

Developing Corporate Taxonomies for Knowledge Auditability: A Framework for Good Practices[†]

Ravi S. Sharma,* Schubert Foo**, and Miguel Morales-Arroyo***

Nanyang Technological University, Wee Kim Wee School of Communication
& Information, 31 Nanyang Link, Singapore 637718,

*<asrsharma@ntu.edu.sg>, **<ASSFOO@ntu.edu.sg>, ***<marroyo@ou.edu>



Ravi S. Sharma is Associate Professor at the Wee Kim Wee School of Communication and Information at the Nanyang Technological University. He was concurrently founding Director (India Strategy) at NTU for a 2-year term. Ravi had earlier been Asean Communications Industry Principal at IBM Global Services and Director of the Multimedia Competency Centre of Deutsche Telekom Asia. Ravi's teaching, consulting and research interests are in multimedia applications, services and strategies. Ravi received his PhD in engineering from the University of Waterloo and is a Chartered Engineer (UK) and a Senior Member of the IEEE. He co-authored a book *Knowledge management: tools and techniques* with Schubert Foo and Alton Chua.



Schubert Foo is Professor and Associate Dean, College of Humanities, Arts and Social Sciences, Nanyang Technological University, Singapore. He received his B.Sc.(Hons), M.B.A. and Ph.D. from the University of Strathclyde, UK. He has published in the areas of multimedia technology, Internet technology, multilingual information retrieval, digital libraries and knowledge management. He co-authored a book *Knowledge management: tools and techniques* with Ravi Sharma and Alton Chua, and co-edited a book *Social information retrieval systems: emerging technologies and applications for searching the Web effectively* with Dion Goh in 2007.



Miguel A. Morales-Arroyo received his Interdisciplinary PhD as a Fullbright Scholar from the University of North Texas. His master's degree in Systems Engineering and undergraduate studies in Engineering were from the National University of Mexico. He is currently a Research Fellow working the area of business models for Interactive Digital Media at the Nanyang Technological University in Singapore. His research interests are related to information management studies, where information, knowledge, and technology are essential to the decision-making process.

[†] **Acknowledgements:** The findings reported in this article are part of the on-going efforts of an informal, irreverent group within the Wee Kim Wee School of Communication & Information at the Nanyang Technological University in Singapore known as the knowledge research factory. The authors are grateful to the funding agency within NTU as well as to their graduate students—Varun Jain, Ekundayo Samuel, Vironica, Mervyn Chia and Angela Wu—for the active engagement. Many thanks are also due to the participating vendors for their useful insights and feedback. The authors gratefully acknowledge the guidance of the editorial staff and the anonymous reviewers in helping to improve the quality of the submission.

Sharma, Ravi S., Foo, Sch., and Morales-Arroyo, Miguel A. **Developing Corporate Taxonomies for Knowledge Auditability. A Framework for Good Practices.** *Knowledge Organization*, 35(1), 30-46. 40 references.

ABSTRACT: The organisation of knowledge for exploitation and re-use in the modern enterprise is often a most perplexing challenge. The entire knowledge management life-cycle (for example – create, capture, organize, store, search, and transfer) is

impacted by the organisation of intellectual capital into a corporate taxonomy or at the least a knowledge map (often incorrectly used interchangeably). Determining the extent to which such an objective is achieved is the focus of what is known as a knowledge audit. In this practice-oriented article, the authors review the fundamentals of creating a taxonomy, the use of meta-data in a necessary process known as classification and the role of expertise locators where the knowledge is not explicit but resides within experts in the form of tacit knowledge. The authors conclude with a framework for developing a corporate taxonomy and how such a project may be executed. The conceptual contribution of this article is the postulation that corporate taxonomies that are designed to facilitate knowledge audits lead to greater organizational impact.

1. Organising corporate knowledge

The organisation of knowledge resources is hardly a novel undertaking. Many in the western world (cf. Woods 2004) attribute to Aristotle the first attempt at information organization, to the Swedish scientist Linnaeus the first system for categorising the natural world, and to the Melvil Dewey the first library cataloguing scheme of medical and scientific knowledge. However, the ancient civilizations within China, India and the Mid-East in fact organized their knowledge, particularly related to philosophy, government and medicine, carefully for the purpose of transfer and re-use. The ancient libraries of Alexandria (circa. 2000 B.C.E.) which comprised world-class methods in their time for collection, storage and retrieval, were predicated on the realization that a system for organizing knowledge was the key to understanding an existing body of knowledge and its repeated exploitation. Today's knowledge-driven economy demands such a strategy more than ever. From public archives and libraries to corporate repositories and individual collections, knowledge that is not organized is often rendered worthless by the sheer velocity of business decisions that need to be made (Nonaka and Takeuchi 1995; Davenport and Prusak 1998; Senge 1999). In other words, the entire study of organising knowledge into a systematic classification of hierarchical categories, which may be labelled and subsequently searched, is all about fulfilling the cliché "knowing what we know". A corporate taxonomy is the interface for all such activity.

An IDC White Paper by Feldman and Sherman (2001) examined industry practices and indeed confirmed the worst fears of practitioners:

Intranet technology, content and knowledge management systems, corporate portals, and workflow solutions have all generally improved the lot of the knowledge worker. These technologies have improved access to information, but they have also created an information deluge that makes relevant information more difficult to find.

The White Paper concluded that knowledge workers need unified, universal access to all information, but they only need that subset of the information base that actually solves the problem at hand and the requisite expertise or skill that would apply it. It identified the search costs of tediously retrieving required information, the cost of repeated knowledge creation (i.e. not re-using existing knowledge components), and the opportunity cost of not applying known knowledge as "the high cost of not finding information."

Knowledge in the modern organization hence has elements of variety and is incongruous and heterogeneous in nature in terms of creation, storage and re-use. In academic research as well as trade forums, the understanding of what constitutes knowledge is often debated because of the multidimensionality associated with it. Only when the knowledge is captured and organised into proper formats can it be made accessible and put to further use. In effect, capturing knowledge is of little use if it is not stored in such a way that it can be understood, indexed, accessed easily, cross-referenced, searched, linked, and generally manipulated for maximum benefit of all members of an enterprise. Hence the organisation of knowledge plays a critical role throughout the knowledge cycle.

One aspect of an organisation's intellectual capital is collective knowledge, which can be viewed in terms of information within the context of the organization. It first involves the process of acquisition from personal knowledge and existing organisational information resources, then sharing and subsequently action-taking by knowledge workers—resulting in new information being added back to the organisational memory. The collective knowledge of an organization is diffused through several processes of knowledge acquisition, sharing, and action initiated as a consequence of new knowledge being created. This flow is illustrated in Figure 1.

Nonaka and Takeuchi (1995) concluded that "The Knowledge Creating Company" cannot create knowledge on its own without the initiative of the individual and the interactions that take place between individuals and groups. The effective design of the organiza-

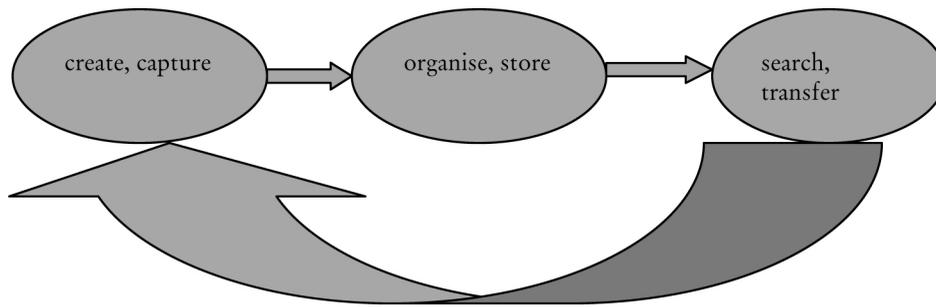


Figure 1: *Knowledge processes in organisations.*

tion makes it possible for the knowledge content of many of these interactions to be captured. Thus, personal and collective cycles of knowledge creation and use are inter-related and as we shall see later in this article, corporate taxonomies serve as useful intermediaries.

