reliance on secondary or archival storage limits DBMSs in two ways: first, in performance because I/O processes are relatively slow; and second, in capacity, because of limits on the size and number of storage devices a system can support. Thus, improvements in access time, data transfer rates, and data capacities translate directly into improvements in DBMS performance and capacity.

Database Machines: Database machines are computer hardware systems that have been specifically designed, modified, or configured to optimize the execution of DBMS code and associated I/O. The promise of database machines is manifested in the commercial product from Teradata, a linearly scalable architecture that can add processors and I/O devices as needed to provide both speed and capacity.

Data Models: Data models provide the mathematical and computational foundation for DBMSs. The relational data model, which has taken over a decade to emerge as a true commercial product, brought many new capabilities to the database world. Work on new and modified data models will likewise produce advances in performance, capacity, and utility. However, the move from research to mature commercial products will take several years.

User Interfaces: Advances in user interfaces translate directly into utility of the underlying DBMS. Historically, user interfaces have been limited to command-line and forms entry. The current trend toward graphical interfaces and advanced navigation tools will greatly increase the utility of DBMSs.

Security: For some applications, data and database security will be mandatory. Historically, such applications have operated in a system high mode. While computer and database security are still in the early developmental stages, research in the area continues. One interesting note is that security has and will continue to be driven more by the government than by the marketplace. Thus gains in database security will likely be made only as a result of government sponsored research.

Distributed Data Management: Distributed data management encompasses the ability to integrate a number of distributed databases resident in heterogeneous database management systems into a virtual single database. The current state of the art is that a number of DBMS vendors allow limited distributed DBMS

functionality across collections of their own DBMSs and some allow communications with selected other products mainly through specially developed gateways. From the user perspective, fully integrated distributed data management is a necessity if universal data access is ever to be achieved. Refer to "Appendix G: Notes from the 4th I/DB Workshop" for a brief description of the MITRE Heterogeneous Information System Testbed developed to evaluate Commercial-Off-the-Shelf (COTS) Distributed Heterogeneous Information System (DHIS) products in terms of strategies, methods and tools, and benchmarking COTS products.

Open systems: in order to achieve interoperability, portability, and scalability the DISA Center for Standards is working on an Open Systems Environment (OSE) and has developed an initial Technical Reference Model (TRM) for Corporate Information Management (CIM) that is based on (among other things): the POSIX operating system standard; the X-windows user interface services standard; Standard Query Language (SQL), Information Resource Dictionary System (IRDS), and Remote Data Access (RDA) data management system standards; GKS and PHIGS graphics standards; and GOSIP plus other network service standards. Work is also proceeding toward establishing an application process interface standard to allow interoperability of 4GLs with heterogeneous DBMS. This currently presents a problem because each DBMS has its own unique non-standard 4GL.

Functional Perspective: Database Technology

FACTORS	EXISTING LIMITATIONS	TRENDS
Database Size	Physics of data storage technologies limits capacity and speed, and affects cost	Improvements in simulation realism through increasing data storage capacity
Database Speed	Hardware Operating systems Database software	Gap in providing real-time M&S support will persist
Database Interfaces	Access technologies Data representation & presentation techniques	Improving navigation methods Use of GUI Use of 4GLs and open system 4GLs
Database Interoperability	Lack of: DD&R standards, and multilevel security (MLS) Immature distributed data management	Development and use of CIM practices Gap in supporting distributed data management technologies is likely to persist Lack of MLS likely to persist

The above table views database technology and its potential contribution to M&S from the perspective of the user, i.e., from the functional perspective. Functionally, the user will be concerned with database size, speed, interfaces, and interoperability.

Database Size: Database size is currently limited by the capacity of data storage devices. There are significant market forces driving improvements in this area. As database size increases, it will become possible to use DBMS technology to, among other things, provide the data needed to increase realism in simulations.

Database Speed: Database speed is limited by CPU, memory, and I/O hardware speeds, by operating system capabilities, and by the software implementing the DBMS functions. Although there are again significant market forces driving improvements in database speed, it is likely that gains in database speed will not keep pace with the desires of the M&S community and that a speed gap will persist between commercial products and the needs/desires of the M&S community.

Database Interfaces: To some extent, users only see a DBMS from the perspective of its interface. The DBMS interface ultimately determines how well the system will support the user. Today's limitations on database interfaces are the result of immaturity of access technologies and of data representation techniques. The marketplace is beginning to demand improvements in these areas, and the vendor community is responding with improved navigation methods and graphical interfaces.

Database Interoperability: If the user community is to ever realize its need for universal access to data, database interoperability must become a reality. Database interoperability is a direct function of distributed data management technology, of database security, and of information resource management tools, standards, and conventions including data directory, dictionary, and repository facilities. Today, these technologies do not provide the level of support needed to achieve true database interoperability. While there are market pressures to achieve database interoperability, the vendor community is not solidly behind interoperability. It is likely that a gap will persist between the needs of the M&S community and the interoperability capabilities of DBMS products. Fortunately, many of the interoperability needs are being addressed by the CIM initiative and by similar initiatives in the Services and DoD agencies.

In "Appendix G: Notes from the 4th I/DB Workshop," Twyla Courtot reported on two other ANSI standards groups: X3.H7 is a new group addressing object information management; and X3.H8, which has been addressing data representation including naming standards, has a subgroup that is concerned with standardizing data.

DB Technology Standards Efforts Related to Interoperability

- Database Management: X3H2-SQL3, X3H2.1— RDA (remote data access) JTC1 SC21/WG3 — data management, SQL Access Group
- Transactions Processing: X3T5 TP, JTC1/SC21/WG5, POSIX 1003.1, X/Open Transaction Processing
- Object Communications and Distribution: X3T5 — OSI (Open Systems Interconnection), X3T3 — (Open Distributed Processing), OMG ORB — Object Management Group's Object Request Broker, JTC1 SC21/WG4 — Management Information Services, X3T1M1.5, OSI/NM Forum
- Data Interchange: X3T2 — Conceptual Schema for Data Interchange
- Domain-specific Data Representation: PDES/STEP (Product Data Exchange using STEP), EDI (Electronic Data Interchange), EDIF (Electronic Data Interchange Format), ODA (Office Document Architecture)
- Repositories: X3H4 — IRDS (Information Resource Dictionary Systems), X3H6 - CIS (CASE Integration Services), EIA CDIF (Electronic Industry Association CASE Data Interchange Format)

The above summarizes the following two pages, which have been copied from the "X3/SPARC/DBSSG/OODBTG Final Report," 17 September 1991, Editors: Elizabeth Fong (NIST), William Kent (Hewlett Packard Laboratories), Ken Moore (Digital Equipment Corporation), and Craig Thompson (Texas Instruments Incorporated).

Recommendations for Standards in Object Information Management

Accredited Standards Committee X3, INFORMATION PROCESSING SYSTEMS
DBSSG/OODBTG Final Report, 17 September 1991

Standards Effort	Description
	Database Management
X3H2 - SQL3	A technical committee responsible for the standardization of database language NDL and SQL. They have, in May 1991, completed specification of SQL2, and are currently working on SQL3, an extension to the current SQL standard which will include object concepts.
X3H2.1 - RDA (Remote Data Access)	A task group under X3H2 on Remote Data Access (RDA). This group is responsible for the specification of a protocol concerned with providing access to data stored at remote sites using SQL.
JTC1 SC21/WG3 - Data Management	An international standards committee responsible for the specification of standards on data management. Projects include data management reference model, database languages SQL, IRDS and RDA.
SQL Access Group	A consortium of users and vendors working to advance the RDA protocol and planning to work on a call-level interface to SQL systems.
	Transaction Processing
X3T5 - TP (Transaction Processing)	A task group under X3T5 (OSI) is responsible for the specification of TP which is an application layer protocol used for exchange of information between two or more distributed systems.
JTC1/SC21/WG5	An international standards committee responsible for the specification of standards on transaction processing languages and bindings, including concurrency, commitment, and recovery (CCR).
POSIX 1003.1	A group working on a profile for transaction processing.
X/Open Transaction Processing	A working group developing the XTP model of transaction processing which includes the XA transaction specification.
	Object Communications and Distribution
X3T5 - OSI (Open Systems Interconnection)	A technical committee responsible for the specification of protocol standards in accordance with the 7-layer Open System Reference Model. In particular, the X3T5.4 Network Management Task Group is responsible for the specification of managed objects using object-oriented technology.
X3T3 - ODP (Open Distributed Processing)	A U.S. technical committee contributing to the international effort JTC1/SC21/WG7. The ODP effort is working on the specification of a standard reference model which will make the complexities of distributed computer systems more transparent. The ODP-RM defines an ODP trader which is a computational object offering services to other objects at service ports.
OMG ORB - Object Management Group's Object Request Broker	A task force within OMG developing technology that performs application invocation and sharing of large granule objects.
JTC1 SC21/WG4 - Management Information Systems	An international standards committee responsible for the definition of the information model of managed objects that corresponds to the information aspects of the systems management model. Although the documents refer to CCITT applications, they define general object management concepts.

Recommendations for Standards in Object Information Management

Accredited Standards Committee X3, INFORMATION PROCESSING SYSTEMS BSSG/OODBTG Final Report, 17 September 1991

Standards Effort	Description
	Object Communications and Distribution
X3T1M1.5	A technical committee responsible for common management information services for managed objects defined in accordance with JTC1 SC21/WG4 documents
OSI/NM Forum	An international forum on OSI network management
	Data Interchange
X3T2 - Conceptual Schema for Data Interchange	A project under X3T2 working on the standardization of conceptual schema mechanism for data interchange. Responsible for ASN.1, a language for data encoding and interchange.
	Domain-specific Data Representations
PDES/STEP (Product Data Exchange using STEP)	The PDES is the U.S. organizational activity that supports the development and implementation of STEP. STEP is the standard for the exchange of product model data. The level 3 product data sharing implementation specifies that multiple user applications access data to a common shared database.
EDI (Electronic Data Interchange)	EDI is an application layer protocol. It is a standard which describes formats for orders, payments, shipments, billings, and other business transactions.
EDIF (Electronic Data Interchange Format)	A format for exchanging CAD chip design data.
ODA (Office Document Architecture)	ODA is a standard for interchange of documents (including text, facsimile and graphics information) which are produced in an office environment. Interchange of ODA documents may be by means of data communications or exchange of storage media
	Repositories
X3H4 - IRDS (Information Resource Dictionary Systems)	A technical committee responsible for the specification of IRDS1 family of standards. This IRDS1 family of standards includes a command language and panel interface specification, a soon to be approved Export/Import File Format standard, and a Service Interface specification. The next family of IRDS standards will utilize object technology.
X3H6 - CIS (CASE Integration Services)	A technical committee working on a family of standard interfaces between CASE environment framework components and tools. Standards are being developed for service and tool registration, change management (versions and configurations), and an object model.
EIA CDIF (Electronic Industry Association CASE Data Interchange Format)	An industry association established to permit interchange of CASE models and data among tools.

DBMS Research Directions

1. **Complex data types**
   e.g., Berkeley, Microelectronics and Computer Technology Corporation (MCC), University of Florida

2. **Object-oriented (methods)**
   e.g., Hewlett-Packard, Texas Instruments, MCC, Brown

3. **Intelligent DBMS (deductive)**
   e.g., Berkeley, MCC, Stanford, George Mason University

4. **Spatial Reasoning**
   e.g., University of Maryland, University of Maine, University of Santa Barbara/Calif., University of Buffalo

5. **Temporal Reasoning**
   e.g., University of Arizona, University of Rochester

6. **Inexact, Fuzzy Reasoning**
   e.g., Berkeley, Princeton, George Mason University

7. **Database Security**
   e.g., MITRE, SRI, George Mason University

8. **Distributed DBMS**
   e.g., University of Alberta, IBM, Princeton, Bellcore, GTE

9. **Extensible DBMS**
   e.g., IBM, University of Wisconsin, Xerox, Berkeley

10. **Physical Data Storage and Compression**
    e.g., MITRE, NSF

11. **Very Large Databases**
    e.g., Syracuse University, NASA, MITRE

12. **Parallel Processing**
    e.g., University of Wisconsin, Teradata, University of Maryland, Thinking Machine

13. **Transaction Processing**
    e.g., Princeton

14. **Statistical/Scientific DBMS**
    e.g., Lawrence Berkeley Labs, UC Berkeley

DBMS research can be organized into many topic areas. The DBTWG agreed on these 14 areas as being of relevance to M&S. A sample of research organizations engaged in relevant work is shown for each area.

1. Complex data types: research addresses how to store, retrieve, and manipulate different types of data such as digitized graphics, images, and voice; matrices; and multimedia objects composed of different typed

components. (Research organizations include UC Berkeley, MCC, U of Florida.)

2. Object-oriented (methods): research topics of major importance are query model and query optimization, user interfaces, design methodologies and tools, view mechanism and performance. (Research organizations include Hewlett-Packard, Texas Instruments, MCC, Brown.)

3. Intelligent DBMS (deductive): research into a declarative rule-based style of representing data and expressing queries and applications on databases, e.g., using predicate logic for representing data and rules, posing queries, and answering queries. (Research organizations include UC Berkeley, MCC, Stanford, George Mason University.)

4. Spatial reasoning: research addresses reasoning about spatial relationships among data objects. The challenge lies in defining abstractions and architectures to implement systems that offer generic spatial data management capabilities and can be tailored to the domain requirements. (Research organizations include U of Maryland, U of Maine, UC Santa Barbara, U of Buffalo.)

5. Temporal reasoning: research addresses developing methods or theories for reasoning about temporal relationships among data such as events occurring before, after, during, within the time span of another event; and expression of relative time (10 minutes, a business day) in a query language. (Research organizations include U of Arizona, U of Rochester.)

6. Inexact, fuzzy reasoning: research addresses expressing and handling imprecise, unavailable, unknown, and missing data. Information about such data may be probabilistically described or described qualitatively (e.g., young, middle-aged, old). A research approach is to apply fuzzy set theory to manipulation of such data. (Research organizations include UC Berkeley, Princeton, George Mason University.)

7. Database security: research is concerned with the ability of a computer system to enforce a security policy governing the disclosure, modification, or destruction of information. (Research organizations include MITRE, SRI, George Mason University.)

8. Distributed DBMS: research into homogeneous/heterogeneous independent or federated DBMSs. Research issues include DBMS autonomy, semantic heterogeneity and schema integration, transaction processing and concurrency control. (Research organizations include U of Alberta, IBM, Princeton, Bellcore, GTE.)

9. Extensible DBMS: research into extending a DBMS by adding new data types and operators at the user interface level, adding new set operators at the query language level, and allowing new implementations of

operators as new algorithms are developed. (Research organizations include IBM, U of Wisconsin, Xerox, UC Berkeley.)

10. Physical data storage and compression: research into developing new compression techniques for reducing data size, and new physical storage devices and technology (including multi-staged storage) for managing large repositories of data. (Research organizations include MITRE, NSF, NASA.)

11. Very large databases: research into storage, retrieval, and manipulation of very large databases. This includes indexing, managing, and accessing data on multi-level staged storage. (Research organizations include Syracuse University, NASA, MITRE.)

12. Parallel processing: research topics include mixing on-line ad hoc queries and on-line transactions without seriously limiting transaction throughput, improved optimizers for parallel queries, and tools for physical database design and on-line database reorganization. (Research organizations include U of Wisconsin, Teradata, U of Maryland, Thinking Machines.)

13. Transaction processing: research includes developing new methods for handling transactions, particularly long transactions and distributed transactions (e.g, nested transactions, semantic knowledge for scheduling transactions, SAGAS, optimistic commit protocols) to improve performance. (Research organizations include Princeton.)

14. Statistical/scientific DBMS: research into support for highly technical and scientific datasets (e.g., meteorological and astronomical data collections) requiring special preprocessing and statistical analyses. (Research organizations include Lawrence Berkeley Labs, UC Berkeley, NASA, NOAA, USGS.)

Outline

Goals and Scope

Perspectives on the current state of database technology

Initiatives identified by the DMSO Database TWG

Status and discussion of current CIM efforts

Summary

Database Technology Issues and Initiatives

- **DD&R: M&S projects need timely and cost effective:**
  - **Access to data (including acquisition)**
  - **Verification, validation and accreditation of data**
- **Distributed Data Management: M&S community requires:**
  - **Interoperability**
  - **Data integrity and consistency**
- **Data Management Product Assessment: M&S community needs:**
  - **Information on applicability of COTS products**
  - **Information on technology gaps**
- **Data Representation: M&S community must represent complex data types/information such as:**
  - **Doctrine**
  - **Human behaviors**

The DBTWG agreed on these four issues as being the most critical and relevant to M&S needs. Each is represented by a following chart showing a possible DMSO initiative addressing this issue.

Directories, Dictionaries, and Repositories (DD&R): The modeling and simulation community has a high-priority need for support in locating, accessing, acquiring, and aggregating data to be used as input to models as well as information about models themselves. There is a need for an infrastructure that will supply the policy, procedures, management, authority, and funding to ensure the design, development, service, and maintenance of DD&R facilities at appropriate places across the M&S community.

Distributed Data Management (DDM): In order for models to interoperate correctly, ground truth data has to be consistently maintained across the models. Research in distributed homogeneous and heterogeneous management of replicated data as well as data security offer insights into problems and solutions that can be applied to distributed simulations.

Data Management Product Assessment: Sharing of data across simulations as well as analysis of output data from simulations could be aided by integrating a DBMS product with the running simulation, thus supporting persistent data or objects. As the trends toward larger, faster, and more cost effective memories and secondary and tertiary storage continue, increased performance may make this a reality. The M&S community would benefit from a community-directed effort to evaluate and assess potential DBMS products as to their applicability in meeting specific M&S performance needs.

Data Representation: Many M&S applications require the representation and manipulation of data not currently handled well by most commercial DBMS products. This includes data about spatial, temporal, fuzzy, behavioral, and doctrinal concepts. Research in this area may not be adequately funded by commercial endeavors and may need support from the DoD community.

Definitions for Directory, Data Dictionary and Repository

Directory: a database of entries each of which identifies the directory object type (e.g., organization, database, model) by name, data, search terms, functions, etc.

Data Dictionary: a specialized type of database containing metadata that is managed by a dictionary system; a collection of data describing the characteristics of other data.

Repository: a container for a collection of such things as directories, dictionaries, models, databases, etc.

These definitions are included as an aid to the discussion.

Terminology to Aid in Discussion

Repository	**repository of dictionaries, directories, etc.**
Directory	**directory of databases, contains a collection of references to individual databases**
Database	**a relational database contains relations or tables**
Relation/table	**a relation or table contains records composed of data elements or fields**
Data element	**a data element or field has values**
Values	**valid data element values belong to a domain**
Domain	**a domain may define a numeric range, define a set of enumerated values, or be defined as a set of values that cannot reasonably be enumerated (e.g., social security numbers, proper names)**

The above table is intended to help the audience understand the relationship between the different terms used. A repository can hold dictionaries, directories, databases, and other electronic objects. A directory contains a collection of references and pointers to other objects. For example, a database directory contains information and pointers to databases and other database directories. A database may be composed of structures such as files or tables that contain records or a flat file database that contains only records. A relational database contains relations or tables that contain records composed of fields or data elements.

A data element or field is what is standardized in a data dictionary as a "standard data element." A data element is an attribute of a data entity. The CIM effort identifies data entities from DoD data modeling activities and incorporates entity names in attribute or standard data element names.

An instance of a data element contains a value (e.g., "smith" is an instance of the data element "employee last name"). Valid data element values belong to a domain of values. A domain may span a numeric range of values, be composed of a set of discrete enumerated values, or be composed of a set of values that cannot be reasonably enumerated such as social security numbers or proper names. Logical links of information across databases are made on the basis of data elements that share the same or similar domains.

DMSO Directories, Dictionaries, and Repositories (DD&R) Initiative

- **DMSO need: M&S projects need timely and cost effective access to, acquisition of, and verification of data for use in simulation models**
- **I/DB Task Group formed from the DMSO Information and Data Base TWGs with broad objective to promote interoperability, sharing, and reuse of databases and models throughout the DoD M&S community**
- **Objectives**
  - **Help define DMSO's policy, procedure, management functions to promote data and model reuse and interoperability**
  - **Define and prototype DMSO Information System**
  - **Provide directories of directories, organizations, databases, and models**
  - **Provide means to handle M&S specific data elements**
  - **Explore and trade off alternative approaches to DD&R**
- **Approach**
  - **Explore/coordinate with DISA/CIM, NIST, CALS, Components, etc.**
  - **Design and prototype alternative DD&R approaches**
  - **Involve DMSO & M&S community in prototype evaluation**

The modeling and simulation community has a high-priority need for support in locating, accessing, acquiring, and aggregating data to be used as input to models. The data must be verified against constraints to locate potential errors, validated to be consistent with test data, and certified for use in particular applications. Directories containing database descriptive and access information or access to other directories that contain such information are needed. Security, proprietary rights and privacy require maintenance of separate directories and strict enforcement of access and dissemination policies. Data technology advances addressing directories, dictionaries of meta information about individual databases, dictionaries containing meta information about standard data elements, objects, higher level concepts, database schemas, etc., and repositories of such data are needed. Efforts at NIST, DISA/CIM, and computer-aided acquisition and logistics (CALS) in Information Resource Management (IRM), information dictionary standards, and data element naming conventions are applicable as well as advances in data base security and data base research in accessing data across heterogeneous databases.

The I/DB Task Group, composed of members from the Information and Data Base TWGs and new members, will help DMSO with these activities. Its broad objective is to promote interoperability, sharing, and reuse of databases and models throughout the DoD M&S community. To accomplish this goal requires data and model administration policies and procedures compatible with those of CIM and the Services as well as the design and development of a DMSO Information System and appropriate tools. The DMSO Information System will be responsive to problems expressed by the M&S community in knowing who is in the community, what data and models are available, where they are, and who is responsible for them. Not only

are there few directories or catalogs of databases and models, but there is no community consensus on definitions of concepts and data elements used in databases and models. The I/DB Task Group recognizes that current DoD CIM, Service, defense agencies, and Joint Staff efforts are addressing similar problems and would like to develop compatible policies and procedures where possible. These would guide M&S organizations as well as individual M&S developers.

More general questions about the architecture of the DMSO Information System could reach out beyond just the DMSO community. Should the DMSO Information System be a repository system that includes a DMSO data dictionary? Should DMSO store and maintain sharable databases and models after projects are completed and there is no other place to maintain them? Should DMSO support the maintenance of repositories by Services and other organizations rather than at DMSO? How should different repositories exchange information? Do we need a directory system of repositories and their wares? Of server systems and their services? Should an information system act as a server front end to users to handle their requests by searching other servers and repositories?

So far, the I/DB Task Group has addressed the services, tools, and resources required by the DMSO Information System. The DMSO Information System is being designed by IDA and will be managed by DTIC. The Oracle relational DBMS will be used to manage the directories.

The DMSO Information System will support the services by providing: (1) M&S special interest groups including bulletin board, email groups, and automatic forwarding of messages to members at their request; (2) M&S related general announcements and event calendar; (3) M&S common definitions, acronyms, and library references; (4) directories/catalogs of M&S organizations, databases, and models; (5) electronic versions of M&S policy and procedures documents and other documents; and if required, (6) a repository of models, databases, and a DMSO data dictionary. The DMSO Information System tools will need to include manipulation of flat files, relational databases, data objects, and multimedia objects, and a federated interface to heterogeneous data collections. Resources include support for communications and possibly extensive storage for models, databases, directories, and the DMSO data dictionary.

Distributed Data Management Initiative

- **DMSO needs: M&S community needs interoperability, data integrity, and data consistency across multiple simulations and across a variety of data repositories**
- **Approach**
  - **Identify DMSO distributed data management (DDM) requirements**
  - **Identify technology gaps in DDM**
  - **Design and develop prototype DDM approaches for evaluation**

In order to use data across runtime models (as in extensions of SIMNET), the M&S community needs to address the interoperability/sharing of data across distributed and diverse M&S applications and products. This involves understanding the use of data replication to provide data reliability, availability, and improved performance and the need to maintain the consistency of replicated data (e.g., ground truth data) across distributed simulations, and the proper separation of data at different classification levels. The DBTWG has laid out an approach to the DDM issues organized into four tasks:

Task 1, DMSO Requirements Assessment for DDM, will address needs including (1) federated/heterogeneous data management; (2) replicated data and concurrent updates; (3) near-real-time DBMS processing requirements; (4) different access classes with different assurance levels and policies; and (5) system high and MLS security requirements. The result would be a white paper containing the analysis results, estimated at two staff-months level of effort.

Task 2, Architecture Description, based on requirements assessment of several architectures, will be developed to meet requirements geared for current environment, five-years out, long-range. The result would be a report containing the architecture descriptions, estimated at a six staff-month level of effort.

Task 3, Near-Term Testbed, for the near-term architecture, the design and development of a demonstration capability to be implemented in a distributed testbed. The result would be used to verify the viability of meeting the defined architecture and to confirm the requirements assessment. The result would be a report describing the near-term architecture, implementation, and demonstration, estimated at 1.5 staff-years.

Task 4, five-year and Long-Range Architecture Research, in parallel with the near-term testbed development, develop a plan for how necessary research will be performed to address technology gaps identified in the five-year and long-range architectures. The result will be a report containing the research plan, estimated at a six staff-month level of effort.

Data Management Product Assessment Initiative

- **DMSO needs: a solid body of data base product reviews relevant to the M&S community and identification of gaps in COTS products**
- **Objective**
  - **Reduce redundant DBMS evaluation in M&S community**
  - **Identify desirable features and problems in COTS products**
  - **Foster capability to do COTS evaluations within government or government-supported lab(s)**
- **Approach**
  - **Identify specific M&S needs relative to database products**
  - **Produce semi-automatic test suite(s) specific to M&S**
  - **Produce independent verification of test results as needed**

Currently there appears to be no DoD organization responsible for evaluating DBMS or most other software products. Very frequently when a DBMS product is needed for an M&S database or model development effort, the contract includes the selection of a DBMS product by assessing a set of commercially available products. Thus frequent repetitive assessments of relational DBMS products continue to occur within different DoD organizations. This initiative is based on the postulation that the M&S community would benefit with respect to cost, interoperability, and perhaps even technology development if the M&S community had an organization that fulfilled a "consumer union" role of assessing existing products based on the requirements of the M&S community. This would be another infrastructure activity that would be established and maintained indefinitely.

The main objective of this initiative is to reduce redundant database product testing and evaluation efforts within the M&S community and as a by-product, to educate the community as to the desirable features as well as potential problem areas of evaluated products. An additional objective is to develop the infrastructure to perform well-defined product evaluations by an M&S selected organization (e.g., an FFRDC, government laboratory, NIST, or ?).

The short-term approach is to: (1) identify specific M&S needs for database products; (2) prioritize needs; (3) collect available relevant information; (3) supplement with additional testing as required on a time/funding permitting basis;

and (4) distribute results. A short-term test (at approximately one staff-month of effort) would be the first cut of a document on the use of relational DBMS in three primary areas (incorporating already available data): personal computer, minicomputer/workstation, mainframe, and supercomputer; identifying the next set of most relevant products for evaluation; and performing initial scoping of M&S test suite.

The long-term approach (if the infrastructure concept is viable) is to establish an organization to perform the product testing that will: (1) continue the approach outlined in the short-term plan; (2) produce a semi-automated test suite specific to M&S needs beginning with one for relational DBMS products; (3) solicit running of suite by vendors and other parties; and (4) produce independent verification of results as required. Long-term products would be a series of reports covering important commercial database products for all major classes of M&S needs; widely distribute reports and incorporate feedback; and encourage vendors to address modeling and simulation needs.

Data Representation Techniques Initiative

- **Objective**
  - **Develop capabilities to manage complex data types**
    - **Spatial/temporal data**
    - **Imprecise/fuzzy data**
    - **Unknown/missing data**
    - **Behavioral**
    - **Military doctrine**
- **Approach**
  - **Assess and prioritize DMSO needs**
  - **Evaluate current R&D efforts**
- **Results**
  - **Identification of needed R&D initiatives**

The purpose of this initiative is to analyze M&S community needs for advanced data representation techniques, evaluate commercial developments and R&D efforts, identify the gaps, and propose DMSO initiatives when it is ascertained that the gaps are not likely to be met by other means.

The M&S community needs to represent and manipulate complex types of data that differ from widespread commercial needs and even from other DoD needs except for selected parts of the C4I community. Thus the community cannot rely on commercial or other government R&D support to develop needed capabilities. The types of complex data include: spatial/temporal data such as terrain and environmental data (e.g., weather, smoke, dust) as noted by the Environment TWG; data expressing military doctrine and force behaviors as noted by the Behavioral Representation TWG; and imprecise/fuzzy data as well as unknown/missing data as noted by most TWGs but particularly by the Instrumentation TWG. Furthermore, the need to collect runtime data from models for analysis, use by other models, replay, etc. could benefit from incorporating historical DBMSs (which is a current research area) into modeling environments.

The research endeavors in these areas are discussed further under "DBMS Research Directions."

