

# The Environmental Information Mall

Michael N. Huhns, Munindar P. Singh, and Gregory E. Pitts

MCC

3500 West Balcones Center Drive

Austin, TX 78759-5398

U.S.A.

{huhns,msingh,pitts}@mcc.com

## Abstract

The *Environmental Information Mall* is being developed for the dissemination of information about the environment, including information derived from the application of analytical tools to environmental data. This paper describes an implementation plan to create an environmental information forum situated on the National Information Infrastructure (NII). The success of this forum depends on the development of an ontology of environmental concepts, which will be used to enable the interoperation of data and analytical tools from a variety of independent sources. We describe tools for the creation and maintenance of the ontology, and then show how it can be used: by information sources to advertise their capabilities, by mediators to combine analytical tools with the data on which they operate, by end users to locate information, and by interfaces to fuse information from several sources. Development of the *Environmental Information Mall* has just begun, so there is as yet no implementation or results to describe.

## 1 The Vision

Environmental awareness is in everyone's best interest. Businesses and consumers all benefit when products are manufactured, used, and ultimately disposed of in ways that are safe and beneficial for the environment. A key to achieving this awareness is to make environmental information widely available. However, much of the available environmental information is not reaching either the large or small business communities.

More than 15 federal organizations, numerous state and local agencies, and many private corporations are now actively engaged in producing or gathering environmental data,

spanning efforts from "design for the environment" (DFE) to remediation. Currently, the data is impossible to monitor, let alone utilize. It is located at hundreds of distributed data sites and resides on numerous operating systems and configurations. Even if the data could be inventoried, wildly varying differences in quality, vocabularies, and operating platforms make it incredibly difficult for individuals and organizations to access and use it. Therefore, individuals and organizations cannot effectively take advantage of this information to minimize compliance costs, design for the environment, or reduce the impact of production processes. Also, without a high-level view of what information is being used and not used, they will frequently make shortsighted, reactive plans for future research.

A comprehensive project to couple the development and collection of environmental data with the dissemination of the information to the actual user groups could change all this. MCC has proposed that an *Environmental Information Mall* be established to make available the information needed for intelligent environmental decisions. The availability of myriad data and analysis tools would enable more sophisticated modeling and DFE processes, thereby reducing lifecycle environmental costs, and would facilitate industry and government access to resources and activities in the environmental area.

The *Environmental Information Mall* would provide an easily accessible source for environmental regulations, for analysis and solution of environmental and remediation problems, and for commerce in environmental technology and services. It will provide on-line, active manuals promoting environmental technologies. The resulting information will enable businesses and individuals to make intelligent environmental decisions. It would provide a means for the EPA to guide its own decision making and to use its resources in helping industry comply with environmental regulations. It would also provide a secure means to track, collate, analyze, and disseminate pollution prevention and remediation information.

## 2 Realizing the Vision

The establishment of an *Environmental Information Mall* is timely, because most of the needed technology components are now in place:

- **Communication Infrastructure**—fostered by the National Information Infrastructure program and implemented by telecommunication companies
- **Environmental Data**—collected by many government agencies, commercial organizations, and consortia
- **Environmental Analysis and Simulation Tools**—developed by many commercial organizations for executing on the above data.

These technology components are necessary for successful establishment of an *Environmental Information Mall*, but they are not sufficient. A remaining challenge is integration of the many repositories of data and varieties of analysis tools, which are mostly heterogeneous and independently maintained. These sources of information will be useful only if they can be easily accessed and readily applied. MCC, in concert with government and commercial sponsors, is creating an information forum situated on publicly available communication facilities, such as the Internet. The project will define a dynamic, distributed “information space” where disparate data, currently kept in myriad formats, can be placed and accessed logically. The information space must be sophisticated enough that information technology tools can be run against the data to make the information usable (both in form and in content). Finally, the operating parameters should be such that the maintenance and structure become

To assist in this, MCC provides the necessary interoperation tools to:

- Incorporate new databases and other information resources into an open information environment
- Develop and maintain environmental domain ontologies
- Browse domain ontologies and available information resources
- Formulate queries, in terms of the domain ontologies, for accessing the resources.

However, this technology must be adapted to the National Information Infrastructure, and then applied to existing resources. The result will be environmental information that is not only accessible, but also usable.

**Application Domain.** The Federal Government is aggressively seeking to assist small- and medium-sized businesses through a variety of mechanisms. Within DoD, the Manufacturing Centers of Excellence and the CALS Shared Resource Centers are attempting to provide information via the emerging electronic highway. Similarly, NASA has established Technology Outreach Centers, and the Department of Commerce, through the National Institute of Standards and Technology (NIST), is establishing Manufacturing Technology Centers (MTCs) and Manufacturing Outreach Centers (MOCs).

