NISTIR 6071

Computer Integrated Knowledge System (CIKS) for Construction Materials, Components, and Systems: Proposed Framework

Thomas Y. Kurihara
Lawrence J. Kaetzel

Building and Fire Research Laboratory
Gaithersburg, Maryland 20899

NIST

United States Department of Commerce
Technology Administration
National Institute of Standards and Technology

NISTIR 6071

Computer Integrated Knowledge System (CIKS) for Construction Materials, Components, and Systems: Proposed Framework

Thomas Y. Kurihara
Lawrence J. Kaetzel

September 1997
Building and Fire Research Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899

U.S. Department of Commerce
William M. Daley, *Secretary*
Technology Administration
Gary R. Bachula, *Acting Under Secretary for Technology*
National Institute of Standards and Technology
Robert E. Hebner, *Acting Director*

Acknowledgements

The authors wish to thank Dr. Geoffrey Frohnsdorff, Chief of the Building Materials Division at NIST, for his long-term support of the CIKS effort. Also we wish to thank Bill Danner, NIST, and Dr. Robert Amsler, Integrated Resource Group Inc. for many conversations to discuss some aspect of the framework, and the many reviewers of this report for their comments.

Abstract

Developing a proposed framework for a Computer Integrated Knowledge System (CIKS) for construction materials, components and systems requires a practical approach in order to be utilized by the building materials industry. The need for practicality has driven the design of the framework to include many commercial technologies, as well as research results where feasible. This report presents the framework from the external, management, and operational views, followed by a conceptual representation discussion in Appendix A. The external view includes the partners, user interface, and network. The management view addresses the application, data, and knowledge management. These views are followed by an operational view, which includes a pilot system, menu of options, and development phases, which outlines the steps for building an example of the framework. Appendix B contains a discussion on the CIKS coating partnership for highway structures. Appendix C , a glossary, concludes the report.

Disclaimer

Trade names and company products are mentioned in the text to adequately characterize the implementation focus of this framework. In no case does this identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily the best available for the purpose.

Table of Contents

1 Introduction

The construction community is represented by a diverse set of private and public organizations and disciplines. The industry lacks consistency in the way it represents, communicates, and interprets information about its products, materials, and systems. There is a need to establish a framework that will improve the service-life and durability of structures and components, and reduce costs associated with construction, operations, and maintenance activities. Industry-scale improvements can result by the adoption and implementation of a framework that includes:

- consistent terms for use by construction industry knowledge users;
- criteria and standards for data quality and formats;
- a standard format for knowledge interchange;
- improved methods for seeking and interpreting knowledge;
- an intelligent interface for users.

This report outlines a proposed framework to address the needs of the construction industry product, materials, and systems knowledge users involved in operation and maintenance activities, such as condition assessment, material failure analysis, and material selection. The report is published in both printed and digital form. Comments on implementation strategies and opportunities can be provided to the authors on-line through the World Wide Web (W^3) site for six months after publication. The framework can be reviewed and comments submitted to the Computer Integrated Knowledge System (CIKS) developers by pointing a W^3 browser to:

www.ciks.nist.gov/framework.html

Refinements to the CIKS framework will be ongoing in the W^3 document and updated periodically in a NIST report.

CIKS is intended to be used for constructed facilities, such as bridges, private and public buildings, etc. Initially, the framework addresses the industrial coating material area. However, the activities and methodologies described will apply to other construction materials, such as cementitious materials, steel, aluminum, composites, and roofing.

Several goals have been established for the implementation of CIKS. These goals provide a context for its development, a long-term vision, and near-term usefulness. It is envisioned that refinements to the framework will occur as user needs change and are better understood by the CIKS developers and partners, and as enabling information technologies emerge over the 5-year development life of CIKS. CIKS will show concept and provide value within a two-year time frame, yet maintain a five-year development life. This will be achieved through the establishment of NIST/construction industry partnerships and the establishment of a test bed whereby partners can test production systems and data for interoperability, and obtain developmental and implementation solutions for incorporation into industry-developed systems.

Goals of CIKS

Universal exchange of knowledge

One goal is to help the construction industry share, exchange, and manage knowledge through a neutral knowledge interchange format. A major focus will involve document-based knowledge sources such as handbooks, guides, standards, and dictionaries. These documents provide a clear reference to knowledge sources that have been refereed by industry leaders. Examples of these include the American Society for Testing and Materials (ASTM), the Society for Protective Coatings (SSPC), and the American Concrete Institute (ACI). Use of guides developed by the public sector will result in the creation of digital libraries, and decision-support aids (e.g., expert systems) that will be available to knowledge users, such as practicing engineers. Commercial guides will be incorporated into the CIKS framework by: 1) an interpretation of the guides (contained within an expert system) to aid in engineering practice, and 2) the application of electronic commerce, permitting the sale of documents in electronic form.

Knowledge format	Example of use	Knowledge source	Examples
Facts and rules-of-thumb	Represented in decision-support functions by a reasoning process, resulting in a computer-generated hypothesis, conclusion, or recommendation.	Domain experts, standards organizations, guides to practice	HWYCON expert system
Documents: standards describing test methods, material specifications, guides to practice, results of relevant domain specific work	Form digital libraries that can be retrieved, viewed, and interpreted by a knowledge user.	ASTM, SSPC, ACI, AASHTO, public works facilities, product and material manufacturers and suppliers.	ASTM C-150 "Standard Specification for Portland Cement", SSPC "Good Painting Practice", ACI "Manual of Concrete Practice"
Data	Entities (databases, spreadsheets, photographs, audio/video clips) that reflect results or provide facts.	Product and material supplier performance measurements, materials research organizations.	Product data sheets, corrosion resistance data sets from accelerated aging tests, photographs of material failures and conditions.

Table 1: CIKS knowledge formats, examples, and knowledge sources.

Additional knowledge formats will be included in the CIKS framework. These include, data from material tests, and condition assessment (both human and machine-readable), visual information, facts and rules-of-thumb from domain experts. Table 1 provides examples of CIKS knowledge formats, examples, and knowledge sources. The challenge brought by the use of multiple knowledge formats involves the need to integrate information to support the decision-making process. Just-in-time delivery of information, coupled with "when-changed" and "when-needed" actions are important considerations. An example of a "when-changed" action can occur when new data, information, and knowledge is created (i.e., a coating formulation has changed that affects its performance). A "when-needed" action occurs when information is needed to make an informed decision, such as detailed information for better understanding or to substantiate a computer-generated hypothesis, conclusion, or recommendation.

The ability to integrate and represent the Table 1 formats will depend largely on the success of establishing common terms, dictionaries, and standard guides for representation. Also, criteria must be established that reflect data integrity or its absence.

Commercial development and implementation tools

An evaluation of commercial off-the-shelf (COTS) products and information technology research will be conducted with an emphasis on reusability. This will accelerate the development process and reduce the level of effort needed to implement CIKS. The intention is not to conduct new research, but to integrate or 'glue' existing information technologies and research results into a functional framework. If voids exist, then the value of building or researching those missing links will be considered.

Development and implementation partnerships

Private and public knowledge centers and digital information networks with specific focus areas (e.g., coating, high-performance concrete) hosting their knowledge repository will serve as examples of the CIKS framework. Industry leaders will collaborate as partners to establish and host knowledge bases, identify user requirements, and provide reusable research results. One factor involved in the success of CIKS will be industry's interest and ability to adopt CIKS strategies and, in some cases, reengineer their business process. User demands for value added knowledge in digital form will serve as a driver for product and material manufacturers and information brokers to maintain a competitive advantage. Upon completion, the various stages for implementing the CIKS framework, will be submitted to a standards setting body for standardization. Two scenarios can be presented for possible standardization: 1) a U.S. national body such as ASTM, or 2) a domain specific body such as The Society for Protective Coatings for the coating industry or the American Concrete Institute for cementitious materials. The standardized framework would establish standards and guidance on terminology, data formats, and models for representing processes (e.g., condition assessment, and material selection procedures) in decision-support systems.

CIKS Proposed Framework

Contained in this report are several perspectives, which include the external, management, and operational views. Section 2, the external view, includes the partners, user interface, and network. Section 3, the management view, describes the application, data, and knowledge management. Section 4, an operational view, describes the pilot system, menu of options, and the different development phases, and outlines the steps for building an example of the framework. Section 5 provides the summary. Appendix A, a conceptual representation view, describes exchange options, Appendix B provides a description of how the CIKS framework could be implemented for a CIKS coating partnership for highway structures, and Appendix C is a brief glossary.

Figure 1 shows the different views and their respective parts. The foundation of the framework is built on the conceptual representation which includes dictionaries, a lexical semantic model (a model of the relationships contained in the dictionary), standard terms, data elements, data models, document models, and knowledge models. All of these items will be discussed in their respective sections. (The square shapes refer to formal or engineered items and the curved shapes refer to natural or human aspects.)

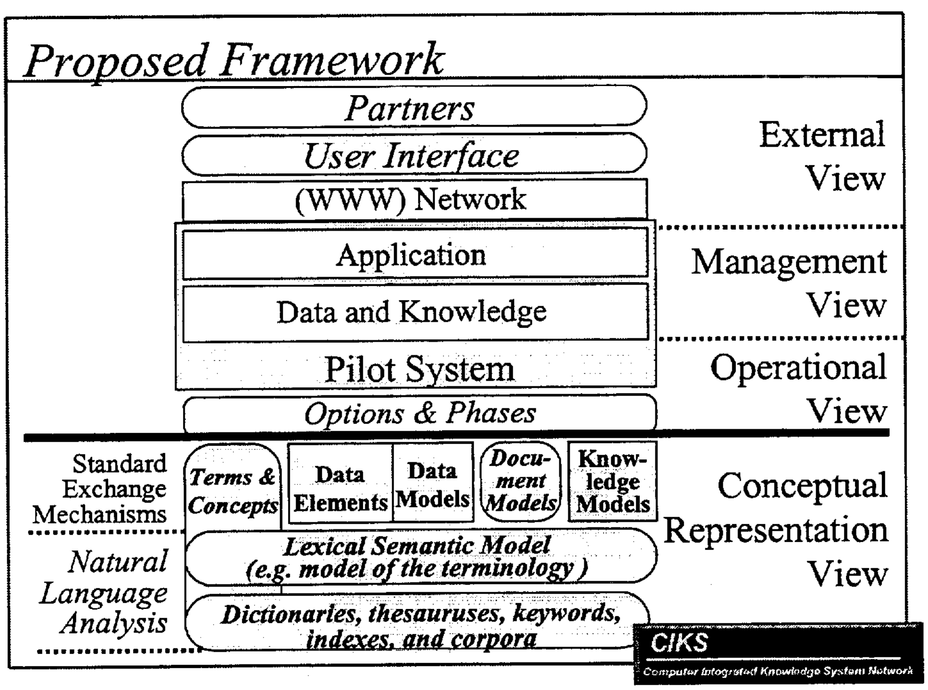

Figure 1

Figure 2 shows that dictionaries, thesauruses, and corpora or machine-readable text are the inputs to creating the lexical semantic model, a foundation of the framework. This figure also shows the relationship of the lexical semantic model to the implementation of data, document, and knowledge exchange mechanisms (through standard models) and its role as the foundation upon which all of the other models are built. The current exchange models provide only a syntactical exchange mechanism, while the lexical semantic model provides the foundation for a much more powerful semantic exchange mechanism. However, the intention is not to conduct natural language research, but to incorporate the research results into a functional framework.

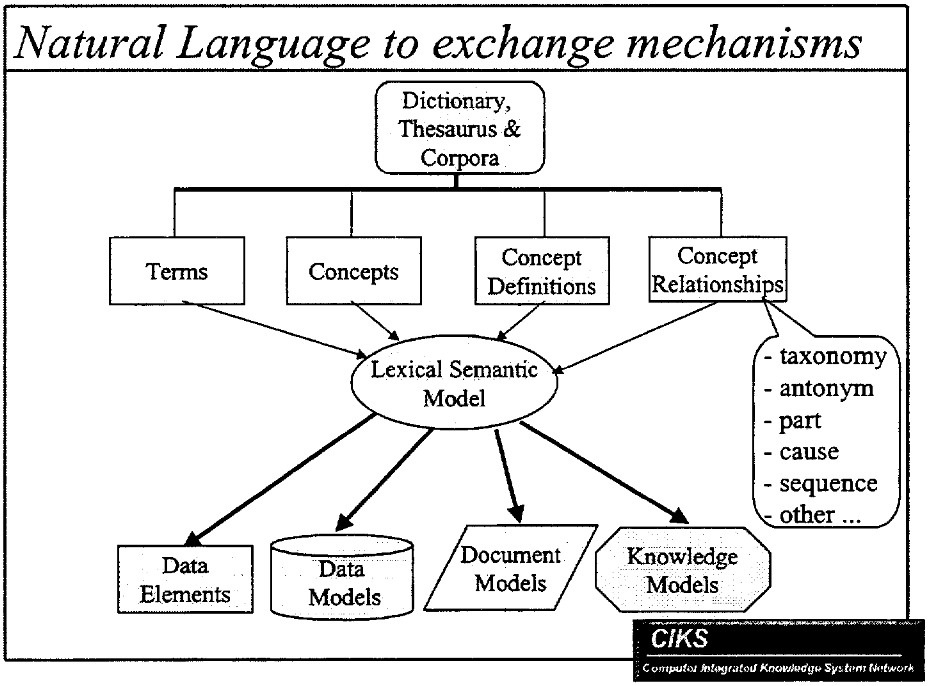

Figure 2

2 External View

The external view is analogous to looking at the outside of a car. This view has been divided into partners, user interface, and network.