Listing of Major Briefings Given at I/DB Task Group Meetings

ARMY ACTIVITIES:
- How the Army is Organized to Support M&S Data: Erwin Atzinger (AMSAA)
- TRAC Automated Data System for Army M&S: Walt Swindell, Howard Haeker (TRAC)
- Army Master Model and Simulation Catalog: Lana McGlynn (Army)
- Data collection efforts of the CCTT program: Bill Johnson, Rob Wright (CCTT)

AIR FORCE ACTIVITIES:
- Air Force M&S Analysis Functional Area Needs: Roy Reiss (AFSAA/SAG)

CIM BRIEFINGS:
- DoD Data Administration Policy and Procedures (including DoD 8320.1, DoD 8320.1-M, DoD 8320.1-M-1, DASP, migration prototype, naming qualifier issue, roles (Functional/Component Data Administrators, DoD Data Administrator, FIMs, software developers)): Bob Molter (DDI/CIM)
- DoD Data Model (including status, role of Model Integration Group, DoD data element standardization): Russ Richards (DISA/CIM)
- Defense Data Repository System (DDRS): Dan Lewis, Jeff Wolfe (DISA/CIM)
- DISA/CIM Data Administration education/training: John Hovell (DISA/CIM)
- Defense Management Report Decision (DMRD-918): Twyla Courtot (MITRE)

COMPONENT DATA ADMINISTRATION BRIEFINGS:
- Air Force Data Administration: Bao Nguyen (AF DAd)
- Army Data Administration: Jim Glymph (Army DAd)
- Navy Data Administration: Rebecca Wade (Navy DAd)
- Joint Staff Data Administration: Janet Barlill (JS DAd representative)

DMSO ACTIVITIES:
- DMSO Analysis Functional Area Needs: Wally Chandler (USA/CAA)
- Complex Data Types and Issues: Stephanie Cammarata (RAND)
- DMSO Information System and Directories: Cy Ardoin (IDA), Dennis Shea (CNA), Iris Kameny (RAND)
- Data VV&C Panel: Iris Kameny (RAND), Dennis Shea (CNA), Howard Haeker(Army/TRAC), Mike Barton (Army), Dave Danko (DMA), Simone Youngblood (John Hopkins/APL)

IDEF:
- John Grobmeier (Army), Bruce Rosen (NIST), Paul Rehmus (OSD/PA&E), Tom Shook (DMSO), Twyla Courtot (MITRE)

IRDS:
- Information Resource Dictionary System: Burt Parker (MITRE), Bruce Rosen (NIST)

JOINT ACTIVITIES:
- J-MASS: Bill Mc Quay (WL/AAWA-1)
- J-8 Data Management: OASIS project, Dan Hogg (J-8)
- A Modelbase Concept for Model Interoperability: Bob Sutter (Argonne Labs)

Listing of Major Briefings Given at I/DB Task Group Meetings (continued)

NAVY ACTIVITIES:

New Navy M&S Organization and M&S Data Support: Jim Weatherly (Navy)
Navy Universal Threat System for Simulators: Gail Coffey (GPSC)

RESEARCH ACTIVITIES

Asset Source for Software Engineering Technology (ASSET) DARPA repository system (part of STARS effort): Chuck Lillie (SAIC)
MITRE Distributed Heterogeneous Information System (DHIS) Testbed: Don Rea, Bill Carpenter, and Miro Medek (MITRE)
DARPA Intelligent Integration of Information Research Program: Gio Wiederhold (DARPA/SSTO)

The above briefings were given at the I/DB Task Group meetings in order to familiarize the group with relevant efforts addressing process and modeling tools, data administration, data element standards, data dictionary efforts, data and model directories/catalogs, repositories, IDEF methodology, projects supplying data to M&S developers and users; etc.

Impressions from the briefings are that there is widespread recognition of the need for data administration and management throughout DoD and efforts are underway in all of the Services and the Joint Staff as well as many of the DoD agencies. Current dictionary efforts include: DISA/CIM, AT&T 3B2 running Oracle/DDRS; Air Force, Microvax 5000/Ultrix running Sybase/MIDAS; Army, IBM running Oracle/ADSS; Joint Staff, VAX 8600 running Oracle/WISDIM; DIA, IBM running M204/IDEAS; and OASIS dictionary effort for J-8, SUN, Ingres/OASIS; and DLA is also working on a system. There appears to be no DoD or Joint effort directed toward distributed exchange among these dictionaries (the DDI/CIM office believes eventually they will all use the Defense Data Repository System (DDRS)). However, the DARPA repository project, ASSET, plans to develop distributed processing capability among repositories and the AF data administration program will address interchange among MIDAS distributed servers.

The DoD/CIM effort is directed toward creating a single integrated DoD Data Model and one DoD Data Dictionary maintained in the DDRS. This is a very ambitious task and Bob Molter, reporting on DDI policy and procedures, estimated it will take ten years to achieve this goal. Although attention is focused on the policy and procedures for the future when there is an established DoD Data Model and DDRS, there are some gaps as to what to do in the interim. Some attention is focused on how to handle legacy systems, including re-engineering and reverse engineering, and DISA is working on prototypes (i.e., the CIM supported OSD/PA&E MIDAS effort is an example) but they appear to be very human intensive undertakings.

Currently, (1) the DoD Data Model exists only at a high level (not detailed enough to yield entities on which to base SDEs); (2) there is approved policy for

8320.1 "DoD Data Administration," and 8320.1-M "Data Administration Procedures" and 8320.1-M-1 "DoD Data Element Standardization Procedures" are undergoing approval; (3) the DDRS software is immature but is up and running and dictionary partitions have been created to allow the component and functional data administrators to enter candidate data elements (generated outside of the DoD Data Model) to get things started; and (4) DISA/CIM is developing reverse engineering approaches to apply to legacy systems.

NIST is in the process of creating draft Federal Information Processing Standards (FIPS) for IDEF0 and IDEF1X methodologies based on the original Air Force documents. These are expected to be published in the Federal Register by first quarter FY93. There are some problems with the proposed drafts; see Appendix G for a discussion.

Outline

Goals and Scope

Perspectives on the current state of database technology

Initiatives identified by the DMSO Database TWG

Status and discussion of current CIM efforts

Summary

Areas of CIM/Component Data Administration Methodology Appropriate to M&S DD&R

1. **Business process models and data models: use of IDEF tools**
2. **Standard data element definition and management**
3. **Repository system for directories, dictionaries, databases, models, etc.**
4. **Legacy systems, re-engineering, reverse engineering**
5. **M&S directories for organizations, databases, and models**

A CIM data administration goal is to provide the road map to be followed by Component Data Administrators (CDAds) and Functional Data Administrators (FDAds) to develop the DoD data model and populate the Defense Data Repository System (DDRS) with the appropriate standard data elements needed to ensure sharing and interoperability. It currently appears that the process will be for CDAds and facilitators to help Component functional area experts in developing business process models of their areas and from those or concurrently with those, data models. The process and data modeling will identify entities that will form the basis for standard data elements (attributes of the entities). NIST has established draft standards for IDEF0, business process or functional modeling methodology, and IDEF1X, relational data modeling methodology. Currently, there are a number of commercial tools designed to support IDEF0 and IDEF1X.

CIM data administration status (as of October 1992) is:

8021.1-M, "Functional Process Improvement (Draft)," August 1992
8320.1 "DoD Data Administration," September 1991
8320.1-M, "Data Administration Procedures (Draft)," September 1992 is under review
8320.1-M-1 "DoD Data Element Standardization Procedures (Draft)," October 1992 is under review

8320.1-M-1 supports the development, naming conventions, approval process, and maintenance for standard data elements. This effort is addressing atomic data elements but currently does not address more complex data elements that are needed by the M&S community.

CIM has developed the Defense Data Repository System (DDRS) mainly to house and manage the standard data element dictionary. Initial procedures, DDRS training courses, and use of DDRS partitions for storing existing Component data elements were available in August 1992.

CIM is exploring the use of several reverse engineering tools to enable the extraction of data entities and data elements from legacy software systems that Components do not plan to re-engineer in the short term.

There seems to have been little attention paid to directories of databases until the recent efforts of AMSAA and CCTT. There have been a number of hard copy model catalogs mainly compiled by the Joint Staff. There is a recognized need for such in electronic form with proper maintenance support. The current effort to produce an Army master catalog of models and simulations is well thought out and was recommended as the basis for the DMSO model directory effort at the October 1992 I/DB Task Group meeting.

Common Issues Across Data Administration Areas

- **Security**
- **Distributed access to multiple DD&R sites**
- **Intelligent search and access**
  - Classification and typing: search structure and terms
  - Resolving/recognizing semantic similarities
- **Maintenance and consistency**
  - Version support
- **Policy and procedures information resource management**
  - Information resource management
  - Access and dissemination
  - Fiscal responsibilities

Security will be an issue for directories, dictionaries, and repositories. For simplicity's sake, the following discussion refers to a dichotomy of classified and unclassified Directories, Dictionaries, and Repositories (DD&Rs), though there may actually be several levels of classification and compartmentation which will require separation.

For directories, protection against unauthorized users gaining knowledge of the existence of a classified model or database will result in keeping the information from appearing as an entry in an unclassified directory. In data element dictionaries, the classified information would most likely be of enumerated domain values and the usage information furnished about a classified database or application that uses the standard data element (again, this protects against an unauthorized user gaining knowledge of a classified database or application). Repositories contain directories, dictionaries, databases, models, etc., any of which may have both unclassified and classified information.

There are two ways in which security can be handled. One is by creating a single multilevel secure directory, dictionary, or repository in which separation is enforced by the system, and the other is by maintaining separate DD&Rs at each classification level. At the current time there is a lack of commercially available multilevel secure products, so maintaining separate DD&Rs may be the best option. In either case, users would need to be made aware that classified entries may exist and that they may have to initiate multiple searches. Another security issue is that of data aggregation. Since DD&Rs contain large amounts of data that did not previously exist in a single collection, it may turn out that the aggregated data may be more sensitive than the classification level of the data it was aggregated from.

DISA/CIM is not currently addressing the security issue; however, the Air Force is addressing security in its MIDAS dictionary effort and this is a reason why Navy Intelligence and the Coast Guard will be using the MIDAS system. This may be an important issue for the M&S community.

Distributed access to multiple DD&R sites: the future will see an abundance of DD&Rs both inside and outside DoD that M&S users will need access to. Examples are: NASA (Earth Observing System Data and Information System (EOSDIS)), the EPA, the DOE, and other nations. So far, DISA/CIM has paid little attention to the need for a distributed architecture that would support easy access to heterogeneous databases (which is what the DD&Rs are).

Intelligent search and access: the key to reuse, sharing, and interoperability is not only to build and maintain the DD&Rs but to make them readily accessible (i.e., via protocols and the internet) and easy to search. Effort has to be put into techniques for developing appropriate search terms and structures to facilitate navigation. These should include aids to naming and aids in recognizing groupings, regrouping, and linking entries on the basis of semantic similarities. Though the DD&Rs will support reuse of models and data, consideration also has to be paid to partial reuse, and to how retailored data, objects, and models can be reintroduced and represented in the systems. This becomes more important as M&S developers utilize object-oriented technology and want to share, reuse, or modify existing objects and their methods.

Maintenance and consistency: replication of DD&R information for M&S use across DD&Rs or for M&S application use requires that attention be paid to maintaining currency and consistency of information. For example, if changes to applications result in changes in the standard data elements they use, that should be reflected in the usage for those standard data elements. If a database described in a directory was maintaining order of battle data for five countries and in later versions this changed to three countries, then there should be proper warning of the change in the new version information.

Policy and procedures: DMSO needs to establish policy and procedures for development, maintenance, and use of the DD&R resources. Three areas of policy have been identified related to developing and maintaining DD&Rs and making them easy to use and available to the M&S community. They are (1) information resource management (IRM), (2) access and dissemination, and (3) fiscal responsibilities.

Information Resource Management (IRM) includes Data Administration and Configuration Control and Management policies. Though it has been determined that DTIC will maintain directories and provide access to the M&S community, more than a library function is needed. A special group with M&S expertise and knowledge is necessary to decide which databases and models to acquire in directories, maintain the directories so that they indicate current versions and releases of databases and models, propagate version changes to users, maintain

multiple directories resulting from security and proprietary requirements and furnish the human link to help M&S users get access to the information they need.

The access to, and dissemination of, information requires clearly stated policy as to how DoD classification and proprietary restrictions will be met. These are needed for access by the M&S community to information about existence of databases/models, dissemination of data/models to M&S community, updating and propagation of changes, and archiving.

Fiscal responsibilities require the development of motivation, policies, and funding to enable the necessary parties to maintain DD&Rs. Policy could require M&S RFPs to have contractors (1) use CIM/M&S data element standards and other IRM methodology in defining new databases and configuration control of database and model development, (2) furnish information, data, code, documentation, etc. on databases, models, organization, points of contact, etc. to DMSO/DTIC, (3) clearly spell out responsibility for maintaining information furnished, and (4) furnish instructions for turning over completed products at end of contract.

The DMSO and the M&S community need to establish how access and use of directories, data, and products will be paid for. For example, when users need access to DD&Rs on remote systems, will DTIC login on their behalf and accept charges, make accounting entries, and later bill users, or will each user need to establish an account and login on his/her own to access DD&Rs not held by DTIC? How will use of proprietary models and databases (for which there may be license fees) be handled?

CIM Business Process Models and Data Models

- Is M&S a CIM functional area?
  - M&S acts as a functional area for organizations/offices that manage M&S activities
  - Developers of simulation models may use functional area process and data models to guide their simulation model design
- Functional process and data models may be located in various repositories
  - Knowledge of existence
  - Security access
  - Maintenance and consistency

One question that has been asked repeatedly is whether or not M&S is a CIM functional area. This is a question that was addressed at the I/DB June 1992 meeting. Although Dr. Kimmel, the FDAd for OSD/Acquisition, considers M&S a functional area of Acquisition, Lana McGlynn, from the Army Model and Simulation Management Office, thinks it is not. Lana has supported the Army M&S groups in using the IDEF0 methodology at a high level to develop process models of their organization's management of M&S activities. This seems a proper interpretation and use of business process modeling methodology for M&S.

However, the use of business process modeling methodology does not necessarily seem appropriate for the developer of an M&S model since the modeler is usually modeling activities from other functional areas (e.g., command and control). If these functional areas have developed business process models and data models as prescribed by CIM, then these models could be made accessible to simulation modelers as design input to the processes and data they need to develop and use in their simulation models. We are suggesting that this be supported by making business process models and data models accessible to simulation modelers through directory and repository access. To get full benefit from the use of these models requires that the functional users maintain the models as the business process changes.

CIM may still require developers of all software systems to furnish business process models and data models of their systems that will include M&S implementors.

Standard Data Element Definition, Naming, Management

- Process for defining standard data elements:
  - Functional area: perform process and data modeling, integrate data model into DoD data model, new entities become prime words in the naming process, attributes of entities become standard data elements
  - M&S developer: defines a new attribute of a data entity that needs to be entered into the standard data element nomination process:
    - Developer may check for existing standard data element in Defense Data Repository System (DDRS), in Component repository, etc. If not found, then nominate new data element to Component DAd
    - Component DAd checks for existing or close matching standard data element; if not found, nominates the data element to DoD DAd, who checks with functional DAd, and if approved, becomes new standard data element

The I/DB Task Group has been very interested in understanding the CIM procedure for determining data elements and nominating them into the standard data element process. The CIM procedures have not been officially released yet, but the flavor is captured above and reflects how it could work once there is a DoD data model and populated DDRS. As mentioned earlier, the M&S developer may want to make use of already existing process and data models developed for the functional area he/she is modeling. The modeler may want to acquire the functional area process and data models and possibly change them to fit the simulation model. In doing so, he/she may create new data elements which should be handled as described above.

The DMSO sponsored Joint Data Base Elements project is taking a bottom-up approach to defining standard data elements for M&S. Their approach is to teach database and M&S application users and developers how to use IDEF1X tools to develop project data models of their databases and applications. After this is done, they will group the data models into subject areas and form subject area information modeling groups composed of the people originating the data models and other subject area experts. Each group would work together to concur on the data model, entities, and data elements for their subject area, which would then be placed in a repository and used as the basis for developing new databases and applications as well as to develop mappings from legacy systems. An additional function would be to coordinate the subject area data models, entities, and data elements with the CIM DDRS effort.

Standard Data Element Related R&D Issues

- **Handling of complex data elements: not being addressed at this time by CIM**
  - **Correctly modeled:**
    composites (e.g., basic encyclopedia number)
    lists/sequences (road network)
    object (weapon and its parts)
    derived data (probability of hit)
  - **Poorly modeled:**
    civilian strength: [count of civilians | validity of count]
    aircraft types: [major family | sub family]
    aircraft capabilities [cap | cap | capl...]
    geographic installation intelligence production specifications
  - **Need to understand how to represent these as complex standard data elements**
  - **Need to map these (where appropriate) to standard atomic data elements**
- **Data Value Domain identification and management**
  - **Ranges, enumerated lists**
  - **Disjoint and overlapping sub-domains**
  - **Standard nomenclature for data values**

The I/DB Task Group has participated in the CIM meetings to review 8320.1-M-1, "DoD Data Element Standardization Procedures." As a result, we realize that CIM has restricted its initial focus to atomic data elements. "Through its name and definition a data element must convey a single informational concept" (DoD 8320.1-M-1, September 30, 1992, page 2–1).

We have also determined that the M&S community will require sharing and reuse of complex (i.e., non-atomic) data elements in order to interoperate across M&S applications. There is a need, therefore, to establish standard complex data elements. This includes correctly modeled (in the I/DB view) data elements which need to go through a nomination process and be entered into a dictionary, and poorly modeled data elements of legacy systems which will, at the least, need to be mapped to standard data elements to ensure their correct interpretation and usage by modelers.

We are beginning a study of complex data elements in order to classify them and develop a standard data element schema that will capture their descriptions.

Some preliminary examples of correctly modeled complex data elements were prepared for the DMSO to use at a CIM meeting. They include composites, such as the Basic Encyclopedia Number, which is a data element composed of a World Area Chart Number and an Installation Number. The CIM focus on atomic data elements would recognize the components of the Basic Encyclopedia Number as standard data elements but would not make the Basic Encyclopedia Number itself a standard data element, yet it is a well-known and used data element across the DoD community. Similar examples include the Army Standard Requirements Code and the JOPES TPFDD Unit Line Number.

Other examples of composites include road networks composed of road segments with an implied ordering and objects that may be composed of other objects and may include methods. Objects are an extremely important composite to address since many in the DoD M&S community are beginning to use object-oriented technology. The desire to reuse and retailor objects raises many questions because objects consist of both data element structures and program code (methods) to be applied to the objects. Among other questions, Should objects be part of a reuse software library and/or be captured in a data element dictionary? Obviously, it is important for reuse and sharing to capture the structure of an object in the dictionary/repository and to be able to determine object similarity and differences. There are many research issues having to do with handling heterogeneous object bases, such as being able to recognize similarities across different object bases even though the structures are different, and to be able to correctly detach an object from a larger object structure.

An example of a complex derived data element is P(h) (probability of hit), which is a general shared concept across the M&S community but varies from simulation to simulation in the way in which it is derived. It is a function of such factors as firing system, gun system, ammo, tank position, target position, target size, range, environmental conditions, etc.--but the combination of factors and the function may differ across simulations. Our intent is to be able to capture, in the standard data element schema, the indication that this data element has a derived value and that the derivation function varies. In the usage information (i.e., the part of the definition that explicitly defines the meaning of the data element in a database or its use by an application), we would want to provide a list of the standard data elements that participate in the derivation.

Domains present another important area of study. In order for simulations to interoperate, it is necessary that the domains of their linking data elements be the same or similar. For example, a simulation of environmental pollution in North America won't be valid if its water pollution input comes from a water pollution model that only uses data about the U.S. and Mexico. Domains need to be defined and represented in the dictionary/repository. A difficult issue is how to represent disjoint and or overlapping subsets of domains and the relationships between them in such a way that they are easily maintainable and modelers can easily understand

them. The standardization of domains also requires defining nomenclature standards for the data values in the domain. A successful example of this is the established standards for names of countries of the world and country codes. Some examples of problems are the many different ways dates, weapon systems, and people's titles, names, and addresses are currently represented.

Legacy Systems: Reverse Engineering, Re-engineering Issues

- Reverse engineering tool methodology and evaluation
  - –Affordability of use with respect to human capital
- Re-engineering
  - –Transitioning from old to new
    - Revolutionary
    - Evolutionary

Since re-engineering an application is very expensive and time consuming, it is obvious that the DoD community will have to live with legacy systems for many years. CIM is interested in finding new methodologies to assist in semi-automatically developing process and data models from legacy systems in order to enhance the DoD data model by including standard data elements used by legacy systems, and also to capture the usage of existing standard data elements by legacy systems. CIM is currently trying out several exploratory reverse engineering tools on two "simple" legacy systems. This effort should be completed by October 1992. At the October I/DB Task Group meeting, OSD/PA&E briefed a project they are embarking on with CIM help to generate the "as is" process and data models for the MIDAS transportation model, which is a more complex example than the other reverse engineering efforts. The real concern with reverse engineering is the amount of human capital that may be required, even with tool support, to really understand the workings of a poorly documented legacy system.

Re-engineering implies a re-design and re-implementation of a legacy system into a new system that may retain the same functionality of the old system or provide additional enhancements. Issues to be addressed are how to transition from the legacy system to the new system, especially if the legacy system is used by many applications, e.g., such as a database system. The new design should accommodate graceful system evolution in the future.

M&S Directories for Organizations, Databases, Models, and Directories

No explicit DISA/CIM plans at this time

- Report on "Concept for Integrated DoD Directory Services," 27 November 1991 (page 6.1) proposed plan with deployment in FY93

Status of DMSO directories:

- Initial relational schema for directory of database directories and databases
- Recommendation to base DMSO directory of models on the "Design Documentation for the Construction of Department of the Army Master Models and Simulation (M&S) Catalog" with a few additions
- Also investigating the directory schema required for environments for developing models and simulations (e.g., J-MASS; RAND: Anabel, Exploratory Modeling, TLC)

The I/DB Task Group is unaware of any definite DISA/CIM plans for directories at this time. We have a DISA sponsored report, "Concept for Integrated DoD Directory Services," dated 27 November 1991, that contains a proposed plan for deployment in FY93 with implementation to have started in FY91. We are also aware of an Interim Report to the ASD/C3I dated June 1991 on a C3I Data Base Survey that lists 164 systems/databases. There was indication that a follow-on survey would also be sent out that would provide a summary about the systems/databases to include name and nomenclature, function or use, names and nomenclature of other systems with which it exchanges data, and point of contact (POC). In an informal discussion, we were told that DISA/CIM believes a DoD database directory to be important but that it probably wouldn't be implemented before 1994.

Since the M&S community has made database and model directories a high-priority need, the I/DB Task Force is building prototype database and model directories. The database directory schema is complete except for the search term relations and will be initially implemented at IDA on the Oracle relational DBMS. CNA investigated the requirements for the model directory and at the October I/DB Task Group meeting recommended basing it on the Army M&S catalog with some specific additions. An issue that arose at that meeting was whether the M&S directory schema would be adequate for describing the new M&S environments that are being developed for designing, implementing, and running models. This issue is currently under investigation.

Verification, Validation, and Certification (VV&C) of M&S Data

- Data VV&C is of great interest to the M&S community
  - Has the potential to enhance verification, validation and accreditation (VV&A) of models
  - Is a controversial subject among M&S community and data suppliers to modelers
- What is data VV&C?
  - Data verification: accomplished by applying constraint tests to data values to ensure they are reasonable
    - data value constraints derived from domain information: range, enumerated list of values, rules
    - data set constraints determined through use of statistical measures
  - Data validation: accomplished by comparing consistency or similarity to existing test data: utilize statistical and other methods
  - Data certification: approval of data set(s) by approval agency for use in an application or a type of application
- Issues include: definitions, methodology, management procedures and guidelines, understanding the VV&C role in relation to model VV&A
- RAND M&S VV&A researchers and data VV&C researchers will be addressing the two areas together

The VV&A of models must include the VV&C of the data for these models. The VV&C issue is always important but becomes even more so when we begin to tie distributed simulations together (e.g., the output of one model becomes input to another). We need to understand, not only the data and its source and preparation, but also the purpose of the models particularly when trying to tie models together of different resolutions that were not designed to run together.

There is some controversy over data certification and what is meant by it. One view is that the "real world" is not real; models are scenario specific and use anecdotal evidence--thus ultimate certification cannot be done against the real world. We can say data are consistent or credible but not certified.

Data used in M&S are sometimes generated and distributed by collection agencies (e.g., DMA, DIA) in a form that is not specifically designed for particular models, but could be used in a number of models and systems. Other data (e.g., TADS, OASIS data) are collected and designed specifically for a group of models or for a particular model and problem the model is addressing. It is extremely

important to understand that the data must be consistent with the purpose of the model. For example, weapon characteristics data may exist in many forms: values of characteristics as found in the specification, values as collected on test ranges under various conditions, and values as collected in combat. The data may be collected and represented in many different ways and resolutions. Experience shows that modelers have, in the past, inadvertently mixed the types of data in a model by not realizing that data are identified not only by name but also by source and method of collection and transformation.

Steps in data verification include: identifying source data, selecting/addressing data version issues, verifying the source data, converting the source data to model formats and reverifying, and verifying data values during model execution. Methods of verification include: use of domain constraints, use of higher order knowledge such as rules, use of statistical or set operators over a dataset, and use of application specific techniques such as superimposing map data on an image and checking for common sense errors.

The difference between data verification and validation needs to be defined, as does certification. The M&S community needs a methodology for VV&C so that degrees of certification and the responsibilities of the certifying agent are well understood and reliable. VV&C like VV&A is expensive, requiring a management structure (probably from the DoD level down to the individual organization) to assign responsibility and to run assurance checks that VV&C is being carried out properly. VV&C must be planned ahead of time and must be an integral part of data preparation for M&S.

Outline

Goals and Scope

Perspectives on the current state of database technology

Initiatives identified by the DMSO Database TWG

Status and discussion of current CIM efforts

Summary

Summary

- Accomplishment of CIM objectives will go a long way toward meeting M&S needs for sharing and reusing data and improving model interoperability, but DMSO needs to:
  - Continue monitoring progress and developing relevant policy and procedures to make use of CIM methodology and products
  - Continue development of DMSO Information System
- But certain areas are not being addressed by CIM:
  - R&D areas
    - Understanding and standardization of non-atomic data elements
    - VV&A of data
    - Advancing technology for classification and typing to aid in intelligent search and access of DD&Rs
    - Domain standardization
    - Security threat from data aggregation
  - Engineering areas
    - Development of database and model directories
    - Distributed architecture for DD&R sites
    - DD&R security

CIM is addressing many of the data related needs of the M&S community—but not all. It is important for the M&S community to be aware of the data needs that are not being met by CIM and are unlikely to be met by commercial or other DoD means. These data needs should be addressed by the M&S community. It is critical that the I/DB Task Group continue to monitor CIM activities and help DMSO develop compatible M&S guidelines and procedures whenever possible while pointing out possible incompatibilities with CIM. It is also critical that DMSO continue the development of the DMSO Information System so that the M&S community can share information about M&S happenings, projects, databases, models and simulations, organizations, etc.

The R&D data areas critical to the M&S community that are not being addressed elsewhere include: (1) developing an understanding, methodology, and standardization for complex data elements (e.g., rules, objects, networks, images, voice, matrices, etc.); (2) developing data VV&C definitions, methodology, management procedures, and guidelines in coordination with the M&S community VV&A needs; (3) developing advanced methods for classifying, locating, and accessing information in distributed DD&Rs; (4) developing methods for standardizing domain values, icons, and graphical representations, and addressing the representation and manipulation of domains; and (5) addressing the security threat resulting from the use of aggregation and inference techniques applied to the large DD&R data collections.

The data engineering areas critical to the M&S community that are not being addressed elsewhere include: (1) developing database, model, and organization directories as part of the overall development of the DMSO Information System (currently being done); (2) addressing the issue of distributed architecture and access to heterogeneous DD&Rs; and (3) addressing the issue of managing and accessing multi-level information in and across DD&Rs.

Appendix A

LIST OF ACRONYMS

4GL	Fourth Generation Language
ADD	Army Data Dictionary
ADP	Automated Data Processing
ADSS	Army Data Standardization System
AF	Air Force
AFIRDS	Air Force Information Resource Dictionary System
AIS	Automated Information System
AMSAA	Army Materiel Systems Analysis Activity
AMSEC	Army Modeling and Simulation Executive Council
AMSMO	Army Modeling and Studies Management Office
ANSI	American National Standards Institute
APL	Applied Physics Laboratory
APP	Applications Portability Profile
ASSET	Asset Source for Software Engineering Technology
ATIS	Atherton Tool Integration Set
BCA	Business Case Analysis
C2	Command and Control
C2I	Command and Control and Intelligence
C3I	Command, Control, Communications, and Intelligence
CALS	Computer-Aided Acquisition and Logistics
CAM	Cc mputer-Aided Manufacturing
CASE	Computer-Aided Software Engineering
CCITT	Consultative Committee on International Telegraph and Telephone
CCR	Concurrency, commitment, and recovery
CCTT	Close Combat Tactical Trainer
CDA	Computer Design Activity
CDAd	Component Data Administrator
CDIF	CASE Data Interchange Format
CFS	Center for Standards
CIM	Corporate Information Management and Center for Information Management
CINC	Commander in Chief
CIS	CASE Integration Services
CNA	Center for Naval Analysis
COE	Common Operating Environment
COMRATT	CIM Repository Architecture Tiger Team
COTS	Commercial-Off-The-Shelf
CPU	Central Processing Unit
DA	Data Administrator
DAC	Data Administration Council
DAd	Data Administrator
DAMA	Data Administration Management Association
DARPA	Defense Advanced Research Project Agency
DASP	Data Administration Strategic Plan

DB	Data Base
DBMS	Data Base Management System
DBTWG	Data Base Technology Working Group
DD	Data Dictionary
DDI	Director of Defense Information
DDI/CIM	Director of Defense Information/Corporate Information Management
DDM	Distributed Data Management
DDN	Defense Data Network
DDRS	Defense Data Repository System
DD&Rs	Directories, Dictionaries, and Repositories
DE	Data Element
DHIS	Distributed Heterogeneous Information System
DIA	Defense Intelligence Agency
DISA	Defense Information Systems Agency
DISA/CIM	Defense Information Systems Agency/Center for Information Management
DLA	Defense Logistics Agency
DMA	Defense Mapping Agency
DMSO	Defense Modeling and Simulation Office
DOD	Department of Defense
DOE	Department of Energy
DON	Department of the Navy
DTIC	Defense Technical Information Center
E-R	Entity-Relationship (Model)
EDI	Electronic Document Interchange
EDIF	Electronic Document Interchange Format
EIA	Electronic Industry Association
EOSDIS	Earth Observing System Data and Information System
EPA	Environmental Protection Agency
EXCIMS	Executive Council for Models and Simulations
FDAd	Functional Data Administrator
FFRDC	Federally Funded Research and Development Center
FIM	Functional Integration Manager
FIPS	Federal Information Processing Standard
GAO	General Accounting Office
GB	Gigabyte
GIS	Geographic Information System
GKS	Graphics Kernel System
GOSIP	Government Open System Interconnection Profile
GUI	Graphical User Interface
I/DB	Information/Data Base (Task Group)
I/O	Input/Output
ICASE	Integrated Computer-Aided Software Engineering
IDA	Institute for Defense Analysis
IDEAS	Intelligence Data Element Authorization Standards
IDEF	Integrated Computer-Aided Definition Language
IDEF0	Activity or process model methodology

IDEF1X	Data model methodology
IGES	Initial Graphics Exchange Specification
IOC	Initial Operational Capability
IRDS	Information Resource Dictionary System
IRM	Information Resource Management
ISO	International Standards Organization
IT	Information Technology
ITWG	Information Technology Working Group
JB	Juke Box
JDBE	Joint Data Base Elements (Project)
JIEO	Joint Interoperability Engineering Organization
J-MASS	Joint Modeling and Simulation System
JOPES	Joint Operation Planning and Execution System
JS	Joint Staff
LAN	Local Area Network
M&S	Modeling and Simulation
MAC	Military Airlift Command
MAISARC	Major Automated Information System Review Council
MB	Megabyte
MCC	Micro Electronics and Computer Technology Corporation
MC&G	Mapping, Charting, and Geodesy
MIDAS	MAC Integrated Data Administration System; also the name of a PA&E simulation model
MIIDS	Military Intelligence Integrated Data System
MIKE	"Team MIKE" Naval Warfare Analytical/Modeling and Simulation Oversight Council (NMSOC) is known as Team MIKE
MISMA	Model Improvement System Management Agency (Army)
MLS	Multi-Level Security
MSRL	(J-MASS) Modeling and Simulation Reuse Library
MSTS	SPAWAR 31 Modeling and Simulation Technical Support
MTF	Message Transfer Format
NATO	North Atlantic Treaty Organization
NDL	Network Data Language
NFS	Network File Server
NIST	National Institute of Standards and Technology
NMSOC	Naval Warfare Analytical/Modeling and Simulation Oversight Council
NSF	National Science Foundation
NTC	National Test Center
NWTDB	Naval Warfare Tactical Data Base
OASIS	Operations Analysis and Simulation Interface System
ODA	Office Document Architecture
ODP	Open Distributed Processing
ODUSA(OR)	Office of the Deputy Under Secretary of the Army for Operations Research
OMG	Object Management Group
OO	Object-Oriented

ORB	Object Request Broker
OSD/PA&E	Office of the Secretary of Defense/Policy Analysis and Evaluation
OSE	Open Systems Environment
OSF	Open Software Foundation
OSI	Open Systems Interconnection
OSRM	Open System Reference Model
PCTE	Portable Common Tools Environment
PDES	Product Data Exchange using STEP
PHIGS	Programmer's Hierarchical Interactive Graphics System
POC	Point of Contact
POSIX	Portable Operating System Interface (for Computer Environments)
PPDB	Point Precision Data Base
R&D	Research and Development
RAPID	Reusable Ada Packages for Information System Development (Army)
RDA	Remote Data Access
RDBMS	Relational Data Base Management System
RFP	Request For Proposal
SAI	Subject Area Information (Model)
SDE	Standard Data Element
SORTS	Status of Resources and Training System
SQL	Standard Query Language
STARS	Software Technology for Adaptable Reliable Systems
STEP	Standard for Exchange of Product Model Data
SUMM	Semantic Unifications Meta-Model
TADS	TRAC Automated Data System
TP	Transactions Processing
TPFDD	Time Phased Force Deployment Data
TRAC	TRADOC Analysis Command
TRADOC	Training and Doctrine Command
TRM	Technical Reference Model
TWG	Technical Working Group
USA/CAA	U.S. Army Concepts Analysis Agency
USD(A)	Under Secretary of Defense (Acquisition)
USMTF	U.S. Message Transfer Format
USTRANSCOM	U.S. Transportation Command
UTSS	Universal Threat System Simulator
VV&A	Verification, Validation, and Accreditation (of models)
VV&C	Verification, Validation, and Certification (of data)
WAM	WWMCCS ADP Modernization
WAN	Wide Area Network
WISDIM	Warfighting and Intelligence Systems Dictionary for Information Management or WWMCCS Information Systems Dictionary for Information Management
WWMCCS	World Wide Military Command and Control System

Appendix B

DOCUMENT REFERENCES FOR I/DB TASK GROUP (UPDATED TO JUNE 1993)

Aerospace Corp., "Panel Review of Aggregate Level Linkage Technology," Authors: Danny Cohen, David Mills, Richard Nance, Paul Reynolds, Mike Sovereign, Donna Vargas, Pat Cheatham, and Robert Weber, El Segundo, California, 1 February 1993.