A key infrastructure component to all these technologies is an electronic link between sources of information and users of information. Once the linkage is established, however, providing or finding the relevant information in a timely

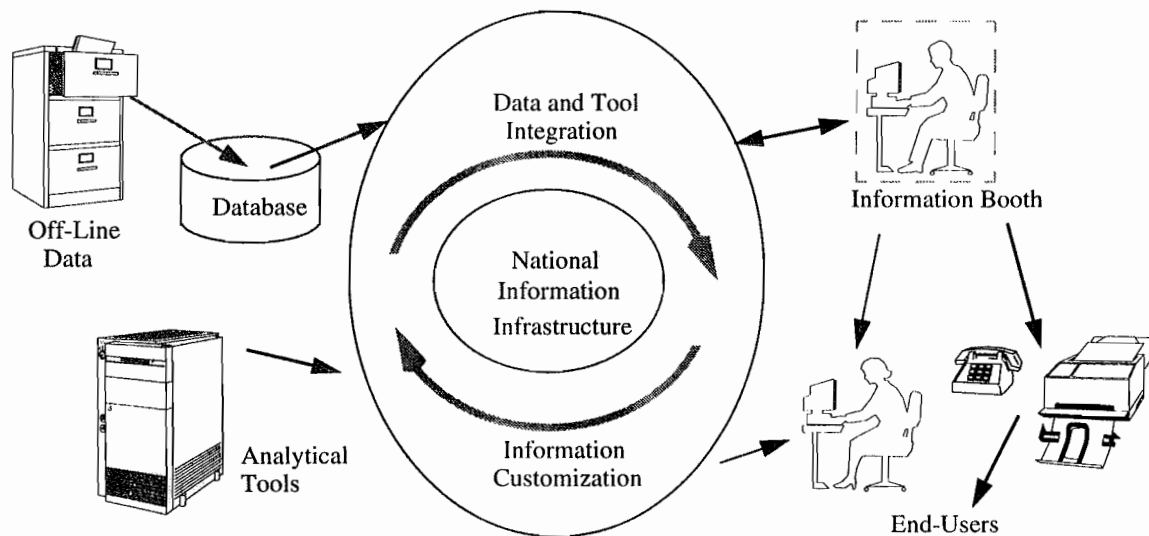
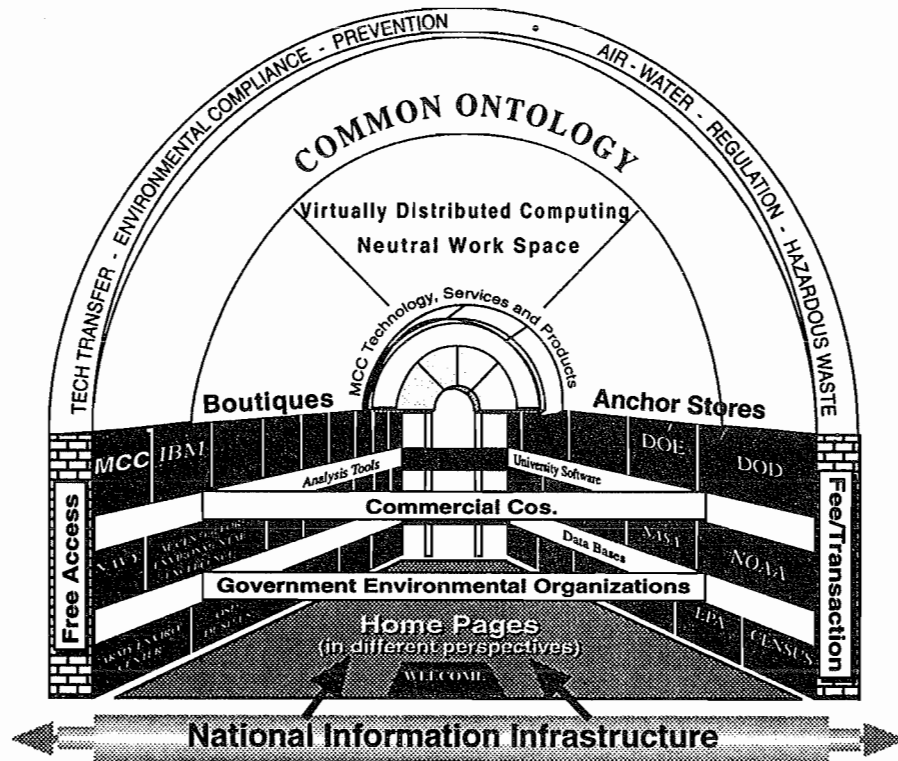


Figure 1. Conceptual View of the Relationship of Users to Environmental Data and Analysis Tools



fashion will be critical to the success of the program. The use of intelligent mediated information agents is needed by such a network. The most important type of information to be exchanged concerns environmental regulations, compliance, and remediation.