2.1 Partners

Partners are organizations with whom we will collaborate. The three types of partner organizations that we plan to work with in building this framework are knowledge providers, user associations, and research organizations. Some organizations may incorporate two or three areas. For example, a trade association (e.g., the Society for Protective Coatings; American Concrete Institute) is a knowledge provider and represents users, while exploring internally how to deploy their knowledge. The three categories help focus our collaborations.

Knowledge providers will become "Knowledge Centers" for specific domain areas. These organizations must determine if it is in their best interest to become a national or worldwide expert in their domain area. With that goal in place, they will be interested in a framework, ultimately including electronic commerce, that helps manage and disseminate their knowledge. Standards, trade, and private organizations are typical organizations that are in different stages of incorporating the CIKS framework.

User organizations that will be identified as potential partners include those that are concerned with facility operations and maintenance. We would look to these organizations to help guide knowledge presentation, and to determine an appropriate user interface. These organizations could help us focus on additional areas that would help their constituents.

Finally, research organizations will provide reusable information technology research results that could be incorporated into the framework. These include educational, government, and private organizations with an interest in research that is freely available to the public. Typically, government funded research programs at universities are of interest since the intellectual property is normally public and available on the W^3.

2.2 User Interface

Using the premise that knowledge repositories will be distributed and accessible from knowledge centers, the design and implementation of a coherent user interface will be challenging. One of the major challenges is usability. Several approaches will be taken, which include considering the user's context, providing a natural-language interface, and/or providing an application software based approach.

The context-based approach would consider a number of perspectives. The following three have been identified:

- Product life-cycle: planning, design, construction, operations, maintenance and de-commissioning;
- Technical discipline: architectural, civil, mechanical, electrical, materials, etc.;
- Industry: commercial buildings, transportation, process plants, ship building, etc.

By reviewing well-known document interfaces (e.g., Sweets or Thomas Register) used in a specific context, a familiar user interface could be provided, and modified by the user. However, the copyright issues within the user interface area must be considered.

A general natural-language based approach would be a challenging undertaking. However, in the near future, we may have an adequate system for attempting to answer simple, very constrained natural-language questions.

Another approach is application software based, with similar application types having similar interfaces. For example, if all expert systems built in this framework used the same product, the interface would be the same assuming the same design specification was applied.

Accurate assessment of user needs will facilitate building customizable user interfaces to handle the distributed knowledge bases. By reviewing commercial vendor products and results from the computer science research community, insight into which enabling information technologies should be considered for implementation will be provided. Areas such as natural-language processing, virtual reality, and distance learning will provide additional insights into how to design a robust user interface. CIKS material and information technology working groups, established to develop pilot projects for their respective areas, will test the feasibility and usefulness of the CIKS user interface. Finally, providing help systems or computer-based training will bridge the gap during the transition stage from an evolving user interface to one that is highly usable.

2.3 Network

The network is the digital communication infrastructure that enables users to connect to knowledge bases hosted at knowledge centers. The W^3 will provide the major networking infrastructure. Where possible, we will focus on the "non-networking" aspects of the framework and adapt to the particular "networking de jour" since information technology product life cycles are getting shorter. For example, a commercial software vendor, Citrix Incorporated, has developed the WinFrame client-server product that allows one to run Windows™ applications using the W^3. Use of this capability simplifies porting of CIKS applications from a Windows™ client-server environment to a totally W^3 integrated application environment.

3 Management View

The management view is a top-down view of the major functional components, divided into application, data, and knowledge management, with each component having additional sub-elements. The distinction between data and knowledge is tenuous. Dictionaries appear to be circular in their definitions of data and knowledge. For our discussion, data is what users store in commercial databases, though some databases could store all of the knowledge management sub-elements. What distinguishes typical data from knowledge in this report, is the ability to inference. Data in this sense is more static, while knowledge is more dynamic.

3.1 Application Management

This section is divided into the user-oriented and the research-oriented applications. Both application types will interface with the data and knowledge management components discussed in the next sections.

3.1.1 User applications

The purpose of this section is to identify tools or technologies to help users improve the performance of work tasks using commercial products. These include information retrieval systems, collaborative environments, computer-based training systems, expert systems, or any other available technology that would improve a user's performance within the context of the framework.

An appreciation of the state-of-the-art in information retrieval can be acquired by searching for content on the W^3 using search engines such as AltaVista, Excite, and Yahoo. Commercial information retrieval products can be obtained from major software vendors like Fulcrum [W^3-24], Microsoft [W^3-17], Oracle [W^3-18], Verity [W^3-25], and many others. Documents may be divided into "information or knowledge chunks" to provide the relevant and appropriate level of detail and avoid retrieving a book when only a paragraph or page is needed.

The goal of using collaborative software is to overcome space and time limitations in order to enable people in different locations and time zones to work together as a virtual organization. Email is a common application, which allows collaboration with others at different locations and different times. Using collaborative technology to allow knowledge workers to collaborate at different locations during different times (e.g. when creating or revising a knowledge document) will reduce the time to completion of a task. Document Management Systems (DMS), discussed in section 3.3.2, should provide the capability to perform collaborative document processing, as will tools from major software vendors that allow one to collaborate across the Internet. A recent Seybold Report [Eccl97] identified a product that "allows authors and editors to enter and edit content through web browsers." Finally, national efforts have been the focus of attention at workshops such as the "IEEE Sixth Workshop on Enabling Technologies:

Infrastructure for Collaborative Enterprises" [W^3-20] and the National Industrial Information Infrastructure Protocols [W^3-8].

Computer-based training has the capability of providing just-in-time training across the W^3 or on CD. This improves human knowledge acquisition and the comprehension aspect of knowledge management. A recent article [Seac96] evaluated different authoring tools including Asymetrix's ToolBook II [W^3-9], which is a W^3 based product. Tools such as these allow end users to build their own customized computer based training products to improve human comprehension, which should be considered within the context of knowledge sharing and exchange.

The previously mentioned technologies may be augmented by expert systems to provide additional power or a better user interface. C Language Integrated Production System (CLIPS) [Giar93] is an expert system tool that is being evaluated for decision support. Detailed information on CLIPS is available at NASA [W^3-10]. CLIPS is of interest because translators exist between it and knowledge interchange formats. CLIPS (developed by NASA's Lyndon B. Johnson Space Center) can represent knowledge as rules, functions, and objects. The National Research Council of Canada created FuzzyCLIPS [KSLI94], which extends the capabilities of the CLIPS rule-based expert system shell and allows the representation and manipulation of fuzzy logic (rules). Fuzzy logic is a function used to express uncertainty, such as conditions that are neither totally true nor false. FuzzyCLIPS can deal with exact, fuzzy, and combined logic, allowing fuzzy and normal terms to be freely mixed in the rules and facts, while using only two basic inexact concepts, fuzziness, and uncertainty. Finally there is the Java Expert System Shell (JESS) at Sandia National Laboratories [W^3-11] which is described as a clone of the core of CLIPS written in Sun's Java by Ernest Friedman-Hill.

Other technologies, such as intelligent help-desks, could be adapted to maintain best practices and provide answers to frequently asked questions. A variety of methods may be used to find similar and relevant past practices. Other commercial solutions are emerging, which will require further investigation. The goal in this area is to integrate commercial or commercial-like technology in order to enable a working CIKS framework for providing user support.

3.1.2 Research applications

Information technology applications that are research oriented include two main areas of interest for CIKS: software agents and automated knowledge acquisition. Nwana's [Nwan96] overview of software agents describes agents that are static or mobile, deliberative or reactive, autonomous, and cooperative, as well as agents which have the ability to learn. As agent technology matures and begins to perform human-like tasks, there will be an ever greater need and benefit for distributed data and knowledge repositories. Agent technology is where "the rubber meets the road" in terms of using and testing the benefits of distributed data and knowledge bases. Commercial software agents (e.g., Autonomy, General Magic, IBM, and Verity) currently exist for this purpose.

The focus of CIKS will initially be on document-based knowledge (e.g., handbooks, practice guides, codes, etc.) instead of human domain experts for our models. Therefore, natural-language processing (NLP) technology is being evaluated for use in automated knowledge acquisition of machine-readable documents. Current academic research [HBCAD91, HS95] indicates the potential of this technology. Advances in this area will increase the potential for building knowledge bases and overcoming one of the major bottlenecks in building smart systems. The main goals are to, 1) create a lexical knowledge base from a machine-readable dictionary; 2) create a Knowledge Interchange Format (see section 3.3.3) representation from a text document using a lexical knowledge base; and 3) create a machine-readable dictionary from text documents. If this were successful, knowledge that is currently locked in machine-readable documents could be incorporated into operational systems.

3.2 Data Management

The data management part is divided into standards, models, and tools. Data management is an integral part of any knowledge base since knowledge-like behavior requires access to large amounts of relatively static data. In fact, some databases could store all of the knowledge management sub-elements as drawn in figure 3 in section 4.1 and as mentioned in the introduction to the Management View section.

ISO 10303 Standard for the Exchange of Product Data (STEP) and ISO 11179 Specification and Standardization of Data Elements will be used in the framework where appropriate. In addition, the ASTM Committee E-49 on "Computerization of Material and Chemical Property Data" established fundamental principles on data elements and standards, and is being used as a basis for developing product data formats for coatings and concrete materials. The efforts of other organizations will also be utilized.

Standard data elements within data models are similar to terminology in documents, and are considered fundamental to data sharing for non-geometry (e.g., non-CAD) type data. The same representational concerns that exist for terms and concepts apply to data elements. Therefore, standard data elements should be developed and refined by defining them based on an analysis of an area's terminology where possible (see the next section). One of the first steps in building the CIKS framework will be identifying and agreeing upon standard data elements for data sharing or exchange. One example of the scale of the effort is the ongoing Department of Defense (DoD) effort [W^3-2] in developing its enterprise-wide models and registering standard data elements, which total between ten to fifteen thousand. Other examples of the establishment of standard data elements and data formats for construction materials involve the Society for Protective Coatings, Committee C.4.10 which is developing a standard for coating-product data format and the American Concrete Institute, Committee 126 which is developing material property formats for the constituents of concrete [Kaet97].

Data models differ in domain or abstraction.

- Examples of differences in domain include engineering versus science, specialized versus generalized, and single stage versus multi-stage life cycle.
  - Modeling restricted to the engineering context is not as complex as when scientific research is included (e.g., materials engineering versus materials science).
  - A specialized model focuses on a certain restricted context such as a particular type of material and its use. In contrast, a generalized model must be designed to accommodate multiple types of materials (e.g., coatings, concrete, steel).
  - A single stage model refers to one stage in the life cycle of a structure. Multi-stages include more than one life-cycle stage, or the whole life-cycle perspective (e.g., planning, design, construction, operations, maintenance and demolition).
- Examples of differences in abstraction include conceptual versus physical, data versus process, and temporal versus non-temporal.
  - A conceptual model attempts to accurately represent the natural world without concern for implementation details, while a physical model is an implementation of the conceptual model using a specific database management system.
  - Data models are in contrast to process models. If we use the example of a cup of coffee, the process model would address making the coffee (put x number of scoops in a filter, fill with water, turn on coffee maker, etc.). The data model would describe the properties (type of coffee bean, amount of coffee, amount of water, etc.).
  - Temporal aspects become important when there is a concern about real-time issues or the need to maintain a history of all of the transactions to roll back to a consistent point in time.

These dimensions must be taken into account as models are built to describe an area of interest and assist in communication among users and applications interoperability. For example, one might specify whether the context relates to a specialized design-stage conceptual data model, or a generalized physical life-cycle process model, as well as the content associated with that model type. One quickly realizes how challenging it is to create this artificial reality within information models.