Figure 1 also shows some of these fundamental knowledge processes within an organization. While there are several models for knowledge cycles, for example Birkinshaw and Sheehan (2002)—capture, store, transfer; Feldman and Sherman (2001)—create, distribute, manage, retrieve, apply; Mohanty and Chand (2005)—create, capture, organise, store, use; and the figure above is nevertheless a reasonable synthesis described in Foo et al. (2007). Knowledge workers create intellectual capital in the course of their work, or more precisely, as a result of their work, and the extent to which this may be captured is a measure of the organisation's standard procedures or structural capital. The knowledge infrastructure within the organisation also drives how value in the form of re-exploitable knowledge is organised and stored. The competitive advantage of the organisation lies in the speed and precision with which its knowledge assets are searched and transferred as and when opportunities arise.

At yet another layer of abstraction, as these processes continue in perpetuity within an organisation, two types of knowledge flow through—explicit, which is codifiable, and implicit or tacit, which is not codifiable but resides as expertise in the minds of knowledge workers. The explicit vs. tacit knowledge dichotomy is often held as being analogous to structured (databases, web portals) vs. unstructured knowledge (e-mail, blogs). However, this analogy is misplaced as both explicit as well as tacit knowledge may be structured or unstructured. In other words, explicit knowledge can be well articulated, especially in the written form, while tacit knowledge might be much less so. While Nonaka and his co-authors were

not the first to popularise these terms, their work nonetheless pointed to the stark differences in which the two types of knowledge are created, captured, organised, stored, searched and transferred for re-use. Hansen et al. (1999) concluded that whilst explicit knowledge sharing, which they called codification, comprises the most frequent in terms knowledge transactions, most value was derived from tacit knowledge sharing, which they called personalisation. This was primarily because tacit knowledge was less common, more difficult to replicate and therefore served as a competitive advantage. They further suggested that superior IT infrastructure such as web portals and wireless LANs were only suited for codification; personalisation requiring opportunities for both traditional as well as IT-based communication and collaboration.

Several scholars have addressed the major challenges facing the sharing of knowledge within an organization. Nonaka and Takeuchi (1995) have also suggested the requirement of a shared workspace that will facilitate knowledge sharing across the organisation. Zack (1999), in an empirical study of knowledge strategies in over 20 knowledge intensive firms, concluded that the misalignment of business and knowledge strategies was a fairly common cause for poor performance—for example, businesses that did not know when value was being created and did not organise themselves to continually exploit this value. Gupta and Govindarajan (2000) performed a more extensive field investigation of knowledge flows within multinational corporations and found that the mismatch between the source and target of knowledge flows—for example, if knowledge was not transferred appropriately—led to severe ineffectiveness. Hence the fundamental motivation for building a corporate taxonomy is the realization that knowledge is of little value unless it can be shared and re-used (as opposed to re-discovered) when opportunities for exploitation arise. There is considerable agreement

that such taxonomy is the basis for interactions and organizational learning (Cheung et al. 2005; Davenport and Prusak 1998; Gilchrist and Kibby 2000; Gilchrist 2001; Nonaka and Takeuchi 1995; Potter 2001).

In the realm of digital resources, Noy and McGuinness (2001) suggest that taxonomies are particularly useful: (1) to share common understanding of the structure of information among people or software agents; (2) to enable reuse of domain knowledge; (3) to make domain assumptions explicit; (4) to separate domain knowledge from the operational knowledge; (5) to analyze domain knowledge.

The remainder of this article describes the building blocks of a corporate taxonomy and presents a synthesis of best practices for developing it so that it may be continually updated so that it stays useful. The development of good and relevant taxonomy coupled with a supportive knowledge management platform and environment is an integral aspect of remaining competitive in the face of the continual deluge of (particularly, digital) knowledge.

2. Principles of corporate taxonomies

The word taxonomy is derived from Greek words (*taxis* + *nomos*)—*taxis* is arrangement and *nomos* law—and can be conjugated to mean “the science of classification.” The Swedish scientist Carl Linnaeus (1707-1778) was perhaps the first to use the idea of taxonomy to classify the natural world. From its origins in the classification of living things, the idea of taxonomy now has universal applications in grouping knowledge so that it can be systematically developed, stored and re-used. In the information sciences, the study of corporate taxonomies has been a subject of considerable and longstanding interest among researchers (Cheung et al. 2005; Geisler 2006; Gruber 1993; Noy and McGuinness 2001; Saeeh and Chaudhry 2002) as well as practitioners (Conway and Sligar 2002; Delphi 2002; Ernst & Young; Gilchrist and Kibby, 2000; Gilchrist 2001; Greif 2001; Lehman 2003; Pepper 2000; Potter 2001; Woods 2004).

A modernist definition of a corporate taxonomy may be found in Lehman (2003 emphasis original):

A taxonomy is a *subject map* to an organization's content. [It] reflects the organization's purpose or industry, the functions and responsibilities of the persons or groups who need to access the content, and the purposes / reasons for accessing the content.

Hence a corporate taxonomy may be viewed as a conceptual map, an information access tool, and a communications and training device at the same time, providing history, expertise and inside information that can assist every business activity. Naturally, as with any other information access tool, a taxonomy has to serve special requirements and purposes before it is developed and exploited. Other classical perspectives widely accepted in the literature include:

1. A taxonomy is a creation of structure and labels to aid location of relevant information. A closer definition might be the arrangement and labelling of metadata to allow primary data or information to be systematically managed and manipulated (Gilchrist and Kibby 2000).
2. A taxonomy is a hierarchical presentation of information that represents a specific knowledge domain. It includes several sub-topics that can contain two types of relations, namely, hierarchical relations where one category is viewed as being above another category, and non-hierarchical relations using links that indicates that a certain category is related to another category. Applications are the navigation tools available to help users find information (Graef 2001).

Taxonomies are often referred to as conceptual knowledge maps in knowledge management. However, they have some distinguishing characteristics in the sense that: (1) they support structure, content and applications (navigational tools); (2) they are customised to reflect the language, culture and goals of particular organisations; (3) they are often created using a combination of human effort and specialised software; they may refer to disparate information resources such as e-mail, memoranda, documents, books, part of books, reports and URLs; (4) they are usually created by multidisciplinary teams; and (5) they are part of a process so that they are constantly refined and updated.

In either case, corporate taxonomies and knowledge maps may be considered the fundamental basis for knowledge sharing in the organization and specific processes such as create, capture, organize, store, search and transfer) in the organization. They provide a common understanding within the organization that link to the knowledge cycle.

Corporate taxonomies are dynamic and need constant refinement and update because organizations need to adapt to a changing environment (competition, threats, etc.), which forces them to modify

their knowledge flows. Grey (1999) suggests posing the following key questions to the knowledge workers within an organization in order to ascertain the major knowledge flows: (1) What type of knowledge is needed? (2) Who provides it and how does it arrive? (3) How is it improved and re-used? (4) What happens to new knowledge that is created? (5) What prevents the organisation from doing more, better, faster? (6) How can knowledge flows (therefore) be improved?