ANSI X3.1772-1990 "American National Standard for Information Systems—Dictionary for Information Systems," 1991 (adopted as FIPS Pub 11-3).

ATCCIS, "ATCCIS Battlefield Generic Hub Data Model," Edition 1.0, ATCCIS Working Paper 5-2, 23 April 1993.

ATCCIS, "ATCCIS Conventional Fire Support Data Model," ATCCIS Working Paper 5-2B, 1993.

COE Working Group, "Joint Service Common Operating Environment (COE) Common Geographic Information System Functional Requirements," January 28, 1992 (prepared by Wayne D. Meitzler, Pacific Northwest Laboratory).

DDI, "Corporate Information Management Functional Economic Analysis Guidebook," Version 1.1, 15 January 1993.

DDI, "The DoD Enterprise Model, A White Paper," February 1993.

DDI, Briefing from "The DoD Enterprise Model Symposium," March 22, 1993.

DDI (Office of), Briefing on Strategic Framework DoD Corporate Information Management (CIM), Paul Strassmann, OASD/C3I Meeting, November 26, 1991.

DISA: "DoD Information Technology Standards Management Plan," DISA Center for Standards, 17 January 1992.

DISA/CIM, "DoD Data Administration Strategic Plan Fiscal Year 1992 (Draft)," Defense Information Systems Agency Center for Information Management, 18 February 1992.

DISA/CIM, "Technical Reference Model for Corporate Information Management," DISA Center for Information Management, Version 1.1, November 27, 1991 (obsolete), Version 1.2, March 1, 1992 (obsolete).

DISA/CIM Center for Data, "Migration Systems Integration and Standardization," received by Jim Shiflett 27 April 1992.

DISA/CIM, Briefing, "Purpose: to present an information reference model for consideration by the Information Standards Management Group," April 1992.

DISA/CIM, "Department of Defense Technical Architecture Framework for Information Management, Volume 1, Implementation Concept," Version 1.1, October 21, 1992.

DISA/CIM, "Department of Defense Technical Architecture Framework for Information Management, Volume 2, Architecture Guidance and Design Concepts," Version 1.1, October 21, 1992.

DISA/CIM, "Department of Defense Technical Architecture Framework for Information Management, Volume 3, Reference Model and Standards Profile," Version 1.3, December 31, 1992.

DISA/JIEO Plan 3200, March 1992, "Department of Defense Information Technology Standards Management Plan," March 1992.

DISA/JIEO/CFS, "Information Standards White Paper (Draft)," DISA JIEO CFS, 31 March 1992.

DISA/JIEO/CFS, Draft Agenda for April 2, 1992 meeting of Information Standards Management Group (ISMG), charter, and list of representatives.

DoD, Directive 5000.11, Subject: Data Elements and Data Codes Standardization Program, December 7, 1964.

DoD, DOD-STD-2167A, "Military Standard Defense System Software Development," 29 February 1988.

DoD, DoD Standard 7935A, "DoD Automated Information Systems (AIS) Documentation Standards," October 31, 1988.

DoD, DoDD 8320.1, "DoD Data Administration," 26 September 1991.

DoD, Draft "DoD Software Technology Strategy," prepared for Director of Defense Research and Engineering (DDR&E) in Partial Fulfillment of the DDR&E Software Action Plan, December 1991.

DoD, "DoD Data Administration Framework, Appendix A, Draft," February 1992.

DoD, Information Paper Interim Department of Defense (DoD) Data Dictionary (given out at the DSMO meeting, March 30-31, 1992).

DoD, DoD Data Administration Overview, briefing at JIEO by Dawn Hughes, April 1, 1992.

DoD, DODD 8000.1, "Defense Information Management Program."

DoD, Directive 8020.1-M (Draft), "Functional Process Improvement," August 1992.

DoD, "DOD Data Repository System (DDRS)," procedures published August 1992.

DoD, "DDRS End User Manual," August 24, 1992.

DoD, DODD 8320.1, "DoD Data Administration," September 26, 1991.

DoD, 8320.1-M draft, "Data Administration Procedures," September 1992, in formal coordination as of November 1992.

DoD, 8320.1-M-1, "Data Element Standardization Procedures," in formal coordination as of October 1992.

DoD, Defense Management Report Decision (DMRD) 918, on Defense Information Infrastructure, September 15, 1992.

DoD, Memorandum, Subject: Defense Information Infrastructure and Integrated Computer-Aided Software Engineering (I-CASE), 7 May 1993.

DoD, Draft DoD Directive 50xx.x, Models and Simulation.

DSB, Defense Science Board, "Impact of Advanced Distributed Simulation on Readiness, Training and Prototyping," Report of the DSB Task Force on Simulation, Readiness, and Prototyping, January 1993.

Federal Computer Week, "White Paper CIM Corporate Information Management," September 1991.

FIPS PUB 11-3, "American National Standard for Information Systems—Dictionary for Information Systems," 1991 (adopted from ANSI X3.172-1990).

GAO, GAO/IMTEC-91-35, "Defense ADP Corporate Information Management Initiative Faces Significant Challenges," April 1991.

IEEE: The Institute of Electrical and Electronics Engineers (IEEE), "IEEE Standard Glossary of Modeling and Simulation Terminology," IEEE STD 610.3-1989, 1989.

IDA, "Clearing House Structure," Cy Ardoin, March 30, 1992.

IDEF Software Products (handout at March 30-31 meeting), brochures from D. Appleton, Inc., META Software Corp., and Knowledge-Based Systems, Inc.

(IDEF), AFWAL-TR-4023 Volume IV, "Integrated Computer-Aided Manufacturing (ICAM) Architecture Part II, Volume IV - Function Modeling Manual (IDEF0)," SofTech, Inc., Waltham, MA 02154, June 19

(IDEF), AFWAL-TR-4023 Volume V, "Integrated Computer-Aided Manufacturing (ICAM) Architecture Part II, Volume V - Information Modeling Manual (IDEF1)," SofTech, Inc., Waltham, MA 02154, June 1981.

(IDEF), AFWAL-TR-4023 Volume VI, "Integrated Computer-Aided Manufacturing (ICAM) Architecture Part II, Volume VI - Dynamics Modeling Manual (IDEF2)," SofTech, Inc., Waltham, MA 02154, June 1981.

(IDEF), AFWAL-TR-86-4006 Volume V, Part 1, "Integrated Information Support System (IISS) Volume V - Common Data Model Subsystem, Part 4 - Information Modeling Manual - IDEF1 Extended," General Electric Company, Schenectady, New York 12345, November 1985.

IDSS, Briefing on "The Interoperability Decision Support System (IDSS)," Major Thomas J. Zuzack, 30 March 1992.

ISMG, "Charter Information Standards Management Group (ISMG)," Draft, April 1992.

ISMG, "Operating Procedures Information Standards Management Group (ISMG)," Draft, April 1992.

ISO/DIS 9594-2 ISO/CCITT Directory Convergence Document #2, CCITT Draft Recommendation X.501 (Version 8), The Directory Model, Gloucester, November 1981.

JCS Pub 1-02, "Department of Defense Dictionary of Military and Associated Terms," 1 December 1989.

Jones, Mark, "Brave New World: A Vision of IRDS," *Database Programming and Design*, November 1991.

J-MASS-TDA-1.0, "Joint Modeling and Simulation System (J-MASS) Terms, Definitions, and Acronyms," Version 1, 15 August 1991.

JTC3A, Memorandum to ASD C3I, Subject: Interim Report on C3I Data Base Survey, June 1991.

Law, A. M., and W. D. Kelton, *Simulation Modeling and Analysis*, McGraw-Hill, New York, 1982.

MITRE, "Concept of Operations for the Joint Information Resources Organizational Memory," W151-M-136 304S, 11 October 1991.

MITRE, "Department of Defense Data Administration Strategic Plan Fiscal Year 1992," Draft, 18 February 1992.

MITRE, Information Resource Dictionary System (IRDS), briefing given by Burton Parker at the DMSO meeting March 30, 1992.

MITRE, MTR-91W00198, "A Functional Economic Analysis Reference Methodology," Draft Report, January 15, 1992.

MORS, Draft Briefing Slides for MORS Workshop Series on Simulation Validation.

MORS: Military Operations Research Society (MORS), "SIMTAX -- A Taxonomy for Warfare Simulation," Workshop Report, 27 October 1989.

NBS Special Publication 500-149, "Guide on Data Entity Naming Conventions," Judith J. Newton, October 1987.

NBS Special Publication 500-152, "Guide to Information Resource Dictionary System Applications: General Concepts and Strategic Systems Planning," April 1988.

NBSIR 88-3700, "A Technical Overview of the Information Resource Dictionary System (Second Edition)," Alan Goldfine and Patricia Konig, January 1988.

NIST, FIPS 119, "Ada," ANSI/MIL-STD-1815A-1983, April 1986.

NIST, FIPS 127-1, "Database Language SQL," ANSI X3.135-1989 & X3.168-1989, February 1990.

NIST, FIPS 156, "Information Resources Dictionary System (IRDS)," November 1988.

NIST Special Publication 500-167, "Information Management Directions: The Integration Challenge," Elizabeth N. Fong and Alan H. Goldfine, September 1989.

NIST Special Publication 500-173, "Guide to Data Administration," Bruce K. Rosen and Margaret H. Law, October 1989.

NIST, FIPS 127-1, Database Language SQL (ANSI X3.135-1989 & X3.168-1989), February 1990.

NIST, "Application Portability Profile (APP), The U.S. Government's Open Systems Environment Profile OSE/1 Version 1.0," NIST SP-500-187, April 1991.

NISTIR 88-3896, "Proceedings of the Federal Information Processing Standards (IFIPS) Workshop on Information Resource Dictionary Systems (IRDS) Applications," Alan Goldfine (ed.), December 1988.

IST-CF-89-1, Summary Report: The First Conference on Standards for the Interoperability of Defense Simulations, 1989.

IST-CF-90-01, Summary Report: The Second Conference on Standards for the Interoperability of Defense Simulations, 1990.

IST-CF-90-13, Summary Report: The Third Workshop on Standards for the Interoperability of Defense Simulations, 1990.

IST-CF-91-11, Summary Report: The Fourth Workshop on Standards for the Interoperability of Defense Simulations, 1991.

IST-CF-91-13, Summary Report: The Fifth Workshop on Standards for the Interoperability of Defense Simulations, 1991.

IST-CF-92-02, Summary Report: The Sixth Workshop on Standards for the Interoperability of Defense Simulations, 1991.

IST-CR/92/12, Military Standard Version 2.0 (Draft), "Protocol Data Units for Entity Information and Entity Interaction in a Distributed Interactive Simulation," Institute of Simulation and Training, Orlando, Florida, September 4, 1992.

IST-CR/92/13, Military Standard Version 2.0 (Draft), Appendices A-J Designation Information, "Protocol Data Units for Entity Information and Entity Interaction in a Distributed Interactive Simulation," Institute of Simulation and Training, Orlando, Florida, September 4, 1992.

IST-PD-92-1, "Rationale Document for Entity Interaction Protocol for Distributed Interactive Simulation," 1992.

IST-PD-92-2, "Military Standard (Draft): Communication Architecture for Distributed Interactive Simulation," 1992.

OASD, Memorandum for Paul Strassmann, Subject: Implementation of Department of Defense Data Standards, September 16, 1991.

OASD, Memorandum for Paul Strassmann, Subject: Roles and Functions in Implementation of Corporate Information Management (CIM) System, July 29, 1991.

RAND, "RMMS: The RAND Metadata Management System," S. Cammarata, I. Kameny, J. Lender, and C. Replogle, July 1991.

RAND, "Report of the Database Technology Working Group (DBTWG) (August - November 1991), January 1992.

RAND, "Database Technology Assessment for Modeling and Simulation Presented 11 August 1992 to the Defense Science Board Summer Study on Modeling and Simulation," Iris Kameny, August 1992. (This document is being finalized and will also include notes for I/DB Task Group meetings held in February 1992, March 1992, June 1992, and October 1992.)

Standards forums: List of Commercial, Federal, and Military Information Standards Forums (received from Jim Shiflett April 1992 with caveat that the list is a year old).

USA, Memorandum to Program Manager for Training Devices, Subject: Data Modeling Requirements, COL Gilbert Brauch, January 16, 1992.

USA, U.S. Army Model Improvement and Study Management Agency (MISMA), "Army Model and Simulation Management Program (Draft)," 18 July 1991.

USTRANSCOM, Draft report, "United States Transportation Command Transportation, Command, Control, Communications, and Computer System Model Catalog," USTRANSCOM Office of Information Resources, March 1, 1989.

WISDIM, UM 1-91, "Warfighting and Intelligence Systems Dictionary for Information Management (WISDIM) Users Manual," 22 April 1991.

WISDIM, TR 277-91, "Warfighting and Intelligence Systems Dictionary for Information Management (WISDIM) Consolidated Release Package," 22 April 1991.

Appendix C

INFORMATION/DATA BASE TASK GROUP MEMBERS (NOVEMBER 1992)

NAME	ORG	PHONE	EMAIL AND/OR FAX
Cy Ardoin	IDA	703-845-6647	ardoin@ida.org
Erwin Atzinger	AMSAA	410-278-6576	erwin@cleo.amsaa.mil
Bob Bishop	DTIC	703-274-7662	rbishop@dgis.dtic.dla.mil
Don Blanton	AMSAA	410-278-3368	blanton@cleo.amsaa.mil
Larry Buchsbaum	Navy	215-441-1534	buksbaum@nadc.navy.mil
Stephanie Cammarata	RAND	310-393-0411	steph@rand.org
Bill Carpenter	MITRE	703-883-5777	wcarpent@mitre.org
Wallace Chandler	USA/CAA	301-295-1692	fax 301-295-1834
Patrick Cheatham	Aerospace	703/318-5403	psc@aero.org
Bill Clydesdale	SPAWAR 312	703-602-5696	fax 703-602-5891
Gail Coffey	GPSC	703-271-7700	fax 703-271-8566
Martin Costellic	Defense TPDC	407-658-5158	tpdc053@tpdc.navy.mil
Twyla Courtot	MITRE	703-883-7343	courtot@mitre.org
David Danko	DMA	703-285-9236	fax 703-285-9396
Tim Doane	GPSC	703-271-7700	fax 703-271-8566
Linda Donaldson	USA/MRJ	703-385-0886	fax 703-385-4637
Ed Fitzsimmons	DMSO	703-998-0660	fitzsim@charm.isi.edu
Charlotte Gross	OSD/ODDI	703-746-7939	fax 703-746-7396 cgross@ddi.c3i.osd.mil
Lucy Haddad	RCI-CCTT	407-282-1451	fax 407-658-9541
Howard Haeker	USA/TRAC	913-684-3030	haekerh@trac.army.mil
Becky Harris	CIM/XF	703-285-5381	fax 703-285-5403 harrisb%cimtys%cim@ cimgate.disacim.osd.mil
Ollie Hedgepeth	Army	804-728-5832 804-727-4349	polsley@rand.org (Attn: O.Hedgepeth)
Dan Hogg	JCS/J8	703-697-8899	fax 703-693-4601
Mike Hucul	J-MASS	513-255-3969	huculmp@p2.ams. wpafb.af.mil
Roseann Hynes	DMA	703-285-9250	fax 703-285-9396

Iris Kameny	RAND	310-393-0411	kameny@rand.org
LtCol Ken Konwin	AFSAA/SAG	703-697-5616	fax 703-697-3441
Jim Lacey	SAIC/TMA233	703-553-6203	fax 703-979-2505
Steve Lawyer	IDA	703-845-6696	lawyer@ida.org
Mike Lilienthal	DMSO	703-998-0660	fax 703-998-0667 mlilient@dgis.dtic.dla.mil
Ernie Lucier	NASA/SED	202-358-0772	nhqvax.hq.nasa.gov
COL. Mike Mancino	OASD(PA&E)	703-695-4290	
Steve Matsuura	USAEPG/JDBE	602-538-4967	matsuura@huachuca-emh7.army.mil
Janet McDonald	USAEPG/JDBE	602-538-4958	mcdonald@huachuca-emh7.army.mil
Lana McGlynn	AMSMO	703-746-8075	fax 703-746-8068
Bill McQuay	WL/AAWA-1	513-255-4429	fax 513-476-4746
Miro Medek	MITRE	703-883-7877	medek@mitre.org
Bob Molter	OSD/ODDI	703-746-7926	rmolter@c3i.osd.mil
Bao Nguyen	Air Force	703-693-8691	nguyen@saf3.hq.af.mil
Jack Nicklas	RCI-CCTT	703-893-6120	fax 703-893-0917
Dale Pace	JHU/APL	301-953-5650	dkp@aplcomm.jhuapl.edu
Don Rea	MITRE	703-883-6249	drea@mitre.org
Paul Rehmus	OSD(PA&E)	703-695-0539	fax 703-695-0707
Roy Reiss	AFSAA/SAG	703-695-5552	fax 703-697-3441
Lauri Rohn	OSD (PA&E)	703-695-1681	fax 703-693-5707
Bruce Rosen	NIST	301-975-3246	rosen@ecf.ncal.nist.gov
Pat Sanders	OSD/PA&E	703-697-3521	fax 703-693-5707
Mike Sarkovitz	NAVAIRSYS COM	703-692-2641 x3006	fax 703-692-1517 michaelsarkowitz @airerew@pma205.navair.navy.mil
Roberta Schoen	DTIC	703-274-5367	rschoen@dgis.dtic.dla.mil
Dennis Shea	CNA	703-824-2352	dshea@pinet.aip.org
Jim Shiflett	CCTT	703-998-0660	shiflett@charm.isi.edu
Tom Shook	DMSO	703-998-0660	tshook@dgis.dtic.dla.mil
Cassandra Smith	MITRE	703-883-6703	mcsmith@mitre.org
Bob Sutter	Argonne Lab	202-488-2427	sutter@anl.gov
Walt Swindell	USATRAC	913-684-3030	swindelw@tracer.army.mil

Peter Valentine	USAEPG/JDBE	602-538-4976	pvalenti@huachuca-emh7.army.mil
Jim Weatherly	SPAWAR	703-602-4540	weatherj@amtp-gw.spawar.navy.mil
Gio Wiederhold	DARPA	703-696-2218	gio@darpa.mil
Ken Wimmer	SAIC/DMSO	703-379-3770	kwimmer@dgis.dtic.dla.mil
Joyce Wineland	NAVINTCOM	301-763-1390	fax 301-763-1359
Andrew Wirkkala	NAVAIR	703-746-1974	fax 703-746-1937
Jeff Wolfe	DISA/CIM	703-536-6900	fax 703-536-7480
Rob Wright	RCI-CCTT	407-282-1451	fax 407-658-9541
Dan Wu	DISA/DFS	703-487-8409	dwu@dgis.dtic.dla.mil
Simone Youngblood	JHU/APL	301-953-5000 x4000	simone@aplcomm.jhuapl.edu fax 301-953-6910

Appendix D

NOTES FROM THE 1ST I/DB WORKSHOP, FEBRUARY 20, 1992

SECTION 1: TABLE OF CONTENTS

List of Attendees	Section 2
Summary List of Tasks	Section 3
Overview of DMSO, and Meeting Objectives	Section 4
Information about Attendees' Organizations	Section 5
CIM Process and Standard Data Elements	Section 6
Main Topics of Discussion	Section 7

SECTION 2: LIST OF ATTENDEES

Cy Ardoin
Bob Bishop
CDR Mike Chase
Twyla Courtot
Carol Dula
Iris Kameny
Steve Lawyer
Miranda Moore
Marjorie Powell
Lauri Rohn
Maj Mike Robinson (didn't attend but is member)
Bruce Rosen
Roberta Schoen
Jim Shiflett
LtCol Charles Snead
Dan Wu

SECTION 3: SUMMARY LIST OF TASKS

This is my attempt at a summary list of what the current task group needs to do for DMSO. Please make comments or add to it:

(1) DB/DD Directory: (a) redo schema; (b) apply examples; and (c) address issues above, in particular: how to support search, develop directory data elements for describing terrain databases if necessary, and see if Schema.2 is adequate to support multiple database versions. (Address (a) and (b) by 30 March.)

(2) Study and get guidance on the IRDS standard and products that support it and which could be used to implement the DMSO enterprise information directories and dictionaries. Select an IRDS compliant product (e.g., WISDIM or INFOSPAN). (Will begin to do this at March 30–31 meeting.)

(3) Draw the pictures, etc. that Shiflett requested to explain relationship of directories to databases and models and to standard data elements. (Do this for 30 March meeting.)

(4) Address model directories and the schema to describe models (no schedule yet).

(5) Address terms and definitions by reviewing those from the database standards community as well as those from the M&S community (no schedule yet).

(6) Develop a list of reference documents (do this for 30 March meeting)

(7) Help Jim Shiflett to fulfill his CIM data administration responsibilities by understanding how to use the IDEF tools to model process and data in order to develop the M&S standard data elements. (Begin to do this at March 30–31 meeting.)

(8) Plans for next meeting at IDA on 30–31 March.

(1) People to invite:

— Tom Lopez

— representatives from the Army, Air Force, and Navy (if possible) working on data standards.

- Twyla has emailed me a few names: Becky Harris, formerly Army Data Management Division, 703–536–6900, Jim Pipher has taken over for Becky Harris, 703–355–7134.

- Jim Shiflett has faxed me a message about army activity and data modeling using IDEF with a POC Maj Grobmeier 703–614–0757.

SECTION 4: OVERVIEW OF DMSO, AND MEETING OBJECTIVES

Jim Shiflett opened the meeting by describing the DMSO mission and organization. He gave an overview of the M&S community data problems, and some insight to his need for help in responding to the CIM initiative in his role as the DMSO Data Administrator.

Jim listed three topics to be addressed during the meeting:

- Data definitions
- Data schema
- List of references

and two goals: (1) to determine what needs to be done, and (2) who will do it.

At the end of the meeting, Jim passed out a two page paper titled "DMSO Information System," which is the first cut on an information system designed by COL Jim Shiflett (DMSO), Cy Ardoin (IDA), and Bob Bishop (DTIC) on February 13, 1992.

SECTION 5: INFORMATION ABOUT ATTENDEES' ORGANIZATIONS

Lauri Rohn (PA&E) was invited to attend after PA&E made a presentation to DMSO about the DAMIS objective which is to come up with a standardized database for modeling and use in preparing the defense budget. They have been asked by the CIM community to develop standard data elements for their functional area (similar to Jim Shiflett being asked to do so for the DMSO functional area).

DISA has a Center for Standards that contains an Information Standards Directorate and an Information Processing Standards Directorate. The mission of the Center for Standards is to support the functional data administrator for C2I. Charles Snead is in the information standards directorate and is concerned with Information Resource Dictionary System (IRDS) standards and data element standards while Dan Wu is in the information processing standards directorate and is concerned with database access and exchange standards (e.g., SQL and schema standards).

The Information Standards Directorate has responsibility to the DDI for specific modeling and simulation deliverables: (1) preliminary assessment of M&S, and then detailed assessment of M&S, (2) a management plan for M&S, and (3) an action plan.

SECTION 6: CIM PROCESS AND STANDARD DATA ELEMENTS

The CIM process for modeling a functional area and developing standard data elements was discussed and it seems the process is not yet well defined. There is a "Technical Reference Model for Corporate Information Management," dated November 27, 1991, which defines a target framework and profile of standards for this infrastructure and the applications it will support. Jim Shiflett sent me a copy of that document. Jim asked how we can get things going in this area and we were told that the Center for Standards has the billets but not the staff to support the DMSO effort now.

Directories: the Center for Standards is supposed to house directories but they have put this off due to lack of personnel and prioritization of needs.

The DoD Standard Data Element Dictionary may be based on WISDIM (from JOPES) and IDEAS (from DIA). Bruce Rosen said that the IRDS standard is described in a 600+ page document. There is a draft technical reference document and NBSIR 88–3700, "A Technical Overview of the Information Resource Dictionary System (Second Edition)," dated January 1988 which are more readable. (Twyla Courtot can furnish a copy of the former and I have a copy of the latter.) It was suggested that we could consider basing the DTIC IRDS support for DMSO on WISDIM or a product from (or called) INFOSPAN.

CASE tools: an RFP is out for ICASE (Integrated CASE) proposals, to be managed by the Air Force. The goal of ICASE is to supply a sort of backplane that would serve to integrate different CASE tools. The Center for Standards and NIST would like ICASE to support an IRDS compliant interface which will require export and import standards. The view is that IRDS should be the glue to make it possible to exchange information between tools.

SECTION 7: MAIN TOPICS OF DISCUSSION

Definitions

We discussed the need for well thought out and agreed to definitions. There was some discussion of standard definitions (e.g., those already defined by DoD Directives) and definitions that may be more appropriate or more specifically defined for the M&S community. I shared with the group my discussion with Paul Davis about the definition for "data" and how we might want to develop more explicit terms used by the M&S community such as "endogenous variable," "exogenous variable," "output variable," "input variable."

To do this task, we have begun to develop a list of relevant documents with glossaries. These are in the reference list.

Database/Database Directory (DB/DD) Directory

The draft schema needs to be redone reflecting the results of the meeting. I will redo the schema taking a more logical higher level approach (as suggested by Bruce Rosen) by defining conceptual descriptions of the information needed. I will distribute this to everyone for comments and after applying the comments will take a few samples such as the J8 and USTRANSCOM/J6 directories, and a sample database from RAND and represent them in terms of the schema. This will be done by the time we meet on March 30–31. Though I will assume responsibility for this, I would like some volunteers to try their hands at the J8 and USTRANSCOM/J6 schemas so that we can compare results.

One issue is how we decide to organize the DMSO directories since the J8 and USTRANSCOM/J6 directories are of models and not databases. Do we want to put directories of models and databases, databases, and models in the same directory? Or do we want a directory of databases and database directories that is separate from a model directory?

Addressing the seven policy issues from the draft schema paper:

(1) Policy for building and maintaining the DB/DD Directory:

- How do we populate the directory? Answer is that when DMSO establishes an M&S proponent for a model, it also establishes that organization as a proponent for the databases needed by the model (or output by the model and needed by others) unless the databases already have a proponent.

- How is the directory guaranteed current information? DMSO will establish a tickler file of POCs and periodically ask for updates; if information has not been updated after a given period of time, the entry will be archived or destroyed.

- How is updating done? The proponent can either enter update information through a user interface at DTIC, or furnish DTIC with electronic or hardcopy. He/she will not be allowed to directly update the directory. The update will be verified by DMSO staff as well as possible (e.g., confirm that organization names are in the DMSO organization list, that dates are legal, etc.). Finding these kinds of errors could be done automatically, though correcting them would have to be done by a person.

- Issue of implementation of DB/DD Directory. Implementation of the DMSO directories should comply with the IRDS standard.
- Examine/compare WISDIM and commercial tools that do this
- Determine if the directory system can be supported by Topic or if an additional software system such as a relational DBMS is needed.

(2) Security issues: Shiflett's direction is to leave this alone for now. If we later find that security poses a problem, we may have to develop a separate classified set of directories.