Implementation products or tools are available to help manage data and metadata (data about the data) over the life cycle of planning, design, development, operations, and maintenance. Implementation products include Computer Aided Software Engineering (CASE) tools to manage planning, design and development, and Database Management System (DBMS) tools that deal with operations and maintenance. CASE tools can be divided into upper or lower CASE tools, where the upper refers to the planning and design stage, while the lower refers to the application development stage.

The DBMS marketplace includes relational, object-relational, and pure object-oriented tools where each technology has programming and performance advantages and disadvantages for different kinds of data. Major database software vendors provide W^3 sites to advertise the ease of use and/or performance of their database management systems for particular applications.

3.3 Knowledge Management

Initially, a major emphasis of the framework will be on document-based knowledge management that will be enhanced by emerging technologies with an emphasis on knowledge sharing or exchange. The justification for this is based upon the fact that documents currently contain most of the engineering know-how or knowledge in the form of handbooks, design guides, standards and other natural-language forms. This section will discuss the areas of terms and concepts, documents, and knowledge, which comprise the CIKS knowledge management area.

3.3.1 Terms and Concepts

It is theorized that the framework's foundation should be built on managing diverse terminology or vocabulary. Precise word meanings are the building blocks for accurately expressing and specifying technical data and knowledge in every discipline. To assist in shared communication, it is necessary that word meanings be evaluated on a common basis. To ensure that the meanings are equivalent for shared communication requires a rigorous model and methodology to resolve differences in terms. Identifying different terminology for the same meaning or the same terminology for different meanings is critical to ensure shared communication.

The Roget's 21st Century Thesaurus [Kifp92] uses a Concept Index, which it refers to as "a semantic hierarchy of the most common concepts we use in American English, as it is spoken and written today." This concept index begins with ten concepts: actions, causes, fields of human activity, life forms, objects, the planet, qualities, senses, states, and weights and measures. All other concepts fall under one of those ten. The CIKS framework is evaluating the use of various research communities (e.g., linguistics, database, knowledge base) to describe the relationships between the concepts that include and extend beyond a hierarchical structure. Familiar relations include synonymy, antonymy, and taxonomy while others include part, sequence, cause, and child relationships [BB89, Evens88, and PB92].

The focus here is to work with industry to engineer the meanings or semantics of the terminology that will be used to specify all of the data and knowledge within the building materials industry. We propose building these relational models of the terminology to lay the foundation for better enabling data and knowledge exchange. There is a need to identify all of the related dictionaries and thesauruses within each industry, align differences in definitions, and establish an industry standard dictionary and thesaurus that would be available on-line. Trade and standards organizations should lead the effort to provide a unified dictionary and thesaurus, so all future data and knowledge products can be built on a consistent and engineered foundation of well defined words.

The Unified Medical Language System (UMLS) project, which has been ongoing since 1986, provides some insight into the scale of the problem for the medical community. The UMLS project was

> " ... a long-term research and development effort with the ambitious goal of enabling computer systems to 'understand' medical meaning. The project was proposed to Congress as essential to the development of advanced health information systems -- and as requiring an initial 5-10 year development phase which would cost $1 to 3 million dollars per year. The Congress responded with generous and faithful support [W^3-3]."

One can find more information by visiting this National Institute of Health project W^3 site [W^3-3]. David Penniman, Professor and Associate Director of the School of Information Sciences at the University of Tennessee [StCl96, Penn93] has suggested using a unified engineering vocabulary, a similar effort along the lines of UMLS for the scientific and engineering community. There are also international conferences held on the subject [W^3-4] that address the language translation issues in harmonizing terminology internationally.

There is also relevant work in the ISO community. An example is, the ISO Technical Committee 37 (TC37), Terminology: Principles and Coordination, which has a Sub-Committee 3 (SC3) for Computational Aids in Terminology. TC37's SC3 has an ISO DIS 12200: Terminology Interchange Format (TIF) – An SGML Application, as mentioned by Alan K. Melby in STP1223 [WS95]. The TIF format "was developed as part of the Text Encoding Initiative, an international consortium intended to produce carefully thought out, reusable SGML Document Type Definitions (DTDs) for various classes of documents and to place these DTDs in the public domain" [WS95]. There is work from TC37's SC3, reflected in ISO CD 12620, which addresses exchanging terminology data between traditional databases, and should be reviewed for the latest version.

There is existing research in integrating or aligning various terminology or lexical models from the research community. Specifically, members of an ANSI Ad Hoc Ontology Standardization Group have investigated integrating different large lexical models (ontologies) [Hovy97]. This work may indicate that the upper levels of different ontologies may agree enough to be considered for standardization. Work such as this is very relevant to the CIKS framework, because the framework's foundation is built on natural-language analysis of terms and concepts, and assumes that meaning can be standardized and aligned across different fields to help enable data and knowledge sharing. (See appendix A for further information.)

3.3.2 Documents

As mentioned previously, a machine-readable document-centric approach for representing major knowledge sources will be used for the industry. This provides a stable published knowledge source, which can be referred to from an electronic implementation of a digital library of machine-readable documents.

One basic premise is that during the document creation or revision phase, considerations of human comprehension could influence the document structure to ease automated knowledge acquisition for the future. Efforts from organizations such as Information Mapping Inc. [W^3-15], the University of Ottawa's ClearTalk [W^3-16], and Formal Object Role Modeling Language [Asym94], as well as research insights from the natural-language processing for automated knowledge acquisition [HBCAD91, HS95], should influence an organization's editing guidelines for document creation or revision. Another long-range consideration is using Standard Generalized Markup Language (SGML) [GR90] for embedding additional structural, semantic or other additional tags for information retrieval or automated knowledge acquisition, and for those tags to apply to both paper and hyperdocument formats.

Assuming that machine-readable documents are an organization's major knowledge assets, making a reasonable investment in an integrated Document Management System (DMS) is justifiable. Typically, these documents include design guides, handbooks, guides to practice, etc. A DMS may address storage, creation, version, retrieval, access control, and routing of an organization's documents. Frappaolo [20] describes three types of DMS: file, library, and compound document managers. Compound document managers treat documents as a collection of other documents. If we imagine the typical desktop suite of office tools (word processing, spreadsheets, databases, presentation, etc.) a compound document could include all of these document types and more, using technology such as object linking and embedding (OLE), Extensible Markup Language (XML) [W^3-23], HTML, SGML, etc. Frappaolo goes on to identify the extent of the collection as desktop, group, or enterprise-wide. Evaluating W^3-enabled enterprise-wide compound document manager products provides insights into this technology. Among the vendors identified by Frappaolo are Documentation, Inc. [W^3-5] and Interleaf Inc. [W^3-6], which provide broad coverage of the technology. Open Text is an example of a quick growing W^3-enabled DMS [AW97].

Finally, we will be investigating the DARPA, NASA and NSF funded Digital Library Initiative [W^3-7] as it relates to implementing a knowledge center for insight into how machine-readable electronic documents could be managed for maximum utility. The initiative's focus is to research the means to collect, store, and organize information in digital forms, and make it available for searching, retrieval, and processing via the W^3 with a user-friendly interface.

3.3.3 Knowledge

One important objective is the identification of a neutral knowledge structure as an interlingua or repository to enable knowledge exchange, sharing, or reuse. Currently, the use of the DARPA-sponsored Knowledge Sharing Effort [PFPMFGN93] is being evaluated. The DARPA effort includes the Knowledge Interchange Format (KIF), Knowledge Query and Manipulation Language (KQML), and an Ontolingua server [W^3-12]. KIF is a first-order logic-based language with extensions, serving as an interlingua between knowledge systems and as a meta-knowledge representation language. The goal of KIF is to provide a thorough language for expressing knowledge while minimizing the constraints of operational requirements. Stanford University, through its Ontolingua server, [W^3-13] has implemented a subset of KIF with extensions for use by collaborating W^3 authors. However, there are other competing interlingua languages such as conceptual graphs [NNGE92, PN93, W^3-14], EXPRESS [Wils96], and higher-order-languages (beyond first-order logic) [GM93]. KQML [MLF95], a language and protocol for exchanging information and knowledge, is being evaluated by the research community to interface with other distributed knowledge bases, and will be tested for applicability in the software agent areas (see section 3.1.2). For these reasons, the DARPA effort is being considered as a practical choice to evaluate at this time.

A major effort will be required to integrate existing knowledge bases. In order to accelerate the construction of a building materials knowledge base, it will be necessary to incorporate existing knowledge bases in these and other related efforts. Consideration may be given to building translators from native formats that the existing knowledge base is in, to a neutral knowledge interchange format such as KIF or a KIF-like structure. The long-term goal is to develop sharable knowledge to demonstrate the concept, initially by creating a central repository and then by identifying solutions to using distributed knowledge bases. However, it must be said that KIF only provides the syntactical solution, not the semantics. The semantics will have to be built using tokens embedded in a KIF or KIF-like structure.

4 Operational View

The operational view consists of the pilot system, menu of options, and development phases. The initial application context will focus on determining the remaining service-life of a coating, performing coating failure analysis, and then identifying the optimal corrective response. The optimal response will be based on life cycle costing and could include doing nothing, providing in-house or an out-sourced repair, or an in-house or out-sourced replacement action. The operational view will focus on addressing this application requirement.

4.1 Pilot System

Figure 3 shows a pilot system layout between the major components, described below.

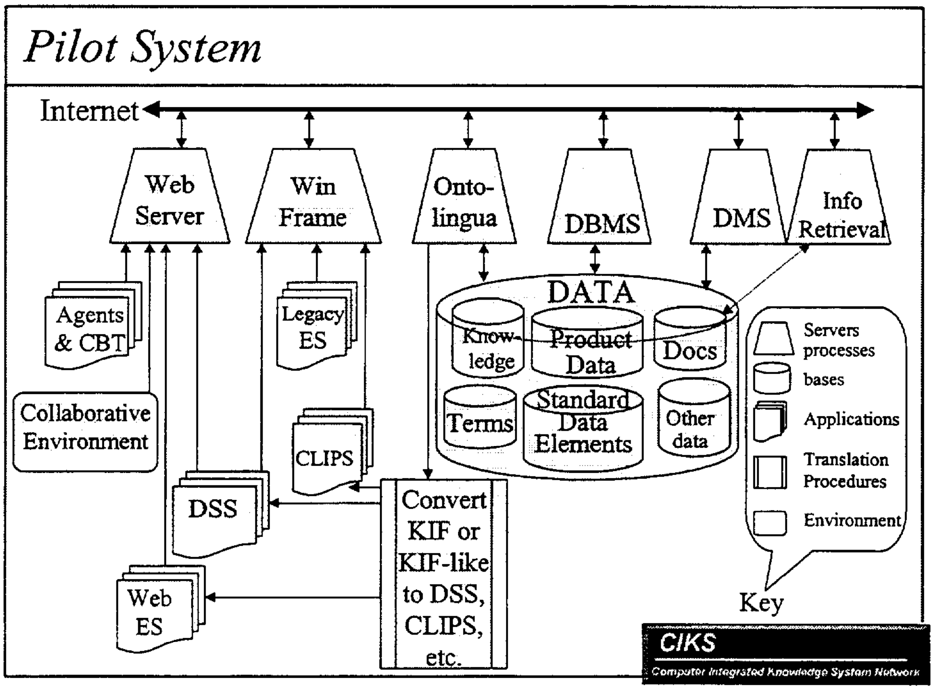

Figure 3

1. The *Web server* provides users with access to the system through the application interfaces (e.g., decision support system *(DSS)*, Web Expert System *(Web ES)*, and computer-based training *(CBT)*). The web server will interact with all of the other servers. It is also considered the front door where all users would enter a web-based knowledge center, therefore it must all fit within an intelligible user interface.
2. *WinFrame* is a server product (as mentioned in section 2.3), which allows Windows™ applications to run through a web browser. This technology will provide an interim capability until existing applications are reengineered to operate completely on the web.
3. *Ontolingua* provides the knowledge repository, and uses first order predicate logic and frames for representing knowledge. Currently, work is underway to test the

feasibility of using the DARPA Knowledge Sharing infrastructure, prior to going full-scale. Assuming it is feasible, we will consider the effort to build the translation procedures to go from Ontolingua into other representations (e.g., Web ES, DSS, and SQL) for operations. The knowledge base shown in figure 3 is the repository for the initial coatings knowledge.

4. *DBMS* is a web-enabled data base management system that could have many databases within the *DATA* item shown.
5. *Docs* is the collection of machine readable documents which will be resident either in the operating system file system, inside a database, or inside a document management system.
6. *DMS* is a web-enabled document management system. It is possible to use either a DBMS or the operating system file system to store machine-readable documents.
7. The *Information* (or text) *Retrieval* server would need to be considered based on what kind of document management system is implemented for the scenarios of the documents being in a file system, a DBMS or a DMS.