This is in effect what is known as a knowledge audit (Cheung et al. 2007; Liebowitz et al. 2002; NLH 2005), which involves identifying what knowledge is needed, what knowledge already exists, where the gaps lie, who needs the knowledge, and how it will be used. Hylton (2002) more formally states that the knowledge audit (K-Audit) is an assessment of how the sum of explicit as well as tacit knowledge within an organisation is exploited throughout the knowledge-cycle and the people and business processes add to such knowledge. More specifically (Hylton 2002, 2):

The knowledge audit process involves a thorough investigation, examination and analysis of the entire 'life-cycle' of corporate knowledge: what knowledge exists and where it is, where and how it is being created and who owns it. It measures and assesses the level of efficiency of knowledge flow. From knowledge creation and capture, to storage and access, to use and dissemination, to knowledge sharing and even knowledge disposal, when the organisation is no longer in need of particular elements of explicit or codified knowledge. With respect to people, the K-Audit measures the efficiency of transfer of tacit knowledge skills, when particular skills or expertise is no longer needed.

It may be viewed as the knowledge management equivalent of the requirements determination phase undertaken during traditional systems analysis and design.

During the course of a knowledge audit, there is a critical first step which leads to the creation of a knowledge map--a visual representation of an organisation's knowledge. Technically, a knowledge map is a logical abstraction of a corporate taxonomy, which includes implementation details such as how knowledge assets are to be captured and indexed. A knowledge map, at first cut, reveals possible answers to the key questions of a knowledge audit (outlined above).

There are two recommended approaches to knowledge mapping (NLH 2005): (1) map knowledge resources and assets, showing what knowledge exists in the organisation and where it can be found; and (2) include knowledge flows, showing how that knowledge moves around the organisation from source to target. In both cases, the key is a diagrammatic schemata of corporate knowledge of the explicit as well as tacit nature and an accompanying realisation of the value-added during the course of the knowledge flows. This may be derived from well-known techniques such as process maps, class diagrams, use cases and organisation charts.

Building a knowledge map looks deceptively simple but perhaps requires more effort and resources than any other phase of developing a corporate taxonomy. It is a profound, soul-search that involves the highest level of strategic management and domain expertise to make judgments on fundamental business and knowledge strategies. One technique for deriving a knowledge map involves the use of the so-called Boston Box suggested by Drew (1999). Figure 2 shows four quadrants of the Boston Box for analysis of a complete coverage of an organisation's

Knowledge Awareness	1. What We Know We Know	2. What We Know We Don't Know
	<i>Emphasis:</i> knowledge sharing, access and inventory. <i>Tools:</i> e.g. benchmarking, communities of practice.	<i>Emphasis:</i> knowledge seeking and creation. <i>Tools:</i> e.g. R&D, market research, competitive intelligence.
	3. What We Don't Know We Know	4. What We Don't Know We Don't Know
	<i>Emphasis:</i> uncovering hidden or tacit knowledge. <i>Tools:</i> e.g. knowledge maps, audits, training, networks.	<i>Emphasis:</i> discovering key risks, exposures and opportunities. <i>Tools:</i> e.g. creative tension, audits, dilemmas, complexity science.
Knowledge Content		

Figure 2: Building a knowledge map. Source: Drew (1999, 134)

knowledge capital. Quadrant 1 asks what the core competencies of the organization are. Quadrant 3 addresses the unexploited seepage in its knowledge capital repository. Quadrant 2 takes the organizational learning impetus which seeks to position the organization to execute its strategic plans for growth. Quadrant 4 refers to the blind spot of hidden opportunities and threats that may not be (as yet) apparent within the organisation’s leadership. Daunting as this analysis may seem, it does not represent a paradigm shift. The point being made in this article is that the organisation of knowledge in the form of a corporate taxonomy carries with it criteria for evaluating possible gaps as well as leaks that need to be plugged. Drew (op. cit.) had captured some of these issues for some time now and the knowledge management community has since developed an entire repertoire of tools for each of these quadrants (cf. Foo et al. 2007 for a textbook coverage of many of these tools).

What then makes a corporate taxonomy effective, extendable and practical? Table 1 below, which is a compilation from Gilchrist (2001), Graef (2001), Lehman (2003) and Woods (2004), offers eight perspectives or families of taxonomic elements, which apply to an organization, although more perspectives do not necessarily translate to greater business effectiveness.

Industry Segments - Marketing / Positioning / Competitive Intelligence Perspective; Industry Segments may overlap with Products and Services.
Organizational Functions - the organization breakdown of a business or organization by function or responsibility
Business Relationships - the intensities and types of other companies or organizations a business deals with; including customers, vendors, regulators, associations, partners etc.
Business Issues & Events - economic, legal, M&A, regulatory, environmental, labour, safety, other government interfaces, etc.
Products & Services - products sold; MRO materials; indirect services, direct materials & services purchased.
Technologies - applicable to the industry or industries in which the firm participates. Basic or applied sciences are also included as appropriate.
Geography – referring to location, particularly region or jurisdiction.
Document or Record Types - this perspective provides valuable reduction of results based upon the document’s purpose and its connection to the information need.

Table 1: *Perspectives of taxonomic elements.*

As a guide to content, a taxonomy has multiple entry points (such as business functions or product types), and will have the same element (lowest level class). The consensus on what characterizes useful elements of corporate taxonomies is the following:

1. Elements are precise and do not overlap—the closer to proper named elements at the lowest level, the better.
2. Elements are independent of the type of content, and the organization structure (be they digital or multilingual or distributed data).
3. Elements reflect the access needs and expectations of every constituency inside or outside the organization; and,
4. Industry standards (such as UN/SPSC for products and services. IBSN and ISSN for published documents) are recognized and applied whenever possible.

To conclude, it is clear that corporate taxonomies have indispensable roles in the organization of business knowledge. The bottom-line for a good taxonomy is whether or not the knowledge sharing process is facilitated. There are methods and tools which help verify and validate that such sharing is indeed taking place. In the next section of this article, some of these building blocks are described.

3. Building blocks for corporate taxonomies

Some of the fundamental challenges on how knowledge (explicit as well as tacit) may be incorporated into a corporate taxonomy are addressed by a variety of techniques drawn from the domains of computer and information science. These include the concepts of directories of domain expertise; classification and clustering; indexing, tagging and the use of meta-data. Classification is the technique used to organise a body of knowledge assets that reside within an organisation. It is supported by meta-data which are used as keywords or descriptors for indexing, storing and searching knowledge assets.

The word “metadata” is derived from Greek and Latin words (Greek: Meta + Latin: Data). Since Meta means along with, next or after, metadata is data about data itself; it contains information about other nuggets of information or knowledge. Metadata is documentation about documents and objects; they describe resources, indicate where they are located, and outline what is required in order to use them successfully. In creating corporate taxonomies, a

practitioner makes use of metadata to describe documents and other resources thereby enabling a richer means of defining the context of the resource and to provide more information access points to support information query and retrieval operations. This is a technique known as “tagging” in contemporary parlance and is very relevant to the idea of describing knowledge assets (whether codified or residing within experts) and cataloguing them for storage and search. In this section, we discuss some of these.

3.1 Classification, clustering and cataloguing

A recent IDC study (Gantz et al. 2007) estimated that the “the digital universe” equals approximately three million times the information in all the books ever written, or the equivalent of 12 stacks of books, each extending more than 93 million miles from the earth to the sun. The amount of information created and copied in 2010 will surge more than six fold, from 161 to 988 exabytes, a compound annual growth rate of 57%, nearly 70% of which will be generated by individuals. Considering the exponential growth of information and knowledge, particularly in the digital and Internet domains, it makes sense to classify content in some order so that search and retrieval becomes manageable (Feldman and Sherman 2001).

