



Taxonomies: The Backbone of Enterprise Content Management

By Cadence Group

We all know the amount of information available to businesses is exploding daily. According to one source, it doubles every eleven hours. An IDC report, *The High Cost of Not Finding Information*, and other studies show that knowledge workers spend at least 15 to 25 percent of the workday searching for information. Only half the searches are successful.

Other estimates show that a company employing 1,000 knowledge workers could lose an average of \$6 million per year on unsuccessful searching. Poor content classification systems contribute greatly to this problem.

A company's content is defined as all of its stored knowledge, from the Internet and intranets to sales presentations, technical documents, and other proprietary materials it creates. Some of its content exists electronically, some in paper documents and other media, and some in the minds of subject matter experts who also need to be readily identified and accessible. The challenge is how to organize, access, and manage the content when it is spread across numerous storage media, global departments, and intranet sites.

The solution recommended for many companies is Enterprise Content Management (ECM) – formerly known as content management systems or CMS – which is a framework for managing information assets across an organization. ECM ties in many platforms and programs including data capture/imaging, electronic document management, electronic records management, business process management, collaboration, Web content management, and digital asset management.

ECM Needs Discipline and Governance

It is not surprising that companies, when deploying an ECM system, look first at establishing a technology platform. However, most of us know by now that a technology solution *alone* is not sufficient for improving content management. Gartner researchers have noted that "to get value from the vast quantities of information and knowledge, enterprises must establish discipline and a system of governance over the creation, capture, organization, access, and utilization of information."

This is echoed by *CIO Update* guest columnist Sam Goldman, "Investing in content management is a valuable undertaking, but without creating a taxonomy as part of the process, companies could find themselves saddled with a next-to-useless solution." Taxonomy is the governance that enables companies to reap the benefits of ECM systems.



According to Goldman, “Creating a taxonomy should be central to any enterprise content strategy. Without that framework, even the best technology may not meet expectations because of the numerous intranet sites and discrete pieces of information it has no way of interconnecting.” Put another way, a company's taxonomy is the skeleton from which all content hangs.

The initial taxonomy development effort can be strenuous but the payoff is enormous. With well-designed classification systems, you have better and quicker access to information, resulting in greater productivity and process efficiency. Other benefits include:

- ▶ Improved search results, giving you higher quality and quantity of information
- ▶ Consolidation of multiple media formats into one repository, yielding better search results with less effort
- ▶ Increased e-commerce conversion rates because customers can find products faster
- ▶ Improved relationships with partners, regulators, and customers who can find what they need more quickly
- ▶ Less duplication of work efforts due to streamlined search results

Taxonomy vs. Thesaurus

Forrester researchers note in *Best Practices in Taxonomy Development and Management* that, ideally, taxonomies “represent agreed-upon terms and relationships between ideas or things and serve as a glossary or knowledge map helping to define how the business thinks about itself and represents itself, its products and services to the outside world.”

Frequently, people confuse the terms taxonomy and thesaurus. Both are content classification “tools” that, when used together, form the foundation for an organization’s information architecture and navigational structure. How are they different?

A taxonomy is a hierarchical list of terms or content categories. It forms a consistent and relevant method of organizing your organization’s stored knowledge so that you can quickly find the information you need.

For example, a sample medical taxonomy might look like this:

- >Diseases & Conditions**
 - >>Heart**
 - >>>Angina**
 - >>>> Aortic Valve Stenosis**
 - >>>> Palpitations**

A thesaurus, on the other hand, is a listing of preferred terms, phrases, and terminology, as well as the relationships between those words and phrases, their variants and synonyms. A sample thesaurus entry for “angina” might look like this:

Angina: angina pectoris, chest pain, coronary artery disease, stable angina, unstable angina

Used together, a taxonomy and thesaurus are essential for establishing discipline, governance, and organization over your information resources.

Seven Signs You Need a Taxonomy

You may be wondering how you can tell when or if you need a taxonomy. If you have any of the following “seven telltale signs,” then you probably need one.

1. You performed a search on a topic and received hundreds of pages of results. Your advanced search got it down to 30 pages. Is this the best to hope for?
2. You know e-mails, news, published reports, and other documents exist on your intranet on exactly the subject you want, but you can’t seem to bring them all together in one comprehensive search.
3. You’ve assigned customized preferences to your portal and subscribed to particular topical e-mail alerts, but you’re still getting junk.
4. You know your products and services are e-commerce supported, but your customers are not buying because they can’t find what they need.
5. You built a great knowledgebase – full of documents, research, information, and the names of subject matter experts. But, people are still having trouble finding answers.
6. Your company’s classification systems are in conflict. Your Webmaster posts information one way; record managers file documents another way; and, your server uses a classification system different from both of these. How can you get everyone to “speak the same language?”
7. You suspect that your company is one of the companies that may be losing considerable time and money due to the inability to quickly find information.