4.2 Menu of Options

Table 2 below provides options within the framework, and identifies the technology that can be used in each row in the table. This table is not meant to be complete, but it provides the general breadth of areas that need to be considered. The conceptual representation will be discussed in the appendix, where parts of it are shown in the context of the implementation. The data column is shown separately, although all of the knowledge management pieces could be considered data and stored within a capable database. As mentioned earlier, knowledge management is distinguished from data, to focus on the knowledge management aspects of terms, documents, and knowledge exchange.

The first row, primary real-world sources, identifies the source for developing the particular column (data, terms and concepts, documents, knowledge). Within the Knowledge Management area, it is easier to use documents as the initial primary source than to consult domain experts. The second row identifies example standards, communities and projects that will be considered as the CIKS framework evolves. The third row (addressed in the appendix) presents examples of conceptual representation alternatives, which though not inclusive, provides a set of choices. The fourth row, exchange, or sharing option identifies known choices for exchanging or sharing data, documents, and knowledge. As mentioned earlier, the exchange structures only provide a syntactical mechanism, without the semantics. The semantics will be built using a KIF or KIF-like structure, first by modeling the terms and concepts, then evolving to the documents and data, and finally to the knowledge or know-how representation. The fifth row identifies the tools that would be used during the development of the data, terms and concepts, and documents and knowledge. Finally, the last row refers to the commercial environment for operational purposes. For example, in the data column a commercial database product would be selected which would influence the choices of available development tools.

Headings	Data	Knowledge Management		
		Terms and Concepts	Documents	Knowledge
1. Primary real-world sources	• Domain Experts	Documents • Dictionaries • Thesauruses • Keywords	Documents • Handbooks • Guides • Standards	Documents • Handbooks • Guides • Standards
2. Example standards, research communities or projects	• STEP • ASTM E49 • DAMA • ODMG • Metadata • IEEE	• ASTM E02 • ISO TC 37 • Terminology • Computational Linguistics • UMLS	• SGML community • ODMA • Digital Library Initiative • Document Management	• DARPA Knowledge Sharing • AAAI • IEEE
3. Conceptual representation alternatives	• ERD • IDEF1X • NIAM • OMT • EXPRESS-G	• Primitives • Concepts • Ontology	Print-based • Logical • Physical Hyperdocument • HAM • Dexter • Tower	• First Order Logic • Frames, semantic nets, production rules • Neural Nets • Uncertainty & Fuzziness
4. Exchange or Sharing options	• Data Elements • SQL • ODL/ODQ • EXPRESS • (Middleware)	• KIF • EXPRESS • Conceptual Graphs • TIF (a DTD)	• Proprietary • SGML DTDs • XML • (ODA)	• KIF • EXPRESS • Conceptual Graphs
5. Development tools	• CASE tools • Database and procedural language extensions	• Ontolingua	• SGML and XML tools	• Ontolingua
6. Operations	Commercial database management systems	Commercial database management system	Commercial document management system	Commercial expert systems

Table 2: Menu of Options

4.3 Development Phases

This section shows a general implementation and the different phases of the CIKS framework. After the material coating has been taken through all of the phases, other materials will be considered to identify the issues of integrating heterogeneous materials (e.g., concrete, steel) or components and systems (e.g., roofing, HVAC).

In Table 3, the column headings identify the different development phases. The column headings of crawl, walk, and run phases are used to represent a progression. The rows identify the type of research applications, user applications, data, documents, terminology, and knowledge management components that would be developed and made available through a pilot system.

Headings	*Crawl Phase*	*Walk Phase*	*Run Phase*
Research Applications		• Software Agents	• Automated Knowledge Acquisition • Intelligent Agents
User Support Applications	• Information Retrieval • Collaborative Environment	• Decision Support • Computer-based training	• Expert Systems
Data	• Web-enabled database • Standard metadata • A few standard data elements	• Many standard data elements • Simple standard data models • Complex data models	• Complex standard data models
Documents	• Web-enabled document management system (or system of systems)	• Simple Standard Document Models (DTDs)	• Complex standard document models
Terminology	• Standard terminology • Adapt existing lexical models	• Existing standard concepts • Thesaurus • Align differences in selected lexical models	• Standard lexical model or ontology
Knowledge			• Standard knowledge model • Knowledge base

Table 3: Development Phases

An abbreviated example that addresses the application described in the opening paragraph of section 4 would initially proceed as described below:

1. The first step is to identify coating dictionaries as the standard for the terminology that is used in the specific application area. If a coatings dictionary does not exist, one could be developed (though this would be very time consuming).
2. The next step is to begin collecting machine-readable documents that address the area of interest. The subjects for this example would include how to: assess the condition of a coating, perform failure analysis, and make economic decisions when the existing coating system must be repaired or replaced.
3. In parallel, decisions about how all of these documents will be managed (e.g., using a file system, an existing database, or a dedicated document management system) must be made and implemented. This would include selecting the web-enabled database and document management system that will be used. If it is not possible to select a single system or type of system for managing all documents, then it will be necessary to develop specialized interfaces between different systems.
4. All documents should be contained within the document management system (or system of systems) and in a format that permits searches using information retrieval

technology. This would be the beginning of a digital library of relevant engineering knowledge or know-how documents concerning condition assessment, failure analysis, and economic decisions.

5. The next step requires reviewing metadata (data about data) standards and deciding which standards will be adapted.
6. After that step, identifying a few standard data elements is similar to agreeing on terminology, but within the database context.
7. Finally, assuming that how-to or knowledge documents are created by the organization, a collaborative environment should be defined and implemented, which could be as simple as using email, some workflow package, or selecting a web-based collaborative document editing environment.

Appendix B includes additional details and provides another example of the implementation details of CIKS.

5 Summary

This report provides an overview of a proposed framework for CIKS that addresses the area of construction industry products, materials, components, and systems. The intent is to provide an implementation solution, which can add value in the short-term, yet is flexible enough to include longer-term information technology research results. It is the authors' intent that this report provide the basis for a blueprint to facilitate exchanging and sharing knowledge for the construction industry. The framework is considered a master plan for the CIKS effort, and therefore will be adjusted to reflect the changing needs of knowledge users and new information technologies. The long-term goal of CIKS is to establish a standard framework for knowledge sharing within the construction industry. Short-term results will involve developing domain specific applications that address specific components (e.g., data formats, digital libraries, decision-support modules).

The framework's foundation is built on standard terminology, which should be used to define standard data elements, formats, and know-how. Standard terminology requires agreeing on dictionaries, thesauruses, and keyword lists, which are written in natural-language. An analysis of natural-language is critical to understand the natural-language issues in data and knowledge sharing that may resurface when data and knowledge is exchanged using formal data or knowledge languages (e.g., SQL, KIF, etc.). A model of the terminology (e.g., lexical semantic model, ontology, or lexical knowledge base) is critical to illustrate issues, obtain consensus, and help create a bridge between our natural and formal worlds and their respective languages.

An indicator of the success of this framework will be its adoption by construction materials organizations that will test the framework in the conduct of their business or practice. The authors are available to consult with organizations which endeavor to implement a knowledge management strategy, and to work with user organizations to help ensure that the framework meets their needs.

Managing an organization's knowledge requires a major effort from the basic methods used by libraries (as the original knowledge repository) to the development of a knowledge management framework. This framework provides a phased approach for an evolutionary process, from developing a digital document library to completing an agent-enabled distributed knowledge base.

Finally, the framework will continually evolve and improve as new research is made available. The authors encourage feedback on CIKS strategies, potential partnerships, and applications.

6 References

[AFQ89] J. Andre, R. Furuta and V. Quint (editors), *Structured Documents*, Cambridge University Press 1989.

[Amsl81] R. A. Amsler, *A Taxonomy for English Nouns and Verbs*, Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics, 1981, pp.133-138.

[Asym94] Asymetrix Corporation, *InfoModeler Guide to FORML*, Asymetrix Corporation, 1994.

[AW97] L. Alschuler and M. Walter, *Intranet vendors shake up the mark at Documation*, Seybold Report on Publishing Systems, v26 n13, April 7, 1997.

[BB89] B. Boguraev and T. Briscoe (editors), *Computational Lexicography for Natural Language Processing*, Longman, 1989.

[BF81] A. Barr and E.A. Feigenbaum, *The Handbook of Artificial Intelligence*, Addison-Wesley Publishing Company, 1981.

[BFRL96] Building and Fire Research Laboratory, *COEX-I Expert System for Steel Protection*, Research Update, v2, n3, National Institute of Standards and Technology, Gaithersburg, MD, Fall 1996.

[BS96] K. Brady and J. Sullivan, *User's Guide for RDA/SQL Validation Tests,* NIST Internal Report 5725, National Institute of Standards and Technology, Gaithersburg, MD, December 1996.

[DOD95] Department of Defense, *Military Handbook: Handbook for Paints and Protective Coatings for Facilities,* MIL-HDBK-1110/1, Naval Facilities Engineering Command, Lester, PA, January 1995.

[DSS93] R. Davis, H. Shrobe, and P. Szolovits, *What is a Knowledge Representation*, AI Magazine, Spring 1993, pp 17-33.

[Eccl97] S. Eccles, *KnowledgeView tackles Web-based editorial system*, The Seybold Report on Internet Publishing, v1 n6, February 1997.

[Evens88] M. W. Evens (editor), *Relational Models of the Lexicon*, Cambridge University Press, 1988.

[Fill84] J. Filliben, *Dataplot Introduction and Overview,* NBS Special Publication 667, National Institute of Standards and Technology, Gaithersburg, MD, June 1984.

[Frap96] C. Frappaolo, *Creating the optimal document database*, Data Based Advisor, Volume 14, No. 4, pp. 89-93, April 1996.

[Furu90] R. Furuta (editor), *EP90 Proceedings of the International Conference on Electronic Publishing*, Document Manipulation & Typography, Cambridge University Press, 1990.

[Giar93] J.C. Giarratano, *CLIPS User's Guide*, NASA, Lyndon B. Johnson Space Center, Information Systems Directorate, Software Technology Branch, May 1993.

[GM93] M. J. C. Gordon and T. F. Melham (editors), *Introduction to HOL*, Cambridge University Press, 1993.

[GR90] C.F. Goldfarb and Y. Rubinsky (Ed.), *The SGML Handbook*, Clarendon

Press, Oxford, England, 1990.

[HBCAD91] J.E. Hodges, L.M. Boggess, J.L. Cordova, R. Agarwal, and R. Davis, *The Automated Building and Updating of a Knowledge Base Through the Analysis of Natural Language Text*, Technical Report MSU-910918, Department of Computer Science, Mississippi State University, September 1991.

[Hend93] P. Henderson, *Object-Oriented Specification and Design with C++,* McGraw-Hill Book Company, New York, NY, 1993.

[Hovy97] E. Hovy, *A Standard for Large Ontologies*, Presented at NSF Workshop on R&D opportunities in the Government, Washington, DC May 1997 (unpublished).

[HS95] F.G.R. Hull, and C. Segami, *Automated Knowledge Acquisition from Encyclopedic Texts*, IEEE International Journal of Artificial Intelligence Tools v4 n3, pp. 349-367.

[IH83] C.G. Interrante and F.J. Heymann (editors), *STP 806 Standardization of Technical Terminology: Principles and Practices*, ASTM, 1983.

[Kaet97] L.J. Kaetzel, *Proposed Format for Data on Cements in a Materials Properties Database*, NIST Internal Report 6034, National Institute of Standards and Technology, 1997.

[KCKS93] L.J. Kaetzel, J.R. Clifton, P. Klieger, and K. Snyder, *Highway Concrete (HWYCON) Expert System User Reference and Enhancement Guide,* NISTIR 5184, National Institute of Standards and Technology, Gaithersburg, MD, May 1993.

[KCS94] L.J. Kaetzel, J.R. Clifton, and K. Snyder, *Users Guide to the Highway Concrete (HWYCON) Expert System,* SHRP-C-406, Strategic Highway Research Program, National Research Council, Washington, DC, 1994.

[Kipf92] B. A. Kipfer (editors), *Roget's 21st Century Thesaurus*, Dell Publishing, New York, New York, 1992.

[KSLI94] Knowledge Systems Laboratory, Institute for Information Technology, National Research Council Canada, *Fuzzy CLIPS Version 6.02A User's Guide*, September 1994.

[Lede91] R. Lederer, *The Miracle of Language*, Pocket Books, 1991

[MLF95] J. Mayfield, Y. Labrou, and T. Finin, *Evaluation of KQML as an Agent Communication Language* in Intelligent Agents II: Agent Theories, Architectures, and Languages, Springer-Verlag, New York, NY, 1995, pp. 347-360.

[NNGE92] T.E. Nagle, J.A. Nagle, L.L. Gerholz, and P.W. Eklund (editors), *Conceptual Structures Current Research and Practice*, Ellis Horwood, 1992.