### 3 Technical Challenges

Existing NII technology for accessing information from open information systems is insufficient because it relies on hard-coded links, does not mitigate heterogeneity, incorporates no semantics, and provides no fusion of results. In developing the current technology, the emphasis has been on finding text, not data or databases. Additional challenges involve the

- Amount of Data (finding needles in haystacks)
- Quality of Data
- Using the Data
  - Mediation to match tools to databases [Wiederhold 1992]
  - Autonomous application of tools against databases.

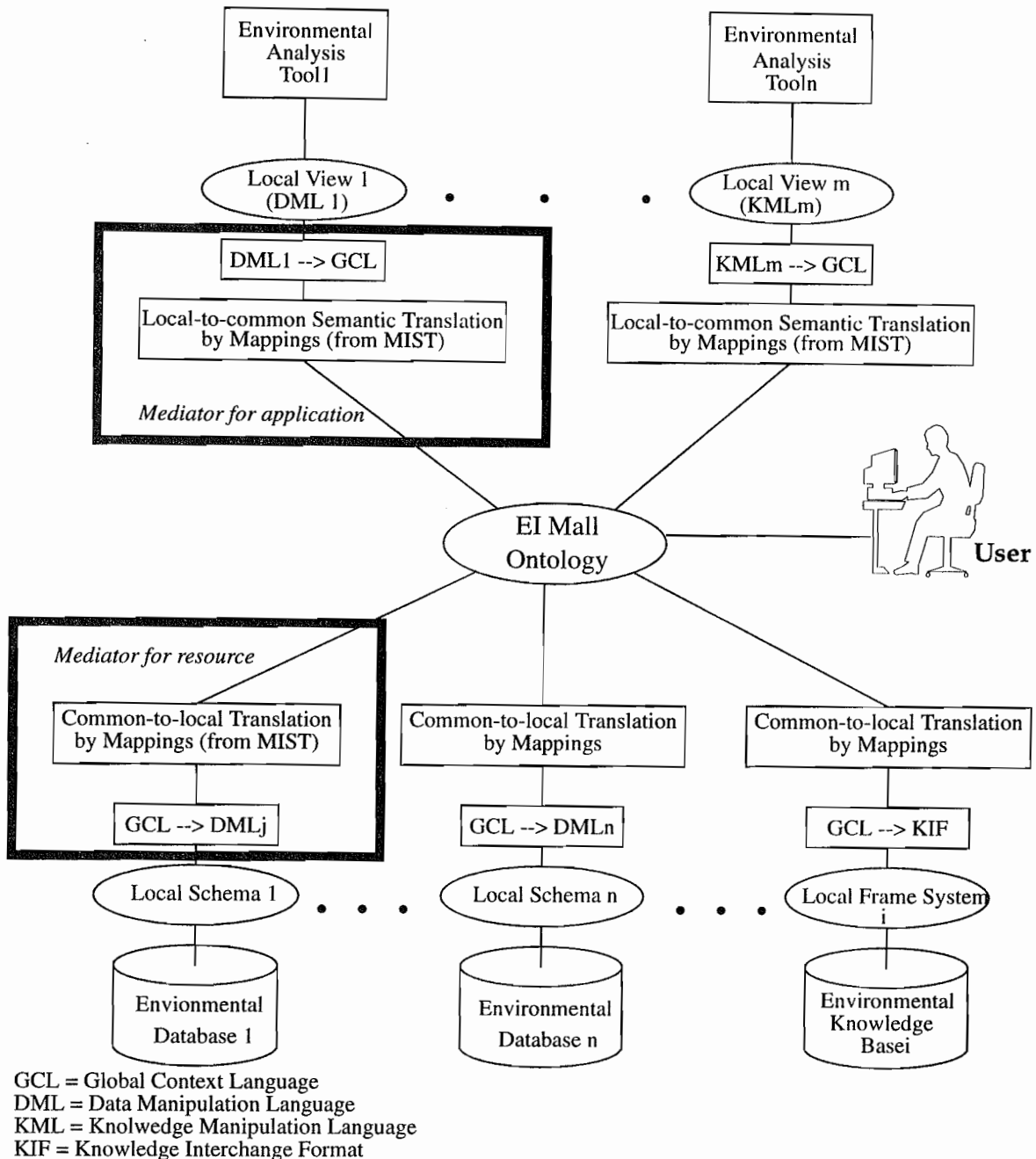
### 4 Project Description

The *Environmental Information Mall* will be assembled in a

series of phases. When fully complete it will be a resource node on the NII that will be accessible via one of several “homepages” oriented toward specific user disciplines, i.e., law, manufacturing, research, etc. Upon entering the *Mall*, users will be able to browse for both analysis tools and databases. An intelligent “sales clerk” or mentor will be available to assist users as they ask questions and browse the domains within the *Mall*. A key feature of the *Mall* will be the apparent ubiquitous nature of the analysis tools and databases. This will be accomplished by creation of a common ontology or context that can be viewed from several user perspectives [Guha 1990]. All of the tools and resources in the *Mall* will be related or mapped to the common ontology [Collet *et al.* 1991]. It becomes the unification, or the “walls” that define the *Mall's* space. The ontology development will use MCC's Carnot technologies [Woelk *et al.* 1992].

As users browse the *Mall*, they will identify sets of pertinent data, and the sales clerk will identify appropriate tools to run against the data to yield the specific information needed. The intelligent mentor will also apply a statistical probability or priority rating to inform the user of the likelihood of receiving the knowledge from the tools and databases accessed. The user will exit the *Mall* with just the information needed, not the torrents of data used for the analysis.

In order to accomplish this, the *Mall* will offer a neutral computing space of high-performance computing (HPC) facilities where the analysis tools can be applied to data from



one or more databases. In doing so, the *Mall* will take advantage of the HPC program, which has put in place the ability for a virtual computation environment. The user will issue only a simple request to process the data using the identified tools, with known costs and temporal constraints.

An important aspect of the *Environmental Information Mall* is the ability to do commerce. While certain sets of data and tools will be available free of charge, others will incur a fee