The Best Approach to Taxonomy Development

There has been much discussion about the merits of automated taxonomy development versus manual classifications. According to Philip Russom, Giga Research Director at Forrester Research Inc., “Even the vendors of tools for automating document classification now recommend significant human involvement in the everyday chores of taxonomy tweaking and text tagging. And, companies with an army of librarians and taxonomists devoted to manual classification also bring to bear automatic classification software in the battle against ‘infoglut.’”

Russom believes that “the new best practice starts with knowledge workers (or more specialized personnel, such as librarians and taxonomists) who create high-level topics

and subtopics, arranged in a hierarchy. Many content management tools today help workers discover topics (via search, query, clustering, or mining technologies), then convert the findings directly into a topic in a taxonomy. Editing tools enable knowledge workers to fine-tune topics and establish business rules for how text is classified to them.”

In most cases, a combination of software automation and manual classification is best, but you still have to decide how much of each is appropriate for your organization. A recommended approach is to find an information management (IM) consulting firm that has professional taxonomists and uses a proven taxonomy development methodology. This firm may have alliances with software vendors but, ideally, does not sell taxonomy software. This gives you a measure of knowledge and objectivity that you will not find with software firms.

Consultative Approach

The IM consulting firm will help you define your organization’s information needs and goals for your taxonomy project. The firm then begins to analyze, design, and build your taxonomy – to your needs and specifications. Typically, the deliverables include:

- ▶ Scoping document
- ▶ Research summary report
- ▶ Requirements document
- ▶ Custom taxonomy design document
- ▶ Completed taxonomy structure

Once the taxonomy is built, the consulting firm establishes processes and procedures for on-going maintenance by your staff or by an outsourced entity.

Example of a consultative approach: A federal government agency wanted to organize 70 years of worldwide investigation files for easy retrieval around the world. Using its meta-tagging and taxonomy expertise, Cadence Group, an information management consulting firm, organized the files, then analyzed, abstracted, indexed, and imaged the historical content. The content is now accessible over the intranet as easily from a laptop in Zaire as from a desktop in the U.S.

Facilitated Approach

Another approach to taxonomy development is a facilitated solution where an experienced taxonomy development consultant designs a custom workshop for your staff. The workshop usually covers several days and brings together an organization’s subject matter and content experts to facilitate the taxonomy development. The workshop is tailored to fit your particular needs and audience.

Prior to the workshop, consultants work with you to prepare a scoping document, requirements document, and conceptual design for the taxonomy. Through a series of custom, structured activities, your content experts use their knowledge to develop the categories, sub-categories, and hierarchical structure of an initial taxonomy.



The result is a taxonomy developed for your organization by your own experts who then know how to maintain it going forward.

Example of a facilitated approach: An international health inspection agency needed to develop an agency-wide taxonomy that would achieve more efficient storage and retrieval of agency content. The taxonomy would also facilitate the flow of information among employees, across business units, and from the agency to the public. The agency wanted to train its own subject matter experts to design, build, and maintain the taxonomy. Cadence Group delivered a two-day, hands-on workshop where attendees successfully developed a taxonomy that unifies the way in which all business units and programs organize and store content.

Conclusion: Value Far Outweighs Cost

Organizations have amassed more information than they can keep up with, locate easily, or make available to users. This information can be just about anything your company values – Web content, electronic documents and records, archives, reference material, catalog entries, sales materials, e-commerce items, database records, and other information on the Internet or intranets.

Experts agree that taxonomies help you significantly improve access to content through better classification, organization, and structure, resulting in improved productivity and process efficiency when searching for information. It's a vital first step in setting up an ECM system.

In addition, implementing a taxonomy enables an e-commerce customer to find a product more quickly. Studies show that when customers find what they are looking for within three clicks of their mouse, they are more apt to buy that product.

To get started, you need to define the goals for your ECM project and then engage an experienced IM consulting firm to guide you through the taxonomy development process. You will find the return on investment far outweighs the initial cost.

About Cadence Group

Cadence Group is a woman-owned, 50-person information management consulting firm based in Atlanta, Georgia. For many years, the information specialists at Cadence Group have helped clients develop ECM strategies and solutions that bring order to content – across the enterprise.

Cadence Group has experience with many information formats: documents, electronic records, e-mails, Web content, historical content, archives, library collections, knowledgebases, and other enterprise content. Regardless of your content source, Cadence Group can help you develop ECM strategies and solutions that give you quicker, more reliable access to information.

For more information about ECM and taxonomy solutions, contact Cadence Group at 404-874-0544, ext 0 or send an e-mail to info@cadence-group.com. You may visit our Website at www.cadence-group.com.