Smart Catalogs and Virtual Catalogs

Arthur M. Keller
Stanford University

Smart Catalogs and Virtual Catalogs

Arthur M. Keller
Stanford University
Computer Science Dept.
Stanford, CA 94305 USA
ark@cs.stanford.edu

Abstract. We present an architecture for electronic catalogs, called Smart Catalogs and Virtual Catalogs. Smart catalogs are searchable, annotated combinations of machine-readable and machine-sensible product data. Virtual catalogs dynamically retrieve information from multiple smart catalogs and present this product data in a unified manner with its own look and feel, not that of the source smart catalogs. These virtual catalogs do not store product data from smart catalogs directly (except when caching for performance); instead virtual catalogs obtain current product data from smart catalogs to satisfy specific customer queries. Customers interact with smart catalogs and virtual catalogs through WWW or other interfaces. Product data is disseminated through the architecture using ACL (Agent Communication Language). In particular, ACL is used to communicate queries and answers among smart catalogs and virtual catalogs.

1. Introduction.

We present an architecture for smart catalogs and virtual catalogs that supports cross-search of multiple catalogs.

First let us consider how electronic catalogs are currently organized. Typically, these catalogs use ordinary World Wide Web (WWW) protocols and clients, such as Mosaic and Netscape. Each document, called a page, contains desired information or pointers to other pages, hopefully getting the user closer to the desired information. Typically, the World Wide Web (WWW) resembles a giant menu system, where each document behaves like a menu of links to other documents that hopefully get the user closer to the information desired. Each vendor maintains its own (collection of) catalogs. There may not be uniformity even among multiple divisions of the same company, let alone among multiple companies manufacturing or selling comparable products.

There are a variety of limitations of this typical approach to electronic catalogs. It is hard to find what you are looking for. You have to figure out where to start. For example, in order to find information about a particular product, you must separately visit each vendor of that product, and then navigate through that vendor's Web space to find the desired product. Making such navigation more difficult is that each vendor organizes its Web space differently.

Few catalogs have the ability to search by content (i.e., reverse-search). Reverse-search is typically by keywords. Keyword searching techniques, such as WAIS, are only of limited help, as they require that the user phrase the query using the terminology of the each data source, and vendors tend to use differing terminology from each other to describe products. A notable exception is Danish International's BingoSearch, which allows parameter-based search of a single catalog. Except for centralized catalog schemes that force a uniform catalog structure among multiple vendors, there is no fielded capability for cross-catalog search.

Our approach is to support reverse-search (i.e., search by content) of multiple vendor catalogs based on a deeper understanding of the contents of these catalogs.