The study of classification is hence re-emerging from a hiatus after the pioneering work of Ranganathan, the reknowned classificationist, who drew much inspiration from the work of Dewey and other pioneers in order to formulate rules on how documents might be classified so that they could be retrieved with sufficient specificity when needed (cf. Ranganathan’s web repository for a retrospective). Several contemporary studies have shown that using either natural language, such as keyword searching, for defined metadata fields, or a controlled vocabulary of subject headings and browsing thesauri, result in superior knowledge retrieval and re-use when the knowledge base is classified or clustered for effective search (Delphi 2002; Feldman and Sherman 2001; Stratify 1997; Williamson 1997). This would be particularly the case when end user tagging is enabled with the use of controlled vocabularies and multi-faceted taxonomies are constructed to facilitate the search effort.

The idea of classification is frequently interchangeably used with the term “taxonomy” but is semantically different—a classification may lead to a

taxonomy (usually a visual representation) but is always described in terms of a method or scheme that groups a set of entities such that elements within a group (or class or cluster) are more similar to each other than elements in different groups. Clustering is the technical term used when these groups (or clusters) are non-overlapping. In both cases, the organisation may be hierarchical and multi-faceted. The proliferation of Internet content and the requirement for optimised search engines has brought this ancient science into yet another domain that sustains its relevance. Hlava and ven Eman (1999) suggest that most cataloguing schemes, many of which within the English speaking world originated from the UK and US library communities, use one of 3 methods for classification: original text or idea; existing vocabulary or topic; or a combination. Given the volume of content to be organised, today much of this has to be automated (and continually refined manually) using idea extraction, keyword counts, or adaptive algorithms that discern document context.

A classification typically also results in what is known as an ontology. Ontology is the term referring to the shared understanding of some domains of interest, which is often conceived as a set of classes (concepts), relations, functions, axioms and instances. In the knowledge representation community, the highly cited definition is adopted from Gruber (1993, 199):

An ontology is a formal, explicit specification of a shared conceptualisation. Conceptualisation refers to an abstract model of phenomena in the world by having identified the relevant concepts of those phenomena. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine readable. Shared reflects that ontology should capture consensual knowledge accepted by the communities.

Noy and McGuinness (2001, 1) offer a more IT-centric definition: “An ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them.”

It should be noted that classification and cataloguing are active, dynamic activities that are never complete, much like arranging the folders of one’s desktop. Hence, there is a continuous process of ap-

pending, updating, pruning, and so on, to keep it relevant and useful. A classification is hierarchical and multi-faceted in order to support multiple perspectives such as user profiles, applications or data models. When recognised that small corporate taxonomies comprise up to a thousand knowledge resource items and large ones greater than twenty thousand (Woods 2004), it is easy to understand why an automatic processor of metadata (often times extractor) and classification rules is necessary. This is discussed next.

3.2. Automatic classifiers and other tools

Automated classifiers typically use various proprietary clustering methods for analyzing the content of each knowledge asset and creating concept folders that contain related items; organize the concept folders hierarchically based on their interrelationships, and sort each document into one or more concept folders that describe the document in whole or in part (cf. Stratify 2006). With codified knowledge, such clustering is typically based on a statistical analysis of all words within a document and extracting keywords or context. Clustered documents are placed within concept folders that contain documents that are “close” or similar, to each other according to a chosen similarity measure which attempts to match the term lists or subject headings (controlled or otherwise) of the documents to a cluster. During search and retrieval, a centroid or representative from each cluster is matched with the requirements of the query, and the entire cluster is retrieved if held similar in the expectation that they are similar and hence must be equally relevant. In this manner the search space is reduced.

Functionalities aside, classifiers and taxonomy builders exhibit a commonality of traits. For example, most tools should be able to measure the depth of a person’s experience on a particular topic, based on relevant prior experiences, roles in projects, peer recognition and other measurements and track relevant end user activity, identifying those individuals who may be best suited to address the task. They should also incorporate automatic learning functions where the program becomes more accurate with continuing usage, hence removing the need to rely on manual user feedback.

They typically allow for customisation such that it is flexible enough to accommodate the different knowledge management needs of different environments in which it will be deployed. To illustrate, in

order to assess the needs of two different kinds of industries, say the healthcare and manufacturing industries, the tool should be flexible enough to cater to both business contexts and vocabularies – categorise medical treatment and prescriptions as well as raw materials, equipment and facilities. Another desirable property of such tools is scalability. It must become more accurate when the organisation gets larger and not see drastic reductions in accuracy when dealing with increasing knowledge resources. Last but not least, if the tool is easy to use and adopted very quickly, it does not require extensive hours of training for employees in order for adoption to take place.

The past decade has seen a proliferation of tools for developing corporate taxonomies, many of which exhibit some or all of the following functionalities:

1. Classification of digital resources, documents, databases, directories, etc.
2. Tagging of knowledge resources during content creation.
3. Site navigation and creating categories for discovery of information through browsing.
4. Identification and retrieval including cross searching of different resource bases.
5. Personalisation and delivery of information.
6. Visualisation of knowledge maps.

One of the leading and authoritative web site on search engines, *SearchTools.com* (<http://www.searchtools.com/info/classifiers-tools.html>) provides a list of valuable and up-to-date resources on “Tools for Taxonomies, Browse-able Directories, and Classifying Documents into Categories.” Vendors specialise in different subsets of the above functionalities and there is really no ubiquitous solution that may be adopted by the organisation developing a corporate taxonomy. The following are some of the leading vendors of taxonomy tools which support the key functionalities and together possess a dominant market-share.

Company	Tool	URL
Autonomy	Expertise Finder	www.autonomy.com
Convera	SAAS Semantic Search	www.convera.com
Cadenza	Knowledge-LEAD	www.cadenzainc.com
Divine	Athena and MindAlign	www.divine.com

Company	Tool	URL
Entrieva (formerly Semio)	Semio Taxonomy and Tagger	www.entrieva.com
Groove Networks (now part of Microsoft)	Groove Virtual Office	www.groove.net
Hummingbird	Hummingbird Knowledge Server	www.hummingbird.com
Kamoon	Kamoon Connect	www.kamon.com
IBM (Lotus)	Intelligent Miner, Discovery Server and K-station	www.lotus.com
Microsoft	Sharepoint	www.microsoft.com
MITi	Readware Knowledge Workshop	www.readware.com
Quiver	QKS Classifier	www.quiver.com
Sopheon	Organik	www.sopheon.com
Stratify	Stratify Discovery System	www.stratify.com
Verity (now part of Autonomy Group)	Verity Knowledge Organiser & Ultraseek Advanced Classifier	www.verity.com
XBRL	XBRL Taxonomy Builder	www.xbrlsolutions.com

Table 4: Taxonomy tools in action.

An examination of the approaches taken in many of these tools confirms that vendors disagree over the choice of techniques and approaches to the classification of information and knowledge. The approach that suits any individual organisation will naturally depend on a mixture of the business requirement, the type, volume and volatility of the information to be managed, the skills available in-house and the time and resources to be expended. But, as Woods (2004, 7) suggests:

The greater the volatility of the information and the categories to be used, and the less the in-house experience available, the more attractive an automated solution will be. For organisations with extensive experience of in-house

taxonomy design, greater manual control may be preferred.

Although these tools serve the common purpose of organising knowledge and support for information seeking and discovery, they all have their own subtle characteristics. The question is not about “natural language or controlled vocabulary,” rather the solution is “natural language and controlled vocabularies” to complement each other for information seeking. Several of these vendors have provided a range of organisation tools such as taxonomies, subject directories, subject hierarchies, topic maps and knowledge mapping tools in order to assist in the effective management of content in the organisation’s Intranet, Internet or knowledge portals. These tools provide hierarchical or decision tree structures with considerable variation in complexity and sophistication.

It is also apparent that many of these tools relate to industry-specific taxonomies. The tool should be such that it should preferably be able to “understand” user requests and derive similar requests in order to return more relevant results. “Knowing what you know” is the central objective of a corporate taxonomy and hence taxonomy building tools.

