Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138.
doi:10.7152/acro.v2i1.12552

Terminology Maintenance for Corporate Information Retrieval

Marlene Rockmore
Massachusetts Institute of Technology
Laboratory for Information Decision Systems
Cambridge, MA 02139

The goal of information retrieval in the corporate environment is to provide management, sales, marketing, and product development with timely information about the changes in external forces that impact strategic and tactical business decisions including what products to build and how to position these products in the marketplace.

The elements of this type of information change rapidly. Mergers and acquisitions change the structure of the marketplace, the vendors and their products; political and economic changes impact production and market decisions from manufacturing plant locations to the location of distribution facilities. Thus, corporate IR is a complex, volatile application which requires tools and processes to allow dynamic modification to content.

This paper will discuss how Thesaurus/Indexing Management System (TIMS) is used at Digital Equipment Corporation to allow for modification of the information retrieval application without modifying or reloading the underlying content. Part of the solution lies in the capabilities of using faceted thesaurus structures, that can serve as an abstracted model of the user's need and the content.

Consistent information retrieval, and its benefit of ongoing funding and support, depends on the continual application of maintenance processes as part of the overall management of a corporate information retrieval program. This paper will present the following topics related to thesaurus maintenance: capturing and structuring of unknown terms, implementation and training, and evaluation.

I. OVERVIEW

Information retrieval (IR) systems based on use of classification have not been widely accepted in corporate information retrieval because of the perceived overhead in labor and costs. Data collected from Digital Equipment's Corporation internal information system shows that A knowledge-based IR system, using faceted thesaurus structures, while requiring some specialization in knowledge engineering, actually shows an substantial return on the investment in terms of reliability of the retrieval.

The application used two methods of retrieval: a free text query based on keywords in the inverted index file and a thesaurus-controlled query, where the keywords in the inverted index file were enhanced by terms and relations from a highly structured faceted thesaurus. The system usage averaged 1,000 sessions per week and contained approximately 15,000 documents.

The system has maintained some interesting performance. Users' search terms are captured through the forms entry screen at user interface input. Of terms entered to the free text query, only 2 out of 5 terms are matched against the index file; while on the thesaurus controlled field 4 out of

ISSN: 2324-9773

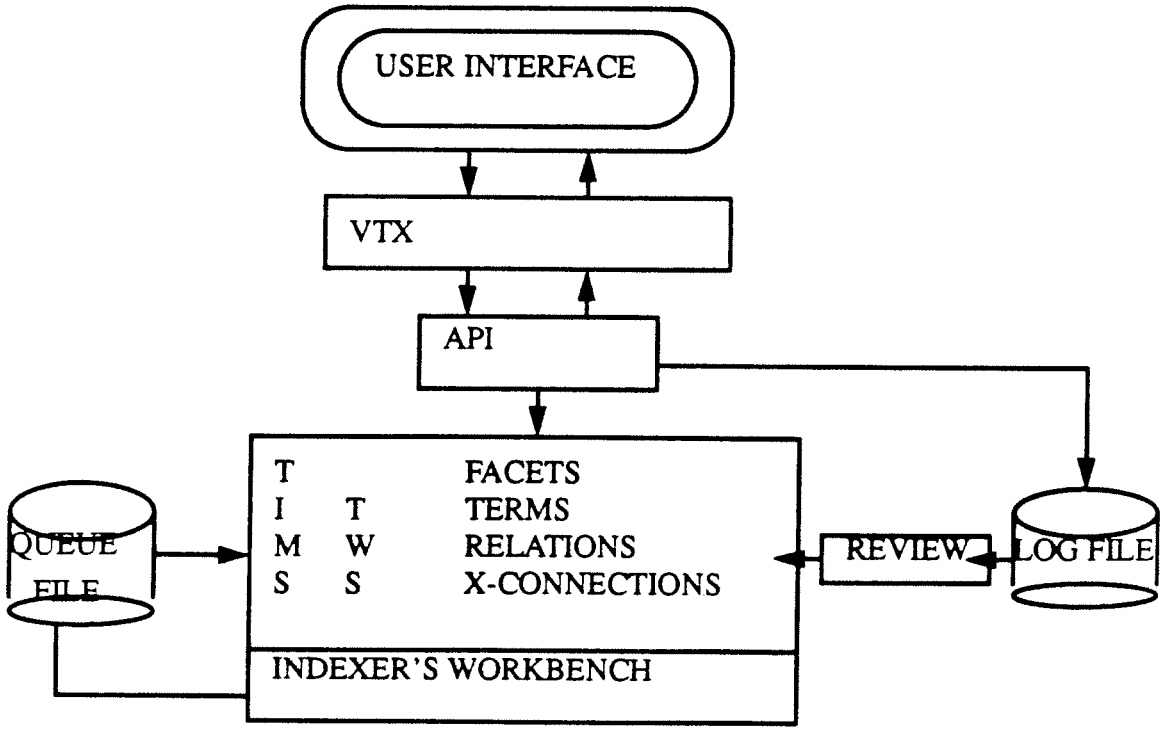
Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