(3) How to develop, structure, and maintain appropriate key word or search information for browsing remains an unsolved issue that needs to be immediately addressed. Should functional groups develop their own network of terms and concepts? What are the useful categories of identifiers such as organization, functional area, community, etc.?

(4) Will one DB/DD schema suffice for all types of databases such as terrain, special weapons effects, etc.? We need to discuss this with Paul Birkel and the Environment TWG especially with regards as to what descriptive information is needed about the geographic or spatial area covered by a database or directory.

(5) Ascertain whether draft schema is adequate to maintain information about database for which there will be many views or versions (e.g., TPFDD databases by date/service, output databases from the same model produced by different runs): we still need to do this

(6) Investigate standards for database exchange so that the CD, tape, floppy disk, etc., formats of databases are readable: the only standard we came up with was delimited ascii files. This remains an issue.

(7) Mapping/translation techniques for DD/DB exchange so that the data would be directly usable in M&S applications. This remains an issue.

Appendix E

NOTES FROM THE 2ND I/DB WORKSHOP, MARCH 30–31, 1992

SECTION 1: TABLE OF CONTENTS

Agenda	Section 2
List of Attendees	Section 3
Action Items	Section 4
Col Shiflett — Introductory Discussion	Section 5
IRDS Briefing Summary	Section 6
IDEF Briefing Summary	Section 7
Standards vs. IDEF Modeling	Section 8
DISA/CIM Report	Section 9
Orlando DIS Work Shop	Section 10
Briefing on DD, DB and Dictionary	Section 11
DMSO IS Discussion	Section 12
Review of Directory Schema	Section 13
Summary of JIEO Briefing	Section 14
Document List Progress	Section 15
OASIS Project Discussion	Section 16

SECTION 2: AGENDA

Agenda for March 30–31 Meeting of DMSO ITWG/DBTWG Task Group At IDA, 1901 N. Beauregard St., Alexandria, VA in Building 1901, Room 118

Monday, March 30th

9:00—11:00	IRDS briefing by Burt Parker (MITRE)
11:00—11:15	Break
11:15—12:15	IDEF briefing by Major John Grobmeier (first half)
12:15— 1:15	Lunch
1:15— 2:15	IDEF briefing by Major John Grobmeier (second half)
2:15— 2:30	DISA/CIM Data Administration update by Dan Wu
2:30— 3:00	Orlando DIS Workshop Highlights by Major Mike Robinson
3:00— 3:15	Break
3:15— 4:15	DMSO Information System discussion by Cy Ardoin
4:15— 5:00	Functional overview (illustrated) of data directory, data dictionary, model directory by Iris Kameny

Tuesday, March 31st

9:00—11:00	Review DB/DD directory schema, go over examples, make changes, etc., led by Iris Kameny
11:00—11:15	Break
11:15—12:15	Discuss IRDS choices: WISDIM, INFOSPAN, etc., led by Twyla Courtot
12:15— 1:15	Lunch
1:15— 3:00	Home in on IRDS compatible course of action for DSMO Information System with emphasis on the implementation of DMSO products (e.g., directories, data dictionary), include discussion of use of COTS products such as relational DBMSs, Topic, led by Cy/Iris/Twyla
3:00— 3:15	Break
3:15— 5:00	Discuss use of IDEF for DMSO data administration activities assuming M&S is a C3I functional area, address: tool choices, what needs to be done, who should participate (e.g., Center for Standards, DMSO office, FFRDC help, etc.), led by Iris/Twyla

SECTION 3: LIST OF ATTENDEES

Cy Ardoin
Paul Birkel
Bob Bishop
Martin Costellic
Twyla Courtot
Iris Kameny
Steve Lawyer
Tom Lopez
Lani McGlynn (represented at meeting by Walton Dickson 703–746–8076)
Alan Peltzman
Marjorie Powell
Lauri Rohn
Maj Mike Robinson
Bruce Rosen NIST

Scott Roth
Roberta Schoen
LtCol Ceasar Sharper
Jim Shiflett
Gio Wiederhold
Dan Wu

Briefers:
Burton Parker MITRE on IRDS
Major John Grobmeier USA on IDEF

SECTION 4: ACTION ITEMS FOR FUTURE

(The responsible party is identified by: CAPITALIZED NAME)

(1) Identify the Data Administrator (DA) for M&S: JIM SHIFLETT

(2) IRDS Actions. The Task Group has decided that the DMSO Information System should be IRDS-1 compatible if possible since this is likely to become a CIM standard. An issue is whether we can pick up existing software to use (the NIST prototype was not recommended for use by Rosen). We also would like not to have to acquire and use a RDBMS in addition to TOPIC if possible. We know that IRDS-1 has shortcomings that may cause problems: it doesn't easily support E-R m-m mappings; it doesn't support objects (e.g., as used in o-o simulations), large matrices or collections of data that appear as an entity (e.g., consumption tables); images; speech; etc.

(a) Evaluate possibilities for DMSO IRDS compatible system:
CY ARDOIN
- Evaluate TOPIC with respect to IRDS Module 1 and 2 compatibility and whether it can serve as a RDBMS
- Examine WISDIM/ORACLE as basis
- Examine Sybase/ORACLE use in AF MIDAS dictionary effort (Twyla will give Cy name to call)
- Examine CIM interim DD on AT&T ORACLE/UNIX

(3) Look into the Defense Dictionary Repository System (DDRS) configuration AT&T/UNIX/ORACLE to get data element attribute schema: TWYLA COURTOT

(On April 1st, Wu and Kameny spoke with Dawn Hughes and Becky Harris and learned that the SDE attribute schema, naming conventions, and instantiated SDEs are respectively in review stage, being addressed by ad hoc NIST/CIM meeting, and are left over from 1965 and not available for us to examine.) Perhaps, you can try to get some examples of WISDIM SDEs?

(4) Update directory schema and develop IRDS based schema: IRIS KAMENY

— Update schema/keywords

— Representation of IRDS schema based on Army/Joint/CIM efforts

— May need extensions

(5) DTIC perform search on IDEF: BOB BISHOP

(6) IDEF Training for Task Group: CY ARDOIN and TWYLA COURTOT.

The Task Group has decided to use the IDEF tools and methodology for development of the DMSO Information System. This will enable us to be functional users and to evaluate the tools and methods before having DMSO recommend they be used in new M&S developments. After using them we will have a better feel for what kind of training and facilitator support is needed.

— MITRE is giving a course and Twyla has already forwarded an email message that is included in section 5: this is for use of IDEF0, the new DoD standard

— They will also look into CIM 2-week courses, and IDEF1X by contacting Mike Yeomans 746–7932. Patricia Cobb and David Stipey are also Army people experienced with IDEF, talk to them for advice

— They will keep us informed as to who is taking the course(s) when

(7) Jim Shiflett requested a picture of the SDE process as we envision it that he can use in talking to CIM. Cy and Iris composed one on April 2 and left it with him.

(8) SDE contacts: IRIS KAMENY

Find out what others are doing, have done about SDEs, wrt naming, deriving, maintaining, adjudicating, etc.

— Talk to Howard Haecker at Ft Leavenworth: AV 552-3030 who is responsible for the Data Dictionary for all TRAC DBs

— Talk to OASIS project people about M&S, SDE, DD (Dick Wyman, is contact): see results of this meeting under Section 16. OASIS Project

— Also receive call from Erwin Atzinger (AMMSA) who will stop by in mid-May to talk about these issues

(9) Database Search Terms: IRIS KAMENY, MARTY COSTELLIC

— Marty will get Iris reference or document from US Naval Institute on key words

— Also contact Carol Jacobson, head of Mil Lab Div Spec Library Association

(10) Need to find out results of SDE naming conventions either from Bruce Rosen or from Bob Molter (703–746–7926) who is working with Strassmann on DA problems and was the one to set up meeting with NIST to address naming conventions: BRUCE ROSEN, IRIS KAMENY

(11) M&S catalogs, Army will be going out and using the J8 catalog and SIMVAL, need to look into what MORS has done also: LANA MC GLYNN

Also ask Lana about the Army 19 functional areas and how they will relate to Army M&S functional area

(12) Need to get CIM Business Process Guide: DAN WU

SECTION 5: COL JIM SHIFLETT'S INTRODUCTORY DISCUSSION

Jim said it is not clear how M&S establishes standard data elements (SDEs) since M&S is not a "normal" functional area. We need to find out who at CIM can clear this up and what the M&S DA has to do to satisfy CIM. (UPDATE: wrt Kameny and Wu visit with CIM people on April 1, they were unaware of M&S as a functional area and had no insight into this problem). There are issues with SDEs as to who nominates them and the process of doing so. Jim suggested that we take the DEs in M&S databases and match these to CIM DOD SDEs, and if there are no matches (i.e., they don't exist as SDEs), then M&S maintain them as new M&S DEs, adjudicate their metadata (e.g., name, description, etc.) among the M&S

community, and propose them as SDEs to CIM. This seems to follow the process laid out by CIM in Draft DoD 8320.1-M, "DoD Data Administration Procedures Manual," 25 October 1991.

SECTION 6: IRDS BRIEFING SUMMARY

Burton Parker (MITRE) gave an informative briefing on IRDS-1 and -2, including the current standards, IRDS related activity standards, and IRDS relevant/related repositories. The briefing charts are available through Cy Ardoin at IDA. Below is a summary of what I considered the salient information.

IRDS is an environment to describe, document, protect, control and access information resources of an enterprise. This includes all meta information about data: who, what, where, why, when, and how. Its goals are to: affect "seamless" application portability, data transportability, and information access within and across organizations, and to provide a common information resources environment. IRDS-1 is a dictionary tool focused on data entity metadata management. IRDS-2 will be a repository tool focused on information resource entity meta-information management.

IRDS-1 has been released as ANSI X3.138–1988 standard on October 1988 and adopted as a FIPS (as described in FIPS Publication 156) effective September 1989 and mandatory in March 1991. IRDS-1 consists of 7 modules, of which modules 1 and 2 are considered the basic modules. Module 1, the core standard, is mandatory and includes basic definitions, command language and panel interfaces, and a minimal IRD schema. Module 2 contains the basic functional IRDS schema. An implementation of FIPS 156 must include Module 2. The other five modules are: security (IRDS access control), schema structure manipulation (control of IRDS through a life cycle), procedures (define and execute procedures to operate on IRD and IRD schema), application program interface (interface to allow applications to access IRD and IRD schema), and entity lists (means for defining and manipulating IRD entity lists).

IRDS-1 includes defining, administering, and maintaining data systems; controlling access to and modification of data; and the ability to exchange IRDS data between IRDS-1 environments. IRDS-2 goals go much beyond this; whereas IRDS-1 is a passive dictionary system, IRDS-2 will be a dynamic one. IRDS-2 is intended to be consistent with ISO IRDS-2, will be available optimistically in 3–5 years, will operate in a heterogeneous environment, will address multi-media objects, and will be a dynamic repository system.

IRDS-1 includes encyclopedic information (how, why, who); directory information (where, who); and dictionary information (standard, proposed, nonstandard data elements and formats, and data types of string, text, and numeric); and DBMS information instantiations (about partitions, files, records, and fields). IRDS-1 includes control of SDEs with no control over proposed and nonstandard DEs.

The Task Group requested a definition of data and information. Parker defined information to be organized data. Gio Wiederhold also offered a definition: information is the result of bringing together stored data and knowledge and performing actions on them. By this definition, the IRDS is seen as storing data and knowledge. Information would be created in the process of using the IRDS by applying knowledge to data.

IRDS near-term evolution includes a family of standards:

— IRDS-1 current published standard
— IRDS-1 services interface to external software: ANSI X3.185–199X draft standard is awaiting committee approval
— IRDS-1 export/import file format for data exchange between IRDSs: ANSI X3.195.199Y draft standard is awaiting committee approval
— IRDS-1 naming convention verification: technical report was approved in 1991 but more work on naming conventions is being done with CIM
— IRDS-2 Reference Model due out at X3H4 committee in 1992

IRDS-related activity standards include:
CALS: Computer-Aided Acquisition and Logistics Support
CDIF: CASE Data Interchange Format (EIA)
ISO/IRDS: IRDS (international)
PCTE: Portable Common Tools Environment (European Computer Manufacturers Association (ECMA))
PDES: Product Data Exchange Using STEP (PDES, Inc.)
P1175: Computer Society Task Group on Professional Computing Tools (IEEE)
STEP: Standard for Exchange of Product Model Data (Europe)
SUMM: Semantic Unifications Meta-Model (PDES, Inc.)
X3H6: Subcommittee for CASE Integration Services (ANSI)
US TAG to ISO/SC7 on software engineering

IRDS related programs include:
CIM: IRDS-2 oriented
Joint Staff Warfighting and Intelligence System Dictionary (WISDIM) of joint data elements, IRDS-1 oriented
EPA: full life cycle management, IRDS-2 oriented
US Army Data Management: Army SDEs, IRDS-1 oriented
US Air Force MAC:MAC Integrated Data Administration System (MIDAS), IRDS-1 oriented

Repositories claiming IRDS-1 conformance:
INFOSPAN based on ORACLE
NIST prototype based on ORACLE

Nonconformant repositories available from:
Computer Associates International: based on Datacom/DB

DEC: IRDS compliant E-R interface, based on DEC's Rdb DBMS
IBM: AD/Cycle Repository, IRDS compliance via INFOSPAN partnership, based on DB2
ORACLE: CASE*DESIGNER, CASE*GENERATOR

SECTION 7: IDEF BRIEFING SUMMARY

Major John Grobmeier gave an informative briefing about IDEF methodology and its use in the Army. IDEF is a government owned/non-proprietary system originally developed by the Air Force. IDEF was made a DoD standard on 22 January 1992. The best experienced IDEF people in Army are Patricia Cobb and David Stipey (can get phone numbers through Grobmeier). The briefing charts are available through Cy Ardoin at IDA. Below is a summary of what I considered the most salient information.

Definitions:

IDEF: The Integrated Computer Aided Definition Language developed in the late 70s by the Air Force. It is a top-down driven method to engineer effective businesses

IDEF0: documents what is currently being done and how it could be done better in the future (as is vs. to be). Resulting models are commonly referred to as "activity" or "verb" models

IDEF1: documents what is needed to support what is being done. An extended set of IDEF0 is called IDEF1X. These models are commonly called "data" or "noun" models

IDEF2: documents when an organization needs to know what it needs to know in order to do what needs to be done (no one has really thought about how to do this yet)

IDEF is used for four things: process improvement, organization improvement, data framework (IDEF1X), and system design and implementation (IDEF2). IDEF0 models inputs, constraints, mechanisms, and outputs apparently in a general way without defining dependencies between these (my comments).

The Army says that IDEF is not a methodology but tools to support the STRAP methodology. IDEF was successfully used during ODS to model a coordinated fire support system process in a few weeks.

IDEF uses: functional analysis of business (process-activity orientation) (primary function of CIM usage); structured documentation of tasks and their relationships to each other and supporting business rules; an apolitical analysis tool to arrive at optimal solutions and plans while building consensus; identification of corporate data, includes logical and physical database design plus SDEs (primary

purpose of Army Modeling Program supporting Army Data Management); and can be used in system design, prototyping, and development.

Observations on IDEF:

— Critical to standardizing data and designing responsive information sharing systems, will also be springboard to move from relational to object-oriented databases

— Value of IDEF for automator is dwarfed by the value which can be achieved through returns on functional efficiencies

— IDEF has proven track record in its use by business

— Used by and applicable to both functionals and automators and is applicable to both the sustaining base and warfighting arena

— Provides sensible, engineering solutions to make organizations and information systems more effective and responsive.

IDEF minuses: uses lockstep procedures; tool sets need to be better integrated; requires time, commitment, and resourcing; must be well focused quickly or can ramble and accomplish little; can be oversold; needs experienced guides/trained personnel to use; there are a limited set of qualified facilitators at current time; and DoD organizational support is currently in the developmental phase.

To be successful, the use of IDEF must be: functionally driven and supported at highest levels; involve subject matter experts; be fully resourced; facilitated by personnel with experience; be focused and scoped; start at top and focus on ROI candidates driving down to implementable projects; have a dedicated project leader and support of all participants; and have right modeling environment/support.

Discussion of who is doing what with IDEF:

(1) Documentation (from Bruce Rosen):

UM110231100, Function Modeling Manual (IDEF0)
This is the activity/process modeling methods description document, which also provides instruction.

UM620141001, Information Modeling Manual-Extended (IDEF1X)
This is the information/data modeling methods description document and it includes instruction.

Cy Ardoin has ordered these manuals and also an IDEF2 manual.

(2) Training (from Twyla Courtot):

MITRE is scheduling IDEF training for April 27–30, inclusive, and can make room for one person from our Task Group.

General information on training: The two most viable training sources are DACOM (Dan Appleton Company) and Wizdom Systems, Inc. (vendors of IDEFine set of tools) for training in methodology, not so much in navigating through the tool set. The WIZDOM contact is Allen Batteau, (708) 357–3000. They will do "Custom On-site IDEF Instruction," which means they will work with some prespecified examples from an area of interest. They do anything from 1 to 5 days of training as follows: 4–6 students: 1 day = $2000, 2 days = $3500, 3 days = $5000, 4 days = $6500, 5 days = $8000. 7–10 students: 1 day = $2600, 2 days = $4550, 3 days = $6500, 4 days = $8450, 5 days = $10400. 11–15 students: 1 day = $3200, 2 days = $5600, 3 days = $8000, 4 days = $10400, 5 days = $12800. Twyla didn't know what the course content consisted of or what is optimum amount of time to spend in a course. MITRE went for the 4-day course, so they could screen it.

DACOM offers a public IDEF modeling workshop in Fairfax on 21 April. This is a 4-day program, @ $995 per attendee. For on-site training, the same 4-day class (I interpret this to mean no customization to your environment) costs $10500 plus travel for the LA based instructor. Their quote on travel is approx. $2500. Course will accommodate up to 15. As of March 20, upcoming available dates were 4 or 18 May. Facilitation of Business Process definition is a 'to be determined' price. Other upcoming public classes from these folks are: LA—12 May, LA—2 June (Sorry, the 12 May is a one day Modeling for Managers Seminar @ $295 pp), Dallas—15 September, Washington—20 October, LA—8 December. The 1-day overview is 12 May, 22 Sept, 10 Nov, all in LA. Local (DC) contact is Ronald Batman (703) 573–7644. They also have an 800 number: 800–322–6614. It is also unclear what the DACOM folks really address—IDEF0 & 1X, or what.

National Defense University is developing an IDEF course, the Army Management College teaches IDEF (course developed by Richard Preston of BDM).

DISA/CIM: Center for Data Administration Information has a task of developing an IDEF course. CIM is working on a draft directive for IDEF. The Functional Integration Managers (FIMS) will be in charge of process and high-level data modeling activities using a generic process model. The DoD standard covers IDEF0 only, Mike Yeoman (746–7932) is the contact to find out more about CIM use of IDEF. DoD has funded NIST to produce the FIPS for IDEF0 and possibly for IDEF1.

Miscellaneous:

Other nations think IDEF is a smart way to get interoperability

Everyone is looking at normalized relational data modeling for the logical model and can then denormalize for implementation

IDEF was oriented toward manufacturing and IDEF2 has been made specific for manufacturing. In the past, the IDEF users group was mostly non-government, now has more government participation.

There are some simulation tools for use with IDEF, CACI is interested in doing this. The Air Force is incorporating IDEF in its ICASE tools.

On Army M&S recent use of IDEF: ODUSA (OR) decided who to ask about Army studies. MISMA concentrated on warfighting models, but didn't get into modelbuilding—stayed focused on model management—this was funded on shoestring.

Jim Shiflett asked how we could use IDEF to determine M&S data elements? Recommendation was to take a CIM two-week course for functional data managers (FIMS).

Others using IDEF to do modeling include: FT Gordon just completed model for ISYSCOM; SDI is using IDEF; someone is using it to model the electronics of the B1 bomber.

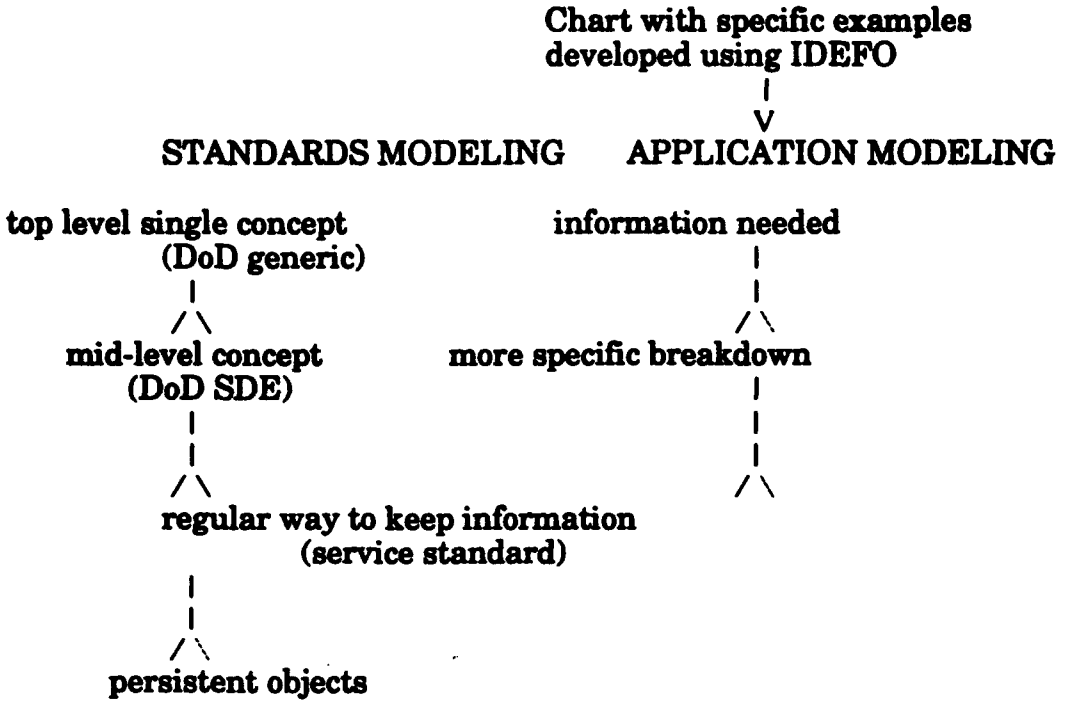
Grobmeier thought the M&S environment is hardest one he has seen to model. The warfighting world is pretty clean and able to use IDEF modeling. People at Leavenworth are interested in using it.

Action items: get Strassmann's "CIM Business Process Guide" maybe from the Wright-Patterson system library, POC is Judd Hudson.

SECTION 8: BRUCE ROSEN'S VIEW OF STANDARDS VS. IDEF MODELING

Bruce will be participating in trying to solve some SDE naming issues such as the Army's choice of making measurement terms part of the name rather than as modifiers to the name. NIST had put forth naming conventions (NBS Special Publication 500–149 "Guide on Data Entity Naming Conventions," October 1987) and CIM has to agree upon naming conventions and formalize policy and procedures.

Bruce drew chart showing the difference and similarities between the standards view of modeling and the IDEF0 view:



An example of a problem is measurements:

— An IDEF application would like to express accuracy of measurement in the name

— From the standards point of view, a more general concept of an SDE is wanted, rather than creating a new SDE for each different expression of measurement accuracy

— An ad hoc group will be addressing this problem and possibly others

Also someone handed out some sheets on IDEF Software Products from D. Appleton Inc., Meta Software Corp., Knowledge-Based Systems, Inc.

SECTION 9: DISA/CIM DA REPORT BY DAN WU

I'll try to summarize the eight report areas briefed by Dan:

(1) DISA DoD Data Administration Council (DAC): first meeting was held Jan. 30, to announce officers and discuss purpose. Denis Brown, DoD DA, is official chair; he assigned responsibility of chair to Belkis Leong-Hong; William Greyard is executive secretary (Leong-Hong/Greyard 285-5380). Council will meet at least quarterly to provide input on matters

concerning DoD DA program and facilitate issue resolution and data exchange.

(2) CIM Repository Architecture Tiger Team (CIMRATT): near term actions are to design, distribute, and analyze survey results in establishing a directory of "collection of objects" in CIM, then prepare decision briefing for management of findings. (They are trying to develop a definition of a repository.) Ms Showalter is chair of survey group, Becky Harris chair of survey analysis (285–5381).

(3) Draft Data Administration Strategic Plan (DASP) was sent to CIM Director on Feb. 10. The DoD DA Framework has been submitted for internal review and validation (Dawn Hughes 285–5381).

(4) Interim DoD Data Dictionary (DD): discussion of information paper on Interim DoD DD included: functional overview, contents of dictionary, and POCs. Paper was distributed to DoD functional DAs and component DAs. (Rebecca Harris)

(5) Business Case Analysis (BCA) program: they are developing a generic, IDEF0-based BCA model for use in BCA, in anticipation of receiving an OSD(C3I) tasking memorandum requesting high-level functional economic analysis for office automation of OASD(C3I), including ODDI/IDASD(IS), Pentagon, and CIMNET, plus I-CASE (Jim Raney, XF, 285–5377) (There is also a MITRE paper on this that Twyla will try to obtain.)

(6) DoD Enterprise Model: Birch and Davis (R&D) has completed analysis of this FY91 task, finding little conflict/overlap between Civilian Payroll and Personnel process models (Ken Fagen, XF, 285–5381)

(7) Strategic Data Model Contract: Principal and secondary entities and other entity types have been identified for nearly all 15 functional ASD/USDs based on mission and function statements. Nearly 500 planning or management statements have been cross-referenced in CASE tool. (Russ Richards, XF, 285–5387)

Project: Mike Yeomans agreed that the DoD Strategic Data Model is needed to provide guidance, integration structure and high-level architecture to the CIM initiative. (Russ Richards)

(8) Model Management and reuse: Model's life cycle and process are surfacing in discussions of DA services/products/interfaces in the software development framework context and use of interim DD. (Judy Albert/R. Harris, XF, 285–5381)

The Task Group asked who is the C3I DA? Dan Wu has since responded by email:

I obtained from Dawn Hughes a list of all DoD functional (acquisition, policy, C3, etc.) and component (services, agencies, cincs, etc.) data administrators. M&S is not listed here. However, I also found DoD Directive 8320.1, dated Sept 26, 1991, which lists M&S as a sub-functional area under acquisition. The functional data administrator for acquisition is Dr. H. Steven Kimmel at 703–695–0598. Jim may be right that he is the sub-functional data administrator. The C3 functional data administrator is Dr. Thomas Quinn, to whom LTC Snead reports—Snead may not be responsible for M&S.

In short, Jim needs to get guidance from Dr. Kimmel on what he wants to do.

Dan Wu also handed out:

— one page Information Paper on Interim DoD Data Dictionary that says that the ADD/ADSS and WISDIM were selected in combination to meet the requirements of the Interim DoD DD which stores approved SDEs and provides logical partitions for the DoD 5000.12 DEs and component DD. Contains 17 class names and definitions, identified in the Draft DA Procedures Manual 8320.M1, and 1,812 data use identifiers and 1,058 DE standardized under 5000.12, dated April 1965. DoD component agencies may request a logical partition to develop and store generic elements, DEs, and application elements prior to submission to DoD approval process. Procedures to do so can be found in the user manual. The Interim DoD DD resides on a Vax 8600 at DISA. (Contact Perry Lyles 703–693–5184)

— Appendix A DoD Data Administration Framework (Draft) 2/17/92: shows framework and shows relationships and responsibilities for DoD DA, Functional DAs, and Component DAs. In essence, CDAds are responsible for coordinating the execution of DA operations within their respective component. User data requirements are captured by the CDAds as SDEs, data models, or other forms of information about data. These products are reviewed for technical adherence to standards. Standard data products (e.g, SDEs and data models) are forwarded for approval to the FDAds in the appropriate functional areas. The FDAd then forwards standard products to the DoD DA for approval and registration in the DoD repository.

SECTION 10: ORLANDO DIS WORKSHOP HIGHLIGHTS BY MIKE ROBINSON

500 people from services, industry, and foreign countries attended. A key comment from each of the four subgroups (land, sea, atmosphere, and communication architecture and security) was that database management and coordination was a top need. DIS is up for acceptance as an IEEE standard. Since it supports the capability to dial into other systems, including simulation, security really needs to be addressed. Jim Shiflett says we need to come to grips with security, secure systems. For example, should message headers be in the clear or encrypted? At what time do you classify what you do? Do you want to have a

standard level of operational security? What about the data aggregation problem? Need to start working on this issue.

SECTION 11: BRIEFING ON DATABASE DIRECTORIES, DATABASES, AND DATA DICTIONARIES

I gave the briefing and didn't write down any particular comments. If someone wants to add something in here, let me know.

SECTION 12: DMSO INFORMATION SYSTEM DISCUSSION

Cy went over the information system design. I think I left my copy with my notes with him, but as best as I remember there was general agreement on the format but some discussion about whether to have a top level choice of "directories" that would have choices at the next level of "database directory," "model directory," "organization directory," etc. rather than showing the detailed list at the top level. There was also agreement that lower leaves need to identify their higher level subject matter at the top of the screen rather than at the bottom. Cy, or anyone else, feel free to add in anything I have overlooked.

SECTION 13: RESULTS OF REVIEW OF DIRECTORY SCHEMA

I will be preparing a later version of the schema by the end of the month. I noted the following changes:

(1) in General Information: add field "distribution comments"
(2) describe "access limitation" as a text field
(3) include "development" in status of database or directory
(4) make "source" a separate section from administration
(5) change "organization name of owner" to "organization name of technical POC"
(6) add fields for a "release POC," organization, etc., and comment field
(7) change search/indexing information to key words, and move this section closer to general information: also try to figure out how people will use these to define structure, predetermined list, etc.
(8) identify documentation as user, technical, overview: may need to add additional documentation field to version section to handle documentation specific to a version
(9) in description field, advise people not to just repeat the key words

Jim Shiflett says DMSO will have someone knowledgeable to review information before it is entered into the system.

Paul Birkel reviewed the schema and doesn't believe we need to add additional fields to describe terrain and environment databases.

SECTION 14: SUMMARY OF BRIEFING AT JIEO (ATTENDED BY KAMENY AND WU)

In order to talk with Dawn Hughes and Becky Harris, Dan Wu and Iris Kameny sat through a briefing given to a British general. We have two sets of briefing charts: one on Team JIEO (which is informative with respect to JIEO missions, etc.), and the other given by Dawn Hughes which is mostly high level but does include the DoD DA Framework in the handout that Dan gave us at the meeting. It discusses where they are and where they are going without much detail.

I was prepared to talk with them about the following things (they didn't have much time and we spent less than 30 minutes with them):

(1) DMSO needs direction from CIM on an IRDS compatible system for DE dictionary, etc. We need: (1) software, (2) attribute set for SDEs, and (3) current SDEs.