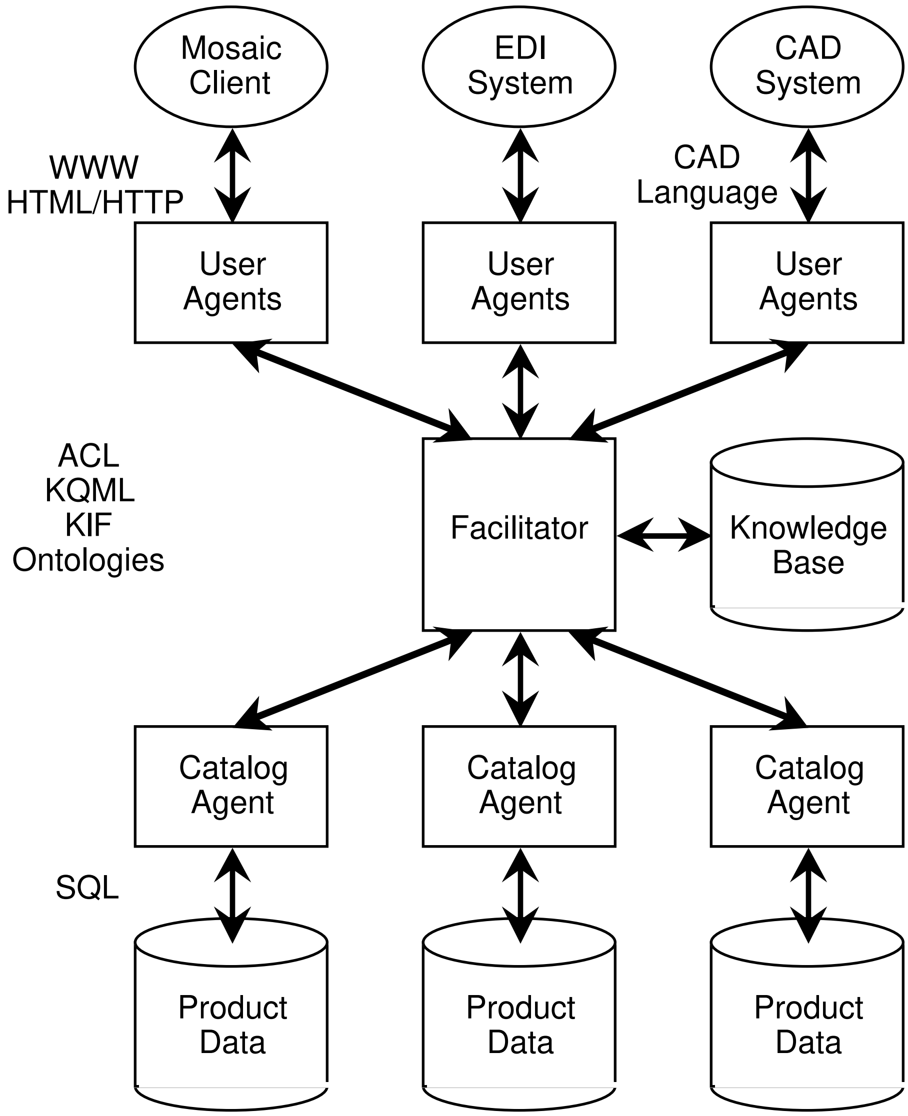

Figure 1. Smart Catalog and Brokering Architecture for Electronic Commerce

The following section describes the overall architecture of a Smart Catalog. Section 3 describes the concept of a Virtual Catalog. Section 4 contains a usage scenario of Smart Catalogs and Virtual Catalogs. Section 5 describes the current status and future work of this project. Section 6 gives our conclusions.

2. Architecture.

Our architecture is illustrated in Figure 1. It consists of agents communicating with facilitators, plus other components, such as data sources and user interfaces, that interact with the collection of agents and facilitators.

The communication language used by the agents in our architecture is Agent Communication Language (*ACL*). ACL consists of the Knowledge Query and Manipulation Language (KQML), the Knowledge Interchange Format (KIF), and a set of Ontologies. *KQML* consists of performatives, such as `ask-one`, `ask-all`, and `tell`, that describe the nature of the action to be taken. KQML has the role of the communication language in CORBA. *KIF* is based on First-Order Predicate Calculus and is the content language we use with KQML. KIF is powerful enough to contain or to encapsulate any other content language, so that any information may be obtained from the information source, translated to the desired format, and transmitted to the requester, assuming the necessary components exist. Each product database is described by an *Ontology* (another term for ontology is "controlled vocabulary"), which defines the database, its structure, the terms used in it, and how they relate to each other. You can think of KIF as the "grammar," KQML as the "verbs," and ontologies as the "nouns" and "adjectives" that comprise the language for interoperating agents in our architecture.

Each Facilitator in our architecture acts as a broker. Facilitators perform routing, reasoning, and translation. A facilitator stores agent-provided advertisements of coverage in a knowledge base along with relevant ontologies. The facilitator uses these advertisements to determine which agents can support a particular request. And it translates requests into the language and terminology used by each responding agent and also translates responses into the language and terminology used by the requesting agent. A facilitator will decompose requests requiring action by multiple agents and then compose the responses for the requester.