[NR96] S. Nirenburg and V. Raskin, *Ten Choices for Lexical Semantics*, Computing Research Laboratory, New Mexico State University, MCCS-96-304, 1996.

[Nwan96] H.S. Nwana, *Software Agents: An Overview*, The Knowledge Engineering Review v11 n3, 1996.

[PB92] J. Pustejovsky and S. Bergler (editors), *Lexical Semantics and*

Knowledge Representation, Springer-Verlag, 1992.

[Penn93] W.D. Penniman, *Electronic Libraries and the Engineering Profession*, American Society for Engineering Education Annual Conference Proceedings, vl, 1993, pp. 889-894.

[PFPMFGN93] R.S. Patil, R.E. Fikes, P.F. Patel-Schneider, D. Mckay, T. Finin, T.R. Gruber, and R. Neches, *The DARPA Knowledge Sharing Effort: Progress Report* in B. Neble, C. Rich and W. Swartout, (editors), Principles of Knowledge Representation and Reasoning, Morgan Kaufmann, 1992, pp. 777-788.

[PN93] H.D. Pfeiffer and T.E. Nagle (editors), *Conceptual Structures: Theory and Implementation*, Springer-Verlag, 1993.

[Seac96] D. Seachrist, *How Multimedia Tools Compare*, Byte Vol 21 No. 11, November 1996.

[Srir95] R.D. Sriram, *Intelligent Systems for Engineering: A Knowledge-Based Approach*, 1995 (to be published).

[StCl96] G. St. Clair (editor), *InfoManage: The International Management Newsletter for the Information Systems Professional*, v3, n12, November 1996.

[SW93] R.A. Strehlow and S.E. Wright (editors), *STP 1166 Standardizing Terminology for Better Communication: Practice, Applied Theory, and Results*, ASTM, 1993.

[SW94] D. Schenck and P. Wilson, *Information Modeling the EXPRESS Way*, Oxford University Press, 1994.

[WB97] S.E. Wright and G. Budin, *Handbook of Terminology Management*, John Benjamins Publishing Company, 1997.

[Wils96] P.R. Wilson, *On the Translation of Kif/Frame Ontologies to EXPRESS*, NISTIR 5927, 1996.

[WS95] S.E. Wright and R.A. Strehlow (editors), STP 1223 *Standardizing and Harmonizing Terminology: Theory and Practice*, ASTM, 1995.

[W^3-1] Http://www.ciks.nist.gov

[W^3-2] Http://ssed1.ncr.disa.mil/srp/datadmn.html

[W^3-3] Http://www.nlm.nih.gov/pubs/factsheets/umis.html

[W^3-4] Http://www.jula.upf.es/riterm/ritxinuk.htm

[W^3-5] Http://www.documentum.com

[W^3-6] Http://www.interleaf.com

[W^3-7] Http://www.cise.nsf.gov/iris/DLHome.html

[W^3-8] Http://niiip.org

[W^3-9] Http://www.asymetrix.com/products/toolbook2

[W^3-10] Http://www.jsc.nasa.gov/~clips/CLIPS.html

[W^3-11] Http://herzberg.ca.sandia.gov/jess

[W^3-12] Http://www.cs.umbc.edu/kse

[W^3-13] Http://www-ksl.stanford.edu

[W^3-14] Http://turing.une.edu.au/~cgtools/cgtools97.html

[W^3-15] Http://www.infomap.com

[W^3-16] Http://www.csi.uottawa.ca/~kavanagh/Ikarus/Cleartalk.html

[W^3-17] Http://www.microsoft.com/ntserver/info/indexserver.htm

[W^3-18] Http://www.oracle.com/products/oracle7/oracle7.3/html/context_opt.html
[W^3-19] Http://www.microsoft.com/netmeeting
[W^3-20] Http://www.cerc.wvu.edu/WETICE
[W^3-21] Http://www.win.tue.nl/2L670/static/architecture.html
[W^3-22] Http://www.w3.org/pub/WWW/MarkUp/DOM
[W^3-23] Http://www.w3.org/XML
[W^3-24] Http://www.fulcrum.com
[W^3-25] Http://www.verity.com

Appendix A: Conceptual Representation View

This section focuses on the conceptual representation research background, documents some of the research, and points the reader to relevant references. It is not, however, intended to be a complete survey of the field. The framework was assembled with an implementation bias and with consideration of a reasonable representational foundation. Included in this section are: background information, discussion about data, term and concept, document, and knowledge representation models.

A.1 Background

It is thought that natural language is the root or foundation upon which we build formal models and that better understanding of natural language problems will provide insight into building better formal models. Relevant work from the computational lexicography field discusses research using machine-readable dictionaries. One study by Wilks et al. [BB89] analyzed a small dictionary for learners of English as a second language, Longman's Dictionary of Contemporary English (LDOCE). The preparers of LDOCE stated that they used a controlled vocabulary of about two thousand words. This controlled vocabulary was used to define all of the other words, and there were on average twelve different meanings for each word. In that study, one of the researchers identified a set of about twelve hundred words called the "Key Defining Vocabulary" which were found to define the two thousand word controlled vocabulary. What remains to define those twelve hundred words are only themselves. Another study conducted by Professor Ohmann at Wesleyan University [Lede91] provides insight into the richness of natural language. In this study, he asked twenty-five students to describe, in one sentence, a single cartoon. He then collected the twenty-five sentences and determined that he could create twenty billion grammatical sentences, just with the words that were used in the initial set. This is an example of the complexity of natural language, and it points to the challenge involved in creating our models.

Models are used as a substitute for what exists in a natural or abstract world. No model can replicate the infinite complexity of our natural world; therefore, no model is complete. Therefore, we must accept that the consequences of imperfect models are imperfect results [DSS93]. As we automate knowledge-like activities involving expert systems and intelligent agents, the computer representation may become more complex and not easily understood by people used to using only natural language. The benefit of this complexity is the ability to automate what previously required human participation. Currently, although computers have problems with processing natural language, they are improving. Perhaps some day in the far future we will be able to converse with a computer. However, for now, we need to build formal models based on formal representations.

A.2 Data Models

This section refers to the STEP (ISO 10303) community to provide a data modeling methodology. There are many alternative representations such as entity-relationship, IDEF1X, NIAM, OMT, Shlaer-Mellor, DAPLEX, GEM, SQL, EXPRESS-G, and EXPRESS [SW94]. STEP has been active in the data representation area and there are many commercially available data modeling products for interested parties to review for more information. Due to the accessibility of information and coverage of data representation by STEP and the commercial community, CIKS will look to these existing efforts for guidance in the data modeling area.

As mentioned earlier, the terms, concepts, documents and knowledge can be represented in a data model and stored inside a database. This report has made the distinction to focus our discussion on the knowledge management aspects. What distinguishes typical data from knowledge in this report is the ability to inference. Data in this sense is more static, while knowledge is more dynamic.

A.3 Term and Concept (Lexical Semantic) Models

The *Handbook of Artificial Intelligence* [BF81] discusses semantic primitives within the knowledge representation chapter, and characterizes them as the building blocks that provide meaning. Wilk's research, cited in that chapter, identified recommended properties (finite, comprehensive, independent, noncircular, and primitive) to develop a set of primitives. This was contrasted with Shank's Conceptual Dependency Theory, which is based upon concepts, instead of words. Shank claimed his representation should be language independent, assuming the set of unique concepts was identified. However, in an environment of syntactic and semantic ambiguity, his concept representation required an unambiguous and unique representation. Both ideas appear similar in that ultimately, senses or concepts need to be identified.

Initial work on dictionaries [Amsl81] provided insight into using computational methods to identify the tangled semantic relationships between lexical items contained in a machine-readable dictionary. This effort and others resulted in development of the field of computational lexicography. Other examples of research which provide background information for modeling lexical semantics are contained in *Relational Models of the Lexicon* [Evens88], *Computational Lexicography for Natural Language Processing* [BB89] and *Lexical Semantics and Knowledge Representation* [PB92].

One view of the options is provided in a recent paper entitled, "Ten Choices for Lexical Semantics," by Nirenburg and Raskin [NR96]. The abstract states:

> The modern computational lexical semantics reached a point in its development when it has become necessary to define the premises and goals of each of its several current trends. This paper proposes ten choices in terms of which these premises and goals can be discussed. It argues that the central questions include the use of lexical rules for generating word senses; the role of syntax, pragmatics, and formal semantics in the specification of lexical meaning; the use of a world model or ontology, as the organizing principle for lexical-semantic descriptions; the use of

rules with limited scope; the relation between static (context-independent) and dynamic resources; the commitment to descriptive coverage; the trade-off between generalization and idiosyncrasy; and, finally the adherence to the method-oriented or task oriented ideology of research. The discussion is inspired by, but not limited to, the comparison between the generative lexicon approach and the ontological semantic approach to lexical semantics.

Historically, generative has meant "a way of enumeration through mathematical derivation," and has been characterized as "world knowledge." The major points of the generative lexicon (as Nirenburg and Raskin referenced Pustejovsky) are that some of the lexical item meanings are characterized in relationships with other lexical items. These relationships result in fewer senses for a lexical item, and the missing senses can be derived with lexical rules operating on the identified relationships with other lexical items. Most common dictionaries contain enumerated lexicons and they have been criticized for lacking relations, inconsistent sense selection criteria, and being incomplete in usage. Enumerated lexicons can be as good as generative lexicons provided they use theory-based procedures and methods.

Other researchers believe that ontologies are an essential part of lexical semantics. Nirenburg and Raskin quote Pustejovsky [NR96], who said, "the meanings of words should somehow reflect the deeper conceptual structures in the cognitive system, and the domain it operates in". Furthermore, the authors state that "the notational elements that are treated as theory within the generative lexicon approach can, in fact, be legitimately considered as elements of semantic theory if they are anchored in a well-designed conceptual ontology. Then and only then, can one justify the postulation of a certain number of theoretical concepts, a certain set of roles and features, and a prescribed range of values." Other researchers characterize these ontological efforts as creating another natural-like language to manage the real natural language while asserting its necessity. The reply of Nirenburg and Raskin to that accusation is that they are building a "language-neutral" representation for some domain, which includes a semantic primitive repository with a semantic and "discourse-pragmatic" relational network between primitives. The ontology's function is to provide the context or world-knowledge for natural language processes. Unfortunately, ontologies are expensive to construct, maintain, and reproduce when they are hand-built. Automated methods should be used to help ensure consistency and the leveraging of intellectual capital for reuse in other domains. The ontology research area deserves attention because it is believed to be a foundation to enable knowledge sharing, software agents, and automated knowledge acquisition. For this reason, CIKS will track relevant organizations (e.g., AAAI) which hold conferences on ontologies in order to continue evaluating their usefulness.

The previous discussion was predominantly from the artificial intelligence and computational linguistic fields. Shifting to the terminology field is like moving from the research to the engineering perspective for doing near-term work. Terminologists can be found in the translation or technical language areas (e.g., ASTM), and therefore, appear to be more pragmatic. The terminology community focuses on the importance of concepts as indicated in several sources [WS95, SW93, and IH83]. In addition, the 1997 *Handbook of Terminology Management* [WB97] begins with the fundamental principles of terminology management addressing concept representation, description, and systems. This book provides a strong case for pursuing the representation and management of concepts as critical for addressing the lexicon, terms, or words used in natural language.

A.4 Document Models

Most print document models discuss the logical as opposed to the physical structure. The logical structure includes elements like chapters, sections, paragraphs, sentences, words, and characters. The physical structure primarily involves the placement of elements on the page. Most of the models are involved in converting the logical structure to the physical. Our interest is primarily in the logical structure, not the translation to the physical.

Furuta [AFQ89] identified element types, structures, constraints, relationship types, and homogeneity as distinguishing criteria in describing other document models. Element types refers to the kind of element (e.g., text, math formulas, tables, or graphs) and level of detail (e.g., paragraph, sentence, word, character, or some sub-part of a character) that they handle. Structure refers to a tree or acyclic graph, and within those structures, there may be constraints to specify the relationships between elements. In terms of relationship types, there may be predominant ones between major elements, and secondary relationships for references or footnotes elements. Finally, there are homogeneous or heterogeneous relationships between element types or within elements. Within this setting, Furuta identified a range of document models from the simple linear to the flat pseudo-hierarchical, as well as environmental with limited nesting, and more capable models such as unconstrained tree-like, constrained heterogeneous tree-like, and truly hierarchical.