In order to select appropriate taxonomy tools that meet specific features or functions, there needs to be a basis for comparison. Table 5 is a list of functionalities and types of features synthesised from the literature. A matrix of the availability of these functionalities and features within some of the leading taxonomy tools is given in Foo et al. (2007) with the objective of rapid prototyping (and subsequent evolution) of an organisation’s corporate taxonomy.

Classification Methods: Rule-based, Training Sets, Statistical Clustering, Manual
Classification Technologies: Linguistic Analysis, Neural Network, Bayesian Analysis, Pattern Analysis/ Matching, K-Nearest Neighbours, Support Vector Machine, XML Technology, and others
Approaches to Taxonomy Building: Manual, Automatic, Hybrid
Visualization Tools Used: Tree/Node, Map, Star, Folder, None
Depth of The Taxonomy: > 3 levels
Taxonomy Maintenance: Add/Create, Modify/Rename, Delete, Reorganisation/Re-categorization, View/Print
Cross-referencing Support:
Import/Export Taxonomy:
Import/Export Formats Support: Text file, XML format, RDBS, Excel file, Others

Document Formats Support: HTML, MS Office document, ASCII/text file, Adobe PDF, E-mail, and Others
Personalization: Personalized View, Alerting/Subscribing
Product integration: Search Tools, Administration Tools, Portals, Legacy Applications (e.g. CRM)
Industry-specific Taxonomy: Business, News, Medicine/Pharmaceutical, Legal, Military, Biotech, Technology, Insurance, Government, Any industry
Access Points to the Information: Browse Categories, Keywords, Concepts & Categories Searching, Topics/Related Topics Navigation, Navigate Alphabetically, Enter Queries, and others
Multilingual Support:
Product Platforms: Window NT/2000, Linux/Unix, Sun Solaris system, and others

Table 5: *Functionalities of taxonomy builders and classifiers.*

The availability of a particular function and existence of a specific feature makes a significant impact on the development of a corporate taxonomy. Besides the obvious requirements fit in terms of Product Platform and Integration, GUI Design, Access Points, Import / Export and Multilingual Support, and so on, there are other nuanced considerations. For example, the easy part of taxonomy implementation is the actual assignment of rules, either from a written “cookbook” for human classifiers or software. There are basically two approaches to software implementation: those packages that accept and execute rules, and those packages that use statistical techniques (“content like this”) to construct their own rules. While the statistical vendors can fairly claim that their products avoid the “rigor” of classification definition, they also miss the precision of rules and the result vagueness that accompanies vague rules. But this may well be the “lesser of the two evils” in instances where there is a voluminous mass of legacy documents and knowledge items. Hence a “middle way” is the possible use of a statistical or learning type classification approach after having classified a large and varied group of document-records, with a manually defined rule-based technique. In either case, the resulting classification is k-auditable and may be changed and improved with time.

Ongoing maintenance of classification rules is another significant but tractable activity, either in-house or from third party specialists. Rule-based classifier software will require very low maintenance. Statistical classifier software needs regular recalibration to address new or modified classification rules. Again, the dynamic environment in which the

organisation operates dictates the more suitable approach. In closing, the selection of an appropriate tool is perhaps a matter of efficiency to the trained and experienced organisation but effectiveness as well to many taking the first steps towards a corporate taxonomy. The final section will note that the use of appropriate (and k-auditable) tools is a salient point in the continual evolution of the corporate taxonomy in support of a learning organisation.

3.3 Expertise locators

Where knowledge is not explicit, pointers to their (tacit) sources are needed. To this end, expertise locators, euphemistically known as electronic yellow pages or skills directories, have emerged in the corporate world as means to identify sources of specific expertise and skills. These can be classified or categorized to be included in corporate taxonomies. Grey (1999) suggests that corporate yellow pages often incorporated into taxonomies give the highest return on investment and it is easy to understand why. The nature of knowledge work is such that most professionals turn to their peers as the first resort (Davenport and Prusak 1998).

Expertise locators are basically lists of the expertise of individuals, usually very highly regarded subject matter experts, in an organisation. Expertise refers to special skills or knowledge of subject embedded in individuals. Hence, an expertise locator becomes a de facto directory of individuals in an organisation with their contact details, designation, and name, along with details about their knowledge, skill sets and experiences. They identify experts in an organisation that are authorities in specific knowledge domains. A subject matter expert may have different levels of expertise in different topics and all this goes into the listing. Such expertise is typically self-reported (on the basis of job responsibilities or qualifications) but sometimes discerned through formal accreditation or using social network analysis (SNA). SNA is a complex yet highly rewarding activity in knowledge organisation. For example, by observation of email trails or posts on discussion groups, it becomes apparent who junior team members turn to for various types of advice during projects or within a community of practice—this is where the re-useable tacit knowledge within the organization lies.

The National Electronic Library for Health (2003), part of the National Health Services of the United Kingdom (www.nhs.uk) which manages a large number of diverse medical facilities and special-

ists, suggests the following tags as meta-data for the identification of (tacit) expertise within an organisation:

Name; Job title; Department or team; Brief job description – current and past; Relevant professional qualifications; Uploaded CV in standard format; Areas of knowledge and expertise selected from a pre-defined list of subjects and terms with self-reported rankings - “extensive”, “working knowledge”, “basic”; Main areas of interest; Key contacts – both internal and external; Membership of communities of practice and other knowledge networks; Personal profile; Photograph; Contact information.

Among other critical success factors for the design, development and maintenance of such an expertise listing, the NELH lists currency, auditability and evolution as the most vital.

Conway and Sligar (2002) remark that capturing such information within the organization may be a little more controversial than doing it in an external Internet environment. The corporate culture will likely influence or constrain the reaction to and acceptance of gathering, mining, manipulating, storing, and making use of what may be regarded as “personal” information. Nevertheless, this approach can be one of the best ways of connecting colleagues and team members and for the corporation to begin to learn more about the tacit knowledge residing within its people.

Expertise locators are useful for large organisations which are spread geographically. While they are technologically simple to develop, it can effectively assist organisations to “know what they know”. As the knowledge community increasingly recognises that the best channel for knowledge sharing is communication between co-workers and the most effective networking protocol is collaboration, an expertise locator provides the identification service to generate the connection and to support relationships building between individuals. It provides a means to access a key component of the organisation’s intellectual assets. Through such connectivity, it supports learning, growing and capacity development. Of course, finding the right person is a means to an end, not an end itself, and the organisation culture and incentives must support knowledge creation, capture, transference and re-use.

Expertise locators are also useful to locate relevant knowledge sets particularly due to the increas-

ing complexity of tasks and work that requires the combined skills of various experts often situated across the globe. The specific benefits of expertise locators are that they are easily implemented; help in locating the knowledge source; allow people to interact and learn efficiently by providing a platform; save time and effort in not “reinventing the wheel.”

As the organisation’s base of information about its subject matter experts grows, so will its ability to harness its knowledge capital. With the information about an individual’s browsing, searching, and posting habits; metadata from the content that has been created; the projects worked on; community involvement; and, preferences and patterns in the data accessed within the corporate intranet—the Enterprise Knowledge Portal (EKP) can manage and describe the knowledge worker’s experience base in the same way it shows the explicit contents of the repository. With the aforementioned understanding of the building blocks of corporate taxonomies, we conclude this article with a framework for building taxonomies in the final section.