5 terms are matched against the index file. This ratio has remained consistent through the four years that the system has been in operational use. [ROCKMORE1]

The major reason for the consistency of the system has been the tools and processes developed to capture and represent user's information needs and queries in a well-structured classification scheme that is designed to represent the external forces (market, business, technology) and overlay that representation as an intermediating structure between the source documents located by the inverted index and the query. The system, TIMS, is a thesaurus/indexing management system consisting of three subsystems:

- **Thesaurus Building Tool** (TWS) that includes functions to create facets, define cross-connections between facets, update thesaurus terms in each facet; update applications, and track postings to thesaurus terms
- **Indexer's Workbench** (IWB) that provides interactive access to terms and documents and allows indexers to assign terms to document.
- **Candidate Term Subsystem** to allow update of terminology from the IWB to TWS

TIMS is used in a retrieval application via an application programmer interface (API) that allows a programmer to write a user search interface customized to an application's requirements. The illustration below presents a high-level overview of the architecture:



ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

It could be argued that in this architecture, TIMS creates an abstracted representation that is overlaid above the contents of the database. Because TIMS utilizes a feature called cross-connections, which are predetermined labelled links between facets, the thesaurus system has characteristics of a semantic network. It also has the ability to capture representations of information into terms, relations, and facets, giving the thesaurus the properties of a knowledge base, a repository of sharable corporate concepts.

The methodology used to create the design is similar to knowledge acquisition processes in artificial intelligence. User queries are used to drive the system design so that the design captures facets of the environment. Continual monitoring of user query terms is used to update the design. Some of the monitoring uses online tools to elicit new knowledge and update the internal representation of terms and relations within facets.

Management of this knowledge base in active applications ensures ongoing quality of the thesaurus and of the application. The processes for maintaining terminology are:

- Capture of input terms from the User Interface
- Encouraging indexers to update terminology through the candidate term subsystem (Queue file)
- Use of thesaurus building tools to allow modification to existing terms and relations and to allow the definition of new facets
- Ongoing analysis and evaluation of the corporate information retrieval environment and user information needs and ongoing representation of those needs in the application design

The maintenance process is managed at the application level; each application is empowered to modify its terms, relations, facet, and application design. At present, the only exception to this process is when two applications agree to share a facet, in which case the applications need to agree as to the facet's relations and editorial policy.

The facet design has proved to be the important element in allowing decentralized application control, while enforcing corporate standards. Facets can be viewed as "lego pieces" that can be combined into one arrangement to create a representation for one application and recombined to create a different representation for yet another application. This is the property of "reusability," which will be discussed later. The role as an intermediating structure between the user's request and the content means that the concept structures (terms and relations) can be dynamically updated to reflect significant changes without impacting the underlying content, other than making the access to the content timely.

Application-level control of the maintenance process has proved to be extremely beneficial. It eliminates bottlenecks caused by centralized processing, and allows the application to meet the business need by responding quickly and efficiently to changes in business information.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

The overall integrity of the thesaurus can be controlled by monitoring the application's decision through the reporting system. Available reports particularly useful to overall management are:

Alphabetical Term Lists by Facet
Hierarchical Term Lists by Facet
Transaction Reports showing terms Added, Modified, Deleted by Date Range

The four terminology maintenance processes will be discussed separately below.