per transaction. Following free market principles, information and databases will be created, perhaps from public information, which address specialized information, or which present the data in particularly easy to use formats. These will have a higher usage rating, will score a higher probability of finding the information, and hence will be able to charge for their use. These "boutiques" will be the beginning of a class of facilities in the *Mall* that will become revenue generators for the developers. This, in turn, will

attract other boutique creators and other listings of environmental databases and tools.

As mentioned earlier, the unifying feature of the *Mall* will be the common ontology. Hence, each database and tool provider will need to pay integration costs for mapping their system into the common ontology for the *Mall*. The revenue generated by users will both encourage further database creation and support the maintenance on the *Mall's* ontology framework.

This concept involves the incorporation of many advanced information technologies, but as stated above, many of these technologies exist today and can be applied fairly rapidly. The *Environmental Information Mall* represents one of the more aggressive and comprehensive approaches to developing both a paradigm and a functional utility for creating a money-making enterprise on the NII and for helping enterprises comply effectively with environmental standards.

## References

- [Collet *et al.* 1991] Christine Collet, Michael N. Huhns, and Wei-Min Shen, "Resource Integration Using a Large Knowledge Base in Carnot," *IEEE Computer*, Vol. 24, No. 12, Dec. 1991, pp. 55-62.
- [Cutkosky *et al.* 1993] Mark R. Cutkosky, Robert S. Englemore, Richard E. Fikes, Michael R. Genesereth, Thomas R. Gruber, William S. Mark, Jay M. Tenenbaum, and Jay C. Weber, "PACT: An Experiment in Integrating Concurrent Engineering Systems," *IEEE Computer*, January 1993, pp. 28-38.
- [Guha 1990] R. V. Guha, "Micro-theories and Contexts in Cyc Part I: Basic Issues," MCC Technical Report Number ACT-CYC-129-90, Microelectronics and Computer Technology Corporation, Austin, TX, June 1990.
- [Neches *et al.* 1991] Robert Neches, Richard Fikes, Tim Finin, Tom Gruber, Ramesh Patil, Ted Senator, and William R. Swartout, "Enabling Technology for Knowledge Sharing," *AI Magazine*, Vol. 12, No. 3, Fall 1991, pp. 36-56.
- [Wiederhold 1992] Gio Wiederhold, "Mediators in the Architecture of Future Information Systems," *IEEE Computer*, Vol. 25, No. 3, March 1992, pp. 38-49.
- [Woelk *et al.* 1992] Darrell Woelk, Wei-Min Shen, Michael N. Huhns, and Philip E. Cannata, "Model-Driven Enterprise Information Management in Carnot," in Charles J. Petrie Jr., ed., *Enterprise Integration Modeling: Proceedings of the First International Conference*, MIT Press, Cambridge, MA, 1992.