Joloboff [AFQ89] discussed the evolution of the document representation models. The first generation involved improving the process of going from the revisable form that the author used to the final form that was printed. The second generation (which includes SGML) provided for richer typographical options, used logical and hierarchical structures, used relationships between elements, and extended element types (e.g., math formulas, tables, references, indexes, etc.). The third generation moved toward the direct manipulation of the elements and structure (WYSIWYG) without knowledge of a particular internal representation. Joloboff identifies Xerox's Interscript as a third generation model. A major goal for Interscript was to have the ability to exchange documents between different systems with different abilities, without needing to know the internal structures of the other system. The approach was similar to a programming language that is ported from one platform to another. Interscript had a script or program to reconstruct the internal representation on any platform. However, this made it more complex, and therefore, difficult to review by normal users. It was also designed so that the attributes of an element are not static, but are resolved dynamically within context as elements are moved around.

There were a number of studies integrating relational, semantic, functional, and object-oriented database technology with document models [AFQ89, Furu90], which at that time were in the study stage. There are document management systems that use commercial relational or object-relational databases, but the documents are normally contained separately. Other higher-end systems handling compound documents, are more powerful, and are similar to hyperdocuments, which will be discussed below.

There are a number of abstract conceptual models of hypertext systems, including: Hypertext Abstract Machine (HAM), Trellis, Dexter, Formal Model by B. Lange, and the Tower model according to De Bra [W^3-21]. Some parts of the HAM and Dexter models have been implemented; otherwise, no other implementation exists for these models. The Tower Model is an object-oriented, extensible data model for hyperdocuments, which was presented in 1992 at the ACM Conference on Hypertext. The model contains basic structural elements, nodes, links, anchors, and objects (tower, composite, and city). Tower objects are used to model the type, storage, structure, presentation, and semantic roles. The composites are composites of tower objects, while the city objects are customized or stylized views of the tower objects. This model allows objects to exist as a result of a function, allowing them to be virtual objects. In addition, browsing has been implemented using petri nets.

The W^3 Consortium is supporting work on a Document Object Model, which is only in the very beginning stages. As stated on their web site, it is "a platform and language neutral interface that will allow programs and scripts to dynamically access and update the content, structure, and style of the documents. The document can be further processed and the results of that processing can be incorporated back into the presented page" [W^3-22].

A.5 Knowledge Models

From a natural language processing perspective, the study of grammar or syntax is considered a knowledge representation. The way words are assembled into phrases and sentences is described by the syntactic approach and the syntax can help define classes of lexical meaning. However, they are usually "very crude, coarse-grained taxonomies of meanings in terms of preciously few features" [NR96]. Other researchers state "there is no way in which meaning can be completely divorced from the structure that carries it" [NR96].

The *Handbook of Artificial Intelligence* [BF81] identified knowledge representation schemes as those including logic, procedural representation, semantic networks, production systems, direct (analogical) representations, semantic primitives, frames and scripts. They appear to be divided into two classes: logic, semantic networks, and direct or analogical representation seem to be fundamental; while procedural, production, and frame-based representations emerge to be based on one of those forms with enhancements or modifications.

Davis, Shrobe, and Szolovits [DSS93] argue that knowledge representation can be understood by the roles it plays. The first role, as a surrogate, is defined as a substitute for the reasoner. Complete models of the natural world cannot be built, something is always ignored, and that something may cause undesired consequences.

The second role, as a set of ontological commitments, could be translated to what is relevant from a domain expert perspective. For example, a materials engineer would model and acknowledge different objects then a material scientist or researcher would. What is relevant, is affected by the selected representation, including what, where, and how you put knowledge into that representation.

The third role, as a fragmentary theory of intelligent reasoning, is based upon the premise that any model is only a partial representation (nothing is complete) and one must understand which characteristics of intelligent reasoning are acceptable and recommended. Davis et al. [DSS93] identified logic, psychology, biology, statistics, and economics as the fields that provided insight about intelligent reasoning. Typical examples of respective approaches include first order predicate logic, frames and production rules, neural networks, uncertainty and fuzzy logic, and rational agents from utility theory. Inferences, which are appropriate, based on the available information are considered acceptable. However, when there are too many acceptable inferences, only those recommended would be the intelligent ones. The authors state, "Representation and reasoning are inextricable and usefully intertwined: A knowledge representation is a theory of intelligent reasoning." They also mention that logic is silent on recommended inferences, motivated by the desire to work on all problems in the same way. "Preventing the representation from selecting inferences and, hence, requiring the user to do so offers the opportunity for this information to be represented explicitly rather than embedded implicitly in the machinery of the representation." However, the opposing view is that the user is responsible for recommending a set of inferences explicitly in the logical language, which places a burden on the user.

The fourth role discusses the trade-off between computational efficiency versus representational richness, which is irrelevant from a representational perspective where the goal is representation over efficient computation.

Finally, the fifth role, as a medium of human communication, refers to the distinction between machine interoperability and human communication. Does this representation make it easy for machines to interoperate, but difficult for human discussion? Alternatively, is the opposite true, making it easy to communicate between humans, but difficult for machines?

In a soon to be published book entitled *Intelligent Systems for Engineering: A Knowledge-Based Approach* [Srir95], model-based representation was discussed in the chapter on knowledge representation. "The models cover what components make up the system, what is the function of the system, how the components are connected, the expected behavior of the individual components..."[Srir95]. This model-based representation can be used for engineering simulation models, with deeper representation based on fundamental theories.

Appendix B: CIKS Coating Partnership for Highway Structures

B.1 Overview of the Partnership

Figure B-1 shows a diagram of a proposed partnership to develop a unified decision-support system for highway bridge coatings using the CIKS framework. Each of the organizations identified in the diagram is currently active in developing information, standards, or computerized systems that address the needs of highway engineers and contractors. With regard to computerization of information, however, there is no unified effort to leverage each organizations work. Recognizing that each organization has its own area of special interest and objectives, it is still possible that this can be achieved and provide consistent methods to enhance decision-making for highway personnel.

One of the main objectives of the partnership would be to adopt standard terms, data formats, and information models so that individual computerized systems or modules would complement each other. Terms and data would be consistent among modules. An example is the coating product data format being proposed as an SSPC guide. This effort should be extended to include additional parameters for representing coating performance properties.

Information models could be useful in the development of modules, initially through the partnership, and ultimately by the entire coating industry (e.g., SSPC standards, and guides). An information model would contain references to standard terms and formats, criteria for data quality, and guidance to developers. It would provide a systematic approach for decision-making. For example, it would outline essential steps and information requirements for a coating management system for highway bridges. Steps could include identification of coating failure, diagnosis and remedial action, coating system selection, maintenance consideration and procedures, and assistance in developing contract specifications.

An engineer would view the system as a single system, comprised of modules (e.g., coating selection, failure analysis, and maintenance) that provide information, guidance, and references. Standards of practice, such as surface preparation, inspection procedures, data formats, and terms would be referenced and obtained from SSPC. Accepted field guides to practice could also be used. Interoperability among the modules could be by user request or automatic. For example, to substantiate a system recommendation involving low volatile organic chemical coatings, the "system" may reference work performed by FHWA that is contained in a database maintained at the FHWA site. Or the system could draw upon an archive (photographs, video) of coating failures that provides the user with detailed information on causes, exposure environments, and remedial actions to be taken based on individual case-specific needs. Emphasis would be placed on information in digital formats and the Internet would be the dominant vehicle for access to the system. Considerations would be made for information of a commercial and proprietary nature.

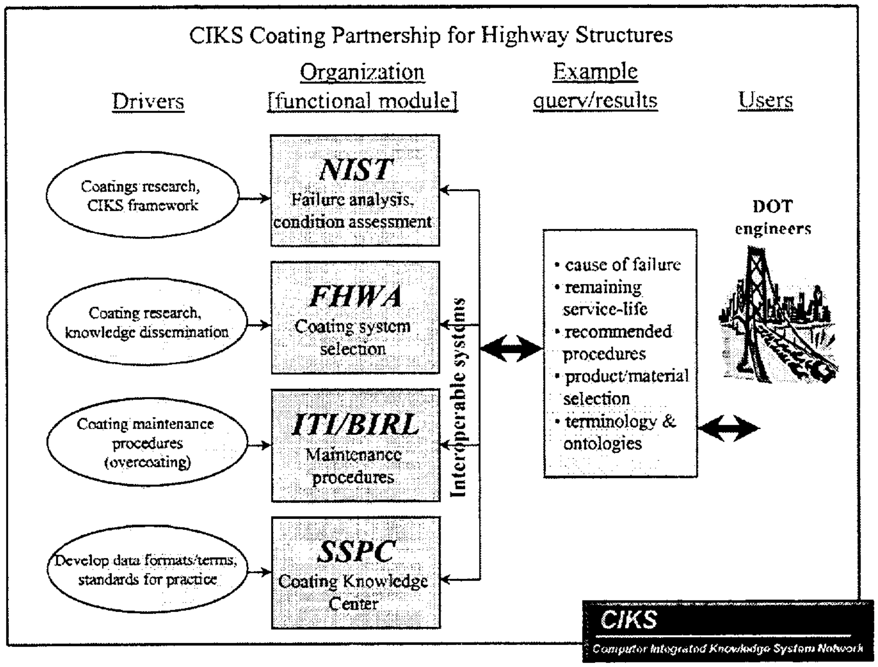

Figure B-1: CIKS Coating Partnership for Highway Structures

B.2 Examples of Coating Industry Knowledge-Based Systems

Computerized Knowledge-Based Systems (KBS) are available within the coating industry. Examples of systems developed for construction industry application are identified in a selected bibliography later in this paper. Many of these systems are prototypes and were developed when information technologies (e.g., expert system shells, neural networks) were in their infancy. Features lacking in these systems include the ability to integrate with the business process and their ability to operate with computerized systems across different disciplines (e.g., designers, maintenance and operation staff) and processes (e.g., material and product selection, and cost analysis).

Many KBS are developed for in-house use. Examples include databases on coating products and services, research data, cost estimating, and decision-support systems. Many of these systems are rarely distributed to others, in formal or commercial formats. The reasons include their proprietary nature, lack of an incentive for distribution (marketability), and costs related to development and maintenance. In fact, many systems are developed and maintained by personnel responsible for developing the knowledge and their existence may even go unrecognized by colleagues.

Computerized systems designed for distribution, such as commercial products, require more effort to develop and maintain. These tend to be more formal systems where marketing strategies, distribution mechanisms, and customer (user) assistance is essential for a successful product and to return the development investment. Often knowledge content is the driving force for these systems. The Journal of Protective Coatings & Linings, "JPCL Archives II" is an example of such a product. Future versions of this product could be made available on-line via the Internet, while still providing income to the publisher/distributor. Mechanisms currently being developed to allow electronic commerce and intellectual property distribution on the W^3 will accelerate this method.

The Internet W^3 has altered thinking on how an organization's knowledge is maintained and distributed. For example, using an Intranet (an internal W^3), access to knowledge can be provided to personnel within an organization (e.g., marketing, testing, and research departments). Use of an Intranet results in benefits such as shorter development times, lower costs, and better data security. Although Intranets are useful within organizations the maturation of new capabilities such as electronic commerce and increased network traffic capacity will enable the Internet to support new relationships among different organizations (e.g., agile virtual enterprises). To remain competitive, Internet-based information is no longer a convenience, but a necessity.

Perhaps the most significant constraint affecting the distribution of organizational and commercial knowledge-based systems is the lack of interoperability. Quite simply, interoperability means the ability to use knowledge and software among computerized systems. Incompatible data formats and computer hardware and software, incomplete or subjective data, inconsistent terminology, and the lack of electronic access are examples of specific factors that prohibit widespread use. Solving these problems will require collaboration by industry and government. Agreement must be reached on a common terminology, standard knowledge formats, criteria for establishing data quality, and common computer interfaces that provide seamless integration of knowledge to the user. The CIKS activity will address these issues.

B.3 Steel Structures Painting Council/Computer-Integrated Knowledge System (SSPC/CIKS) Joint Coating Working Group

One group that is currently addressing coating industry knowledge issues is the SSPC/CIKS Joint Coating Working Group. The Working Group comprises members from the SSPC Committee C.4.10 on "Knowledge-Base Systems for Coatings," and a CIKS Coating Industry Working Group formed during the CIKS June 1996 workshop. SSPC's Executive Director, Bernard Appleman, chairs the group. Members of the group include public and private sector organizations representing coating formulators, consultants and engineers, facility owners, and researchers. Figure B-2 shows a diagram of the group's interaction with the CIKS test bed and its role in the development of CIKS.