II. TERMINOLOGY MAINTENANCE

A. Capture and Structuring of User Query Terms

User query terms are captured at the user interface. This interface allows a user to specify from a menu whether he or she are seeking information about a product, company or subject and to name the relevant field. At the next level the user is offered a screen in which he or she can input a keyword or phrase

Product _____

If the user does not have a specific term in mind or would like to see a list of the available lead terms from the thesaurus, he or she can type a? to generate the list. The online thesaurus display is currently arranged alphabetically and displays only the lead terms. We have not checked to determine how many times users have used a term from the online display against the number of times they enter a term without constraining their own term selection. The system is designed to encourage users to supply unconstrained term selections; that is to use their own terms [ANDERSON 4].

All search sessions are tracked on user session logs, which are used for performance monitoring and overall statistics on usage. A batch job is run each weekend that reads in the logs, and writes all the user search terms by field. It also tallies total number of
query sessions for the week.

This report is reviewed by the vocabulary specialist, who checks for all search terms that retrieve 0 items (unknown term). An unknown term can occur for several reasons:

1. The most common reason is that the term is actually a synonym to a preexisting term but the user has come up with an innovative spelling or label for the term. This problem has been identified by Furnas, et. al. as the "vocabulary problem" [FURNAS3].

2. A term could also be a new concept that needs to be included into the vocabulary.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

3. The term could be a concept that falls outside the scope of the source database. In this case, the term could be a pointer to another database; we have investigated this feature as a way of providing an integrated end-user interface that provides transparent access to multiple remote distributed sources outside our current application and organization.

4. Last, we may have the term in the vocabulary, but do not have any content posted to the term. In that case, content is prioritized for collection and acquisition.

It should be noted that 0 items in our application occur in 1 out of 5 searches, which means a user term will be matched at a frequency of 80 percent. The evaluation of the system will be discussed later in this paper.

Problems associated with the first two types of unknown terms (that is, new but synonymous variants and new concepts) are managed by structuring such terms into the vocabulary as a synonym to an existing term, or as a new lead term. The first step in the process of structuring terms is to identify the appropriate facet[s] to which a term belongs. To do this, the specialist needs to understand the underlying vocabulary design.

In TIMS, terms are categorized into facets. Terms can be concrete objects such as a product name (VAX 6000) or abstract such as computer concepts (midrange systems; fault tolerant). Facets are collected into thesauri (lists of facets) which become a field name in the user interface. Thus the term "thesaurus" has two meanings. A thesaurus is used in the traditional sense to mean terms arranged with equivalent, associative, and hierarchical relations. However, the facet design requires every term to be assigned to a higher-level categorical relation (the facet code). Additionally, these facets (categories of terms and relations) can be grouped into larger groups. Collections of facets, for lack of other terminology, we also choose to call thesauri. The concept of grouping facets into thesauri will be explained in the section "Reusability."

Facets have several qualities in organizing and managing terminology that are critical to information retrieval management:

1. They can be linked at the facet level and the properties of the facet inherited
2. They can be reused
3. They can be used to disambiguate terms and concepts.

Each of these qualities will be discussed separately below.

1. Linking Property. Facet levels are deterministically linked to other top level facets through a facility called cross-connections. This gives the faceted thesaurus structures properties of a semantic network. A semantic model of a business application, for instance, may show that products are made by a company; the cross-connection feature ensures that this relation (i.e., product/producer), which is not strictly associative, equivalent, or hierarchical, is incorporated into the design.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

For example, the most concrete object represented in our application is a product such as a VAX6000

The facet level design for PRODUCT is:

PRODUCT	made by	Companies
	have	Hardware Classification
	used in	Targeted Application

Cross-connections can be used to replace the verbs with relations that symbolize the linkage between concepts; therefore "made by" can be replaced by MB, "have" by GP, and "used" in by GA [ROCKMORE4].

Thus the specific thesaurus representation for the term VAX 6000 would be as follows:

LT VAX 6000
NT VAX 6210
MB Digital
GP Multiprocessing
GA Online Transaction Processing

2. Reusability. In the test application, several fields were available for user searches including:

Product Name
Company
Subject

The subject field in our application is a collection of seven separate facets although this is transparent to the user. This design has several advantages. First, the facets within the thesauri are reusable in other applications. Second, the design allows new topics to be added to an existing application without any respecification of the software.