Product data is stored in databases, for easy search and maintenance. Such data includes structured information, parameters, text, pictures, sound, video, etc. Each product database communicates with a Catalog Agent using its native language, such as SQL. The Catalog Agent performs 3 roles: It advertises the coverage of a product database; it understands queries and translates them into the language of the product database, and it packages answers from the product database in a standard format.

Someone using a WWW client, such as Mosaic or Netscape, will connect to a User Agent using the ordinary WWW protocols HTTP and HTML. The User Agent will present the user with an HTML form, either static or dynamically created. The user will describe the desired object using an HTML form. When responses come back from the facilitator, the User Agent will prepare dynamically created HTML documents with those responses.

We define four types of ontologies. *Base ontologies* are used to define common terms, such as engineering math, legal terminology, standard terms and conditions. Base ontologies are shared among all uses of this approach and are created by universities and research laboratories. *Domain ontologies* contain terms common to all or most vendors in an area, such as CPU speed, RAM size, or disk storage capacity. Typically domain ontologies are created by standards bodies and trade associations. *Product ontologies* contain company-specific terminology and refer to domain ontologies, such as NuBus cards for the Apple Macintosh. Individual companies create product ontologies, although other companies may refer to them. *Translation ontologies* are used to translate specific terms used in one ontology or information source to related terms used in another ontology or information source.* For example, a translation ontology may describe how to use an SVGA monitor (a PC standard) on a Macintosh. Individual companies create translation ontologies to enable them to compete in other markets. We expect there to be service organizations that create and maintain product ontologies and translation ontologies on behalf of other organizations.

3. Virtual Catalogs.

One important problem with using the WWW for product catalogs is the interaction between manufacturers and distributors or retailers. Consider a retailer that sells products from multiple manufacturers. The retailer will want

* The term "articulation axiom" is sometimes used to refer to entries in a translation ontology. Translation ontologies may be used to create an ontology algebra supporting translation across multiple ontologies.

to include product information from each manufacturer in its product catalog. Replicating all this product information in the retailers catalog would incur a considerable storage and maintenance cost. The natural approach using the WWW is for the retailer to hyperlink to each manufacturer's catalog so that the customer may obtain detailed product specifications.

There are several problems with the hyperlink approach. First, the customer may get "lost" within the manufacturer's webspace and not know how to get back to the retailer. Second, the manufacturer does not know the context of the customer's interactions with the retailer. Third, the customer may stumble upon a how-to-order page provided by the manufacturer, and wind up ordering from someone other than the original retailer. Fourth, if the customer does make it back to the original retailer by using the "back" button, no information determined at the manufacturer's site is carried along with the customer, such as the desired product configuration. Fifth, if the customer gets back to the retailer through the manufacturer's how-to-order page, the retailer does not know the original context of the interaction with the customer (e.g., other products selected for order in this same session).

One approach to multivendor catalogs is the integrated approach, where all the catalogs are stored on one site using one implementation. A notable example is by Open Market. An alternative is to provide some mechanism for manufacturers to respond to specific queries by retailers to satisfy customer requests for product information. This approach can be based on a business relationship between the retailer and the manufacturer. This approach is the one we take in a Virtual Catalog.

Virtual catalogs allow retrieval of product data using a distributor's catalog by combining information from multiple manufacturers' catalogs. This retrieval is performed dynamically, upon the user's request, based on the user's search criteria, using the terminology of the distributor or any connected vendor, and will retrieve data from any relevant connected vendor. Therefore, the distributor's virtual catalog is always kept up-to-date and in synchrony with each manufacturer's smart catalog. The distributor can choose to display all of a manufacturer's products or only a subset of those products.