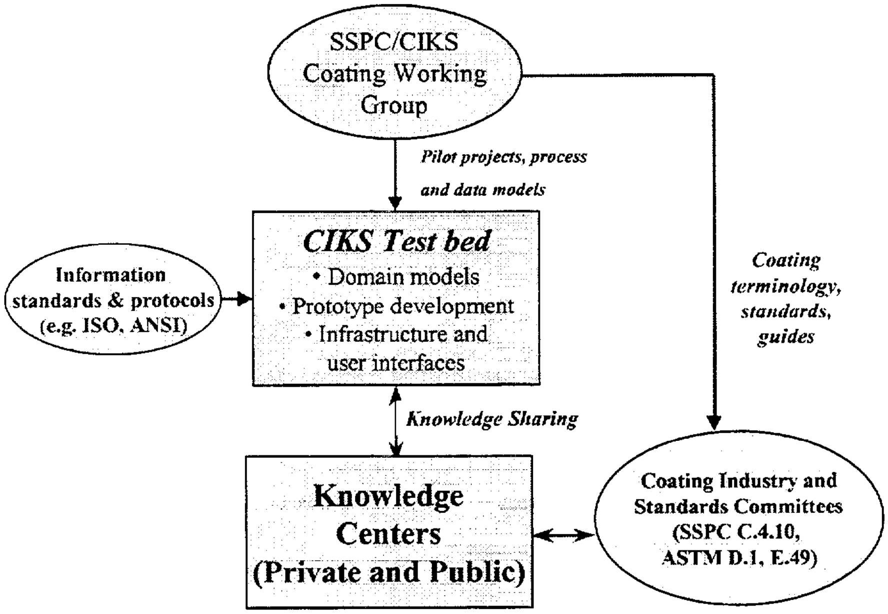

Figure B-2: SSPC/CIKS Coating Working Group role in CIKS development.

Collaboration with other industry and standards setting committees such as ASTM Committees D01 on "Paint and Related Coatings, Materials and Applications, " E02 on "Terminology," and E49 on "Computerization of Material and Chemical Property Data" will be necessary. These collaborations will result in consistent terminology and standards for identifying, representing, and sharing coating material knowledge.

Specific focus areas of the SSPC/CIKS Joint Coating Working Group include:

- developing guides to assist coating industry knowledge users in using and integrating the Internet in business practice;
- improving the communication of computer-based information such as messages, computer-stored files, and access to application systems;
- developing standard formats for representing and exchanging coating product data;
- developing state-of-the-art reports describing current practice and enabling information technologies that have application within the coating industry.

Future projects requiring longer lead-times (2-years) to implement include the development of case-based reasoning (decision-making based on documented observations of coating performance), data dictionaries, and expert systems. Products from the Working Group will be disseminated in the form of SSPC Technology Updates, Guides, and through the SSPC Coating Knowledge Center. The draft "Guide for the Identification and Use of Industrial Coating Material in Computerized Product Databases" exemplifies the Working Group's effort to develop standard guides and methods for representing coating material knowledge. Table B-1 shows examples of the data elements proposed in the guide.

Data segment	Example data elements
Product description	Product name, product identification, generic type, system component
Intended use	Common application, substrate type, exposure environment, compatible undercoat and topcoat
Physical properties	Volume solids, solids by weight, mixed density, test methods
Mixing and application	Minimum and maximum dry film application thickness, theoretical coverage per volume, dew point, induction time, pot life
Key performance parameters	Corrosion resistance, weathering, abrasion resistance, test methods
Safety	NFPA health hazard, flammability
Manufacturer supplemental information	Manufacturer comments

Table B-1: Example of product data segments and data elements.

The draft document is being proposed as a SSPC Guide and would be used by coating manufactures, specifiers, and users (facility owners) for communicating coating product data. Figure B-3 is a diagram of the use of coating product data using the W^3. Product data sheets are now used to communicate this information. However, variations exist among manufacturers when describing product data content, such as terms used and the type of data reported. As more companies use computerized systems such as the W^3 to disseminate information, standards such as the proposed coating product data guide will provide the mechanism for improved understanding (through consistent terminology and data elements). This will enable integration of product data among diverse computerized systems, such as company and facility owner project databases. Another benefit of the guide will be an increased understanding of coating material performance and data quality through the reporting of data elements such as those described in the "Key Performance Parameters" data segment. These parameters will be substantiated through the identification of test methods used to develop parameter values.

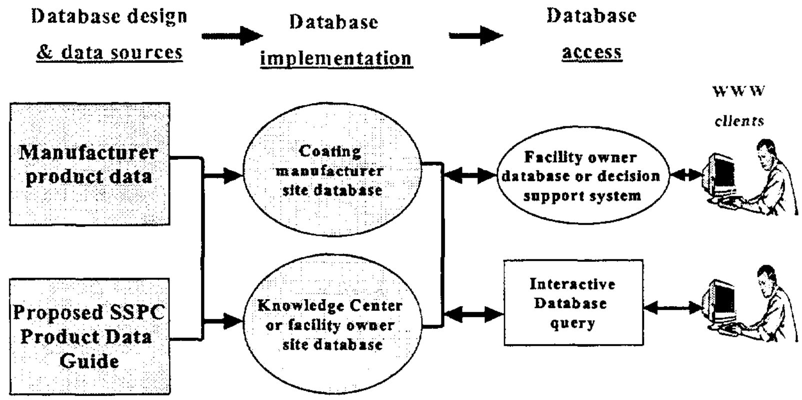

Figure B-3: Diagram showing the generation and use of coating product data.

B.4 Current NIST Effort to Establish the CIKS Test Bed

NIST is seeking to improve the delivery of its research results. The CIKS test bed will be a useful tool in providing construction material knowledge. The AMRL Paint Proficiency Sample Program is an example of the use of the test bed to provide access to technical data. Technical databases and decision-support systems are methods that have received the greatest attention thus far. Material property databases for cement and coatings can now be accessed through the W^3.

Two decision-support systems have been developed by NIST during the past several years. The first of the two systems is the Highway Concrete Expert System (HWYCON). This system is designed to assist highway engineers in the diagnosis, selection, and repair and rehabilitation of highway concrete structures. It includes knowledge related to concrete pavements, bridges, and support structures. Several universities as part of their material science and civil engineering curriculum are also using the system. The Transportation Research Board sells HWYCON. The computerized system requires the Microsoft operating system, Windows version 3, or Windows '95. The distribution package contains a set of floppy diskettes and a report describing the design, installation, and operation of the system [KCS94, KCKS93]. The second system called the Coating Expert Advisory System-I (COEX-I) [BFRL96] contains coating material knowledge and is designed to assist in the analysis of coating failures and selection of coating systems for stationary military structures. An overview of the COEX-I is described later.

B.5 Technical Databases

Technical databases contain many different types of data (see Figure B-4). Examples include: product databases that describe material properties and manufacturer data, laboratory performance measurements, and outdoor exposure test results. It was stated earlier that significant differences occur among databases due to designer/developer preferences. These differences take the form of: inconsistent field names and contents; choice of computer hardware; and software that does not allow interoperability. The proposed SSPC Guide on coating product data formats is only the first step in providing compatible databases that can be used among computerized systems. To realize the full benefit of distributed database exchange, standard methods must be used to implement and disseminate the data. The steps in developing interoperable distributed databases include:

- establishing consistent database formats and terminology;
- establishing a logical schema to represent the physical data in the database;
- implementing the database (acquire, computerize) using a database management system;
- providing electronic access through media distribution or electronically, via the Internet W^3.

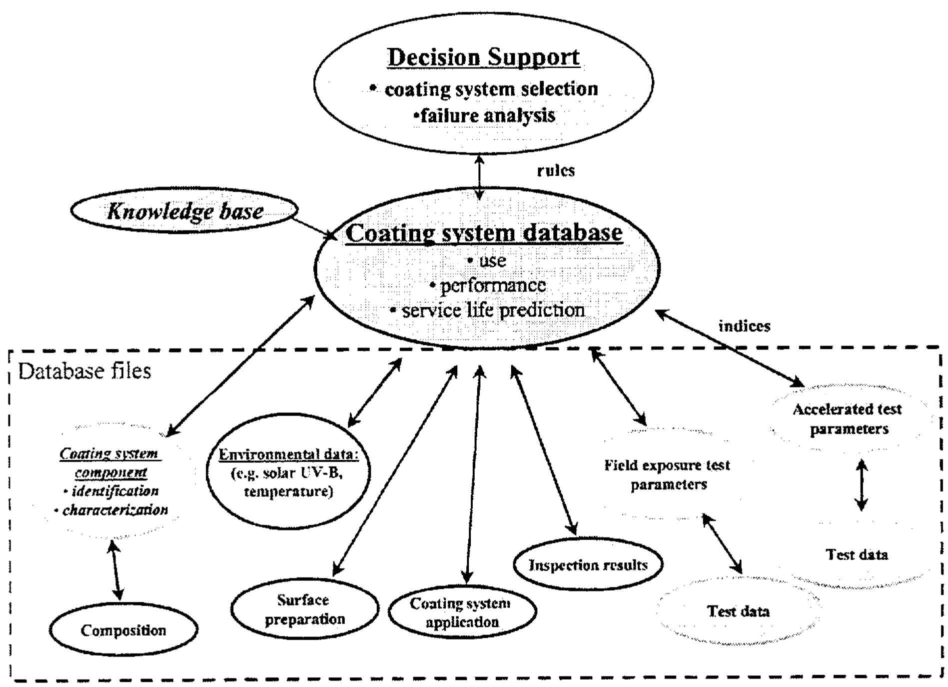

Figure B-4: Examples of technical coating databases.

Significant operational enhancements can be achieved through electronic publication of databases using the W^3. Interfacing database standards such as SQL (Structured Query Language) and the W^3 client interface (e.g., WEB browser) reduces the need to develop multiple user interfaces for different computer platforms and can significantly reduce software development and maintenance costs. Access to databases can be provided in a more timely and convenient manner. Figure B-5 contains a diagram of the components of a distributed technical database system designed for W^3 access. This model can be applied to virtually any type of database. The client (user) is provided the functionality of submitting queries (questions) to the database in an interactive mode. Typically, this is done using a W^3 browser program. The process of converting input from the client involves converting the query into a SQL statement. In the instance shown in Figure B-6, this is accomplished using the ISO Standard 9579 [BS96], "Remote Data Access" (RDA). This standard is a generic model providing database access and has been implemented at NIST for interfacing W^3 clients to SQL databases. The RDA standard was implemented using the C Programming Language. After receiving the SQL statement, the database management system retrieves the data from the physical database and produces a table containing the data elements (fields) and their values. This information is returned to the RDA component that formats the information for output in the Hyptertext Markup Language (HTML) and displays the information using the W^3 client's browser. Added capabilities have been developed to also produce graphical plots. An interface has been developed to the NIST Dataplot Statistical and Graphical Analysis system [Fill84].

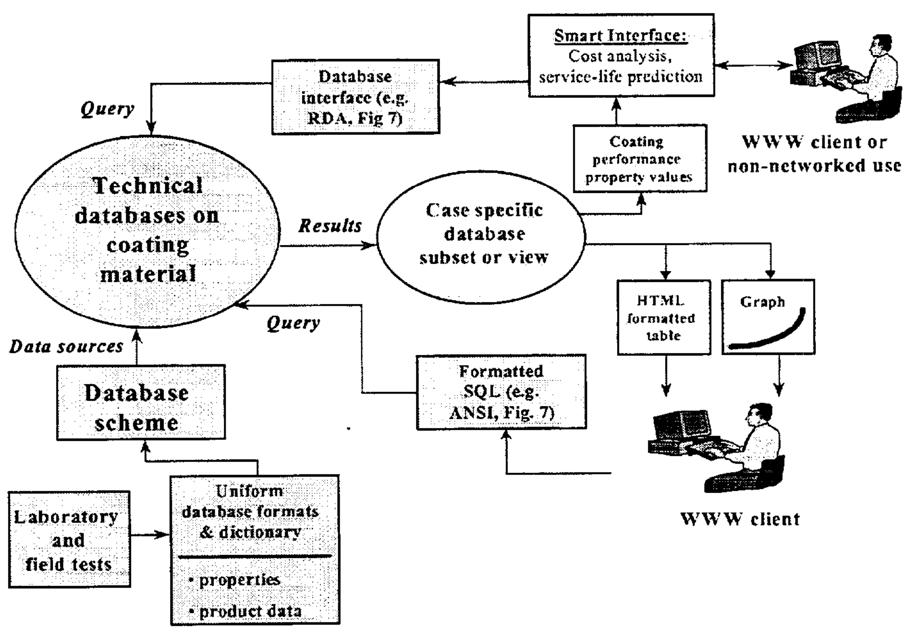

Figure B-5: W^3 implementation for technical databases.

This permits the graphical display of database information, interactively. Since the construction industry is comprised of companies with varying degrees of personnel and funding resources. It is necessary to test knowledge system development using multiple platforms. Figure B-6 shows a method can be used in the CIKS test bed to implement distributed technical databases using ISO, ANSI and de-facto standards for database storage, query, and retrieval. The use of newer, more flexible de-facto standards in the form of Microsoft Internet Information Server and the "Access" or "SQL Server" database management system, provides an opportunity for less costly hardware and software resources.