The complexity of the underlying facet design is transparent to the end user. Fields can be composed of multiple facets, as illustrated in the chart below and each of those facets can have specialized thesaurus relations.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

USER SEARCH FIELD	THESAURUS NAME	FACET CODE	SAMPLE HEAD TERM
PRODUCTS	Hardware products	PROD	VAX 6000
	Software products	SOFT	VAX Rdb
	Peripherals	PERI	RA90
COMPANIES	COmpetitive COmpanies	COCO	AT&T
	Accounts	ACCT	AT&T
SUBJECTS	HARDware concepts	HARD	Parallel Processing
	SOFTware concepts	SOFT	Database Management
	SERVice	SERV	User Education
	APPLication	APPL	Banking
	MARKeting	MARK	Pricing
	COmpany INtelligence	COIN	Earnings
	GEOGraphy	GEOG	Germany

The property of disambiguation applies to codes in the system as well as to terms. Because terms are clearly marked by categories, their definition is represented. Thus, I can have a nonunique code SOFT representing a facet and a thesaurus, because the code is tagged by its category. The property of disambiguation will be discussed next.

3. Disambiguation. Machine-aided elicitation of terms within user interface fields aids in disambiguating concepts. By allowing users to enter terms and to select the field in which the user enters the term, the system provides machine-aided knowledge elicitation and term-role disambiguation. The search report provides some information to help the specialist make this decision. The report shows into what search field the user had placed the term. Thus, if the user puts the term "ASK" in a company field, the specialist knows that the user thinks ASK is a company and not a product or general subject or system command.

The specialist can use the information supplied by the user in making judicious decisions about how to place the term in the vocabulary including the facet to which the term should be assigned and the placement in the thesaurus hierarchy. Capturing user-supplied terminology ensures currency of vocabulary; well-structured vocabulary encourages indexing quality.

The second source of terminology is from the indexing process. This process will be described next.

B. Candidate Term Subsystem

The second process for maintaining terminology is through the candidate term subsystem. The candidate term subsystem or Queuing system is used to manage the interface between the vocabulary and the indexer's workbench (IWB). The IWB provides indexers with interactive access to the vocabulary and documents so that index terms can be assigned to documents.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138.
doi:10.7152/acro.v2i1.12552

The system encourages indexers to add two types of terms

1. Add a new term, which is a term not in TIMS. A new term is automatically recognized.
2. Clarify an Ambiguity: assign an existing term to two or more facets.

Item 2 requires some explanation:Disambiguating terminology is an important aspect of the system, as it encourages indexers to address subtle but important ambiguities in articles. The system is designed to index a document as precisely as possible by encouraging indexing specificity and precision. Precision is encouraged by assigning index terms to the appropriate facet.

For example, in the following text

> Company A recently announced that it will acquire 51 percent of the stock of Target Company, Incorporated. The announcement makes Company A the largest vendor of software for tracking mergers and acquisitions.

there are two important yet ambiguous concepts: An index term such as "merger and acquisitions" would not reflect that there are two distinct concepts, which are the merger and acquisition of two companies and application software for tracking mergers and acquisitions. The faceted thesaurus structure assist in disambiguating these concepts. Indexing for this text would be on the term "mergers and acquisitions" in the Company Intelligence (COIN) facet and the term "Investment" in the facet "Applications" (APPL). The facet names and codes were discussed in the table in the previous section.

This functionality has been used to clarify relations between companies. The article above, for example, makes Company A a possible business competitor. Therefore the indexer may alert the knowledge base to this by classifying Company A as a competitive company. The representation for this concept would be (COCO A).

TIMS indexing tools were designed to encourage this precision in selecting conceptual representations for documents through the indexing and vocabulary. Indexers are discouraged from using their paper tools and are encouraged to use online tools to freely add terminology that they deem significant to the document, while at the same time applying appropriate indexing guidelines. For our system, the guidelines are:

> Specificity: index to narrowest term
> Exhaustivity: index to all facets as assigned in the application design
> Precision: index unambiguously

It should be noted, as an aside, that exhaustivity is enforced through the facet design. The indexer knows that a document should have an index term for each relevant facet heading in the design.