4. Framework for taxonomy building and knowledge-auditability

To recap, corporate knowledge taxonomies play a critical role in knowledge management and the exploitation of corporate intellectual capital with direct impact on the organisation’s performance and growth. They provide a platform that assists employees to seek knowledge residing within the organisation and sometimes beyond, facilitate collaboration and interactions, and support the iterative process that may be necessary to develop new knowledge. At the onset, the corporate taxonomy can provide a “knowledge map” to enable navigation of and access to the intellectual capital of the organisation. Corporate taxonomies work at the level of information management by connecting people to documents, and at the knowledge level by connecting people to people (Gilchrist and Kibby 2001). It is critical that the design of such taxonomy must allow itself to be audited so that the impact of knowledge management may be assessed.

Based on the discussion in the previous section, we now present a framework comprising four major steps required to build and maintain a corporate taxonomy. These are:

1. Conduct a knowledge audit that results in a first-cut knowledge map;

2. Perform a more-intensive requirement analysis that formalises tagging, storage and search;
3. Select the appropriate set of tools that facilitates the creation of the taxonomy and the classification process, and,
4. Refine and update the taxonomy as the organisational context changes.

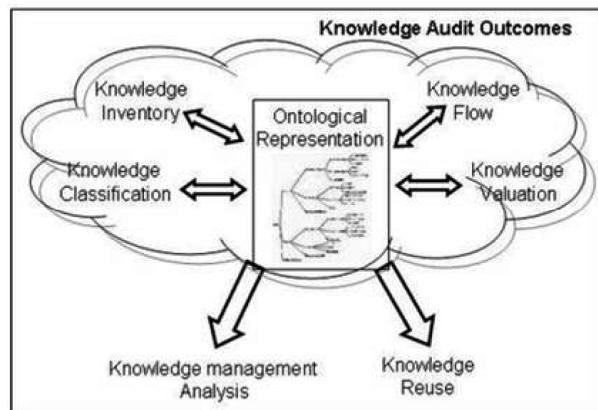


Figure 3: Presenting knowledge-audits as ontologies. Source: Perez-Soltero et al. (2006, 47).

The nexus between a knowledge audit and creating a corporate taxonomy is indeed a symmetric one. Perez-Soltero et al. (2006) have presented a methodology which results in a knowledge inventory comprising knowledge maps and knowledge flows that identify inefficiencies reflected in duplication of efforts, knowledge gaps, knowledge barriers and knowledge-bottlenecks. They show the feasibility of using ontologies as representational schemas in order to formally present the results of a knowledge audit which address the problems of knowledge leakage and additionally the benefits of re-using valuable knowledge. Figure 3 is an abstraction of their methodology. Whilst such an approach makes sense from the point of schematic representation of the results of a knowledge audit for the purpose of communicating with stakeholders, we argue that the opposite is even more critical. Cheung et al. (2007) in fact adopt such a methodology for knowledge audits but do not formally specify the link between building a knowledge taxonomy and auditing the repository of knowledge within an organisation. The design of a corporate taxonomy must necessarily take into account the ease of auditing knowledge inventories, flows, leakages and gaps, and must facilitate the continual growth of the knowledge or learning organisation.

Synthesising the numerous approaches from research and practice on taxonomies, classification and ontologies, we have developed a knowledge-cycle

driven framework for first understanding and then developing corporate taxonomies for effective exploitation of an organisation's valuable knowledge resources. The net result of such integration is a dynamic and relevant corporate taxonomy—what Gruber (1993) calls a “portable ontology.”

Figure 4 is such a framework which incorporates the key concepts discussed with respect to developing corporate taxonomies. The framework also maps classical knowledge flows (create, capture, organise, store, search, transfer and re-use) and inventories (documents, expertise directories, learning communities) with the design of the corporate taxonomy using automated builders and knowledge mobilisation (search and re-use). In such a scenario, it is obvious that neither the knowledge flows nor the inventories would be static. Hence it becomes crucial that a methodology for creating knowledge taxonomies adequately supports the notion of continual growth and consequently, auditability. From this framework, we have derived the four major steps that need to be undertaken to build corporate taxonomies with the design objective of knowledge auditability.

It is worth repeating that although many practitioners (cf. Hylton 2002; Lehman 2003; Pepper 2000) refer to corporate taxonomies and knowledge maps interchangeably, it should be clear by now to the discerning reader that they are indeed distinct in the level of detail they carry. At its simplest, a taxonomy is a rule-driven hierarchical organisation of categories used for classification purposes with the appropriate subject headings and descriptors. However, such a simple definition hides the many challenges to be faced in building and maintaining an effective and usable taxonomy for the organisation (Woods 2004). Corporate taxonomies are particularly used by the various enterprise information systems to permit instant access to appropriate information, where there are voluminous data, and information needs to be managed carefully. Knowledge maps are at best visual aids that help the search and retrieval process. They are the result of what has been described in an earlier section as a knowledge audit – the technical details of which are beyond the scope of this article.

Nevertheless, *Step One* of developing a corporate taxonomy is therefore to conduct a knowledge audit which clearly identifies the creation, capture, organization, storage, search, transfer and re-use of knowledge in the critical business processes of the organization. Conceptually, this is indeed the most complex step as the design is not easily amenable to validation. Grey (1999) suggests that the key to devel-

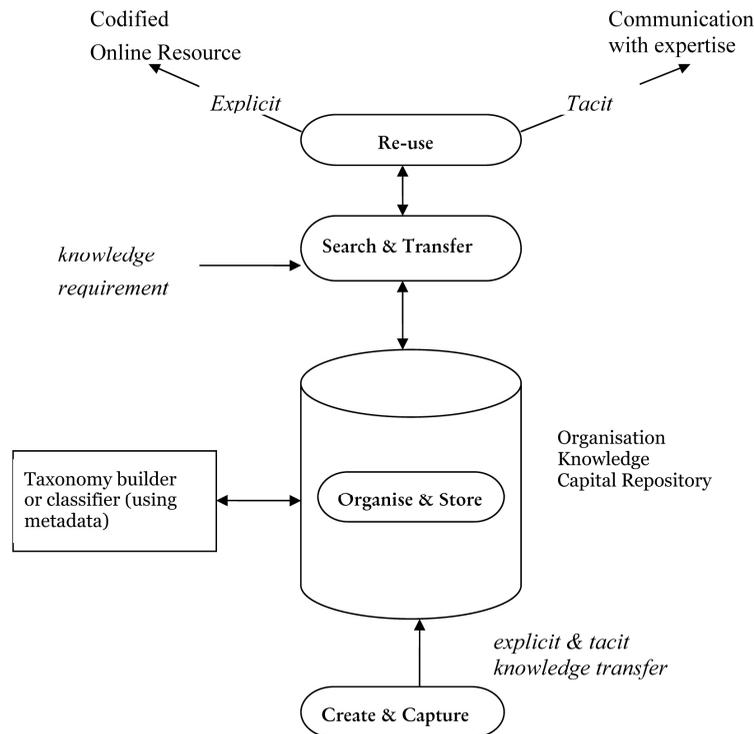


Figure 4: Integrating knowledge flows with corporate taxonomies.

oping validated corporate taxonomies is to: “understand that knowledge is transient” and to “recognise and locate knowledge in a wide variety of forms: tacit and explicit, formal and informal, codified and personalised, internal and external, short life cycle and permanent.” Hence it is also imperative to locate knowledge in processes, relationships, policies, people, documents, conversations, links and context, suppliers, competitors and customers.

Foo and Hepworth (2000) and Hepworth and Foo (2000) have illustrated this complexity in their knowledge audit of a large public organisation in Singapore. They utilised a range of data collection instruments in their proposed methodology based on the theoretical model of the user of information. This encompasses the process of task analysis and information requirements gathering through interviews with top management, focus groups with middle management, a series of observations and talk-throughs, and finally through an organisation-wide quantitative knowledge audit survey. A knowledge map (information architecture as it is referred to then) was subsequently derived showing categories of information (potential taxonomy elements), criticality of information, priority of information, frequencies of use, information flows within and outside the organisation.