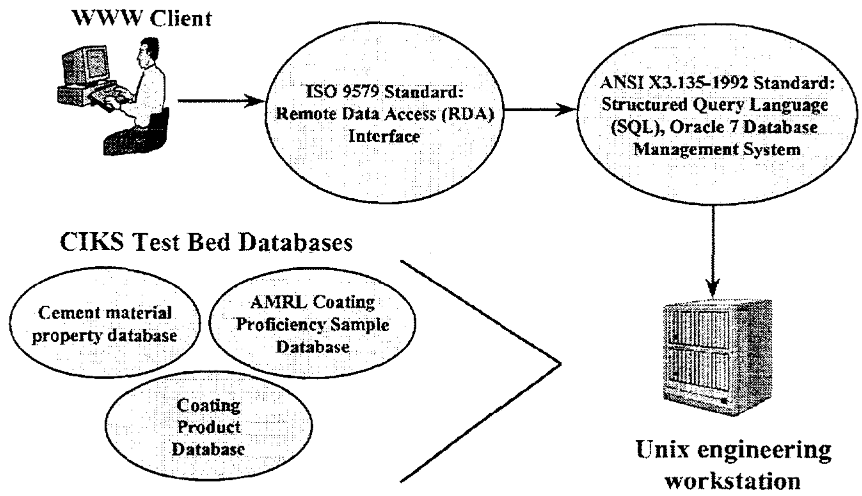

Figure B-6: W^3 database implementation, using an engineering workstation platform, ISO, and ANSI standards.

B.6 Decision-Support Systems

For the purpose of this article, coating decision-support systems (or expert systems) are defined as, computerized systems that can contain virtually any type of coating data and knowledge, such as technical databases, photographs, expert guidance, video and sound, computer-based models, and a logic module to operate (direct the logical instruction sequence) on the knowledge and provide an interface to the user.

Many attempts to develop decision-support systems have occurred in technical areas during the past 15 to 20 years. Examples of decision-support systems developed for construction materials users can be found in the bibliography at the end of this article. There are relatively few commercial decision-support systems available today. The most successful are operational within organizations that are committed to develop, maintain, and operate them within the organization. Historically, complex systems that cover a wide-interest area and involve high-level expert knowledge are costly to develop and maintain. However, advances in decision-support system development tools during the mid-1980's provided more cost-effective development tools. These tools use the object-oriented programming [Hend93] architecture and advanced techniques for the representation and use of knowledge, such as video, sound, and hypertext links. The result is a significant reduction in development time. Additional benefits in using object-oriented development tools include:

- ability to reuse data, knowledge, and procedures;
- relationships can be established among data and knowledge (inheritance);
- efficient graphical user interfaces are included in the tool;
- improved interfaces to external knowledge and programming modules.

B.7 COEX-I: An Object-Oriented Decision-Support System

One benefit of implementing decision-support systems is their ability to provide a systematic approach to problem-solving and knowledge dissemination. Incorporating the knowledge of experts in coating materials and practices can improve levels of decision-making. For example, experts residing in a central location can extend or replicate their knowledge to field staff who need to evaluate the condition of coated structures, and perform repair and maintenance duties. Guidance for these individuals is typically found in manuals, guides, and standards. Computerizing the knowledge to include photographs, sound and video, and guidance on the use of the guide significantly enhances knowledge understanding, resulting in cost savings, and improved facility performance.

The COEX-I expert system was developed by a team of coating experts who had previously written the Military Handbook, "Handbook for Paints and Protective Coatings for Facilities" [DOD95]. The group includes representatives from Department of Defense facilities who are involved in maintaining coated facilities. The group decided to develop a prototype decision-support system to assist military staff in analyzing coating failures that occur on stationary military structures such as water towers, buildings, and bridges. Section 11 of the guide covers the Analysis of Paint Failures and includes a decision tree that is designed to assist the user of the handbook. From this decision tree, rules (logic) were computerized in the form of questions-and-answers. The rules provide a hierarchical structure to the knowledge and guide the user in problem solving, from the identification of visual observations found on the structure to recommendations given by the system. Recommendations include the identification of the coating failure, its cause, and guidance on remedial action(s). An additional capability was added to allow the user to specify criteria for coating system selection where total replacement is necessary for structural steel that shows blistering to the substrate. Figure B-7 shows a diagram of the COEX-I system.

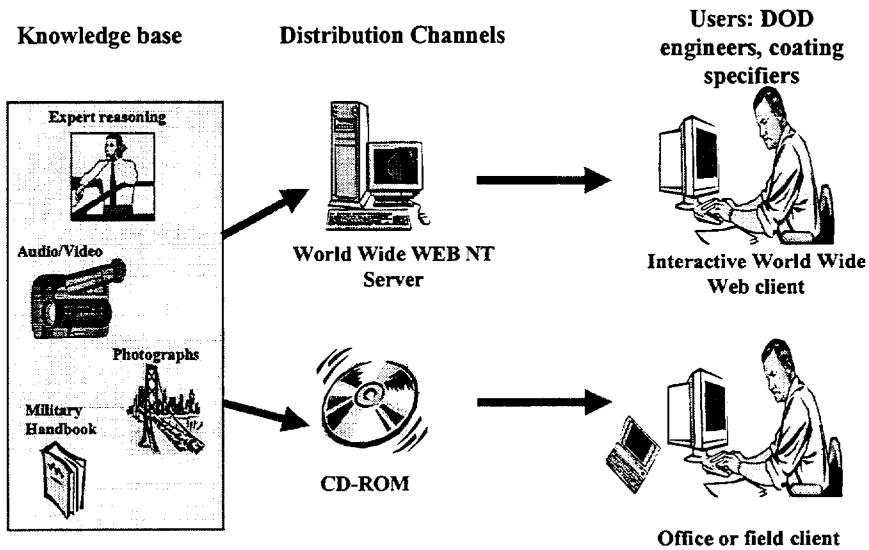

Figure B-7: COEX-I system diagram.

COEX-I is a prototype decision-support system that is being distributed for review and comment. Although the system is based on military structures, parts of the knowledge base apply to structural steel that is present in highway bridges. Examples include corrosion failures and videos contained in the system that provide guidance and inspection procedures for blistering and corrosion causes. Tools used in the development of COEX-I are being applied to a new system designed to assist Federal Highway Administration (FHWA) and State Department of Transportation Engineers in the selection of coating systems for highway steel bridges. Knowledge contained in the system was developed through various FHWA projects during the past decade. FHWA and NIST are developing the system jointly under FHWA sponsorship. It will be operational in late 1997.

Selected Bibliography of Knowledge-Based Systems

H. Adeli (editor), *Expert Systems in Construction and Structural Engineering,* Chapman and Hall, New York, NY (1988).

C. Arnold, R. Drisko, J. Griffith, J. Neal, T. Nguyen, and J. Yanez, *An Expert System Application for Paints and Coatings: Painting Advisor,* Proceedings of SSPC '87, SSPC 87-07, pp. 1-6, Steel Structures Painting Council, Pittsburgh, PA (1987).

S.K. Boocock and L.J. Kaetzel, *Coating Industry Knowledge Base Systems: An Introduction to the SSPC Knowledge Center and the NIST Computer Integrated Knowledge Systems Network,* Proceedings of SSPC '96, pp. 144-149, Steel Structures Painting Council, Pittsburgh, PA (November 1996).

J.R. Clifton and L.J. Kaetzel, *Expert Systems in Concrete Construction,* Concrete International, American Concrete Institute, Farmington Hills, MI (November 1988).

L.F. Cohn and R.A. Harris, *Knowledge Based Expert Systems in Transportation: A Synthesis of Highway Practice,* National Cooperative Highway Research Program Synthesis 183, Transportation Research Board, Washington, DC (September 1992).

Federal Highway Administration, *Developing Expert Systems,* Publication No. FHWA-TS-88-022, U.S. Department of Transportation, McLean, VA (December 1988).

L.J. Kaetzel and J.R. Clifton, *Expert/Knowledge-Based Systems for Cement and Concrete: State-of-the-Art Report,* Strategic Highway Research Program, SHRP-C/UWP-91-527, Transportation Research Board, Washington, DC (1991).

L.J. Kaetzel, J.R. Clifton, P. Kleiger, and K. Snyder, *Highway Concrete (HWYCON) Expert System User Reference and Enhancement Guide,* NIST Internal Report 5194, National Institute of Standards and Technology, Gaithersburg, MD (May 1993).

L.J. Kaetzel, J.R. Clifton, K. Snyder, and P. Kleiger, *Users Guide to the Highway Concrete (HWYCON) Expert Systems,* Strategic Highway Research Program, SHRP C-406, Transportation Research Board, Washington, DC (1994).

L.J. Kaetzel and J.R. Clifton, *Expert Systems for Building Materials and Structures: Expert/Knowledge Based Systems for Materials in the Construction Industry,* RILEM Journal of Materials and Structures, Volume 28, pp. 160-174, RILEM (1995).

L.J. Kaetzel and M.E. McKnight, *Enhancing Coatings Diagnostics, Selection, and Use Through Computer Based Knowledge Systems,* Proceedings of SSPC '95, SSPC 95-09, pp. 287-295, Steel Structures Painting Council, Pittsburgh, PA (November 1995).

S.S. Kuo, D.A. Clark, and R. Kerr, *Complete Package for Computer Automated Bridge Inspection Process*, Transportation Research Board, Washington, DC (1988).

O.S. Marshall, Jr., *Development of the PAINTER Engineered Management System,* U.S. Army Corp's of Engineers, Construction Engineering Research Laboratory, USACERL Special Report FM-94, Construction Engineering Research Laboratory, Champaign, IL (January 1994).

J.H. Pielert and L.J. Kaetzel, *Cement and Concrete Materials Databases and the Need for Quality Testing,* Materials Science of Concrete III, pp. 337-358, American Ceramic Society, Westerville, OH (1992).

R. Sharpe, B.S. Marksjo, F. Ho, and J.D. Holmes, *WINDLOADER: Wind Loads on Structures Advisor,* Building Construction Engineering, SP-012, CSIRO Division, Australia (July 1989).

U.S. Army Corp's of Engineers, Construction Engineering Research Laboratory, *Knowledge Acquisition for Expert Systems in Construction,* USACERL Special Report M-89/10, Construction Engineering Research Laboratory, Champaign, IL (September 1988).

Appendix C: Glossary

Abbreviation, word or phrase	Description
AAAI	American Association of Artificial Intelligence
ASTM	American Society of Testing and Materials
CASE	Computer Aided Software Engineering
CD	Committee Document
CIKS	Computer Integrated Knowledge System
COTS	Commercial Off The Shelf
DAMA	Data Administration Management Association
DARPA	Defense Advanced Research Projects Agency
Data or datum	Something given or admitted; a fact or principle granted; that upon which an inference or an argument is based (in our loose sense used to describe what can be stored in a typical commercial database)
DIS	Draft International Standard
DTD	Document Type Definition
ERD	Entity Relationship Diagram
FHWA	Federal Highway Administration
HAM	Hypertext Abstract Machine
HTML	Hypertext Markup Language
HTTP	Hypertext Transport Protocol
IEEE	Institute of Electrical and Electronics Engineers
Inference	The process of deriving from statements (affirming or denying something that can be characterized as true or false) either the strict logical conclusion or one that is to some degree probable;
Interlingua	An artificial language for international communication intended mainly as a common language for scientists.
ISO	International Organization for Standardization
KIF	Knowledge Interchange Format
Knowledge	The act or state of knowing; understanding; being intelligent; (in our loose sense used to include the ability to reason or inference)
Knowledge-base	A repository of knowledge
Lexical	Of or pertaining to a lexicon
Lexicon	the vocabulary of a particular language or field arranged alphabetically
NASA	National Aeronautics and Space Administration
NIAM	Nijssen's Information Analysis Method
NSF	National Science Foundation
ODA	Office Document Architecture
ODL	Object Definition Language
ODMA	Open Document Management API
ODMG	Object Database Management Group

Abbreviation, word or phrase	Description
OQL	Object Query Language
OMT	Object Modeling Technique
Ontology	A model of concepts; relational model of the lexicon;
Semantics	The study of language meaning; an approach for assigning meanings to symbols and expressions;
SGML	Standard Generalized Markup Language
SQL	Structured Query Language
STEP	Standard for the Exchange of Product model data
TCP/IP	Transmission Control Protocol/Internet Protocol
Terminology	The terms actually used in any business, art, science, or the like
Terms	A word or group of words designating something, especially in a particular field
TIF	Terminology Interchange Format
UMLS	Unified Medical Language System
W^3	World Wide Web
Web	World Wide Web
XML	Extensible Markup Language