With virtual catalogs, the distributor maintains control over the interaction with the user. The user never interacts directly with the manufacturer's catalog. Instead, the manufacturer's information is retrieved on demand and is presented to the user with the distributor's look and feel, not the individual look and feel of each manufacturer. The relationship between the user, virtual catalog, and smart catalogs is shown in Figure 2.

There are two key business relationships in the Virtual Catalog world, and many supporting business relationships. The two key business relationships are between the customer and the retailer (the virtual catalog), and between the retailer and the manufacturer (the smart catalog). Processes may exist in the virtual catalog to bridge these relationships. For example, orders from customers may trigger resupply EDI transactions to the manufacturer. Or the order may be forwarded to the manufacturer for drop shipment of the product directly to the customer. In addition, there are supporting business relationships. For example, if credit cards are used for payment, the customer has a relationship with the issuing bank, and the retailer has a relationship with the acquiring bank, and the two banks have a relationship for clearing the credit card charge. Similarly, there are relationships for order fulfillment, shipment, etc.

Virtual catalogs are appropriate for retailers and distributions, but they also enable new business models. A virtual distributor may operate using a virtual catalog and business relationships for order fulfillment, shipment, etc. The virtual distributor may not even have

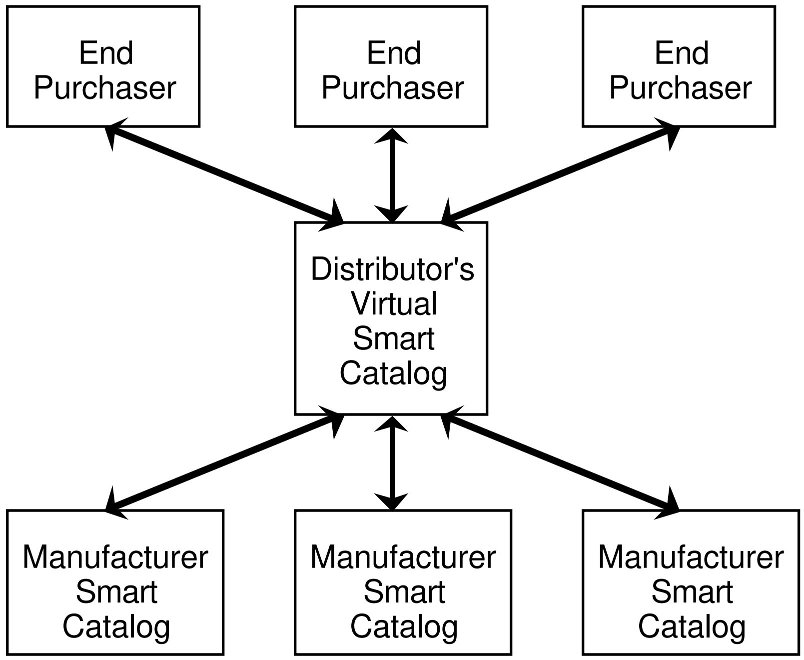

Figure 2. Virtual Catalogs.

any inventory, warehouse, etc. The virtual distributor could be a completely computerized setup, automatically providing product information using manufacturers' catalogs, taking orders and arranging for order fulfillment and payment.

4. Scenarios.

Consider a customer's request for color PostScript inkjet printers for the Macintosh costing under $1000. The User Agent will translate the query into KIF and submit it to a Facilitator. The Facilitator handling the query will consult its knowledge base for the facilitators or agents that can handle this request. For example, the Facilitator may transmit the request to the Catalog Agents for Apple and for Hewlett-Packard. The Catalog Agent will then interrogate the product database and translate the answer into KIF. For example, the Hewlett-Packard Catalog Agent may respond with the description of the HP 560C printer. The Apple Catalog Agent may respond with the Apple Color StyleWriter Pro. The facilitator will then collect these responses for the User Agent, which will package the responses in HTML for the Mosaic client. The user agent and facilitator are provided by the virtual catalog company. The catalog agents and product databases are provided by the manufacturers as part of their smart catalogs.