It follows that *Step Two* of developing a corporate taxonomy is thus an intensive requirements analysis which brings together all stakeholders (users, contributors and managers of knowledge) so as to determine the appropriate vocabularies, term lists, subject headings, classification, search and dissemination. This may be done using modelling techniques from systems analysis such as process mapping, class diagrams, use cases and prototyping knowledge maps. Sharma and Chowdhury (2007) have undertaken action research at five enterprises of various complexities which demonstrates the usability of structured interviews, record reviews, focus group sessions and object-oriented modelling in order to first describe the existing (*as is*) and then prescribe an intended solution (*to be*) to the design of a knowledge map

The selection of meta-data and rules for classification of corporate knowledge (explicit and tacit) and their accompanying structure is a complex and critical activity in developing corporate taxonomies. It is, with rare exceptions, consultative, collaborative and iterative along the lines proposed by Noy and McGuinness (2001)—a knowledge engineering methodology for developing taxonomies that are AI centric:

1. There is no one correct way to model a domain since there are always viable alternatives. The best

- solution almost always depends on the application that is in mind and the anticipated extensions.
2. Ontology development is necessarily an iterative process.
 3. Concepts in the ontology should be close to objects (physical or logical) and relationships in the domain of interest. These are most likely to be nouns (objects) or verbs (relationships) in sentences that describe the domain.

Lehman (2003) concludes after much introspection that the key to end use success is precise classifications that are explicitly, completely and accurately defined. Without such classifications, most natural languages (vocabularies) and the context of their usage will defeat all the good intentions of a taxonomy. He posits that classification should be able to evolve into perfection. "Relative" quality, some resources could be misclassified or missed, which will destroy user confidence. He states with some conviction: "Avoid vague, qualitative or descriptive subjects in your taxonomy. Stay with subjects that are simpler, and are able to be represented by proper names, identifiers or other unique evidence. Create more and simpler classifications, rather than fewer and sophisticated classifications;" noting that the results of 20 years of cooperative research into better textual query have yet to produce techniques and languages that consistently find a large percentage of correct results, and simultaneously avoid a large percentage of incorrect results.

Now that the knowledge domain has been mapped and expertise as well as experts ordered into a structure, *Step Three* of developing a corporate taxonomy would be the selection of an appropriate combination of tools that would automate some of the activities of the previous step, particularly the extraction of keywords and subject descriptors (in order to design the metadata), classification rules, search strategies and dissemination flows. Discerning the vocabulary and index terms and using these as metadata conforming to standards, often with the use of automated tools, for the purpose of classifying and cataloguing and extracting subject headings and rules for the classification of knowledge and expertise is the "easy part". The art of mapping knowledge structures to content (codified as well as residing within experts) and growing this conceptual mapping is far more challenging. The resulting taxonomy structure is usually shown as a two-dimensional tree, similar perhaps to the folders-subfolders-files parent-child hierarchy common in many operating systems. At times, they

can be presented using a visualisation tool to show relationships between content in as a graphical map, star, tree or ring. The use of such a myriad of tools also assures some level of auditability which in turn allows improvements to the implementation. It helps that the use of tools allows what-if design changes at the click of a mouse which provide analytic measures of various parameters that may be improved, for example, ease of access, success of search, storage redundancies, identification of gaps, and so on.

Prabha et al. (2007) found that knowledge workers are motivated to seek satisficing rather than optimal search results for the knowledge needs; that is, they seek sufficiency rather than perfection. In the reported research, they designed a series of stopping rules for information access, given by both qualitative parameters such as requirements, time, coverage etc. as well as qualitative parameters such as the extent to which the search results are trustworthy, representative, current, exhaustive etc. The implication of this is that when modelling knowledge search and transfer, it is (as a first cut subject to subsequent refinements) sufficient to create a design that is usable and stopping rules for the search and transfer should clearly be implemented for such a scenario. Once again, such stopping rules are auditable and hence consciously implemented for the purpose of continual refinement. Hence, *Step Four* of developing corporate taxonomies has to do with its stepwise refinement.

The notion of taxonomy diagnostics is an area which is fast gaining research and development interest after an initial hiatus. Some of the latest tools come with such diagnostic functions (Delphi 2002). The central idea here is that since needs as well as operating environments change, it may not make sense to over-design the first taxonomy hoping that it will be used over a long period of time. Instead, interactive refinement and growth are key elements, and there are several parameters which may be used to ensure that this proceeds in the right direction. An important by-product of such a design objective is that the corporate taxonomy is amenable to mini audits, perhaps part of the much larger in scope knowledge audits. The point being, a corporate taxonomy reveals what an organisation knows and does not know. It also reveals through an external intervention what the organisation knows but does not utilize; and what it does not know at all. Such a design for auditability is roughly analogous to the widely accepted practice of design for testability that is prevalent in software engineering where placing

checkpoints and tracing logic flow demonstrably leads to superior design.

Thus, in order to stay relevant and useful, an evolving taxonomy must address fundamental requirements. In addition, a number of other evaluation criteria such as these that have been drawn from Gilchrist (2001), Graef (2001), Grey (1999), Gruber (1993), Potter (2001) and Woods (2004), among others, determine the effectiveness of corporate taxonomies and help the iterative refinement as shown in Table 6.

Audience	Are the significant knowledge professionals likely to have the need to use it?
Applicability	How broadly does the information and knowledge apply to critical tasks?
Transferability	How easy is it to impart knowledge to others using the taxonomy?
Richness	How much meaning will be lost in simplification?
Currency	How old or timeless is the knowledge and the structure?
Trustworthiness	Does the knowledge come from a reliable source? Is it verifiable?
Item Re-Use	Are corporate intranet site statistics helping to monitor reuse? Is the taxonomy helping in the search for reusable knowledge?
Usability Test	Are knowledge workers able to find relevant information on time by navigating the repository?
Satisfaction	What is the feedback (via surveys or interviews or usage) from knowledge workers?

Table 6: *Evaluation Criteria for Corporate Taxonomies.*

Admittedly, such a diagnosis is more qualitative and judgmental than a quantitative measurement. Pepper (2000) had suggested a more formal model using concepts of topic, occurrence, association, identity, facets and scope as diagnostic dimensions for organizational taxonomies but as yet these have not been developed fully. Nevertheless, the onus is on the stakeholders to grow and learn with usage and help refine the vocabulary, metadata, clusters and other functionalities. After a certain period of use, there may be a need to re-design and re-classify (and perhaps use a new selection of meta-data) in order that the corporate taxonomy continues to serve the organisation. At times, there may even be a need to re-architect and migrate to another (more appropriate)

tool. In all such cases, diagnostics help ensure that the corporate taxonomy remains a competitive weapon for the knowledge organisation rather than a legacy albatross.

In conclusion, the information overload problem has made it difficult for many knowledge professionals to find relevant and up-to-date information on the web, corporate databases or other digital repositories in order to be maximally effective in their work and to make timely decisions. As a result, taxonomies have become an important part of the suite of tools that help users locate information using controlled vocabulary or natural language search and providing powerful browsing capabilities based on structured content organisation and access through directory structures. The corporate taxonomy is potentially an authoritative intermediary that provides the terms and relationships an organisation will use in order to create, capture, organise, store, search, transfer and re-use its knowledge resources – explicit and tacit. Its return on investment should be an auditable improvement in the organisation's knowledge management.