ANSWERS: (1) they are using as an interim DD ORACLE/UNIX on an AT&T 3B2 computer that has been upgraded to include the old SDEs under 5000.12 (1965) in one partition and ability to furnish other partitions when asked for and had no recommendations as to whether to use their interim system, (2) this has not been decided yet, (3) the only SDEs available are those from 1965 and they don't want us to see them.

(2) DMSO will support a data element inventory mechanism
- — attribute set for SDEs
- — voluntary agreement by Army, Air Force, Navy on DEs to build consensus in community
- — some types of DEs are currently not addressed by IRDS-1:
  - objects (such as in object-oriented simulations)
  - matrices and subsets of matrices (e.g., consumption tables) where one doesn't want to make each cell an SDE
  - images
  - telemetry data off satellite
  - composite DE (e.g., basic encyclopedia number, unit line number)
  - aggregate DE

ANSWERS: how to handle attribute set for SDEs has not been determined yet; process for dealing with Components to implement the CIM is determined by the Functional DA, they did not think M&S was a functional area, did not recognize Jim Shiflett as an FDAd, and said that needed to be addressed first. Suggested talking to Kimmel, Sharkey, and Gary Hurd, 614–8985.

As far as the data types not being handled: they have changed the single-concept, homogeneous requirement so that composite and aggregate DEs can be handled; they have no plans for dealing with objects, matrices,

telemetry data, or images and will be very interested in what we come up with.

Some notes from JIEO Briefing

DISA has no control over an information system wholly owned by a Service with no interoperability requirements. For those systems over which they have control, they will be requiring mandatory testing at FT Huachuca to make sure of interoperability.

C3I is one functional area out of 16 with a functional integration manager (JIEO). Data is a mess; everyone has his own DE dictionary and doesn't want to change. The Army is trying to bring the Army systems together and look at SDEs as central, but the cost of doing this in terms of people and time appears to be very great.

Data aggregation and security is a fundamental system problem. They have two programs going: Defense Information Security Program (first they will put policy in place and then an architecture), and, the short term, MLS Technology Insertion, which requires working with users today to provide (1) secure integrated workstations, (2) secure MLS services, (3) trusted software agents, and (4) intelligent routers.

JIEO is coming up with SDEs for MTFs; they are mapping MTFs into data models as they are developed and will end up sending changes into the NATO process. The longer view is that with standardization there will be less of a need for standardized message formats; users will be able to pull data from databases using standards like SQL. But we will be careful to maintain NATO agreements.

Asked for differences between data and information architecture: answered that data architecture is based on E-R model, but information architecture has processes linked to the data it needs.

Asked about object-oriented and response was that they can't do this without a data element base.

Asked about symbology of meaning and Norton Bragg says he has a recently written white paper on the subject.

Discussed how one would model C2 mission areas (there are nine of these functional areas though no one could name them). The process would be to gather data from the CINCs on how systems are used in the battlefield, where they go, attempt to lay out relationships by function to see where information has to flow and in what form, and this can be used to see what you need to do to satisfy interoperability—see what needs fixing and what needs to be done for future systems.

Need to use acquisition system to drive requirements modeling of data elements, develop a body of policy which will determine how ISs will be built, compliance with CIM, JS will decide if system doesn't need to support interoperability, the IS will be reviewed at each milestone, MAISARC process will apply to most ISs but level of reviewing body will vary with cost of system.

Asked about IDEF modeling support and was vague about whether CIM will consult out on this. Said that they are about to use IDEF to relook the WAM requirements.

SECTION 15: DOCUMENT LIST PROGRESS

I put this together in Washington and went over it with Cy and promised him I would prepare an updated email version. Cy will make copies available.

SECTION 16: DISCUSSION WITH OASIS PROJECT

Dick Wyman suggested we talk with Major Dan Hogg J8 (703–697–8899) about the OASIS project and I did on April 2 and these are my notes. To summarize, there may be a lot we have to learn and could use from them, but the contractor seems to be a more knowledgeable source and Hogg did not want us to talk with him before they make their deliverable in the June time frame.

I have four briefing charts that he gave me indicating that OASIS is an application using Ingres RDBMS and the Ingres Windows 4GL Development tool; the contractor is Westinghouse. OASIS is intended to provide data support for current modeling requirements and remain flexible to support future analytical requirements, and will provide data standardization in compliance with DoD directives (though they were coming to DMSO to find out what that is). They are using a client/server architecture where the data dictionary supports both model preprocessor and post processor data needs.

Within J8 there are 20–30 models, each requiring lots of input data from everywhere. Currently, all of these divisions build their own data. The data support branch is responsible for supporting technical operations with their data needs. So they are developing a centralized data management system whose primary goal is to satisfy the J8 data requirements. This includes satisfying the CINCs. Gave example where CENTCOM uses TACWAR model and it is also used at J8 but at a different resolution level. For example, CENTCOM needs company level and current threat data while Conventional Forces Analysis Division (CFAD) needs higher resolution data beyond the current time. If J8 needs to use CFAD, then J8 will be source of the data and maintain it.

They support source collection: by providing a schema for directory for sources, and have developed a source tracking system. The outside files are brought into the reference DB; the data dictionary is used to run error analyses and check the data. The outside file is mapped to an OASIS folder which will put the verified data into a

file. The system supports three layers: files, class, object (list, detail), but when I started asking questions about these and folders, he became a little confused and said it would be better if we spoke to the contractor.

I think a list contains a record and instances. When the user clicks on a record, the system brings up 10–12 main attribute identifiers and then the user can go to the detail frames via those identifiers. The user can build his own folders from the reference data and then modify them to fit his model needs. There are tools provided to build the database from the time it arrives from the source until it is used by the model.

They are developing their own naming conventions and SDE formats and are two months away from getting their first deliverable. They are trying to identify common source databases used in many models and to establish MOIs with other agencies to furnish data to them.

We discussed object-oriented data for a short time and he said there is a Force Structure Accounting System that uses object-oriented data. The data they will use will also be in a flat file for OASIS to use. Contacts are Tom Lyttle 505–667–9596, and Joel Holland 505–667–9596, that system is being developed in C++ openware.

He did not have much time as he had another meeting (I only spent about 40 minutes with him), and he did say that the contractor is located at Tysons Corner and that we could meet with them later in the year.

I think we should follow up on this to coordinate the data sources and models they are using and to help them/us understand how to be DoD data dictionary compliant.

Appendix F

NOTES FROM THE 3RD I/DB WORKSHOP, JUNE 4–5, 1992

SECTION 1: TABLE OF CONTENTS

Agenda	Section 2
List of Attendees	Section 3
Summary of Issues	Section 4
Action Items	Section 5
Jim Shiflett's introduction	Section 6
Session on DoD Data Administration, Data Model, Repository, and Training	Section 7
Session on Data Administration in Services, Joint Staff, and ARPA	Section 8
Session on Army Support for M&S Models	Section 9
Session on Data Collection, Simulation Models	Section 10
TOPIC and IRDS report (action item from last meeting)	Section 11
IDEF Update and Discussion (action item from last meeting)	Section 12
Database Directory Schema.3 discussion (action item from last meeting)	Section 13

SECTION 2: AGENDA

THURSDAY, JUNE 4

8:00—8:30 Status of action items from March meeting, overview of where we are going, goals for this meeting

SESSION ON DOD DATA ADMINISTRATION, DATA MODEL, REPOSITORY, AND TRAINING

8:30— 9:00	DoD Data Administration Policy and Procedures (will cover DoD 8320.1, DoD 8320.1M-1, DoD 8320.1M, DASP, migration prototype, naming qualifier issue, roles, Functional/Component Data Administrators, DoD Data Administrator, FIMs, software developers): Mr. Bob Molter
9:00— 9:30	DoD Data Model (will cover status, role of Model Integration Group, relationships to DoD data element standardization): Mr. Russ Richards
9:30—10:00	Defense Data Repository System (DDRS) (will cover overview, status, approval process, access, and user manuals): Mr. Dan Lewis
10:00—10:15	Data Administration education/training: Mr. John Hovell
10:15—10:30	Break
10:30—11:00	Questions for CIM panel

SESSION ON DATA ADMINISTRATION IN SERVICES, JOINT STAFF, AND ARPA

11:00—11:30	Air Force: Bao Nguyen 11:30 – 12:00 Navy: Rebecca Wade
12:00—12:30	Joint Staff: Janet Baralli
12:30— 1:00	Lunch
1:00— 1:30	Army: Jim Glymph

SESSION ON ARMY SUPPORT FOR M&S DATA, AND ARPA REPOSITORY

1:30— 2:00	How the Army is Organized to Support M&S Data: Erwin Atzinger
2:00— 3:00	Data Development System for Army M&S: Captain Walt Swindell
3:15— 3:30	Break
3:30— 4:30	Asset Source for Software Engineering Technology (ASSET) ARPA repository system (part of STARS effort): Chuck Lillie (SAIC)
4:30— 5:15	Discussion of what needs to be done by DMSO and how the above efforts/products could be utilized: led by Iris Kameny

FRIDAY, JUNE 5

SESSION ON DATA COLLECTION, SIMULATION MODELS

8:30— 9:00	Report of reuse library of J-MASS objects: Iris Kameny (based on information from Mike Hucul)
9:00— 9:30	Data collection efforts of the Close Combat Tactical Trainer program: Major Bill Johnson
9:30—10:30	J-8 Data Management: OASIS project and Planned Enhancements: Major Dan Hogg
10:30—10:45	Break
10:45—11:30	A Modelbase Concept for Model Interoperability: Bob Sutter
11:30—12:00	Discussion of DMSO model directory and repository actions, including introduction of Dennis Shea (CNA), who may be assuming responsibility for developing the model directory: led by Iris Kameny
12:00—12:30	Lunch
12:30— 1:30	Update of DMSO Information System, including TOPIC, and IRDS actions (action item 2 on notes from March meeting): Cy Ardoin and Robert Schoen
1:30— 2:30	Discussion on use of IDEF tools (action item 6 on notes from March meeting): Twyla Courtot with contributions from Erwin Atzinger and Lana McGlynn on experience in use of IDEF
2:30— 3:00	Update and discussion on database directory schema.3 and IRDS-1 (action item 4 on notes from March meeting): Iris Kameny
3:00— 4:15	Discussion of how we should proceed with respect to: (below is a possible set of things to discuss, please add topics and I plan to collect them during the preceding day and 1/2)

1. setting up DB/I Task Group bulletin board, etc. under DMSO Information System and accessing and using it
2. use of IDEF tools for DMSO Information System
3. IRDS-1 compatibility
4. database and model directories

5. pros and cons of DMSO needing to establish own repository for data elements, databases, models

6. Arrange time for next meeting: 4:15 – 6:00. Meeting of core group: Cy Ardoin, Robert Schoen, Twyla Courtot, Dennis Shea, Iris Kameny, Jim Shiflett, Tom Shook

SECTION 3: LIST OF ATTENDEES

Cy Ardoin
Erwin Atzinger
Janet Baralli
Bob Bishop
Don Blanton
Stephanie Cammarata
Twyla Courtot
Martin Costellic
Cynthia Dengler
Ed Fitzsimmons
Jim Glymph
Dan Hogg
John Hovell
Terry Janssen
Bill Johnson
Iris Kameny
Dan Lewis
Charles Lillie
Lana McGlynn
Bud McNeil
Bob Molter
Jack Nicklas
Bao Nguyen
Russ Richards
Maj Mike Robinson
Lauri Rohn
Roberta Schoen
Allen Sears
Dennis Shea
Jim Shiflett
Tom Shook
Cassandra Smith
LtCol Charles Snead
Bill Surles
Bob Sutter
Walter Swindell
Rebecca Wade
Ken Wimmer
Dan Wu

SECTION 4: SUMMARY OF ISSUES

The DMSO DDI Task Group meeting on 4 – 5 June had a very full agenda. It began with DDI/DISA/CIM briefings on DoD data administration policy and procedures, DoD data model, Defense Data Repository System (DDRS), and the training program. This was followed by data administrators from the Joint Staff, Air Force, Navy, and Army discussing their programs and a brief on Asset Source for Engineering Technology (ASSET), the ARPA repository system. Several programs were briefed that provide data support to M&S: the Army initiative in M&S data management including the TRAC Automated Data System (TADS), the J-MASS M&S Reuse Library (MSRL), which is intended to become part of the DDRS program, Close Combat Tactical Trainer (CCTT) program data collection efforts, the OASIS project which is addressing J-8 M&S data management needs, and a long-range project from Argonne Labs on a model base concept for model interoperability. The meeting concluded with short updates on the DMSO Information System, IDEF tools, and the data directory and model directory efforts.

Impressions from the meeting are that there is widespread recognition of the need for data administration and management throughout DoD and efforts are underway in all of the Services and the Joint Staff as well as many of the DoD agencies. Current dictionary efforts include: DISA/CIM, AT&T 3B2 running Oracle/DDRS; Air Force, Microvax 5000/Ultrix running Sybase/MIDAS; Army, IBM running Oracle/ADSS; Joint staff, VAX 8600 running Oracle/WISDIM; DIA, IBM running M204/IDEAS; and OASIS dictionary effort for J-8, SUN, Ingres/OASIS; and DLA is also working on a system. There appears to be no DoD or Joint effort directed toward distributed exchange among these dictionaries (the DDI believes eventually they will all use the DDRS software). However, the ARPA repository prospect, ASSET, plans to develop distributed processing capability among repositories and the AF program will address interchange among MIDAS distributed servers.

The DoD/CIM effort is directed toward creating a single integrated DoD Data Model and one DoD Data Dictionary maintained in the DDRS. This is a very ambitious task and Bob Molter, reporting on DDI policy and procedures, estimated it will take ten years to achieve this goal. Although attention focused on the policy and procedures for the future when there is an established DoD Data Model and DDRS, so far no one has focused on what to do in the interim. This was brought up at the 10 June Pentagon meeting to review 8320.1-M-1 (attended by most of the CDAds who attended the DMSO meeting on 4 June) and is a task that Jerry Cooper will address.

Currently: (1) the DoD Data Model exists only at a high level (not detailed enough to yield entities on which to base SDEs); (2) there is approved policy for data administration but no approved supporting procedures; (3) the DDRS software is immature and incomplete; and (4) DISA/CIM is developing reverse engineering approaches to apply to all legacy systems. Given the current state, it will be difficult for the CDAds and FDAds to plan on how they will carry out data administration. However, DISA is requiring such a plan by 31 July. As soon as a

draft 8320.1-M-1 can be agreed to and the DDRS software is stable, dictionary partitions can be created for the components and functional areas and candidate data elements can be entered that have been generated (outside of the DoD Data Model) for later approval just to get things started.

Several issues were brought out:

(1) The DISA/CIM data dictionary development is focused on defining atomic data elements at this time. The M&S community has a need to define more complex data structures, such as derived data (e.g., P(k)), composite data (e.g., BE Number), and objects;

(2) There appears to be too few trained people in the components to perform the process and data modeling tasks, data element standardization, reverse engineering, etc., since many of their people have been recruited by DISA/CIM;

(3) Security of the DDRS has not been addressed; however, the AF is addressing security in MIDAS and this is a reason why Navy Intelligence and the Coast Guard are using the Air Force MIDAS system; and

(4) There seems to be a need for an organization that will support jointness. Since the CINCs are their own CDAds, it appears that each CINC, Service, Agency, Joint Staff, etc., is making its own plans, and the DDI will bring them together in integrating their data model into the DoD Data Model and having their nominated data elements reviewed across functional areas and component boundaries. One wonders if this will be enough.

Bob Molter later (on Wednesday) expressed some concern with the DMSO plans for the Information System, directories, and even a suggestion of another dictionary. He thinks these are DDI/DISA/CIM concerns and DMSO should be a user of their products. I would agree if they had products to use (which they do not now have) and near-term plans to build directories. It seems that since they have so much on their plate now, if somehow the DMSO effort could be coordinated to contribute to their long-term goal, it would serve DMSO in the short term and DoD in the longer term.

SECTION 5: LIST OF ACTION ITEMS

(1) Documents needed: KEN WIMMER should get DMSO on Jerry Cooper's (DISA/CIM) document distribution list so that we get new versions automatically. We need: DoD 8020.1, AF 42–9, JCS Pub 13.5, MIDAS (from Bao Nguyen), INFOSPAN (from Bao Nguyen).

(2) DMSO Information System: CY ARDOIN AND ROBERTA SCHOEN

Includes:

— firming up screens

— getting T1 speed to RAND so we can move the I/DB Task Group information onto the system and begin using it

— addressing the email package issue as discussed (e.g., select one package)

— using the process definition example for email for definin[illegible]ner relevant processes

— DMSO Information System documentation, write: system description/functional description; user manual (about five pages); top level system drawing

— develop a system evaluation and acceptance plan IDA and DTIC

— develop schedule dates with respect to development, testing, acceptance, transitioning (IDA and DTIC)

(3) Data Administration policy and procedures for M&S community: TWYLA

— identify and determine how to address DA areas (e.g., M&S as a functional area, M&S community guidelines wrt data modeling, handling of M&S atomic DEs and M&S non-atomic DEs, legacy systems, etc.)

— attend and contribute to relevant DDI/DISA/CIM meetings such as Bob Molter's meetings addressing 8320.1-M-1

(4) Establish requirements for DMSO Data Dictionary from M&S user perspective: TWYLA

— interface on non-atomic data needs with IRIS AND STEPHANIE

— if a DD is required, then evaluate options making use of and commenting on study for DISA/CIM by GMU and NIST on the many different data dictionary implementations

(5) Directories:

— database directory: IRIS AND STEPHANIE from schema.3: develop an E-R model, a relational model, define and name the relations and fields, develop preliminary search terms/structures, and implement the directory at RAND for evaluation and review.

— model directory: DENNIS SHEA interface with Lana McGlynn on study for Army model catalog, J-8 catalog, J-6 catalog, and those directory/catalog schemas retrieved from DTIC search to develop schema for model directory (estimated time needed is two months)

— organization directory: CY AND ROBERTA

— directory of directories: IRIS (after database directory is finished)

(6) Study and develop representation for extending SDEs and standard domains to non-atomic data types (e.g., models, matrices, derived data such as P(k)). IRIS AND STEPHANIE - address representation and derivation in models (e.g, P(k)) - address data administration procedures to support new representations: evaluate the old/existing, make modification, or develop new

(7) POCs for I/DB Task Group: IRIS contact Commander Ted Blackwell for Navy POC information, look for reps from Air Force and JS

(8) Get Navy results on state-of-practice in data administration survey: TWYLA

SECTION 6: JIM SHIFLETT'S INTRODUCTION

Dr. Steve Kimmel is the Functional Data Administrator (FDAd) for Acquisition, which includes DMSO. M&S standards will be developed in coordination with JIEO. For example, DMSO is nominating DIS as a M&S unique standard, but some other organization such as JIEO should establish a standard in this area as part of networking standards. Security issues should be addressed by the Joint Staff. Jim will be leaving the DMSO sometime in July or August to become the CCTT PM in Orlando.

SECTION 7: SESSION ON DOD DATA ADMINISTRATION, DATA MODEL, REPOSITORY, AND TRAINING

BOB MOLTER: DOD DATA ADMINISTRATION, POLICY AND PROCEDURES

Bob Molter is responsible for data administration policy for DoD. Under the policy and procedures being established, he would expect the entire DoD data model, data definitions, etc. to be available in 10 years. The solution to data administration is data modeling and one DoD dictionary.

Implementation status of documents:

Process Modeling: DoD 8020.1 draft will be available in June

Data Administration Strategic Plan (DASP), draft April, published August
DoD 8320.1 approved Sept 1991
Data Administration Plans and Procedures DoD 8320.1-M, draft May in formal coordination now
Standard Data Element Development, Approval, and Maintenance Procedures DoD 8320.1-M-1, formal coordination, May 1992
DoD Dictionary System (DDRS), operational acceptance test June, procedures published in August 1992
Draft DoD Cross-Functional Data Model, April 1992
Draft DoD Data Model, October 1992
Data Migration Prototype:
draft and identify systems, May 1992
begin reverse engineering of data, June 1992
complete prototype, October 1992
Data Administration training, class preparation: March—September 1992

Jerry Cooper (703–285–5383) should be contacted for documents, and we (DMSO, Ken Wimmer) need to call about getting on their distribution list.

The DASP will cover the period from now until the first plan is in place. The responsibilities of the Functional Information Managers (FIMs) are laid out in 8020.1. At present only three FIMs have been identified.

To facilitate the coordination process for DoD 8320.1-m-1, two meetings will be held: 10 June at 9:00 in 5C1040, and 17 June at 9:30 in 4E327 to finish up the first meeting and discuss metadata.

A goal of data administration is to develop standard data elements through data modeling by developing classification of data and using formal names. The DoD data model will be part of the DoD enterprise model. (Twyla suggested talking to Bill Kenworthy (X3L8) about standards for data representation, domain independent taxonomy.)

The objective is to name all data elements that cross components, which is common use data, and does not include component unique data. CIM is not recommending trying to stop existing component data dictionary efforts but believes that over time the components will end up using the same software CIM develops for the DoD repository. Strassmann believes the DoD dictionary will be so good that there will be no need for other dictionaries and others will naturally migrate to the DoD system. However, (my note) there currently seems to be no distributed architecture (e.g., client/user) design for distributed use of the DoD dictionary or any interim architecture for distributed use across existing and planned component and DoD dictionaries.

RUSS RICHARDS: DOD DATA MODEL

Phase 1 of a three phase project to develop a DoD-wide data model (a cross-functional model) has been completed. Phase 2 will include integrating models from the components and CINCs, and Phase 3 will create the high level DoD data model. A number of major DoD data models (component, major command, and functional) have also been developed. A subcommittee of the Data Administration Council is proposed to serve as a forum to build consensus in the process of: identifying commonalties among models, integrating the models into the DoD data model, and supporting the evolution and standardization of the DoD data model for data standardization. The DoD data model will be supported by a CASE tool, I.E. Expert, and will be available for distribution electronically as well as in hard copy. The current version supports 700–800 management statements that are cross-referenced in the model and this is used to explain the data model to users. However, the CASE tool representation needs to be converted to the IDEF representation and this is time consuming. They plan to maintain the data model in I.E. Expert and then convert to IDEF when necessary. A data exchange format is needed to go from I.E. Expert to IDEF. Also need standardization of graphical and semantic conventions for representation of data models. NIST is working on a FIPS for process and data models.

So far twelve prime object names have been identified. The example IDEF diagram looked rather busy for users, and Army people agreed that real examples become too complex to be followed and so they use the text rule presentation format instead. Russ explained that there is a need for reverse engineering to add data elements that are not defined by standard data or migrating data but exist in legacy systems that will not be migrated in the near term. This was labeled as boundary data.

DAN LEWIS: THE DEFENSE DATA REPOSITORY SYSTEM (DDRS)

The DDRS is a centrally controlled DoD-wide repository that manages and stores standard data elements, definitions, and associated metadata.

The DDRS runs on AT&T 3B2, using Oracle RDBMS Version 6, written in ETIP language generating C or Ada, and connectivity over DDN, local dial-up, using VT100 terminal type. ETIP is an AT&T 4th Generation Language (4GL). They are currently looking at other platforms such as Sun and Intel PC.

(My comments: the use of ETIP and similar use of 4GLs by other dictionary systems can reduce portability of moving dictionary applications from one RDBMS to another. Using 4GLs enables one to build application systems much more effectively and at less cost, but may lock one into a particular system. There is probably no answer to this, as the commercial vendors try to define a niche by offering different (better) functionality and tools than the next vendor but it will be a problem until "standards" are defined for interchange across the dictionary systems or for 4GLs.)

There is a Component Data Administrator (CDAd) for each service or agency. The data element approval process begins when a DE is submitted to a CDAd. The CDAd reviews and proposes it to the DDAd, who forwards it to the FDAd (Functional Data Administrator) (of which there are 78), who reviews it. Each proposed DE must have generic and prime object names in accord with accepted standards. Initial Operational Capability (IOC) of DDRS is August 1992 and a set of enhancements will be added by October. The DDRS supports: development and approval process for data elements; query of DDRS; electronic mail; bulletin board; report production; and maintenance of AIS and ODA information. The DDRS is based on the Army Data Dictionary (ADD) system and WISDIM. They are planning to port it to ORACLE PC runtime systems so it will be less expensive for people to get copies. The enhancements include the integration of data models and the translation to/from different tools.

Each SDE is associated with a single FDAd steward.

JOHN HOVELL: DATA ADMINISTRATION EDUCATION AND TRAINING

DISA/CIM has plans to develop a comprehensive training plan, establish a professional training organization, and operate a professional, accredited training facility. They will be accrediting the training facility, coordinating with other DoD schools and higher learning institutions, using video taped training, maintaining a training catalog, and becoming a professional training center.

CIM PANEL

(1) Question: How is naming and representation of units and accuracy of SDEs handled? Answer: FDAd will decide unit of measure if appropriate and the naming and accuracy representations. Functional managers make decisions on representation. Discussion of P(k) (probability of kill) used frequently in M&S community: it is a shared concept and would most likely be handled as an SDE for that reason, but since it is not an atomic data element its definition may not be handled in the near term by the DISA/CIM data administration standards.

(2) Question: How will CIM use the results of the Army data model activity? Answer: Russ Richards, DISA/CIM, will be bringing all people developing their models in to discuss the extensions required to the DoD data model in order to integrate the new models. A subcommittee of data administrators will be looking at the models. So far, the only M&S model is in the Army. It is one of 17 functional areas.

(3) Aside, either Becky Harris or Russ Richards will represent DISA/CIM at I/DB Task Group meetings.

(4) Question: How is the accuracy of representation determined? Answer: standards use the highest level of accuracy.

(5) Question: How is DISA/CIM addressing representation of objects, matrices, etc.? Answer: DISA/CIM is now focused on data elements residing in RDBMSs, and the need to manage at the atomic data level – not addressing other data types now.

(6) Question: Are SDEs defined ONLY for data to be shared across components and/or functional areas? Answer: they should be defined within components when they cross major commands or cross functional boundaries (e.g., logistics and operations). FDAd's need training to understand their responsibilities. There are already classes set up for their training.

(7) Question: How are the data administration policies and procedures being integrated into DoD software development policies and procedures? Answer: DISA/CIM is responsible for integrating this into the DoD software directives. MITRE has written a document about what needs to be done.

(8) Question: What is the formal approval process for a data model? Answer: It hasn't been determined yet what form of approval and validation will be required. Right now, the DoD data model is just stored in the DDRS.

(9) Question: What is the formal approval process for an SDE? Answer: User looks for an SDE to fit his/her DE in the DDRS; if an adequate SDE is not there, he/she defines the requirement; submits it to his/her CDAd who reviews it, etc.; if new, it is submitted to the DDAd who has it reviewed by relevant FDAd who may pass it off to other FDAd's; if approved, it is entered into the DDRS, else is returned to user.

(10) Question: How large is the set of prime object names? Answer: As the data model detail increases, the list of prime object names will expand.

(11) Question: Can contractors attend DISA/CIM training classes? Answer: now catering to DoD but contractors can be registered through a DoD component if they are working for them.

(12) Question: How labor intensive will reverse engineering be? Answer: they are beginning with two straightforward prototypes and then will be moving to more complex one. Lauri Rohn said that the more complex one may be a PA&E mobility model.

SECTION 8: SESSION ON DATA ADMINISTRATION IN SERVICES, JOINT STAFF, AND ARPA

JANET BARALLI, JOINT STAFF DATA ADMINISTRATOR REPRESENTATIVE

VADM Macke is the data administrator for the JS. In accord with 8320.1, the JS is not the CDAd for the CINCs; rather each CINC has its own CDAd, but currently, the JS CDAd is coordinating the CINCs with respect to concurrence/non-concurrence of DISA/CIM documents. JS data arise in the (1) joint reporting structure of WWMCCS and JOPES; (2) C4I for the warrior area: data elements for joint and combined arena including joint tactical warfighter, JOPES, SORTS, USMTF program; and (3) Intelligence: MIIDS and IDEAS. She mentioned limited resources to pursue efforts.

BAO NGUYEN, AIR FORCE DATA ADMINISTRATOR

The AF data administrator office was established in 1990 and is working closely with the JS and Army. AFIRDS (Air Force Information Resource Dictionary System) has a three phase development plan: data dictionary, IRDS, and multiple repositories. Phase 1: (1) automatically builds standard names; (2) supports MLS security (includes protection of location of classified data); (3) holds legacy data that has been standardized in gateway files until it is modeled; and (4) provides low cost/convenient access (when update at Pentagon main server will automatically update dictionaries at distributed locations). An AF directive is to introduce use of standards when modernizing; in interim, map legacy data to SDEs.

Documents about data administration	Air Force AF 42–9
	Army A25–9
	JCS Pub 13.5

They have used data modeling to integrate TAC and MAC data. They will be developing a data model in an evolutionary fashion and there are questions as to whether this will work and the extensibility of data models.

Phase 2 will use an INFOSPAN repository that is supposed to conform to FIPS IRDS and cover process model to source to goals. There should be an interface with IDEF. Russ Richards is integrating the AF and DoD model using I.E. Expert. Phase 2 will include multiple sources of information, CASE tool integration, vendor independence, tools to view information and flexible categorization.

Phase 3 will integrate with other dictionaries and repositories.

INFOSPAN: June 11 user's group meeting to explore IRDS repository compatibility. DDRS will have INFOSPAN in front of it. The AF wanted to work together with the Army on a compatible data dictionary but the AF is based on an open system and the Army on IBM. WISDIM version 1 ran under VMS, had single security level, and didn't fit the AF purpose. Current architecture is MLS using Sybase on a Sun and also supports referential integrity, triggers, and client/server architecture. A server costs $36K and client software only $400.00. The Coast Guard will be using MIDAS on a VMS system, Navy Intelligence on a Unix base and AF under Ultrix. To interface the AF data dictionary to JOPES, the Air Force format is translated into JOPES format, can also go from a data model into a relational table.

Bao will send us a copy of the MIDAS and INFOSPAN documents.

Planned: need to integrate (IDEF) CASE tools and other tools to do things IDEF can't do. Difference between MIDAS and DDRS is that MIDAS is operationally oriented and DDRS is high level.

The Secretary of the Air Force for Logistics is strong on MIDAS: thus wants to make sure the DoD selection for DDRS is best of MIDAS and others. There is a study being conducted by GMU Information Systems Department and NIST (Sibley, Goldfine, and Rosen) to evaluate the capability of MIDAS, RAPID, WISDIM, and DDRS – should be available next week.