Adding a New Term. A new term is automatically recognized by the system, and the indexer is asked to add the term to queue by setting the "Add term to Queue" flag to "Yes." An

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

indexer could also suggest an alternate structure for a concept such as placing an existing term within a different facet. The system supplied an online term template in which the indexer could fill out term values.

The online term templates are reviewed by the vocabulary specialist who decides whether to ADD term, DELETE, or HOLD. A HOLD is placed on terms that appeared to need additional research before deciding whether they should be entered into the vocabulary.

ADD Terms, that is new lead terms, can be added automatically from the queueing system. Terms that already exist in the vocabulary and need modification must be updated directly by using Thesaurus Maintenance Functions. These functions will be described next.

C. THESAURUS MAINTENANCE FUNCTIONS

The Thesaurus Maintenance Functions allow the application's vocabulary specialist to directly update the vocabulary. A vocabulary specialist decides that the vocabulary needs modification based on analyzing the weekly user search log, the candidate term subsystem, and his or her own knowledge of the application and industry.

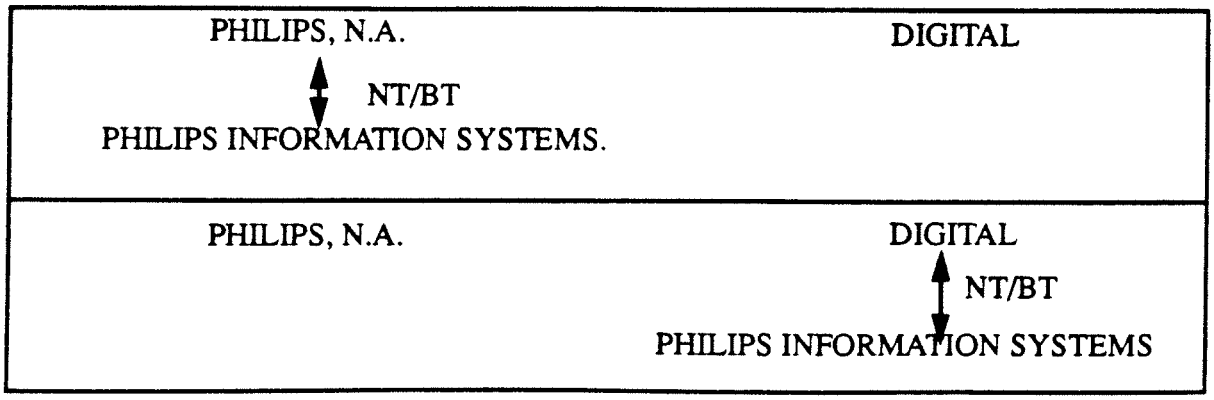
There are three major functions for direct update of the vocabulary as follows:

UPDT: Add or Modify Terms
UPDL: Merge, Delete, Inactivate Terms
UPDA: Add or Modify Applications

Each will be discussed separately below.

1. UPDT: Add or Modify Terms

This function allows modification of terms and relations. It is where the vocabulary specialist adds synonyms, builds and extends hierarchies, and adds associative relations. Both the term and the relation can be modified. This system allows relations to be deleted. For example, in the case of a corporate restructuring, one might delete a subsidiary from one company and add it to another.



ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

While the NT (and its reciprocal BT) are eliminated from one hierarchy.The lead term (Philips Information Systems) and its relations, such as its synonyms and scope notes remain; the hierarchical NT/BT to Philips is deleted.

The system ensures the integrity of hierarchies and synonyms and flags possible problems with interactive messages; integrity checks include:

1. Synonyms must be unique per facet. A duplicate synonym will be flagged and forced into a specific relation type: USA, that holds synonyms that are alike together in a given facet. The USA relation is used to collocate ambiguous terms in a facet. A common use has been to collocate all products with the name 3000.

2. Hierarchies are tested to ensure that a broad term does not exist at the same lower-level in the hierarchy as a narrower term.

3. Related terms are checked to ensure that they are not pointing to siblings in the same hierarchy; related terms must point to other hierarchies in the same facet.

