Infomaster: A Virtual Information System

Donald F. Geddis
Michael R. Genesereth
Arthur M. Keller
Narinder P. Singh
Stanford University
Computer Science Dept.

Abstract. Infomaster is a virtual information system that allows users to access a variety of heterogeneous, distributed information from multiple perspectives. Infomaster accesses information stored in databases or knowledge bases using Agent Communication Language (ACL), a combination of KQML, KIF, and Ontologies. Facilitators are used decompose, route, and translate requests and assemble, route, and translate responses. Information may be queried and presented in a variety of formats. Infomaster is an operational system that currently presents information on rental housing and on people at Stanford.

1. Introduction.

Infomaster is a virtual information system. A virtual information system is one that allows its users to access common information from multiple perspectives. In the extreme, it should allow its users to store and retrieve information by opening "windows" on different portions of information in different conceptual frameworks and in different presentation formats.

Virtual information systems are especially useful in application areas where users with different points of view must access common information in a fully integrated fashion. Typical application areas include command and control, logistics planning and support, multidisciplinary engineering, and health care. For example, Infomaster is being used for electronic commerce by the CommerceNet Smart Catalog project including rental housing information.

Like any other information system, a virtual information system must provide the usual information management capabilities. An information supplier should be able to make statements to the system, confident that the system will distribute the information to interested parties and store the information for future use. An information consumer should be able to ask questions, confident that the system will provide answers on the basis of stored information and questions it asks of other users.

What differentiates virtual information systems from other kinds of information systems is their support for disciplinary, conceptual, and notational diversity. A virtual information system should allow its users to focus on a portion of information, confident of the system's ability to integrate that information with other information. A virtual information system should allows its users to interact in terms of different conceptual frameworks (i.e., vocabularies, data models), confident of the system's ability to relate information from one framework to that from another. A virtual information system should also allow its users to interact in a variety of different formats (e.g., text, tables, graphs, drawings, and so forth), confident of the system's ability to map information between these formats in a content-preserving manner.

2. Infomaster Architecture.

At the abstract level, Infomaster consists of one or more user interfaces, a mesh supporting routing, reasoning, and translation, and a variety of information sources storing information in various formats. More specifically, Infomaster has a World Wide Web (WWW) interface that can be used to enter queries using menus, SQL, or ACL. The queries are then converted into ACL and passed to the nearest facilitator. This facilitator may call on the resources of information sources and other agents and other facilitators it knows about. Information sources handle queries over the data they store. Other agents may handle such tasks as doing specialized translations, such as currency conversion. Facilitators may specialize in particular domains of interest. A virtual purpose facilitator may route requests to a specialized facilitator knowledgable about a particular domain.

We have experimented with several approaches to providing a WWW interface. One approach is to use a standard http daemon along with CGI scripts that interpret the input, convert it to ACL, and pass it on to a facilitator. Dynamically loading the CGI scripts can be slow and several context switches are required

before the data gets to the local facilitator. We have also developed our own specialized http daemon in LISP that bundles in the ACL converter and a facilitator. Because this facilitator may communicate with other facilitators and agents, this latter approach is faster without affecting the kind of information accessible to the user.

There are several approaches to connecting information sources to Infomaster. The information may be represented as a KIF knowledge base loaded directly into an agent. The information may be stored in a database with an agent operating as a wrapper between a facilitator and the database. Or a facilitator may directly understand the database's native language and protocol. The information sources notify one or more facilitators of their willingness to respond to requests. These "advertisements" are used by the mesh of facilitators to route requests to the various information agents.

Information agents may vary in how they represent, describe, query, and present their information. Translation is required in both directions: queries and responses. Some translation may occur in a facilitator, such as handling naming differences. We have also developed converter agents to perform other translations, such as currency conversion, distance, area, and volume.

3. Example: Rental Housing.

The first information available through Infomaster is rental housing in the San Francisco Bay Area. We extract classified advertisements from the WWW sites of several newspapers. These are then separated into individual advertisements, parsed into a structured format and then loaded into a KIF knowledge base. This knowledge base has advertised that it can handle queries for rental housing.

Every morning, an agent accesses the web pages of each of the newspapers and obtains the latest classified advertisements for rental housing. Each of the newspapers organizes its web pages differently. Some of the web pages are static, and can be fetched and parsed. Some of the web pages present a query form, and specific queries are needed to extract the relevant data. Although the actual newspaper advertisements are common in format: a specialized abbreviated English description of the rental unit(s) available, additional information is also available. For example, the advertisements are organized into classifications specific to the newspaper. This classification scheme is relevant, as it can distinguish an apartment from a house from a shared rental situation. Also, other information can be found on some of the web pages, such as the starting and ending dates of the advertisement.

There is a menu interface and an ACL interface. The menu interface allows the user to determine what kinds of information is returned; a checkbox indicates whether each particular kind of information is desired. To describe the search, the user can use menus, such as for the type of housing or desired amenities, or fill in boxes, such as for the minimum and maximum number of bedrooms, bathrooms, or monthly rent.

Once the user has specified a query, the user may determine how many rental advertisements satisfy the constraint and further constrain the query iteratively. At any time, the user may display a table of the rental advertisements satisfying the current query. The user may also look at a particular rental housing advertisement, including the original advertisement before it was parsed.

The rental housing information has been made available just in time for the new Stanford students to arrive. In the first day of availability, Infomaster has handled 3000 queries for rental housing.

4. Future work.

We plan to improve Infomaster in several ways. First, we plan to support many additional sources of information. The information sources we have released are rental housing and Stanford people. In addition, we have multiple information sources on workstations and other products. Others are planned. Second, we plan to extend the capabilities of Infomaster. For rental housing, Infomaster currently uses rental housing advertisements from newspapers. We plan to allow landlords to list their own rental housing advertisements with Infomaster. We also plan for renters to register their interest in housing, so that they will be notified about housing advertisements that satisfy the criteria they have specified. These criteria will be specified through the same query interfaces.

Infomaster can support a variety of query interfaces. When searching for products, a user may want products similar to a specific product. We plan to allow the user to take a particular result, convert it into a query screen, and then alter the query as the user sees fit.

5. Conclusion.

We have developed and fielded Infomaster, a virtual information system. Infomaster supports a variety of user interfaces to various information sources. Infomaster uses Agent Communication Language (ACL), a combination of KQML, KIF, and Ontologies, for communication among facilitators, agents, information sources, and user interfaces, along with the languages native to those systems, such as SQL and HTML.

6. Acknowledgments.

This work was funded in part by a subcontract from the CommerceNet Consortium, a Technology Reinvestment Program project, and by the ARPA Intelligent Integration of Information (I3) program.

Several other people have worked in support of the development of Infomaster. These include Vishal Sikka, Winton Davies, Bill Vrotney, and Mustafa Syed.