The AF has 800 DEs for AF C2 systems that are supposed to be converted to JOPES data element definitions.

REBECCA WADE, NAVY DATA ADMINISTRATOR

The Navy Data Administration Program is part of the Navy Information Systems Management Center (NISMC) and reports to DASN C4I/EW/Space of the ASN RD&A. Rebecca is the DON CDAd and under her is Major Dave Duff, the CDAd for the Marine Corps. They are in the process of establishing USN and MC data stewards, and have organized a data administration action team for providing training and spreading data administration ideas around. The Navy has a high level data model for functional areas of the Department of Navy (DON). They are currently putting in place the management structure and plan for funding a C3I data model next year.

They are expecting the Navy and Marine Corps data stewards to coordinate within the DoD functional areas, though there may not be a one-to-one correspondence. For example, for the one DoD personnel area, the Navy may have three areas: civilian, Navy, and Marines. The principle they are following is to tailor the DoD policy and procedures to suit their organizations. They will require that a proposed data element be coordinated between the Navy and Marines.

The Navy has participated heavily in 8320.1 but currently find it difficult to stay abreast of 8320.1-M-1 as there have been so many versions. They have prepared a draft DON policy and would like it to be in synch with DoD, but if they haven't seen good DoD progress by the end of FY92 they will need to proceed with DON policy, though they intended it to be supplemental to and consistent with DoD policy.

They are expecting NIST (Judy Newton) to issue a DAMA Data Administration Guideline soon. The Navy intends to use this guideline. Rebecca noted that CDAds and FDAds have been instructed to send DA plans to DISA (Dennis Brown) by end of July but without the DoD procedure documents, this will be difficult. She noted that critical success in DA is dependent on the FDAds but the CDAds are the ones with the functional knowledge and expertise that is needed.

The Navy is planning to have a contractor come in and train people in use of IDEF models. They are trying to increase the awareness of functional managers to standard data and data administration and to motivate them, but without teeth to enforce standards it may be difficult. However, the Navy is developing a self-evaluation guideline that managers can use. Bao Nguyen said that the Air Force is enforcing use of standards by auditing programs and withholding funding from those found to be non-compliant.

Rebecca offered a set of definitions for a repository:

> "A software tool for defining, storing and managing all the information and objects needed to accomplish the corporate missions and functions of the DoD. A repository is a much broader concept than a data dictionary in that it must also provide extensive support for modeling and have interfaces to a variety of external applications, including CASE tools." from Final Draft of DoD Information Resources Repository Requirements Definition of 8 November 1991

Definitions from Draft DoD Manual 8320.1-M-1

> "A specialized type of database containing metadata that is managed by a data dictionary system."
>
> "A repository of information describing the characteristics of data used to design, monitor, document, project, and control data in information systems and databases."
>
> "An application of a data dictionary system."
>
> "Provides a central repository of information about data, such as meaning, relationships to other data, origin, usage, and format."

The same software package could be used for the repository and the data dictionary. They are using DATAMANAGER and trying to decide what kind of data dictionary software/hardware the Navy should get. Do they need a separate Navy component data dictionary or can DDRS provide the service and performance they require? They are waiting to see if DDRS meets its August schedule and also watching the DDRS architecture.

Data Model integration within a functional area: they have identified 19 Navy personnel data models, and it takes lots of analysis and human effort to integrate these. Mention was made that Russ Richard's chairs a DoD Data Model Integration Oversight meeting once every two weeks for anyone interested in participating.

Rebecca suggested Iris Kameny get in touch with Commander Ted Blackwell, a C3I POC, for a M&S POC from the Navy to represent the Navy in future I/DB meetings.

The Navy has been conducting a DoD survey on state-of-practice in data administration and will share their survey questions with us. They are building a database of the results for DISA and will be able to share the database format later in June and should have the analysis completed by July. She showed the existing Navy models going into DATAMANAGER and DESIGNMANAGER tools and into I.E. Expert. This includes the DON strategic data model and subject area data models and then these would be passed to the DDRS. She stressed that the Navy has limited resources to support these data administration activities.

JIM GLYMPH, ARMY DATA ADMINISTRATOR

The Army DA program has 30 government people, no contractor funding, and some DoD support for Army data management. The Army Data Dictionary has been developed and maintained by the government. They have produced two main documents: the Army Capstone Information Model and the Army Capstone Data Architecture. The earlier Army Capstone Model doesn't look so good now and they have extended the functional areas to produce the new Army Data Model. The Capstone model had 74 information classes and 71 processes. The real key to success is data stewards: there are a total of 19 data stewards responsible for the information classes, and a data steward is assigned to each information class. They have used about 10 modeling methodologies. They had not done functional area data modeling for the Capstone Model. Now they are truly decomposing 19 functional areas and have finished 16 of the 19 data models.

They are using IDEF0 and IDEF1X tools. His list of tools that supported IDEF included: Leverage, PC Modeler, Accelerator, and Oracle DBMS.

The Army Model has been put in the Corps of Engineers repository. Jim gave an example of an activity node tree approach that they use in modeling.

There are many models in the Army and not all are business models. The Army model is in an Oracle database and can be downloaded to modelers on PCs, by being passed as business rules. There are many different tools that can be used to provide different presentations of the model information.

AR25–9 describes the Army standard data element schema, and defines reference elements, generic domains, specific domains, etc. There are 88 reference elements. The format of a data element is a right hand reference element composed of a class word (e.g., weight) that can be qualified (similar to DoD CIM current thinking) and is mandatory, and on the left hand side, a prime term composed of at least one prime word and any number of prime modifiers.

Two major cross functional Army areas are logistics and personnel.

The DoD DDRS approval process is based on the Army Data Standardization System (ADSS). Jim has five people working full-time in reviewing proposed Army standard data elements. So far the Army has defined: 88 reference elements, 650

data elements, and has 247 users including the Navy, Air Force, and contractors. The ADSS supports query, electronic mail, reports, batch load, alias collection, information system use, audit, etc. They offer 1 day training classes in ADSS.

SECTION 9: SESSION ON ARMY SUPPORT FOR M&S MODELS

LANA MCGLYNN, ARMY MODEL AND SIMULATION MANAGEMENT OFFICE (AMSMO)

The Army's path to data standardization is to fill in the gap between the current reality of no integrated M&S data support structure and the vision to arrive at visibility of data requirements and standardization of data elements. This is being managed by organizing the AMSMO as a policy organization with an Army M&S Executive Council advising and technical coordination with HQDA, TRADOC, AMC, OPTEC, and other modelers.

Lana does not believe that M&S is a functional area, but rather a user of data and needs to participate as users to determine the functional data models. She believes there will be few proposed M&S data elements that do not have data stewards in other functional areas. She sees the M&S community efforts as providing information to the data modeling and data element definition efforts but not as data proponents. The Army Data Model is the result of activity modeling (using IDEF0), data modeling (using IDEF1X), and data standardization with data elements stored in the Army Data Dictionary. The payoff is in supporting the move to open systems environment, cost reduction for system development, and improved data management.

DON BLANTON: HOW THE ARMY IS ORGANIZED TO SUPPORT M&S DATA

The Army M&S community needs information to do their work. There needs to be communication between the data suppliers and what the models require. Modelers are motivated toward an efficient way to perform their studies and not to build information systems.

The AMIP Data Management Committee is chaired by the Director of AMSAA and members include representatives from CAA, TRAC, INSCOM, AMSMO, AMSAA, and others. They are concerned with the development of communication protocols, nomenclature, and data structures. Responsibilities: CAA for force/theater level modeling; TRADOC for corps/division level modeling, and combined arms and support task force modeling; AMC for item level system performance databases; and AIA for threat and allied characteristics databases. Item level databases include: C2, communications, IEW, CSS, combat support, air defense, maneuver aviation, maneuver infantry, fire support, and maneuver armor.

In the model VV&A process it is critical to know how data will be used in different models. This started a discussion about data "certification," what was meant by it, whether it was too hard or impossible to do, etc. No conclusions were reached.

WALT SWINDELL: TRAC AUTOMATED DATA SYSTEM (TADS)

TADS is a method to electronically request, receive, authenticate, graphically display, mathematically transform, and reformat data from data providers into TRADOC's combat development models. Its benefits include tightened quality control over data, and faster, improved analysis. Its current scope is the TRADOC combat development community and it currently manages over 7 billion bytes of technical data. The kinds of data include: terrain, weather, performance, operational, characteristics, P(k)s, etc. They recertify the data for new studies. The sources of data are CAC, AMC, WES and the customers are TRAC models, TRADOC, CAA, RAND, PEOs, etc.

In the future they will have a data dictionary of standardized data element definitions. They are using Ingres 6.3 and the database is classified. There are several steps (1) automated data request: defines scenario, year(s), weapon system, theater, etc.; (2) system produces list of weapon pairings and sends to AMSAA and BRL for relevant raw data; (3) checks raw data coming in to identify anomalies; and (4) if it checks OK, then goes into functional area database from which data is preprocessed and transformed to meet the model requirements; and (5) then provides the data to the modeler that requested it.

They have a standardized nomenclature document and the data in the databases are normalized somewhere between 3rd and 4th normal form. They will be using standardized names in the future, and will be submitting new entities to the Army Data Dictionary. They are currently proposing five additional entities to the Army data model.

LANA MCGLYNN: ARMY INITIATIVE IN M&S MANAGEMENT

The AMSMO is developing an Army M&S Master Catalog of models that will be online, electronic, centralized, and support interactive searching, and uploading and downloading of catalog information. The current task is to field surveys and look at existing catalogs to determine the software and hardware requirements. They have a sample schema. The draft report on this effort is due 15 June and the final report 15 July. During phase II, they will develop the system (Jan 1993) and phase III implementation is scheduled to be complete by Sept 1993. The catalog will contain VV&A information on models and they are developing the methodology needed to implement AR5–11 policy in concert with ongoing MORS efforts (SIMVAL).

SECTION 10: SESSION ON DATA COLLECTION, SIMULATION MODELS

IRIS KAMENY: JOINT MODELING AND SIMULATION SYSTEM (J-MASS) PROGRAM (INFORMATION FURNISHED BY MIKE HUCUL)

The objective of J-MASS is to develop a standardized digital M&S capability with which to develop, test, and assess the capabilities of weapon systems in a

simulated operational environment. Two parts to J-MASS: a simulation support environment to enable users to create models, configure scenarios, execute simulations and analyze results; and a modeling library which contains models and model components of weapons systems from RDT&E community, threat systems from Intelligence community, and environmental effects from scientific community. The Modeling and Simulation Reuse Library (MSRL) is currently using the Army's Reusable Ada Packages for Information System Development (RAPID) and will be moving to the DISA/CIM Defense Software Repository System (DSRS) program.

CHUCK LILLIE: ASSET SOURCE FOR SOFTWARE ENGINEERING TECHNOLOGY (ASSET)

The goal of this ARPA effort is to establish a distributed support system for software reuse. Short term: implement a software reuse library and become focal point for software reuse; long term: help stimulate a U.S. software reuse industry. Activities include: asset acquisition, categorization (faceted classification, want seven now have four), and distribution; asset configuration management; recall; setting up local reuse programs and repositories; and "yellow pages" for reuse goods and services. They are trying to take a windows approach to user interface. Interoperability between repositories is about 5 years away in developing necessary standards and protocols to move software between repositories. The STARS Reuse Library Facility store is based on AI techniques.

Technology interests include: distributed networking of repositories, interchange of assets among repositories, confidence indicators, and "seamless" integration with local environments and repositories. Confidence indicators are developed through asset evaluation at four levels: documented, audited, validated, and certified.

Some problems: legal documentation to protect from software problems. In government sponsored work, the contractor has copyright and needs legal protection in case of software bugs.

ASSET is chartered by ARPA and expected to be self-sufficient in five years but hope to make this by 1997. They are not chartered to do R&D or to develop repository technology.

From STARS they get a distributed network and ASSET Library Open Framework (ALOF) to exchange information. The ANSI repository standards group is working on exchange: two groups, RIGS and ALOF.

Their market analysis revealed that people are interested in using a repository but are not willing to take the risk. There are "not invented here" barriers, need for tailoring to suit application, and fact that people don't want to pay for the reuse, etc.

Related efforts: COSMIC is a NASA reuse project at Georgia Tech that supports many platforms; they have given assets to COSMIC to validate and certify.

CARDS is an Air Force reuse architecture with a command center domain. They may connect to CARDS, which is using an AI-based system. The STARS Reuse Library Facility store is based on AI techniques. AMIX is a commercial venture (takes about $50 to get an account) and they provide consulting services also. SAIC in McLean has a Simulation Reuse Library to reuse their own simulations.

BILL JOHNSON: CLOSE COMBAT TACTICAL TRAINER (CCTT) DATA COLLECTION PROGRAM

CCTT is a group of fully interactive networked simulators and C3 workstations that portray supporting combat, combat support, and combat service support elements and operate on a simulated realtime electronic battlefield. It furnishes platoon to company training and could go up to battalion level.

CCTT is a follow-on to SIMNET-T with more battlefield effects, greater field of vision, more realistic data package, open systems architecture, configuration management, and higher resolution terrain. The data collection program is intended to provide contractors with certified, accurate, and reusable data in a timely manner by establishing: an assistance office, data support network, CCTT data library, and performance data working group. RCI is contractor who will conduct data collection effort and establish an easily accessible but controlled database repository.

AMSAA is working the data issue of how to go from classified to unclassified data so they can satisfy the CCTT need for unclassified data. There are six CCTT data requirement categories: weapon system/equipment characteristics, weapon system/equipment performance, doctrine and tactics, occupational information, crew/force configuration, and environment. Data are collected from all over: engineering drawings, specs, firing tables, models, design documents, service bulletins, etc.

They need to perform VV&A on the data to verify it is complete, identify voids, identify discrepancies, validate that data are correct, and certify by CCTT TRADOC manager that information is acceptable.

DAN HOGG: OPERATIONS ANALYSIS AND SIMULATION INTERFACE SYSTEM (OASIS)

Program is located under J-8 Director of Force Structure, Resources, and Assessment under Deputy Director for Technical Operations, Automation Support. The mission is to develop a system which will significantly improve data collection, access, verification and validation, analysis, reporting, management, and documentation for J-8 studies and analysis processes.

The tasking for a centralized DMS began in 1990 when they saw a Westinghouse prototype at CAA. Contract began in spring 1991. They have a 100 page data management plan. They are looking at the overall system with a #1 priority to satisfy J-8 needs and at same time to be compliant with broad data

needs. Hogg is interacting with M&S people in CINC organizations. OASIS has been given directions to be in conformance with WWMCCS data element names but they don't want to stop the development now to do so.

The analytical suite is Sun/Unix/Ingres, Vax cluster/VMS/Ingres, and Network File Server. They have used E-R diagrams for modeling. The data files/folders can be reference or study type folders where a reference folder contains data that cannot be changed by the user. The user may copy the data he needs into a study folder and can make modifications to it there. There are three levels: folders, class, and list level attributes. Examples of a class are a blue target base class which has objects of installations. There can be sets within a class.

Their data contents change frequently so OASIS supports change of format mapping at runtime. They have a "transfer" screen to change the mapping of files: external to internal, internal to internal, internal to external.

Their data dictionary defines the data tables within a database system. OASIS contains an on-line dynamic data dictionary, data dictionary system tables, and data dictionary access screens. OASIS is based on relational data modeling.

With future development they will try to include the requirement that the model must execute within OASIS. They think they will take over sourcing of conventional force databases at CINCs.

In their own way they are doing code reuse. A future system enhancement may be use of Ingres knowledge management tool for better V&V of data.

J-8 concerns: running at TS system high and need MLS boxes, need to look at move to a trusted MLS environment; need to comply with DoD data standardization efforts—they want to but it needs to be something easy for them to do.

They are willing to furnish DMSO with database entries for the database directory.

SYSCON is putting together a J-8 M&S catalog.

BOB SUTTER: A MODELBASE CONCEPT FOR MODEL INTEROPERABILITY

The purpose is to reverse the trends of building large, proprietary models, to provide a methodology for model interoperability, and to develop an architecture to support future model development. Today they are introducing development concepts for storing model information, illustrating connections to database efforts, and proposing model information be separated from conventional databases. The issues are model interoperability and how to build a model-based information system. They have a concept of a modelbase management system that would interact with a modelbase of models and a model dictionary containing data about model capabilities (based on mission essential task lists), model characteristics, and model data requirements. Sponsors include J-8 and Air Force XOR.

Their approach is to address concept model development methods, illustrate the interoperability problem with a simplified approach, and emphasize what/how to store model information (vs. model construction).

Future efforts include developing prototypes of model object interaction and data object interactions; model selection methodology; and intelligent assistance in defining the modeling task.

SECTION 11: CY ARDOIN: TOPIC AND IRDS REPORT (ACTION ITEM FROM LAST MEETING)

Main issue presented was that TOPIC is geared for search and retrieval only and IRDS also requires functions of adding and deleting data and information.

DMSO will need a relational model to support the data element standardization process. It may also need a relational DBMS to support the database and model directories.

My comment: Though it hasn't been determined that DMSO requires its own data dictionary at this time, the briefings during the meeting indicate that DISA/CIM will not be addressing complex or non-atomic data element representation in the near term. Since this appears to be very important to the M&S community, it may be necessary that some organization, such as DMSO, develop and maintain a data dictionary that can support these kinds of data elements for the M&S community.

Another point is that IRDS requires data element management but a relational data management system may not be enough based on the presentations given above. Each organization seems to be building its own data dictionary using different hardware/software platforms, and implementing application software to perform the IRDS-like and other functions mainly using a system 4GL. The user interface and functionality may then not be portable though the data and the data structures may be. Also, these systems are more or less compliant with IRDS-1 but there is no "accredited" IRDS compliant product on the market now. If DMSO needs to have a data dictionary facility, then some time needs to be spent evaluating what that should be. One action item should be to look at the outcome of the current GMU and NIST study that is evaluating the different systems for DISA/CIM.

The unresolved issues presented were: (1) relational DBMS or object-oriented DBMS, in the sense of whether relational will be enough; (2) screens for the DMSO Information System need firming up; (3) testing schedule has to be agreed to and also what is being tested; (4) target date for transition of the system from IDA to DTIC; (5) plans for the TWGs to begin using the system.

Cy said the initial concept of letting unknown users on for limited usage has been changed for security reasons and will not be allowed. He is planning for initial

use of the system to have available: A model catalog, glued together from the J8 and Transcom models; definitions file composed from the JCS pub on definitions, an SAIC list of terms, and an IDA list of acronyms

He reported that the software source code is 95% complete and is 95% compiled and runs (has passed some amount of unit testing?) but very little has been installed. There have been modem problems due to flooding and power failures.

The NFS link is scheduled to be up on June 10. DDN access will still be possible but is slow and expensive, and the plan is to establish T1 speeds to RAND and other networks. (My aside to Cy: today is June 24 and the link to RAND is still very slow and not very usable.)

SECTION 12: TWYLA COURTOT: IDEF UPDATE AND DISCUSSION (ACTION ITEM FROM LAST MEETING)

Twyla presented a briefing prepared by Elaine Ward titled "The Latest on IDEF" that includes: purpose of IDEF0, its benefits and delimiters, what it is not, purpose of IDEF1 and 1X, IDEF model types, tool support, consulting support, key users/uses, and issues.

Purpose of IDEF0 is to support strategic planning and business process reengineering and to perform high-level process modeling. It is not a structured analysis technique and should not be used for detailed requirements analysis. Purpose of IDEF1 and 1X are to model data and information within the system; can be used independently of IDEF; and if used with IDEF, IDEF1 or 1X data can be mapped to IDEF0 model inputs and outputs.

Model types and uses:

IDEF0	—	process/activity modeling
IDEF1	—	information modeling (obsolete)
IDEF1X	—	relational data modeling
IDEF2	—	dynamics modeling (not used much)
IDEF3	—	simulation modeling (being worked on at Wright-Patterson)
IDEF4	—	object-oriented design (being worked on at Wright-Patterson).
		The rest are not defined although the names continue to IDEF14.

Twyla suggested we look at the following tools more carefully: ACCELERATOR, Leverage, I.E. Expert, and INFOSPAN.

WIZDOM conducted the MITRE training. It covered a four-day period, and contained three-days worth of content; there was not enough hands-on experience (and it included activity-based costing?). MITRE will be using it in their CIM and NIST support activities. MITRE is also interested in a new IDEF tool for the MAC produced by Triune Systems that costs around $500 and appears to be as good as the WIZDOM $8K tool. The contact is Douglas Bernard 513–237–0762.

The Army has used D. Appleton Company for IDEF training. They trained identified proponents and experts, and acted as facilitators in the Army modeling activities.

Key users of IDEF are CIM (business process), CACI (analysis and modeling of Desert Shield/Storm, GM (management improvement), and Imperial Bank of Canada (banking and finance processes).

There were four issues. The Air Force standard is out-of-date and has not been maintained or updated (there may be errors in the examples). There is no standard definition for each model or integration between models (NIST is working on producing FIPS standards). Purposes of models from IDEF2 on have not been firmly defined. Currently there is no interoperability between tool vendors (an exception may be between WIZDOM and Knowledge Based Systems by Oct. 92).

SECTION 13: IRIS KAMENY: DATABASE DIRECTORY SCHEMA.3 DISCUSSION (ACTION ITEM FROM LAST MEETING)

We went over the changes from schema.2 quickly. It was OK with everyone, or it was too late in the day, or people hadn't reviewed it, but there were no suggested changes and so Iris and Stephanie will do an E-R model, a relational model, define and name the relations and fields, develop preliminary search terms/structures, and implement it at RAND for evaluation and review. Stephanie has been exploring the feasibility of implementing it under TOPIC, but this doesn't seem likely. Implementing it in a RDBMS at RAND will make the database portable to whatever RDBMS we decide to use for DMSO, which will most likely be driven by the data dictionary requirements and decision.

Bob Sutter has offered possible data taxonomies from JMETALS and SYSCON and Marty Costellic furnished a copy of the U.S. Naval Institute Military Database and a TPDC evaluation of it to use as insight into the data search terms.

Three other suggestions were made with respect to the database directory:

(1) Include the DTIC database as an entry in the directory.

(2) Agreement to not include database schemas in the directory but rather get at that information through search of the CIM data dictionary or other means. This will make it easier for people to be responsive to furnishing entries, since schemas can be very lengthy and take time to produce in a required format, they are subject to change, etc.

(3) Add a new directory to the DMSO directories: a directory of directories.

Appendix G

NOTES FROM THE 4TH I/DB WORKSHOP, OCTOBER 7–8, 1992

SECTION 1: TABLE OF CONTENTS

Agenda	Section 2
List of Attendees	Section 3
Summary of Issues	Section 4
Action Items	Section 5
Tom Shook's Introduction	Section 6
Session on Update on Organizational Activities	Section 7
Session on Complex Data Types, Data Dictionaries, and Data Support for M&S	Section 8
Session on IDEF Related Discussions	Section 9
Session on Research Issues	Section 10
Session on DMSO Information System Reports	Section 11
Session on Data Verification, Validation and Certification	Section 12
Workshop Session to Facilitate Exchange of Goals, Approaches, Tools, etc. Among Groups Supplying Data to M&S Users	Section 13
General Areas of Interest for Complex Data and Common Tools	Section 14

SECTION 2: AGENDA

Agenda for October 7–8 Meeting of DMSO I/DB Task Group
at IDA, 2001 N. Beauregard St., Alexandria
Building 2001, Room 118

WEDNESDAY OCTOBER 7, 1992

8:00— 8:30 Overview of meeting and goals, Introduction of new people: Tom Shook, Iris Kameny

UPDATE ON ORGANIZATIONAL ACTIVITIES

8:30— 9:00	DISA/CIM: update on data administration, DoD Data Model, DDRS, standard data element schema and naming, migration systems, use of IDEF: Bob Molter
9:00— 9:30	Defense Management Report Decision (DMRD-918), Subject: Defense Information Infrastructure: Twyla Courtot
9:30—10:00	Update on Army M&S data activities: Lana McGlynn
10:00—10:15	Break
10:15—10:45	Overview of new Navy M&S organization and M&S data support activities: Jim Weatherly
10:45—11:15	DMSO Analysis Functional Group data needs: Wally Chandler
11:15—11:45	Air Force M&S analysis functional area needs: coordinated by Roy Reiss
11:45—12:30	Lunch

COMPLEX DATA TYPES, DATA DICTIONARIES, AND DATA SUPPORT FOR M&S

12:30—12:45	Overview of complex data elements and issues: Stephanie Cammarata
12:45 — 1:15	TADS complex data elements effort: Howard Haecker
1:15 — 2:15	Briefing on J-MASS: Randy Brown and Mike Hucul
2:15 — 2:45	Joint Data Base Elements (JDBE) for M&S: Steve Matsuura
2:45 — 3:00	Discussion of Friday workshop with groups furnishing data support for M&S (TADS, JDBE, OASIS, and Navy TIDES): Iris Kameny
3:00 — 3:15	Break
3:15 — 3:45	Overview of major DoD data dictionary efforts and comparison studies: Twyla Courtot
3:45 — 4:15	Discussion of what should be done, where shall we go? leader, Twyla Courtot

4:15 — 4:45	MITRE Distributed Heterogeneous Information System (DHIS) Testbed: Bill Carpenter and Don Rea

THURSDAY OCTOBER 8, 1992

IDEF RELATED DISCUSSIONS

8:00— 9:00	NIST update on IRDS, and standards for IDEF: Bruce Rosen
9:00— 9:30	PA&E report on use of IDEF1X on MIDAS model for force projection: Paul Rehmus
9:30—10:00	DMSO report on IDEF training and plans for using IDEF: Tom Shook and Ken Wimmer
10:00—10:30	MITRE's experience with IDEF training and use: Twyla Courtot
10:30—10:45	Break
10:45—11:15	Discussion of IDEF and next directions for M&S community and DMSO: leader, Twyla Courtot
11:15—11:30	Reports on: object-oriented standards meeting and CIM 8320.1-M-1 meetings: Twyla Courtot

RESEARCH ISSUES

11:30—12:00	DARPA Intelligent Integration of Information research program: Gio Wiederhold
12:00—12:30	Discussion of DMSO call for proposals for complex data and common tools: Iris Kameny
12:30— 1:00	Lunch

DMSO INFORMATION SYSTEM REPORTS

1:00— 2:00	DMSO Information System prototype and use: Cy Ardoin
2:00— 2:30	Discussion of use of system, suggested beta test users: Cy Ardoin
2:30— 3:00	Discussion of DMSO model directory: Dennis Shea
3:00— 3:15	Break
3:15— 3:45	Discussion of DMSO database directory: Iris Kameny

VERIFICATION, VALIDATION, AND CERTIFICATION

3:45— 4:00 **Overview of VV&C issues: Iris Kameny**

4:00— 5:00 **Panel and discussion on VV&C: Dennis Shea, Howard Haeker, Erwin Atzinger, Iris Kameny...**

SECTION 3: LIST OF ATTENDEES

Cy Ardoin	IDA
Erwin Atzinger	AMSAA
Bob Bishop	DTIC
Don Blanton	AMSAA
Stephanie Cammarata	RAND
Bill Carpenter	MITRE
Wallace Chandler	USA/CAA
Bill Clydesdale	SPAWAR 312
Gail Coffey	GPSC
Twyla Courtot	MITRE
David Danko	DMA
Tim Doane	GPSC
Linda Donaldson	USA/MRJ
Charlotte Gross	OSD/ODDI
Lucy Haddad	RCI-CCTT
Howard Haeker	USA/TRAC
Dan Hogg	JCS/J8
Roseann Hynes	DMA
Iris Kameny	RAND
LtCol Ken Konwin	AFSAA/SAG
Jim Lacey	SAIC/TMA233
Steve Lawyer	IDA
Mike Lilienthal	DMSO
COL. Mike Mancino	OASD(PA&E)
Steve Matsuura	USAEPG/JDBE
Janet McDonald	USAEPG/JDBE
Lana McGlynn	AMSMO
Bill McQuay	WL/AAWA-1
Miro Medek	MITRE
Bob Molter	OSD/ODDI
Jack Nicklas	RCI-CCTT
Don Rea	MITRE
Paul Rehmus	OSD(PA&E)
Roy Reiss	AFSAA/SAG
Lauri Rohn	OSD (PA&E)
Bruce Rosen	NIST
Mike Sarkovitz	NAVAIRSYSCOM
Roberta Schoen	DTIC
Dennis Shea	CNA

Tom Shook	DMSO
Cassandra Smith	MITRE
Bob Sutter	Argonne Lab
Walt Swindell	USATRAC
Peter Valentine	USAEPG/JDBE
Jim Weatherly	SPAWAR
Gio Wiederhold	DARPA
Ken Wimmer	SAIC/DMSO
Joyce Wineland	NAVINTCOM
Andrew Wirkkala	NAVAIR
Jeff Wolfe	DISA/CIM
Rob Wright	RCI-CCTT
Dan Wu	DISA/DFS
Simone Youngblood	JHU/APL

DMSO I/DB Task Group Members Not Attending:

Larry Buchsbaum	Navy
Patrick Cheatham	Aerospace
Martin Costellic	Defense TPDC
Ed Fitzsimmons	DMSO
Becky Harris	CIM/XF
Ollie Hedgepeth	Army
Mike Hucul	J-MASS
Ernie Lucier	NASA/SED
Bao Nguyen	Air Force
Dale Pace	JHU/APL
Pat Sanders	OSD/PA&E
Jim Shiflett	CCTT

SECTION 4: SUMMARY OF ISSUES

GENERAL: The I/DB participants said they found the meetings very useful, wanted them to continue, and agreed that every four months is often enough. Based on that, the next meeting will take place in the February-March time frame. The subgroup that met on Friday decided that they there was no need to form a subgroup of people from projects furnishing data to modelers and they will informally keep in contact with each other as desired.

This was the first I/DB meeting in which we had participation from the Navy and Air Force M&S communities and from a DMSO functional working group, the Analysis WG, and we benefited greatly from their participation. We also were interested in the OSD/PA&E briefing on how CIM will be helping them in applying IDEF1X to the MIDAS model and will want updates at future meetings.

PROGRAMS SUPPLYING DATA FOR M&S: The TADS and OASIS programs have accomplished much in data standardization in serving their communities of M&S modelers. They serve as good examples to the CCTT and

UTSS programs that are just starting up. Also the JDBE effort to develop subject area information models is a promising bottom-up approach that could be complementary to the CIM top-down approach.