4. Reciprocal relations will be added automatically; adding the NT Philips Information system to Digital will automatically add the BT Digital to the Philips Information Systems entry in the lead term file.

5. If a term is added as a BT, NT, or RT and does not preexist as a lead term, the new lead term entry will be added. A minimum record is automatically created with the LT relation and the reciprocals.

This function does not allow deletion of lead terms; to delete or merge lead terms, the UPDL function must be invoked.

2.UPDL: Merge, Delete, and Inactivate Terms

This function allows lead terms to be merged, deleted, or inactivated:

Merge. The Merge function makes the current lead term a synonym to another lead term in the same facet. This function has been particularly useful in capturing corporate acquisitions and mergers. For example, the LT Philips Information Systems could be merged into LT Digital. The effect would be to make all articles indexed to the term Digital retrievable by the term Philips Information Systems.

Delete. The Delete function eliminates the lead term, all its relations, and postings in a facet. If the term exists in another facet, that term will not be deleted. Before deleting, the system displays the postings for the term; the system will allow a term to be deleted even when there are postings, since documents, typically, have multiple index terms, and the system assumes that deletion is being done by an expert who understands the reasons for the deletion and the system ramifications. The system will then display a message to confirm the deletion.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

Inactivate. A third option available for terminology management is the option to inactivate a term. This option makes the term unavailable as an index term through the indexer's workbench but accessible at retrieval. In our application, this function has been very useful for controlling number of documents on any one topic and for managing this knowledge base about available products For example, if a product is announced as discontinued, we will inactivate that product term. No more articles about the product will be indexed into the database, but the preexisting base of documents can be retrieved as before, with full functionality.

3. UPDA: Ongoing Analysis and Application Update

The most important function of the professional information management staff that administers the application is to use the tools wisely. based on their knowledge of the user base and the audience. Ongoing analysis is critical to this process including occasionally repeating the user analysis and query sessions that created the original design.

An application design can be extended through use of the Update Application function. This function allows facets to be assigned to applications in an *a posteriori* manner. This process is analogous to the process of composition in object-oriented systems as the facets can be composed into a public application interface, and facets unplugged, substituted, interchanged and extended without affecting the user.

III. IMPLEMENTATION AND TRAINING

In addition to the tools that support ease of maintenance, dynamic modification, and extensibility, maintenance also requires that a process exist that eases the transfer from design to implementation and provides the staff with firm grounding in principles and practices to apply in their ongoing responsibilities.

We have been using the following staffing model for our vocabulary based applications:

Role	Pre-implementation Role	Post-implementation Role
Knowledge Engineer	Plan, analyze and create facet design; Train specialists	Consulting
Vocabulary Specialist	Build terminology	Maintain terminology Train indexers
Indexer	Initiate content collection and indexing	Ongoing indexing

Based on our experience, an application requires one vocabulary specialist to manage the classification scheme and maintain the reliability of retrieval. This person's responsibilities

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

included determining structure of terminology in this classification scheme (synonym, new term within existing facet, new concept), and providing ongoing analysis of user needs and the representation of these needs in the scheme and content.

Training requires the indexers and vocabulary specialist to master the following skills:

1. Understand the classification design
2. Understand the underlying search system
3. Understand impact of indexing decisions on retrieval
    a) Specificity: index to narrowest term
    b) Exhaustivity: index to all concepts
    c) Precision: indexing ambiguous concepts to the appropriate facet
4. Master indexing tools in the system (Indexer's workbench)
5. Encourage indexers to suggest terminology through the queuing system
6. Train vocabulary specialist in updating of classificatory knowledge base from the queuing system.

During the implementation of TIMS in early 1987, we held weekly meetings of the design team and the application staff; these were eventually dubbed facet meetings. They were held weekly for the first three months of the system as we worked through issues and policies. These meetings were substantial discussions of information retrieval issues and evolved a set of policies and philosophies that are utilized today. They provided the staff with a deep understanding of not only what to do, but why and gave them the necessary depth of background understanding to ensure a high level of professional judgment in indexing and vocabulary decisions.

Given interactive tools that allowed maintenance of the classification scheme, the time required for terminology maintenance was decreased as compared to our previous system. Document throughput increased 100 percent per staff member and we saw improvement in the reliability of the retrieval performance of the system for the end user; this will be discussed next.