References

- Birkinshaw, Julian and Sheehan, Tony. 2002. Managing the knowledge life cycle. *MIT Sloan management review* Fall: 75-83.
- Cheung, C.F., Lee, W.B. and Wang, Y. 2005. A multi-facet taxonomy system with applications in unstructured knowledge management. *Journal of knowledge management* 9(6): 76-91.
- C.F. Cheung, M.L. Li, W.Y. Shek, W.B. Lee and T.S. Tsang 2007. A systematic approach for knowledge auditing: a case study in transportation sector. *Journal of knowledge management* 11 (4): 140-58.
- Conway, Susan and Sligar, Char. 2002. *Unlocking knowledge assets*. New York: Quantum Books, New York. Available at: <http://www.microsoft.com/mspress/books/sampchap/5516.aspx#SampleChapter> accessed on 22-May-07.
- Davenport, Thomas and Prusak, Laurence. 1998. *Working knowledge: How corporations manage what they know*. HBS Press, Boston.
- Delphi Group White Paper. 2002. *Taxonomy and content classification*. Available at <http://www.autonomy.com/content/downloads/> access on 22-May-07.
- Drew, Stephen. 1999. Building knowledge management into strategy: making sense of a new perspective. *Long range planning* 32: 130-36.

- Ernst & Young. *What is a taxonomy?* Available at: http://www.ey.com/global/content.nsf/International/XBRL-What_are_Taxonomies accessed on 22-May-07.
- Feldman, Susan and Sherman, Chris. 2001. *The high cost of not finding information*. An IDC White Paper. Available at <http://www.idc.com>.
- Foo, Shou Boon, Sharma, Ravi. and Chua, Alton. 2007. *Knowledge management tools and techniques*. Singapore: Prentice-Hall.
- Foo, Schubert and Hepworth, Mark 2000. The implementation of an electronic survey tool to determine the information needs of a knowledge based organization. *Information management & computer security* 8(2): 53-64.
- Hepworth, Mark and Foo, Schubert. 2000. Defining an information architecture to support a knowledge based organisation. Business Processes Resource Centre (BPRC) Conference on Knowledge Management: Concepts and Controversies, University of Warwick, Coventry, UK, February 10-11.
- Gantz, John. 2007. *Expanding digital universe: A forecast of worldwide information growth through 2010*. IDC White Paper. Available at: www.emc.com/about/destination/digital_universe accessed on 22-May-07.
- Geisler, Eliezer. 2006. A taxonomy and proposed codification of knowledge and knowledge systems in organizations. *Knowledge and process management* 13: 285-96.
- Gilchrist, Alan and Kibby, Peter. 2000. *Taxonomies for business: Access and connectivity in a wired world*. London: TFPL.
- Gilchrist, Alan. 2001. Corporate taxonomies: report on a survey of current practice. *Online information review* 25: 94-102.
- Graef, Jean. 2001. Managing taxonomies strategically. *Montague Institute review*. Available at: <http://www.montague.com/abstracts/taxonomy3.html> accessed on 22-May-07.
- Grey, Denham. 1999. Knowledge mapping: a practical overview." Smith Weaver Smith online article. Available at: <http://www.smithweaversmith.com/knowledge2.htm> accessed on 22-May-07.
- Gruber, Thomas R. 1993. A translation approach to portable ontology specifications. *Knowledge acquisition* 5: 199-220.
- Gupta, Anil K. and Govindarajan, Vijay. 2000. Knowledge flows within multinational corporations. *Strategic management journal* 21: 473-96.
- Hansen Morten T., Nohria, Nitin, and Tierney, Thomas. 1999. What's your strategy for managing knowledge? *Harvard business review*: 106-16.
- Hlava, Marjorie and ven Eman, Jay. 1999. Online resources on intranet taxonomies. Available at: <http://www.accessinn.com/ilib99/index.htm> accessed on 22-May-07.
- Hylton, Ann. 2002. *Measuring & assessing knowledge-value & the pivotal role of the knowledge audit*. Available at: http://www.providersedge.com/docs/km_articles/Measuring_&_Assessing_K-Value_&_Pivotal_Role_of_K-Audit.pdf accessed on 22-May-07.
- Lehman, John. 2003. Corporate taxonomies 101. *New ideas engineering enterprise search newsletter* Issues 1, 2 & 3 – April, June & July. Available at: <http://www.ideaeng.com/pub/entsrch/issue02/article02.html> accessed on 22-May-07.
- Liebowitz, Jay, Rubenstein-Montano, Bonnie, McCaw, Doug, Buchwalter, Judah, Browning, Chuck, Newman, Butler and Rebeck, Rebeck . 2002. The Knowledge Audit. *Knowledge and process management* 7: 3-10. Online version available at: <http://www.library.nhs.uk/knowledgemanagement/ViewResource.aspx?resID=251510&tabID=289> accessed on 22-May-07.
- Mohanty, Santosh and Chand, Manish. 2005. 5iKM3 knowledge management maturity model. Tata Consulting Services White Paper. Available at: http://www.tcs.com/NAndI/default1.aspx?Cat_Id=7&DocType=324&docid=419 accessed on 22-May-07.
- National Library for Health (NLH) KM Specialist Library. 2005. *Conducting a knowledge audit by C De Brun*. Available at: <http://www.library.nhs.uk/knowledgemanagement/ViewResource.aspx?resID=93807&tabID=290> accessed on 22-May-07.
- Nonaka, Ikujiro and Takeuchi, Hirotaka. 1995. *The knowledge creating company*. New York: Oxford University Press.
- Noy, Natalya F., and McGuinness, Deborah L. 2001. Ontology Development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001.-0880. Available at: <http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html> accessed on 22-May-07.
- Perez-Soltero, Alonso, Sanchez-Schmitz, Gerardo, Barcelo-Valenzuela, Mario, Palma-Mendez, Jose Tomas and Martin-Rubio, Fernando. 2006. Ontologies as strategy to represent knowledge audit

- outcomes. *International journal of technology, knowledge and society* 2(5): 43-51.
- Pepper, Steve. 2000. The TAO of topic maps. *Presented at XML Europe 2000*. Paris, France. Available at: <http://www.gca.org/papers/xmleurope2000/papers/s11-01.html> accessed on 22-May-07.
- Potter, Scott. 2001. Building an enterprise taxonomy. *KMworld*, Nov-Dec 2001.
- Prabha, Chandra, Connaway, Lynn Silipigni, Olszewski, Lawrence. and Jenkins, Lillie R. 2007. What is enough? Satisficing information needs. *Journal of documentation* 63: 74-89.
- Ranganathan, Shiyali Ramamrita. 1924- present. Dr. S.R. Ranganathan's fifty years of experience in the development of *Colon Classification*. Available at: <http://www.isibang.ac.in/library/portal/Pages/chp1.pdf> accessed on 22-May-07.
- Senge, Peter M. 1999. *The fifth discipline: the art & practice of the learning organisation*. London: Random House.
- Sharma, Ravi S. and Chowdhury, Naguib 2007. On the use of a diagnostic tool for knowledge audits. *Journal of knowledge management practice* 8(4). Available at: <http://www.tlinc.com/jkmpv8n407.htm>. accessed on 30-Dec-07.
- Stratify Whitepaper. 2006. *Electronic discovery, search and concept organisation*. Available at: http://www.stratify.com/stratify_resources/resources_top.html accessed on 22-May-07.
- Williamson, Nancy J. 1997. Knowledge structures and the Internet. Knowledge organisation for information retrieval: Proceedings of the 6th International Study Conference on the Classification Research. The Hague: FID, pp. 23-27.
- Woods, Eric. 2004. Building a corporate taxonomy: Benefits and challenges. *Ovum expert advice*. Available at: http://www.metier.dk/downloads/kms/20041019_Ovum01.pdf accessed on 22-May-07.
- Zack, Michael H. 1999. Developing a knowledge strategy. *California management review* 41: 125-45.