CIM UPDATE: The report from CIM was encouraging in that 8320.1-M-1 is in better shape than four months ago, we received some sample data elements described by an example DDRS schema, and CIM has offered to support DMSO in investigating complex data types and in proposing extensions to the DDRS schema to accommodate them. DISA/CIM also offered a future view of a combined DDRS and DSRS repository supported by CASE tools that indicated integration across process and data models, and dictionary and software reuse codes. However, we really need to get together and work with CIM to be able to guide the JDBE, UTSS, and CCTT programs in their database, data dictionary, data modeling, and data standardization efforts so their architecture decisions will be such that they can easily accommodate to CIM requirements in the future. TADS and OASIS would also like to know what CIM expects of them in the future. This is a critical issue that needs addressing now in order to avoid wasting valuable resources.

An interesting aside is that there is no IRDS1 conformant product available to base a data dictionary on, so that the approach taken by OASIS and TADS to support the data dictionary within the same relational DBMS as the databases seems to be a good choice. Also, the issue of distributed repositories crops up again, since it seems in the M&S community if others follow the OASIS and TADS approaches, we will want to exchange data element standards and data among the different data support systems. If all the systems follow the open systems approach and use relational DBMSs which support the SQL standard, then both dictionary data and domain data can be exchanged given the systems provided mappings between their dictionary schemas.

SECURITY: Also, security has reared its head again, since of these five systems: OASIS is TS for now, TADS is S, CCTT is unclassified, UTSS will be multi-level, and JDBE may be classified or have some classified portions. Tom Shook said that DMSO is organizing a security WG to address security issues.

IDEF: The IDEF session offered insights into the fact that IDEF is more of a notation than a methodology or technology, and that the FIPS isn't a final product but a way to get something out quickly before the election. Also that a FIPS cannot standardize on a proprietary solution, therefore it can be standardizing on less than the best technology available. Bruce Rosen said that the frontispiece to the IDEF FIPS says that if one has decided to use IDEF modeling technology, then the FIPS is the standard to be followed. Some serious problems with IDEF are: there is no integration between IDEF0 models and IDEF1X models; there is no standard representation for exchange of IDEF models among tools; IDEF1X currently doesn't address M-N cardinality; it doesn't include representation of business rules; and it is limited in flexibility for use in reverse engineering of hierarchical and network data models. We need to get M&S data suppliers involved in the IDEF Users Group.

DATA VV&C: In the session on data VV&C there was general agreement that this is an important issue, difficult to address, and a management as well as technical issue. Most of the panel participants hit on the fact that there can be many types of data that could be selected for a given model (e.g., spec data, test data, combat data) and a type such as test data needs to have many caveats stated about the purposes and conditions under which it was collected in order to determine what data are proper for a model addressing a particular problem. There were questions about where in the process VV&C is done, at the source, after preprocessing (and who VV&As the preprocessor), both places, and how is data directly produced by one model for input to another model VV&Cd? It was generally accepted that although a source can do some VV&C on "standard" data that it makes available, the data need to be considered as part of the model VV&A process.

SUBJECTS FOR FUTURE MEETINGS: Suggestions were made that there be a special classified I/DB meeting to discuss intelligence data needs. Roy Reiss said he would give Iris Kameny names of appropriate people to invite.

There was also general agreement that terrain and environmental data are important to almost all M&S and that data standards in that area should be a topic for the next meeting. We were fortunate in having two people from DMA attend this meeting and they offered to share their standards and VV&C procedures with us before the next meeting.

SECTION 5: LIST OF ACTION ITEMS

(1) Iris and Twyla need to work with CIM to better understand how the efforts supplying M&S with data from centralized databases (using data dictionaries) fit into the CIM scheme. A critical need is to give guidance to those projects (JDBE, CCTT, and UTSS) that are just starting out. Twyla will continue to keep current with the CIM DDRS and related activities and the relevant standards activities (since most of these take place in the Washington area).

(2) Bob Molter said that DISA/CIM would like DMSO help in the area of security and complex data types. Since DMSO is organizing a security working group, I/DB should concentrate on complex data types. Iris will be working on this and will be contacting people (or you can contact me) for more complex data type examples.

(3) Participants thought it would be beneficial to have a classified meeting to discuss M&S needs for intelligence data. Roy Reiss will contact Iris with names of people in the intel community who would need to participate. Iris will take the lead on this.

(4) A need was expressed to categorize data into classes (e.g., threat data, terrain, weather, weapons characteristics, etc.). Howard Haeker is trying to organize a Mini-Mors symposium to address this and other issues and

the AMSEC Data Subcommittee will be addressing this as well as JDBE. Iris will be focal point on I/DB for this interest area.

(5) We need to better understand use of the IDEF0 and IDEF1X tools so we can advise the M&S community. We also need to get some representatives from the M&S community into the IDEF Users Group. Twyla will be working with Tom Shook on this.

(6) We need to get an update on the OSD/PA&E MIDAS use of CIM facilitators and IDEF tools at our next meeting. Iris will do this.

(7) Cy Ardoin and Ken Wimmer need to talk to the TECHNET people and JDBE about coordinating information system developments.

(8) Iris needs to give the model directory schema to the DMSO architecture WG, J-MASS, Paul Davis (and others at RAND) to see if it will be adequate to describe modeling frameworks, environments, and architectures.

(9) Iris and Stephanie need to update the database directory schema to accommodate alias names.

SECTION 6: TOM SHOOK'S INTRODUCTION

Tom Shook, the DMSO Director of Technology, welcomed the I/DB task group and said how important it was to have diverse M&S groups work together toward data sharing and exchange and to influence their common destiny. He emphasized that DMSO is working with the CIM office in a complementary manner. It is important that we, the M&S community, understand CIM policy, how we can execute it, and to bring to his attention any problems we may have in implementing CIM policy so that he can work them out with CIM. He went on to say how the DoD M&S world is changing, using the recent Reforger Exercise as an example, because it was conducted without tanks at a savings of about $20 million dollars. We are doing business differently on a big scale and saving dollars as a result.

SECTION 7: SESSION ON UPDATE ON ORGANIZATIONAL ACTIVITIES

BOB MOLTER AND JEFF WOLFE: CIM UPDATE ON DATA ADMINISTRATION, DOD DATA MODEL, DDRS, STANDARD DATA ELEMENT SCHEMA AND NAMING, MIGRATION SYSTEMS, USE OF IDEF

Bob Molter gave us the status on policies, standards, procedures, and tools.

Policy: 8320.1 on data administration was approved September 26, 1991

Standards: FIPS 127–1 SQL database language, and FIPS 156 for Information Resource Dictionary System (IRDS)

Procedures: 8320.1-M, Data Administration Procedures, is two weeks old and in informal coordination. A new version of 8320.1-M-1, DoD Data Element Standardization Procedures (Draft), dated September 30, 1992 was distributed to functional and component DAds on October 6. This appears to be much improved over the previous draft and Molter said Strassmann seemed to like it and he expected Strassmann to sign off on it very soon.

Tools: the I-CASE RFP is on the street and is expected to provide an interface to the Defense Data Repository System (DDRS).

Currently process models (IDEF0) and data models (IDEF1X) are being stored in the Army Corps of Engineers repository. These are models produced through DACOM IDEF tools. Judy Alberts, 703–285–5383, is DISA POC for the DACOM IDEF tools.

By the end of October, they expect to have draft DoD Enterprise high level process and data models. The data model will have 37 entities suggested by services and JS. They have finished Phase 2 of the data model and will align it with the enterprise model. Examples of prime words (really super entities are: action, agreement, location, organization, person, plan, resource).

A team has been formed, mediated by Russ Richards, to harmonize the Army and JS models (JOPES) with the DoD model. POC is Annett Ivy, 703–285–5381.

DoD CIM is evaluating the possibility of joining MCC to take advantage of the work they are doing in enterprise modeling.

Ultimately, CIM intends for DISA to be the furnisher of data for all of DoD. A data user would tell DISA what data they need and DISA would supply it. There was some discussion about just what this meant, especially in terms of timeliness, verification, validation, and certification of data for M&S.

Molter said that C2 was about to receive a command from ASD/C3I, Duane Andrews, to begin business and data modeling.

DISA/CIM would like DMSO help in the areas of security (data aggregation, multi-level requirements, and access criteria), and complex data types. Molter said that CIM could help fund the DMSO effort in identifying the requirement for complex data types, and extensions to the DDRS.

Jeff Wolfe reported on the DDRS: IOC 24 August, platform is AT&T 3B2, access via DDN and local dial-up, and there are 65 registered users. POC is Pam Boylan, 703–536–6900. As of 30 September: there are 349 developmental prime words; 17 approved class words and generic elements; 1856 developmental data elements; and 2054 migration data elements. The developmental prime words are composed of super entities and JOPES entities. The DDRS, in the "waiting for

approval" partition, has the capability to link non-standard data elements with associated standards.

Jeff went through the data approval process and described how they are currently dealing with SDEs and standard software components (bottom level of DSRS repository objects). They are trying to develop a conceptual view of linking all of this in one repository (data administration, data models, functional process models, software reuse (codes), etc.). He showed us a conceptual picture of how I-CASE tools could support a DRS (DDRS/DSRS) Repository to do all of this.

Jeff, as requested, brought an example data model and example filled-in data element documentation worksheets for three data elements in the model. These are used in the DDRS training program and demonstrate use of an example metadata schema for the DDRS.

Copies of the example DDRS data element worksheets and 8320.1-M-1 were made available to participants.

TWYLA COURTOT: DEFENSE MANAGEMENT REPORT DECISION (DMRD-918), DEFENSE INFORMATION INFRASTRUCTURE

Issue being addressed is: should defense information infrastructure be managed through central technical control and configuration management with decentralized execution to assure an end-to-end information transfer capability which is protected, interoperable, and cost effective?

Solution approach is: effective 1 November 1992, DISA will be central manager of defense information infrastructure. Functions identified in the resources analysis will be transferred to DISA and program resource adjustments will be made. Estimated cost savings for FY93–99 is $12 billion dollars.

DoD will centralize activities in following areas: security, interoperability/standards, communications, data processing installation consolidation and central design activities, procurement, training, and configuration management. Exempted are: embedded systems whose costs are normally included in costs of weapon systems and IT resources dedicated to support strategic and tactical C2I missions and wargaming. (But exempted areas remain subject to IT standards.) Military Departments retain acquisition authority for procurements of service specific federal information processing resources integral to a weapon system or which are in direct support of a critical warfighting mission. (We discussed what all this meant to M&S without coming to any conclusions.)

DISA will provide operational and functional staffs with a single DISA technical POC that can get them what they need to resolve computing and communication problems. DISA will absorb assets from other areas to support its new scope and will be reimbursed for its activities. DISA's reach is beyond MIS systems and will affect C2I systems directly through consolidation activities or indirectly as a result of those activities.

A question arose with respect to the reference to an IT Standards Program Office within DISA. What is it? Tom Shook made available an unofficial organization chart for DISA that is used within DMSO, but there is no reference on it to an IT Standards Program Office.

Tom Shook noted that there is a joint DMSO and Strassmann (DDI/CIM) funded effort for DMSO to see where there are overlaps or disagreements between the CIM reference model and the DMSO plan. IDEF0 will be used. Various people may be asked to come in as subject area experts to help in the modeling. The C3I folks are also doing an architecture study to determine the overlap. DISA is studying the DSI and it will probably be transferred from DARPA to DISA over the next two years so that DISA can support M&S.

Copies of DMRD-918 and the DISA organization charts were made available to participants.

JIM WEATHERLY: NAVY M&S OVERVIEW

The Navy doesn't have a STRICOM like the Army, but the Navy is establishing a Navy focal point for M&S in the Pentagon, N-812. Captain Bruce McClure is the N-812 POC for Naval M&S activities. The Navy is developing a Navy M&S Master Plan that will support the DoD M&S Master Plan.

The Navy vision is that M&S must provide tools to warfighting CINCs for analysis, planning, exercise, and training; to assist DoD budgetary and policy making decisions; and to ultimately tie together the warfighter and the DoD decisionmakers. The Navy course is to continue to support joint open architectures and DIS standards; build upon Battle Force Tactical Trainer/Tactical Combat Training System initiatives; increase funding for development and demonstration of M&S tools for the fleet; emphasize distributed simulation as a tool for the user; centralize support for Navy M&S in N-812; and support operational use of M&S developments to ensure applications provide real value added to users.

The Team MIKE objective is to improve M&S tools and wargaming to support the Navy in all its activities and to coordinate with DoD efforts to enhance M&S in all the DMSO identified functional areas. The SPAWAR 31 M&S technical support (MSTS) responsibilities are: to coordinate execution of tasking with Team MIKE; provide planning for investment of M&S resources required under DMSI; review and coordinate DoD and Navy M&S plans; propose cooperative M&S developments with other DoD components; develop Navy Master Plan; support development of standardized databases, tools, and methodologies; recommend VV&A guidelines; and establish Navy M&S information and data clearing center linked to DoD center established by DoD directive. It was noted that CNA has performed VV&A on 40 Navy models over the past five years. In FY92, the Navy did a document survey and review of a limited set of models used in the assessment process. No formal procedures for VV&A have been established.

JOYCE WINELAND: NAVAL WARFARE TACTICAL DATA BASE (NWTDB)

NWTDB is a management process that will evolve to a common tactical data base that meets the needs of the Composite Warfare Commander and supports joint and combined operations. Process components include: information requirements and user validation, data standards and structure definition and validation, tactical naval warfare system implementation, reference database production, and operational database management. A view graph showed the NWTDB extensive relationships with other efforts. In addition, the NWTDB is coordinated with the Copernicus Navy communication architecture. Copernicus provides C4I for the warrior, real-time tactical use. The NWTDB works on a pull concept, pull what is needed from the source as it is needed. There are four major categories of data: weapon systems, MC&G, forces and facilities (DIA), and cryptologic. The NWTDB data will be in a repository which should be available next summer through dial-on and CD ROM distribution. They are currently in the process of bringing aboard a contractor who will be doing IDEF modeling.

COMNAVINTCOM responsibilities include: act as NWTDB standards and structure administrator; identify conflicting and redundant data; coordinate data set design; develop and manage data dissemination; also data flow, database structure and data transfer formats, and information resource directory.

LANA MCGLYNN: UPDATE ON ARMY M&S DATA ACTIVITIES

Army Regulation 5–11, "Army Model and Simulation Management Program," was published 10 June 1992 and copies are available. It includes a section on cataloging of Army models and simulations and chapters on the Army simulation technology program, VV&A, and data management. VV&A of M&S is included in AR 5–11 but VV&C of data is not included in the model procedures or in the chapter on data management.

At the last meeting, Lana had briefed us on the Army M&S Master Catalog and she brought us up to date on that effort. (The DMSO M&S directory will be based on the Army M&S Master Catalog effort.) The objectives are to provide a centralized M&S catalog for the Army community that is up-to-date, available on-line through interactive standard query and retrieval, and the source for Army input to other catalogs (i.e., the DMSO M&S directory would interface to it for Army M&S directory information). The catalog design was the result of surveying the Army M&S community, collecting M&S data and information about other catalogs, analyzing hardware and software requirements, etc. During the current phase, they will develop a fully operational on-line system including all user and maintenance documentation. Milestones include: populating the information base by 30 March 1993, on-line system by 30 May, and debugged, operational system available by 30 September 1993.

She noted that the Army M&S systems having entries in the J8 M&S catalog, will also be entered in the Army catalog but extended to include required VV&A information.

As an update on the Army model: Lana reminded us of the five week M&S and studies data modeling activity that had taken place previously. She noted that the first round of Army data modeling was completed and the Army M&S community will continue to participate as much as possible. The Army is integrating its 19 identified functional areas into a single Army data model.

Erwin Atzinger announced that a meeting of the Army Modeling and Simulation Executive Council (AMSEC) Subcommittee on Data will be held on October 21st. A report on the DMSO I/DB Task Group activities will be briefed at that meeting and Erwin will report back to us on the AMSEC data subcommittee activities at our next meeting.

ROY REISS: AIR FORCE M&S ANALYSIS FUNCTIONAL AREA NEEDS

Roy went over analytic requirements showing us a pyramid with the weapons system level at the base, supporting analyses next, followed by program alternatives analyses, and mission analyses, and finally at the apex, campaign analyses.

Activities at the base level include trying to determine what a weapons probability of kill or hit is. Studies at one level become data fed to the next level of analyses. The lower level analyses have a longer shelf life than the higher level analyses.

Roy also showed a data hierarchy from systems effectiveness and capabilities on the bottom to a middle level of rule sets (e.g., strategy and tactics), order of battle, and area of interest (e.g., natural and cultural features), to the top level scenario. He explained that a hard problem in analyzing systems effectiveness and capabilities is in developing consistent assumptions and in tracking those. (The same problems we have noted in Pks and Phs, and the reason why we would like to treat them as complex data and capture more about the assumptions or dependencies that are involved in their computation.) As he noted, starting with a high level scenario and then fleshing it out is difficult.

There isn't a single source in DoD for order of battle data; thus integration and consistency are needed. There is a big problem with intelligence data gathering – because DIA doesn't have the resources, the services do much of their own collection and are not coordinated. There is also a requirement to share scenarios and threats at more than one level. Jim Weatherly agreed with Roy's view of the problem with multiple intelligence data sets and inconsistencies. Also noted that although DMA does a good job at elevation data, there are inconsistencies or errors in cultural feature data.

The Air Force has been spending dollars on joint models and can use the models and make changes to accommodate use of different aircraft in the models. Roy also raised the question as to the need in joint models for different Pks or Phs for allies' use of weapon systems since they have different amounts of training,

different doctrine and tactics, etc. We also need better tools to estimate Pks for both sides.

Participants agreed that it would be beneficial to have a classified meeting to discuss M&S needs for intelligence data.

WALLY CHANDLER: DMSO ANALYSIS FUNCTIONAL GROUP DATA NEEDS

Wally discussed analysis functional group wargaming needs. Army analyses require system performance data; geographic, terrain, culture environment data; force data; cost data; and scenario data. These are common categories of data needed by all services and include data about allies and opposing forces. Other types of data include deployment data, command structure, control measures, order sets, model parameters, etc. To interoperate across models, we need to be able to share simulation data outputs.

There is a need to categorize data (Atzinger mentioned that this will be a topic at the AMSEC Data Subcommittee meeting on Oct. 21) and a need to easily find the "right" data. For example, JMIN has a handbook to help in looking up 50 different types of Pk; the Army has about 11 different databases of Army force data. Also we need ways to summarize, roll up, and abstract high detail data into the right resolution for use in models.

There is a need to address security issues: M&S needing data at different security levels, needing to interface models at different security levels, etc.

Also pointed out that weapon system data are often complex—not single point data but composed of vectors, arrays, or probability distributions.

What can we do about these problems in the I/DB task group? We can support taking one or more classes of data, e.g., threat data, terrain data, weather data, and looking at each in a more organized way.

Lt Col Ken Konwin said that DMSO project 13 is trying to fit together a hierarchy of models. Campaign analyses need to be done by understanding what kind of data is needed, identifying the sources, and funding the sources to prepare the data.

SECTION 8: SESSION ON COMPLEX DATA TYPES, DATA DICTIONARIES, AND DATA SUPPORT FOR M&S

STEPHANIE CAMMARATA: COMPLEX DATA ELEMENTS AND ISSUES

Traditionally, a "data element" (DE) identifies an atomic (non-decomposable) piece of data. More recently, some non-atomic concepts (which are commonly used and well understood) are serving as data elements. We refer to these non-atomic decomposable concepts as "complex data elements." Complex DEs occur in well-modeled DBMSs and applications and also in poorly modeled legacy systems.

Examples of well-modeled complex DEs include: basic encyclopedia number, Army standard requirement code, probability of kill or hit (which can be represented as a matrix of values), image data elements, terrain elevation data, list/sequence data elements (e.g., road network of line segments), object data elements (e.g., a weapon system and its component parts), graphs (e.g., CCTT example of test data given them in graph form). Issues for future data modeling include: expressing complex DEs as standard DEs (SDEs) so they can be referenced, accessed, and manipulated automatically; representing components of complex DEs as individual DEs when the components have independent meaning; and modeling the relationships between a complex DE and its component DEs to facilitate data verification, derivation, and consistency maintenance. There are also issues in handling poorly modeled complex DEs in legacy systems: mapping them to SDEs, explicitly representing dependencies that are currently expressed implicitly in legacy data fields, and in decomposing "overloaded" data fields to minimize schema modifications.

HOWARD HAEKER: TADS COMPLEX DATA ELEMENTS EFFORT

The TRAC Automated Data System (TADS) is a method to electronically request, receive, authenticate, graphically display, mathematically transform, and reformat data from data providers into TRADOC's combat development models. Currently, TADS handles large volumes of technical data (i.e., approximately 7 billion data instances). M&S builders can order data electronically from TADS, and TADS electronically orders the data from its suppliers (about 20 including AMSAA and SLAD), the data are delivered electronically in specified formats to TADS and then preprocessed into the form required by approximately 12 models (e.g., VIC, Eagle, CORBAN, Janus, CASTFOREM), and electronically fed to the models.

TADS uses standardized data structures to define exact naming in order to automate (utilizes Army Data Dictionary SDEs, naming conventions, and standardized nomenclature, e.g., M1-A1 tank, not M1A1 or Abrams tank); standardized data files (exact content, format, order, etc.); and standardized transformation process (defines exact mathematics used to reduce or compress data).

TADS has a great need for mass storage, and is now using optical disk, juke boxes, and multiple media. Haeker's example showed how direct fire data was furnished directly to the models mentioned above and he gave us an example of a complex DE for a munition including probability of hit given probability of target acquisition, probability of kill given probability of hit, probability submunition will verify target given it has been acquired, target disposition, mode, state, etc.

BILL MC QUAY: J-MASS BRIEFING

Bill showed us a new video on J-MASS.

The objective of J-MASS is to develop a standardized digital M&S capability with which to develop, test, and assess the capabilities of weapon systems in a

simulated operational environment. Two parts to J-MASS: a simulation support environment to enable users to create models, configure scenarios, execute simulations and analyze results; and a modeling library which contains models and model components of weapons systems from the RDT&E community, threat systems from the Intelligence community, and environmental effects from the scientific community. The Modeling and Simulation Reuse Library (MSRL) is currently using the Army's Reusable Ada Packages for Information System Development (RAPID) and will be passed to the DISA/CIM Center for Software Reuse Operations.

Some J-MASS information:

— There are 100 people working on the J-MASS program.

— There are 25 beta sites, 12 developer sites, 11 pending beta sites.

— By January 1993 the unclassified reuse library will be available from Wright-Patterson via WAN, DDN, dial-up lines.

— They have used IDEF0, and have made it into an object-oriented design tool, providing object decomposition.

— Three hundred people attended the recent J-MASS meeting in Colorado Springs.

— The RASPUTIN project is providing rapid generation of scenarios for J-MASS.

STEVE MATSUURA AND PETER VALENTINE: JOINT DATA BASE ELEMENTS (JDBE) FOR M&S

Goals of the JDBE are to increase integration and data sharing by providing a standard for the exchange of information in the M&S community. This effort includes developing standard definitions for data elements commonly used within the community and developing standard information models for various areas of information. Goals also include mapping existing and future databases to the standard and maintaining a directory of all databases that have been mapped to the standard.

The JDBE process is to:

— Train technical working group (TWG) members in use of modeling IDEF1X tools)

— Have the TWG members develop IDEF1X project data models of their databases and applications

— Group the resulting data model entities into subject areas

— Form a subject area information modeling group from TWG members and other experts across components and organizations that will work together to concur on the SAI data model

— Place the approved SAI data model into a repository

— Use the approved model to develop new systems and to develop mappings from legacy systems to it

The TWG members could be database developers or maintainers, M&S developers or maintainers, or members of standards groups.

The plan is to select at least one SAI this first year of the project, and demonstrate proof of concept. The original intent was to base the SAI selection on J-MASS needs, but since J-MASS is just starting out it may be that JDBE will decide on an initial SAI area that they have internal competence in.

The JDBE intends to work with CIM in order to conform to CIM naming and SDE nomination process policies. It has been very important for this group to get some issues settled as to IDEF1X standards and future directions, and CIM naming conventions, DDRS metadata schema, and how the approved SAI models may be integrated into the CIM repository or DoD data model and/or whether they will be separately maintained by the M&S or C2 community.

To use IDEF1X for reverse engineering will require techniques to flatten hierarchical and network data models.

DON REA AND BILL CARPENTER: MITRE DISTRIBUTED HETEROGENEOUS INFORMATION SYSTEM (IDHS) TESTBED

DHIS is a MITRE IRAD project to develop a capability to evaluate COTS DHIS products. The purpose is to develop a MITRE-wide testbed to support sponsor-specific investigation and analysis of COTS products for database and file system integration through schema and semantic reconciliation, legacy system integration and possible migration, and modern (SQL-based) system integration; and distributed security mechanisms. They will be testing MITRE DHIS and related research products in terms of strategies, methods and tools, and benchmarking COTS products. They are working in collaboration with NIST and leveraging the work against real government problems.

They selected a counter narcotic (CN) application as a prototype for the testbed since the GAO identified a need for interoperable information systems among 35 agencies in the CN area and MITRE is supporting many CN projects. The CN prototype is addressing three technical problems: pointer indexes, multi-database query capability and security packages. Two products were chosen for evaluation from a group of possibilities: Uniface as the only mature product and Heterconnect as the only product supporting an SQL interface to M204.

They have an ambitious plan to complete and document their findings in 18 weeks. They plan to hold a workshop in six months to bring together people from government, technical, and federal organizations, with the research and vendor community to review the results and possibly form a consortium through which to feed government requirements to the R&D and vendor communities. NIST is participating because they have an interest in the standards and security issues involved.

SECTION 9: SESSION ON IDEF RELATED DISCUSSIONS

BRUCE ROSEN: NIST UPDATE ON IRDS, AND STANDARDS FOR IDEF

IRDS UPDATE: The IRDS is meant to be a standalone dictionary product and was issued as a FIPS standard in FIPS Publication 156, April 1989. Copies are available from ANSI as X3.138–1988 IRDS or from NTIC as FIPS PUB 156. There is only one product that currently claims to be compliant with FIPS 156 and that is the INFOSPAN IRMS product. A contact for the INFOSPAN IRMS Users Group in the Washington area is:

Ms. Veena Bhatia
Department of Education
Information Resources Management Service
Room 5624, ROB-3
400 Maryland Ave., S.W.
Washington, DC 20202
(202) 708–9279

The IRDS services interface allows other software to interface with the dictionary while ensuring the integrity of the database. The services interface standard specifies a generic low level external software interface that has been completed by the X3H4 Standards Committee and is now being published by ANSI as X3.185–1992 and may or may not become a FIPS.

The IRDS export/import file format specification standard that specifies the format for bulk data exchange between IRDSs uses ISO ASN.1 and has been documented by ANSI as X3.195.1991. The document also includes some minor changes to IRDS core standard commands. Rosen thinks the FIPS will allow CASE tools to exchange data if they are using the right file format, but to correctly exchange data between IRDSs requires having to first structure the schema before exporting the IRDS data.

There are two IRDS standards: the X3.H4 standard and the ISO standard. The ISO standard uses SQL as its base, and was approved in 1991 as an international standard for a DBMS dictionary.

Next generation dictionary products: IRDS2 and ISO/Atherton Tool Integration set (ATIS). In 1991 ATIS was proposed to ISO as a basis for an IRDS services integration specification. It is an object-oriented paradigm, provides an

extensible set of services (within the service interface of the architecture) and provides a call interface in which each part can be isolated to be worked separately, or each service (or the IRDS) can be processed separately. The call interface would be appropriate to all entity types, and work across vender platforms and repository standards. ATIS would provide version management, configuration management, work flow management, context management, tool registration, and role specification. ANSI would like IRDS2 to become an international standard and so they will have to work to include ATIS object-oriented concepts. Rosen guesses it will be around 2000 before IRDS2 becomes a standard because of the time it takes to get U.S. agreement and then take it to the international community for agreement. Bruce Rosen will be the new international representative to the ISO IRDS standards group.

IDEF UPDATE: NIST, at the request of DoD, is in the process of establishing two FIPS, one for IDEF0 and another for IDEF1X. The draft FIPS has been completed, is based on the original Air Force IDEF documents, with the assistance of the IDEF Users Group, and is expected to be published in the Federal Register by first quarter FY93. After publication in the Federal Register, there is a 60–90 day period for public comment, after which NIST will respond to all comments and the final FIPS will then be published.

NIST has had two main problems with developing the IDEF FIPS. First is that the Air Force IDEF documents were not clear about the differences between what IDEF is (i.e., a methodology and language) and how it is to be used. Second is that the IDEF Users Group was not used to acting as a standards group or committee. As a result NIST made a decision to split the FIPS into a normative section (the specification of the standard) and an informative section (how to use the standard). They hurried to get it out so it could be signed before the election.

Bruce went through the FIPS approval process: first NIST approves a FIPS, then the lawyers approve it and it is printed in the Federal Register, the public reviews and writes comments about it over the next 60–90 days, the comments are returned to NIST and they must respond by changing the document, explanations, etc. NIST then issues the FIPS again usually without another round of public review. In developing the FIPS, NIST establishes non-proprietary criteria usually through committee recommendations, publishes the criteria in the CBD, and gets responses from the public.

A question was asked as to whether the IDEF FIPS was standardizing a methodology instead of a technology. Bruce's answer was that the frontispiece says that if the user has decided to use IDEF modeling technology, then the FIPS is the standard to be followed for its use. Another question asked was whether there will be FIPS for other modeling methodologies? The answer to that question was no, if the methodology is proprietary. A FIPS cannot standardize on a proprietary solution—it can be less than the best technology because it must be non-proprietary.

NIST would like the IDEF Users Group to control the FIPS and take the lead in future updates and changes. Bruce hopes that material in the informative

section can be agreed upon and moved into the normative section. NIST is trying to work with the Users Group to make the group understand what is required to be a standards group and to establish standards, e.g., can't discuss things forever, need to coalesce and make decisions.

Tom Shook asked about the composition of the IDEF Users Group and said that we need to get input to the group, possibly have a representative in the group, and get more DoD people in the group. Peter Valentine said he is a member of the IDEF Users Group and Dan Wu said that anyone in DoD can nominate a DoD person to the group. To do so, just contact Dan Wu.

Copies of the Draft IDEF1 FIPS were made available to all.

PAUL REHMUS: PA&E REPORT ON USE OF IDEF1X ON MIDAS MODEL FOR FORCE PROJECTION

This project is of great interest to the I/DB Task Group as it represents an example of how CIM can assist the M&S community in process and data modeling, and data standardization. This project also represents PA&E's first foray into the CIM world and the result will be evaluated before pursuing other PA&E functional areas. The CIM people helped PA&E to organize the project.