A customer may instead be interested in color PostScript laser printers for the Macintosh for under $3000. As of this publication, such printers cost around $5000, but prices are dropping. So the customer may request notification when any such printer is newly announced, or is lowered in price. This request will be stored in the relevant facilitator's knowledge base. When a manufacturer announces a new product, it will have a catalog agent send a message with the announcement and any changes to the advertisement of the agent to the facilitator. The facilitator will then send appropriate notifications to those parties who have expressed interest in this news.

Notice that the facilitator notification scheme is symmetric. Catalog agents express interest in being given product data queries. User agents express interest in being given product data announcements. The same notification scheme is used for both types of activities.

Operators of virtual catalogs have several alternative models for paying for their operation. The operator may take and process orders and use the markup on the transaction. Alternatively, the operator may make money merely by providing information. (When information is free, search is valuable.) The operator may charge manufacturers, either a fixed fee or per referral. The operator may charge customers, either by subscription or per search. The operator may sell demographic information on customers or on searches to market research firms, manufacturers, retailers, or trade associations. Of course, an operator may obtain revenues from multiple of these approaches.

5. Current status.

The architecture of smart catalogs and virtual catalogs is an application of the facilitator architecture long in use in Logic Group of Stanford University's Computer Science Dept. The facilitator architecture has been demonstrated in the domains of software interoperability and concurrent engineering. It is now being applied to the domain of electronic commerce as part of Stanford's Center for Information Technology's (CIT) efforts on CommerceNet.

Several smart catalogs have been built in collaboration with several companies in the domains of workstations, test and measurement equipment, and semiconductors.

We are extending this work by creating a collection of smart catalogs for several manufacturers, a virtual catalog for a retailer, and a domain ontology for a trade association.

Please contact Arthur Keller by e-mail at `ark@cs.stanford.edu` or by WWW at `http://logic.stanford.edu/cnet.html` for more information.

6. Conclusion.

We have described an architecture for electronic catalogs for multiple companies that interoperate. Companies create smart catalogs of searchable, machine-sensible product information. Retailers and distributors create virtual catalogs that provide customers with product information dynamically requested from manufacturers' smart catalogs. Virtual catalogs provide a new degree of interaction between manufacturers and retailers or distributors. Virtual catalogs enable new business relations and new business models.

Acknowledgments.

This paper has benefited from feedback of numerous people who have heard presentations of this paper, read earlier drafts, or participated in discussions. In particular, Michael Genesereth designed the overall agent communication architecture used by smart catalogs and virtual catalogs, Narinder Singh designed the facilitator used here, and Mustafa Syed programmed some of the other components in our initial prototypes. Note an earlier version of this paper, co-authored by them, appeared at the Workshop on Electronic Commerce following CIKM, December 1994. Several other people have worked on the development of smart catalogs. These include Felix Chow, Bob Engelmore, Wanda Pratt, and Rupert Murdock.

This work was funded through a subcontract from the CommerceNet Consortium, which in turn derived its funding from a cooperative agreement with the U.S. Technology Reinvestment Program, as well as over 100 companies and organizations. In addition, several companies have provided funding of this project towards the construction of pilot smart catalogs.

References.

Danish International’s BingoSearch can be seen at URL `http://www.danish.com`.

Michael R. Genesereth and Steven P. Ketchpel, “Software Agents,” Comm. ACM, Vol. 37, No. 7, July 1994.

Thomas R. Gruber, “Toward principles for the design of ontologies used for knowledge sharing,” in Nicola Guarino, Ed., International Workshop on Formal Ontology, Padova, Italy, 1992.

Open Market’s URL is http://www.openmarket.com

William T. Wong and Arthur M. Keller, “Developing an Internet Presence with On-line Electronic Catalogs,” Workshop on Electronic Commerce, December 1994, available from the URL `http://www-db.stanford.edu/pub/keller/1994/cnet-online-cat.ps`.