IV. EVALUATING MAINTENANCE

To the end user, the system has a high degree of reliability as 4 out 5 search terms are matched no matter into what field in the user interface the user inputs his term.

The specialist who monitors this performance, however, needs to examine performance on each facet to a finer degree of granularity in order to make the system appear reliable.

The user session log, discussed earlier in this paper, in fact, shows which concepts were searched and whether the user found a match in the inverted index. This match of the user term to the term in the inverted index was called index performance.

Interestingly, the data collected from the system indicates that a predictable model is possible based on the level of concreteness of a facet class. I will use the definitions of concreteness/abstraction as used in object-oriented systems, that is a concrete class can be instantiated, whereas abstract

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

classes cannot instantiate themselves. This model can be used to predict when searches are more likely to fail (that is no match on search terms) and to prevent failure.

Terms in concrete classes tend to need modification at a more volatile rate than abstract terms. Product name, for example, is the most concrete object in our facet design. Product name is a single facet containing approximately 1300 computer products. Index performance on this facet regularly degrades to 70-73 percent because of the volume of new product announcements and introductions in software. Vocabulary in this class is updated daily and quickly returns to the 80 percent level.

Company terms also tend to degrade but at a less volatile rate. Finally, Subject terms also degrade but again the levels are much slower. In fact, index performance on subject areas has been as high as 90 percent for some sustained periods. Abstract concepts were also searched less frequently. Nevertheless, abstract concepts also needed maintenance and review to reflect dynamic changes in technology, business dynamics, geographical and political conditions.

Based on this analysis, the vocabulary specialist can manage a planned program to maintain and upgrade the vocabulary, based on review of the facet classes.

V. CONCLUSIONS

Knowledge-based approaches to information retrieval, particularly based on thesauri, have been viewed as expensive largely due to perceived overhead in labor costs. However, when placed in a context of new paradigms, such as semantic networks and object oriented design, a faceted thesaurus system reveals a value-added networked design that can be overlaid onto existing content to hide the underlying complexities and mediate between the user's requests and the content.

In this context, the use of faceted thesauri offers the following advantages in terms of maintaining and managing information retrieval applications:

- Abstracted representation of the application
- Ability to have an eminently modifiable application as the modification is done to the model and not to the content
- Reusability
- Extendibility
- Reliability

The key to success is engaging information retrieval engineers, who can continue to build and extend tools that seamlessly elicit knowledge from the users and that exploit the power of the deeper thesaurus representations and training and motivating information retrieval professionals, who can utilize these tools to the maximum benefit and profitability of the corporation.

George Orwell wrote many years ago in his remarkable essay, "Politics and the English Language" that unclear language was indicative of imprecise thinking. Surely, a nobler goal of IR systems is to create and maintain underlying conceptual representations that encourage precision in language that leads to clearer definition of the questions and ultimately, to excellent, profitable solutions.

ISSN: 2324-9773

Rockmore, M. (1991). Terminology Maintenance for Corporate Information Retrieval. 2nd ASIS SIG/CR Classification Research Workshop, 125-138. doi:10.7152/acro.v2i1.12552

VI. REFERENCES

[ROCKMORE 1] **Marlene Rockmore,** "Computer-aided knowledge engineering for corporate information retrieval." In: Humphrey, S.M. and Kwasnik, B.H. (eds.), *Advances in Classification Research: Proceedings of the 1st SIG/CR Classification Research Workshop. Medford, NJ: Learned Information*, 1991: 137-145.

[FURNAS 2] **G.W. Furnas, et al.** "The vocabulary problem in human systems communication," *Communication of the ACM*, 30(11) (Nov. 1987): 964-971.

[ANDERSON3] **James D. Anderson.** "Information organization based on text analysis (IOTA): Instructional programs for database design." In: Intner, Sheila S. and Hennigan, Jane Anne (eds). *The Library Microcomputer Environment: Management Issues*

[ROCKMORE4] **Marlene Rockmore,** "Structuring a flexible faceted thesaurus record for corporate information retrieval." *Proceedings of the 5th International Conference on Classification Research*, Toronto, Canada, June 24-28, 1991.

ISSN: 2324-9773