The OSD(PA&E) Projection Forces mission is to advise PA&E on programs for mobility, prepositioning capabilities, wartime medical support, and C2 of mobility forces and to provide leadership in promoting and improving analytic tools and methods for analyzing national security planning in these functional areas.

The mobility community includes many organizations, and there are many ADP problems such as disparate models, data quality issues, diverging definitions from organization to organization, and redundant data building efforts.

The project objectives are to: identify management efficiencies; improve standardization, quality, and consistency; build an integrated data model for PA&E mobility community use; and reduce the cost of database O&M.

There are four phases to the project:

A. Survey of processes, data, and models 9/1/92—2/28/93
B. Generation of "as is" process and data models
C. Generation of "to be" process and data models
D. Documentation, registration, and archiving

The mobility community data elements were estimated by Rehmus at approximately 10,000 based on 10 orgs x 10 models x 100 data elements per model.

The MIDAS model owned/used by PA&E/PF, OJCS/J4, and TRANSCOM/J5 was selected. MIDAS is about 15,000 lines of PL/1 code and its functions include selection of mode and intertheater lift, dry cargo according to TPFDL, and resupply

and sustainment. It utilizes many databases including data about: deployment, logistics factors, TAEDP, ships, aircraft, geographic locations, and the TPFDD.

We will be getting updates on this project in later meetings.

TOM SHOOK: DMSO REPORT ON PLANS FOR USING IDEF

Tom showed the group a viewgraph depicting the business re-engineering process: model the "as-is" (show current processes), model "what-ifs" (compare process alternatives), model the "to-be" (develop new process), and document and implement the derived integrated solution.

On the "as-is" level, DMSO is working with JIEO to look at the major classes of simulation in order to relate them to the CIM technical reference model for commonalty and to identify differences. The differences would be translated into funding needed to bring the M&S architecture standards into concurrence with the CIM technical reference model. The three classes of simulation are: constructive (standalone wargames, man may be interactive with game); virtual (networked, man-in-the-loop, e.g., SIMNET), and real or substantive (e.g., NTC, Blue Flag).

The DMSO plans to do some trial runs with IDEF0 to get experience and evaluate the state of the IDEF0 modeling technology to see how doable and beneficial it is before recommending its use in the M&S community. If the initial results are successful, then DMSO would like to use IDEF in describing distributed interactive simulation, and also to produce a functional description of Reforger to make next year's Reforger exercise better.

To get from an "as-is" model to the "to-be" will require use of tools for "what-ifs." Suggested tools are SIMFACTORY (written in SIMSCRIPT) or IDEFINE tool. The idea is to manipulate the "as-is" model in a simulation to show that the description is valid and then to try out "what-ifs" on the way to determining the "to-be." They are also looking at a METADESIGN IDEF0 product that uses color Petrinets and produces graphics that are easier to understand and follow. They would like a tool to also help manage the transition from "as-is" to "to-be"—to be an iterative management tool.

Tom believes they really need a true object-oriented IDEF modeling methodology to represent the complexities of M&S applications and right now those tools don't exist. He doesn't believe the current generation of IDEF tools will get them more than 60% of the way.

Tom is seeking lessons learned with specific IDEF tools, products, advantages, limitations, etc. from anyone with information; just send him an email or give him a call.

TWYLA COURTOT: MITRE'S EXPERIENCE WITH IDEF TRAINING AND USE

MITRE found IDEF training to be very important but it requires careful selection of a reputable trainer. One needs to investigate potential trainers' claims, examine their syllabus/briefing charts, and let them know exactly what your organization needs from the training. It is important to realize that the tools and the training are not dependent on each other. Emphasis on notation and understanding IDEF is more important than knowing how to use the tools and must precede tool use. MITRE chose to have non-tool specific training. But if you choose to get training in the use of the tools, you need to do it with a real application (not the training examples). A comprehensive understanding of the IDEF manuals is not necessary. MITRE found that a 3–5 day training course for 10 students cost around $20K and included free rights to copy the course materials. Costs vary with number of days and number of students in the class.

IDEF tools at MITRE include: Wizdom (PC), MetaSoftware (MAC), and AutoSADT for desktop use. They have found the MetaSoftware tool easier to use than the Wizdom tool. However, none of the tools were adequate for their needs. Shortcomings included: repository was inadequate and it was easier to maintain the glossary in a word processor file, drawing capabilities were not that good, the tools required entering the glossary twice, and the user interface could be much better.

MITRE has used IDEF0 on several projects, to enable MITRE to better analyze and organize their thinking and approach to a client's task, and to use this to communicate with the client about their problem and the MITRE approach to it. On the other hand, their clients, who were also serving clients, were reluctant to use IDEF0 in furthering communication with their clients.

Lessons learned: IDEF0 is a notation, not a methodology, but it is an excellent tool for organization and analysis; IDEF is not a flow charting tool and it may require practice to learn to think IDEF0; if an IDEF0 model seems overly complex, try to rethink it; need to take care to keep a consistent viewpoint throughout the modeling effort; a modeler should always have his/her work reviewed by others; a set of models will be necessary to model an entire process because of the need to support different viewpoints; strawman models may be a good way to stimulate thinking and involve the user in the modeling process; and professional training is important, the "train a trainer" concept is inadequate.

General observations include: IDEF is easy to use and misuse; there is never one right model; modeling requires iteration and refinement; models should be maintained over time; and IDEF models can be done quickly (at least the easy parts).

IDEF HELP AND DOCUMENT:

Charlotte Gross volunteered the CIM business improvement, BPIP, hotline number: 1–800–828-BPIP, which can be called for information about IDEF tools and usage.

In addition, our DTIC participants announced that an IDEF document, "Corporate Information Management Process Improvement Methodology for DoD Functional Managers," is available by calling DTIC at 703–274–7065. You can ask for the document by name or by asking for the CIM or IDEF gray book. Bob Bishop passed out a few copies at the Friday meeting.

TWYLA COURTOT: IDEF DISCUSSION

Pete Valentine said that the JDBE project has been using ERWIN by Logic Works to do IDEF1X modeling. It is user friendly, does top to bottom modeling—E-R model to physical database implementation. However, it does not directly support business rules, these have to be entered as annotations; it doesn't support instance examples; and doesn't map directly into the NIST FIPS 156. They have not found any IDEF0 tool they like, and there is no tool that automatically maps from IDEF0 models into IDEF1X models. IDEF1X needs to be able to support reverse engineering as well as new models.

Gio Wiederhold offered some suggestions of improvements to IDEF1X. IDEF1X needs to: address M-N cardinality; include business rules; and for reverse engineering needs to allow flexibility to model semantics of network and hierarchical models as well as relational. Gio said he has references discussing reverse engineering using IDEF1X.

TWYLA COURTOT: REPORT ON OBJECT-ORIENTED STANDARDS MEETING AND CIM 8320.1-M-1 MEETINGS

Four relevant ANSI standards groups:

X3.H4 addresses IRDS
X3.H6 addresses CASE
X3.H7 addresses object information management
X3.H8 addresses data representation

There seems to be much overlap in what the first three groups are doing. X3.H4 is composed of potential users and vendors and is addressing IRDS2 and trying to come together with the ISO effort. X3.H6 has had about 5–6 meetings to deal with CASE and is still trying to decide where they are going. There is a CDIF group that is dealing with the CASE tool interface and they overlap with IRDS, and have published a preliminary standard. Twyla attended the second X3.H7 meeting and they are trying to understand what it is they will try to standardize. X3.H8 has one subgroup dealing with standardizing data and they are presenting their document to ISO; the group as a whole is addressing naming standards and other standardization issues. It was observed that there is an existing object standard for use in labeling objects for automatic scanning.

SECTION 10: SESSION ON RESEARCH ISSUES

GIO WIEDERHOLD: DARPA INTELLIGENT INTEGRATION OF INFORMATION RESEARCH PROGRAM

The objective of the Intelligent Integration of Information Program is to get useful information to any user in the appropriate form, amount, and level of detail at the right time, exploiting the many data and computing sources available and emerging. Useful information includes data in databases including legacy databases, experience and knowledge bases, and simulations.

The problems being addressed include data overload, information starvation, system rigidity (built-in direct linkages), and complex management (multi-user, multi-task, multi-source, and multi-maintenance).

Architectural alternatives include: integrated architectures, efficient when designed but inflexible to change and inefficient after 10 years; federated architectures, fast to implement, flexible, but costly in dealing with change and never efficient; and mediated architectures, the new choice, which are fairly easy to implement, flexible, basically not highly efficient but provide for data reduction, local caching, and optimization of access, and avoids maintenance by committee.

The roles of mediators are to: combine data by selecting relevant data, resolving mismatches, and reducing volume forwarded; abstract data by summarizing detail, searching for exceptions, and harmonizing for fusion; and processing data to gain information by projecting past to futures, ranking and pruning alternatives, and caching to retain history. In a sense, the mediators will perform functions similar to those performed by command staffs who summarize and fuse data, making use of historical data and developing different what-ifs or possibilities as input to the commander for decision making.

The program has a nine-month project to develop an F-22 integrated weapons system database demonstration prototype. Phase II will be to support variability reduction analysis in manufacturing and Phase III to support tolerance management. They will be talking to people at SEI and STARS to get a disciplined approach to reuse, and will be using the same basic data to generate different views for different user needs. The mediators will do the data sifting and filtering. Gio explained that "shared ontologies" referred to terms and concepts that needed to be defined and shared among the different groups on an "as needed" basis. The view is that not everyone needs to understand everything, understanding is only necessary at the intersection of user groups.

Gio gave a view of how mediators will advance from handcrafted programs using wrapper techniques to access heterogeneous databases today, to comprehensive mediators tomorrow built from standard modules and using formal domain descriptions of the applications and the databases. Research supported by his program includes: wrappers, knowbots to find information, interface formalisms,

enterprise integration technology, persistent object bases, fuzzy algebras, and description of disjoint domains.

An important goal of this effort is to establish a new field of integration science because integration concepts are as important to large systems as the concepts within the component modules. They are defining a path out of the current impasse of having to build systems now for a world that changes more rapidly than resources allow systems to be replaced.

IRIS KAMENY: DMSO CALL FOR PROPOSALS FOR COMPLEX DATA AND COMMON TOOLS

DMSO has just put out an FY93 focused call for projects that are to be submitted through appropriate POCs in the joint staff, OSD, services, and DoD agencies. DMSO plans to allocate 40% of funding to continuing projects, 40% to new projects, and 30% to infrastructure support. The call includes six research areas: Architecture for Dynamic Scalability; Complex Data and Common Tools; Environmental Representation; Human Performance for Distributed Systems; Materiel Acquisition; and Pre-/Post-Crisis Action Missions.

Just in case you didn't get a copy of the handout, Section 14 includes the general areas of interest for Complex Data and Common Tools.

SECTION 11: SESSION ON DMSO INFORMATION SYSTEM REPORTS

CY ARDOIN: DMSO INFORMATION SYSTEM PROTOTYPE AND USE

Cy discussed the functions that the prototype system will support: mail, news, help, announcements, library documents and glossary, and catalogs or directories (i.e., POCs, organizations, models, databases). The effort has recently canceled the TOPIC acquisition and plans to acquire the Oracle DBMS in the near future. The system will be unclassified and it has not been determined what to do to support classified catalogs or directories if such are required. The immediate schedule is: Alpha version currently running at IDA; approved Alpha version running at IDA by Oct. 25; approved Beta version at DTIC on Nov. 9; and IOC at DTIC on Nov. 16.

Peter Valentine asked a question about the need for this system to interface with other systems, noting the possible duplication of effort between the DMSO Information System, the JDBE planned information system, and the T&E Community TECHNET System which is currently up and running. It was suggested that Cy talk to the TECHNET people about possibly extending it to include M&S or reusing their software. Peter will coordinate with Cy so there will be sharing or little duplication between the DMSO Information System and the JDBE effort (which DMSO is supporting).

DENNIS SHEA: DISCUSSION OF DMSO MODEL DIRECTORY

CNA reviewed seven existing model catalogs and the design documentation for construction of the Army Master Models and Simulation Catalog. They concluded that the JCS/J8 and the Army catalogs contained the information needed by most users and were similar in format. The major difference was that the Army catalog contained additional information on VV&A. The recommendation is to adopt the Army schema with qualifications. The qualifications are to add the following fields or words:

(1) under proponent, add: model development organization (if different than proponent);

(2) under proponent, add: available documentation on model methodology;

(3) under description/limitations, add words: (to include critical assumptions and run time considerations);

(4) under input, add: type, form, or specific parameters required;

(5) under input/available databases, add words: (include source and date);

(6) under input, add: graphics interface requirements;

(7) under model construction/time processing, add words: (if time step, specify time step size);

(8) under user(s), add: list of studies where model was used;

(9) under average length of game, add: provide time and complexity of scenario, e.g., number of forces, length of campaign, etc.;

(10) under verification and validation(s), add: available documentation of V&V procedures and results.

The qualifications are mostly in reaction to user concerns with the model directory, including: level of detail/quality of inputs required to run the model need to be better specified; subjective element makes model comparisons difficult; model description often vague; capabilities are often exaggerated; degree of VV&A and results missing from most catalogs; need glossary of terms used in schema; model developers rarely disclose limitations; need to develop proper indexes and key words to aid search; need reference/source of "available" model documentation; cite studies where model was used; and more.

We discussed how the model directory does not include M&S frameworks and architectures and Cy expressed the belief that we will need a different catalog for those.

An action item for Iris is to send the model directory schema to the architecture group for review and to Paul Davis. In particular, we would like them to address whether the model schema is adequate to describe: J-MASS environment and models; ALSP; RAND JANUS-A group of models, Anabel, and TLC model environments; CASES (Navy architecture of models); and SDI architecture of models.

Status: the DMSO Model Directory schema has been turned over to Cy Ardoin for inclusion in the DMSO Information System. An M&S directory (E-R model or IDEF1X) model will need to be made from Shea's schema and then a relational model created to be managed under the Oracle DBMS.

IRIS KAMENY: DISCUSSION OF DATABASE DIRECTORY

Iris reviewed schema.5 for the database directory. The following suggestions were made and will be implemented: add the capability to furnish a group of alias names for the database in addition to the acronym; and make explicit that "organization" for source, technical, and release mean major organization such as a service.

Pete Valentine would like to see the directories (POC, organization, model, and database) tightly coordinated so that the database directory contains identifiers of models using the database; the model directory contains identifiers of databases used as input or produced as output; and the organizations and POCs be consistent with those directories. The general consensus seemed to be that this is a worthy goal, but may be very hard to implement in the near term. Also, the person entering a database (e.g., DMA terrain data) may not know all of the models using the database.

Status: the DMSO Database Directory schema with the modifications will be turned over to Cy Ardoin to be included in the DMSO Information System.

Both the M&S and database directory schema data elements will, sometime in the future, have to be mapped to CIM SDEs and possibly renamed. Also both directories should be represented as IDEF1X models.

SECTION 12: PANEL ON VERIFICATION, VALIDATION, AND CERTIFICATION

The panel session consisted of Mike Barton, David Danko, Howard Haeker, Iris Kameny, Dennis Shea, and Simone Youngblood.

Iris Kameny led off the session by discussing what she briefed to the DSB: that data VV&C is of great interest to the M&S community, has the potential to enhance VV&A of models, is a controversial subject among the community because it is so difficult to do and control, verification is pretty well understood but validation and certification are not, and it should be addressed as part of the model

VV&A technology and methodology. Some steps in data verification include: identifying source data, selecting/addressing data versions, verifying source data, converting source data to model formats and reverifying, and verifying values of data during model operation. Methods of verification include verifying using domain constraints, using higher order knowledge such as rules, and statistical or set operators for verifying over a dataset, and using application domain specific techniques such as superimposing map data on an image and checking for common sense errors.

Howard Haeker said that VV&C requires a management structure to assign responsibility for VV&C. All involved must be aware of the requirements of the scenario and the model. Current problems that go against good software methodology and configuration management, such as when the model doesn't correctly represent the object, trick it by changing the data without reprogramming the model, need to be addressed. Another example given by Dale Pace is that the model data have to match the purpose of the model. The example is that weapon characteristics data exist as specified, as collected on the test range, and as collected in combat situations (e.g, NTC)--the choice of data has to depend on the purpose of the model.

There are also questions as to where and who does VV&C. It can be performed all in one shop with or without independent teams. Others separate functions and use a different quality assurance team than the development team, others may call independent sources in to do VV&C. There needs to be good configuration control of data. The bottom line is VV&C needs to be managed up front, ahead of time. This includes addressing the future, handling of legacy systems, and giving people time and support for carrying out VV&C.

Verification must ensure that data are transmitted and reassembled for input to the model such that they look OK and without typos; validation involves association of the data values with the methodology used; and certification of data has to be done in conjunction with the model and its usage.

Mike Barton posed the questions, "How does data interact with the M&S?" and "Is all data created equal?" He stressed the different types of data that have to be handled and selected on the basis of the model and purpose: test data off the range, performance data that have been heavily processed, hardwired data, and standard data. He asked the question "Certification vs. accreditation, who is responsible?" and gave us an interesting graphic in which he showed a certification circle of data intersecting an accreditation circle of M&S applications. The intersection within which were shown system specific data, preprocessed data, and hardwired data is the problem area. We need to address how data are handled in the intersection.

He presented a list of issues:

— What are the types of data?

— If data production model is accredited, is the data certified or accredited?

— What are the requirements to certify "standard" data? Does responsibility end before or after preprocessing by the user?

— When does certified data become model specific data?

— Who is responsible for certification and VV&A of
  - system specific preprocessors?
  - system specific preprocessed data?
  - proper use of data producing models/output embedded in larger models?

Dennis Shea also discussed the VV&C problem of dealing with different types of data. Many Pks for the Navy come from test results based on pilots who fly over the same target every day and know it very well. The Army Pks are computed by simulating hits at different angles and conditions over hundreds of cases. However, test data are all treated the same—we really need to require that data be identified by source, assumptions, constructs, limitations, etc.

Twyla suggested that we need to manage and control data and need procedures and guidelines to do so. Need to: define terms, define how we apply VV&C, procedures, and management.

Simone Youngblood said that John Hopkins University Applied Physics Lab (JHUAPL) is producing an outline for the Navy Team MIKE M&S support group to support model VV&A, where "A" includes data and the purpose of the model. They are looking at a broad spectrum of models and four levels of accreditation which they have defined to be (1) inspection of model, (2) VV&A (six-month effort), (3) robust VV&A (two person years), and (4) guaranteed for simulation. It takes dollars and time to do VV&A and the output is applicable only to certain uses.

David Danko: we need aggregation techniques to produce data for variable resolution of models. This requires that the data provider understands what the model needs and the modeler knows what the data represent.

David suggested we get some input on data VV&C from Michael Goodchild at the UC Santa Barbara National Center of Geographic Analysis. He will also give us procedures from DMA as they do product maintenance, do better than they have in the past, and deal with currency.

Another voice said that the Army expects certified data from their approved sources.

Another comment: The real world is not real, models are scenario specific, they use anecdotal evidence, the real world is not the ultimate certification. We can say data are consistent or credible but not accredited or certified.

In a later conversation, between Iris and Tom Shook, Tom expressed the view that a model can be verified and validated to handle a certain type of problem and put on the shelf. A source organization can prepare functional area data it is responsible for, perform VV&C on it, and make it available to an appropriate community of M&S users. When the model needs to be used to address a particular problem it is taken off the shelf, the VV&Cd data are selected and preprocessed to meet the model's needs and again verified and validated by someone (i.e., the source, the modeler, or an intermediate), and the model and data are certified together in the context of the problem being addressed.

SECTION 13: WORKSHOP SESSION TO FACILITATE EXCHANGE OF GOALS, APPROACHES, TOOLS, ETC. AMONG GROUPS SUPPLYING DATA TO M&S USERS

The immediate objective of this session was to facilitate the exchange of goals, approaches, tools, etc. among groups supplying data to M&S users. The approach was to bring people together from five programs having common interests to exchange information and help them form a longer term group. The desired near-term result: prevent building individual stove pipes by establishing ongoing exchange of: concepts of data administration: top-down/bottom-up, relationship to DDI/CIM, DISA/CIM and CFS, and omponents; lessons learned; data modeling techniques; data element standards and dictionaries; representation and use of complex data; data verification, validation and certification; and tools. The long-term objective: to develop a common approach that will form the basis for data modeling, data elements, data dictionary, etc. in the M&S community.

Five groups discussed their efforts:

Army program: Automated Data System (TADS) Participant: Howard Haeker

JCS/J8 program: Operations Analysis and Simulation Interface System (OASIS) Participant: Don Hogg

Joint program: Joint Data Base Elements (JDBE) Participants: Steve Matsuura, Janet McDonald, Peter Valentine

Navy program: Universal Threat System for Simulators (UTSS) Participants: Mike Sarkovitz, Gail Coffey

Army Program: Close Combat Tactical Trainer (CCTT) Participants: Rob Wright, Lucy Haddad

The TRAC Automated Data System original statement of problem: data procurement process varies between functional areas; data arrive in a variety of forms; data require extensive processing to make them model ready; and the data procurement and provision process is not automated. The goals were to build a centralized, automated data system for TRAC, develop and maintain computerized

databases with user friendly graphical interfaces, and begin electronic transfer to all customers and providers.

TRADOC has 80,000 people and TADS is responsible for providing them with classified weapons systems performance data and characteristics, operational data, DMA terrain data, and TACWAR with weapon systems data. The methodology is conducted by functional area and is limited to input data for TRADOC's combat development models. The implementation of each area takes seven steps: (1) standardization (define requirements in exact language); (2) determine transfer media; (3) build relational databases (Ingres currently used); (4) develop software to process data from standardized files to Ingres; (5) build graphics capability; (6) develop software to process data from Ingres to models; and (7) develop maintenance procedures. TADS uses standardized nomenclature, data files, and transfer process. They will be using Suns and develop user friendly forms-based applications using Ingres 4GL and SQL. Initial graphics will display selected fields from the functional area database but eventually will include visual depiction of model-generated data by the back-end software. The functional area experts will be trained to maintain the databases. It takes about 12 months to develop a functional area database.

Problems: there are data voids that need to be handled by agreed upon procedures; some of the data preprocessors are very complex and cannot be implemented in 4GL; and they need better graphics. Howard urged that techno-stress suffered by staff sitting at computers all day is a neglected problem that needs attention. He believes the future way to go is object-oriented, and TADS plans to propose to DMSO to do research in developing icon-based pictures of standardized weapon systems.

The OASIS mission is to develop a system which will significantly improve data collection, access, verification and validation, analysis, reporting, management, and documentation for J-8 studies and analysis processes. J-8 does assessments and analyses over land, sea, air, and nuclear. They also use PPDB data in two J-8 directorates in different ways. The OASIS goals are to build a framework, first for the nuclear force analysts and then for the conventional force analysts. There are only two J-8 people working on the system but Westinghouse has around seven people supporting the project.

OASIS uses centralized data management based on Ingres and the Ingres 4GL, and has an on-line data element dictionary. They are running a Top Secret shop using Unix on Vax clusters and Sun workstations and using the Network File Server (NFS). They have around 50–60 Sun Sparc stations and run windows 4GL on the Suns with the databases on the Vax cluster and an EPIC jukebox for archiving. The data sources are DIA, CIA, and services, and their reference database is not model specific. Oasis maintains a three-level hierarchy of files, classes, and objects. The program has a force mix working group who decide on the data based on the performance characteristics of the weapon systems being modeled and use the 4GL interface to create the screens the analysts will use. The 4GL is very easy to use and is used for fast prototyping. OASIS contains an on-line,

dynamic data dictionary defined on the basis of E-R diagrams and uses data dictionary access screens for entering the data dictionary information in the database. It took about a month to develop the dictionary screens.

The main difference between the TADS and OASIS philosophies is that TADS has analysts that preprocess and maintain functional area databases while OASIS has one centralized database for nuclear that is maintained in reference files. The analysts copy the reference data they need into study files and do data verification and validation from their workstations. There is no centralized V&V. Errors found in study file data by analysts are handled by analysts correcting them in the study file and entering a note to that effect in the relevant place in the reference file. The use of graphics for quality control and other techniques all reside with the analysts on their workstations.

The database is updated about every three months. OASIS concerns are with connectivity issues, DoD data standardization, and the fact that some of their data isn't modeled well using the relational model.

Dan just signed a general order for JCS to acquire Ingres products and could furnish products to others through that order.

Haeker and Hogg agreed that when Dan gets to conventional forces he will want to use the TADS data.

STEVE MATSUURA AND PETER VALENTINE: JDBE

In the long term JDBE will be developing a database directory and dictionary and plan to document their methodology in milstandards. As said earlier, they plan to use IDEF1X to develop data models in subject areas and had hoped that the J-MASS program would help them determine which to do first (which currently doesn't seem likely). JDBE's goal is to develop a common methodology that is compatible with CIM. They want to develop the methodology training and support for doing the modeling, integrating the functional areas, and representing the overall functional models to CIM for integration with the DoD enterprise model. Training of the TWGs will be: 50% IDEF1X training, 25% reverse engineering training, and 25% JDBE methodology training. Their initial test cases will be of internal database systems, RASPUTIN, and reverse engineering of OPFAC rules. Their first priorities are J-MASS and electromagnetics but they would accept other SAI suggestions from volunteers. They commented that Oasis is dealing with a very hard SAI because of the data aggregation problems.

Howard Haeker said that TADS has gone down three levels in IDEF1X, and asked if he could send a contractor to JDBE training. One of the problems in doing so is that TADS and JDBE are using different IDEF1X tools and since there is no standard for exchanging IDEF1X data models there is no easy way to move the TADS IDEF model into the JDBE tool or bring the output model from the JDBE tool back into the TADS tool.

ROB WRIGHT: RSI SUPPORT FOR STRICOM AND THE CCTT

RSI is tasked to provide the CCTT contractor with certified, accurate, usable, unclassified data in a timely manner. To do so they have established an assistance office, a data support network, a CCTT data library (hard copy), and a performance data working group to review data. They have performed a data requirements analysis, identified potential data sources, collected performance data for 34 weapon systems, created the DOCATS document catalog system, created a hypertext query system to DOCATS, created the parameters database system (PC based to contain performance data by weapon system), and maintain and operate the CCTT library. Right now the library maintains hard copy manuals and documents but AMC is in the process of deciding how to automate the documents.

They gave us a good example of complex data which they need to maintain as part of the parameters database. It was a curve of full-load fuel consumption representing test range data for the Bradley fighting vehicle. Since that is how they get the data, they would like to present it to the user in the same way and have the user take whatever data values they need from the curve. As changes in equipment occur, there would be new curves.

Since several of the 25 different blue force weapon systems will be leaving the Army inventory before CCTT is delivered in 1995, they will only collect information for weapon systems in the 1994–2000 inventory. They need data about ordnance and affects of ordnance, unclassified Pks and Phs (from AMSAA) for blue and red forces, will use threat doctrine from FM-101,2,3, and will need red and blue SAFOR data. They also need to deal with the data voids that Haeker mentioned earlier. They have a data modeling subgroup to approve the data calls.

The CCTT contract is expected to be awarded in the next two weeks. Having a central data source to furnish all contractor needs rather than having the contractor and subs going out and collecting their own data and information is a new approach.

IRIS'S QUESTION: How large and complex will their parameters database that started out in Paradox and will be moving to Foxpro be? If it will be large, then a PC may not be adequate. In either case, they should consider whether building a PC database that is non-conformant with CIM standards (POSIX open systems) is a good idea in the long run. Other DoD M&S modelers may also want to use their data and it may be more assessable to them if maintained in an open system.

GAIL COFFEY: UNIVERSAL THREAT SYSTEM FOR SIMULATORS (UTSS)

This is a new joint service program that has a joint technical coordinating group to NAVAIR who will run the program with input from the services.

The concept is to eliminate duplication, reduce development time, decrease costs, provide validated data/models, standardize systems, increase capabilities, enhance training, and promote reusable software/hardware among the simulator device community by furnishing them with a master database of all the services'

validated orange, blue and grey threat databases. Currently, the Air Force must use validated DIA data; the Navy, NTIDS; J-MASS, DIA; etc. UTSS needs to get all of the threat players together (as sort of an SAI) to agree on the threat data and to help define the threat database at three levels: unclassified, secret, and top secret.

The working group needs to identify threat models existing in current simulator devices and understand what their data needs are (maybe by reverse engineering them?) and at the same time the program needs to be looking at database modeling and standards. Although the project began seven years ago, they have really started afresh six months ago and have R&D funding for a four-year effort.

DISCUSSION:

It was agreed not to make the people in this group of projects into a separate subgroup of the I/DB Task Group. Projects will interact with each other where it makes sense.

There was some discussion about the DMA terrain modeling problem for which everyone agreed we need a standard. Danko discussed the problem of geographic modeling at a number of different levels: raster/vector/b-tree storage structures, hierarchical/network/relational/object data models, polygon/line/point primitive objects, and higher level conceptual objects such as elevation contours, trees, roads, bridges, etc. We agreed to devote some time at the next I/DB meeting addressing the terrain modeling issues.

SECTION 14: GENERAL AREAS OF INTEREST FOR COMPLEX DATA AND COMMON TOOLS

- Develop metadata extensions or new concepts of "standard data elements" to represent complex data types such as images, networks, objects (including methods and rules as objects), derived data, etc. Metadata extensions should be based on existing data dictionaries to include the Defense Data Repository System.

- Develop/identify/prototype tools or techniques that could prove useful to management of repositories including collection, storage, retrieval, and dissemination of complex data types.

- Develop methods and support for standardization of data element domains and domain values nomenclature (e.g., standard names for aircraft types, standard descriptions of icons, etc.). These should be compatible with ongoing DISA/CIM efforts for nomenclature and symbology.

- Develop approaches and methodologies for verification, validation, and certification of M&S data for specific applications, within an application area, and for general usage (e.g, DMA terrain databases) with specific attention to complex data, particularly objects, rules and datasets.

- Develop approaches to capture and management of historical data (i.e., simulation data as they are generated for later analysis or rerun).

- Define and/or develop search capabilities across heterogeneous repositories:

  — develop techniques for classification and typing (e.g., use of domain hierarchies, facets, etc.) to aid in insertion into, and intelligent search and access across, repositories;

  — develop M&S search terms to find, access, and retrieve database and model directory data, standard data elements, and complex data types including objects from distributed heterogeneous repositories.

- Define and/or develop mechanisms for finding, accessing, and interchanging data including complex data types among distributed passive, active, and dynamic repositories (including data and model directories, and data dictionaries).

- Develop an approach to the security data aggregation issue for large and distributed repositories including: "need to know" for legitimate users of the existence of information at higher security levels; inferring data at a higher classification level from large amounts of data at a lower level; and inferring the existence of higher level data from missing data.