There's a new vision of the Web--the Semantic Web--that will dramatically improve Web-based services and products. It creates a setting where software agents perform everyday jobs for end-users. Deploying hierarchies, metadata, and structured vocabularies, the Semantic Web expands basic Internet functions. According to Tim Berners-Lee (writing with James Hendler and Ora Lassila in the May 7, 2001 issue of Scientific American) and the World Wide Web Consortium (W3C), the Internet has the potential to act as a treasured valet or lady's maid [www.sciam.com/2001/0501issue/0501berners-lee.html], an all-knowing, trustworthy source of practical information.

The Semantic Web is about making people's life easier by answering a whole host of familiar ready-reference queries. Berners-Lee imagines a natural language interface for the Semantic Web. For example, a user could type, "What is the best graduate program in business in the New York City area?" An intelligent agent would scurry out onto the Web, compare university rankings such as The U.S. News and World Report or the BusinessWeek Guide to the Best Business Schools, and return a list of names. The intelligent agent would then fetch the university applications and assorted financial aid information for the top five graduate programs.

Some of the traditional skills of librarianship--thesaurus construction, metadata design, and information organization--dovetail with this next stage of Web development. Librarians have the skills that computer scientists, entrepreneurs, and others are looking for when trying to
envision the Semantic Web. However, fruitful exchange between these various communities depends on communication.

Commonalities exist—as do differences—between librarians who create taxonomies and computer scientists who build ontologies. Mapping concepts, skills, and jargon between computer scientists and librarians encourages collaboration. Speaking the language of other disciplines and professions helps librarians remain a vital part of the Web development community.

**TAXONOMIES: AN IMPORTANT PART OF THE SEMANTIC WEB**

The Semantic Web entails adding an extra layer of infrastructure to the current HTML Web. Metadata and structured vocabularies make it easier for databases to communicate with each other. A major problem with the Internet today is data fragmentation. With the Semantic Web, computers understand the meaning of a Web page by following hypertext links from Web documents to topic-specific ontologies. For instance, ontologies offer cross-references so a computer understands that "movie," "film," "flick," and "motion picture" are different expressions of the same concept.

While intelligent agents do the visible labor of the Semantic Web, taxonomies will be facilitating communication among machines behind the scenes. For computers flung around the world to work together, a common set of terms—vocabularies—is needed and then rules that lay out how those terms work together. Taxonomies are an important part of what makes the Semantic Web "intelligent." Vocabularies and the relationships that exist between selected terms help machines to understand conceptual relationships as humans do.
Computer scientists--along with librarians--are working to solve problems of information retrieval and the exchange of knowledge between user groups. Ontologies or taxonomies are important to a number of computer scientists by facilitating the sharing and reuse of digital information. According to Tom Gruber, an artificial intelligence scholar at Stanford University, the ultimate goal for computer scientists is agreeing upon an authorized set of ontologies that can be reused and applied across multiple disciplines [www-ksl.stanford.edu/kst./what-is-an-ontology.html].

**DEFINING ONTOLOGIES AND TAXONOMIES**

Ontologies and taxonomies are, in functional terms, often used as synonyms. Computer scientists call hierarchies of structured vocabularies "ontologies" and librarians deploy the term "taxonomy." Ontology is a name borrowed from philosophy and concerns the study of reality. A philosopher who is an ontologist explores the fundamental nature of reality. That is, an ontologist asks the following types of questions: How can we know what reality really is? What is the connection between appearance and reality? According to Stanford University scholars Natalya F. Noy and Deborah L. McGuinnes in their article "Ontology Development 101: A Guide to Creating Your First Ontology [http://ksl.Stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html]", computer scientists use the term ontology to describe a topic-specific hierarchy that identifies specific vocabulary terms and then lays out the relationships that exist between those words.

This is familiar ground for librarians in terms of the information structure itself (a hierarchy composed of Broader Terms, Narrower Terms, See, and See Also references). Cataloging theorists and
practicing catalogers argue that taxonomies contain an implicit interpretation of reality. Library scientists agree that taxonomies arrange the world according to specific criteria. For instance, one could categorize "lipstick," "fire engine," or "grass" in a taxonomy by the different characteristics that compose each item. Arranging the terms according to color might result in "lipstick" and "fire engine" being grouped together because they are (often) both red. Ordering objects by texture might lead to a group that consists of "lipstick" and "grass" since they are both soft. Taxonomies define a world view by specifying which characteristics that compose each item count as important.

**STANDARDIZED LANGUAGE AND CONCEPTUAL RELATIONSHIPS**

Both taxonomies and ontologies consist of a structured vocabulary that identifies a single key term to represent a concept that could be described using several words. While Roget's Thesaurus guides users from one key term to multiple synonymous terms, taxonomies work in the opposite direction, helping end-users locate documents about a single concept that can be described using various terms. In addition to helping end-users deal with synonyms, taxonomies also illustrate associative and hierarchical relationships among concepts.

Taking an example from the AAT (Art and Architecture Thesaurus) [www. getty.edu/research/tools/vocabulary/aat], "Lolling Chairs" is the preferred term that describes l9th-century American chairs with upholstered, low seats and high backs. The term "Martha Washington Chairs" is listed as a synonym. "Lolling Chairs" is also associated with the concept "Armchairs," and all these terms are found within the "Furnishings" hierarchy. Illustrating these kinds of conceptual relationships is one of the most important functions of hierarchies for both computer scientists and librarians.
DIFFERENT POINTS OF EMPHASIS: INHERITANCE

In general, those in computer science (CS) are concerned with how software and associated machines interact with ontologies. Librarians are concerned with how patrons retrieve information with the aid of taxonomies. Software developers and artificial intelligence scholars see hierarchies as logical structures that help machines make decisions, but for library science workers these information structures are about mapping out a topic for the benefit of patrons. For librarians, taxonomies are a way to facilitate certain types of information-seeking behavior. It would be a mistake to overemphasize this point since one can point to usability experts in the CS camp who advocate user-centered Web design or librarians who are fascinated with cataloging theory to the exclusion of flesh-and-blood patrons. Yet, as an overarching generalization, software developers focus on the role ontologies play in the reuse and exchange of data while librarians construct taxonomies to help people locate and interpret information.

This difference is illustrated by the concept of inheritance. Computer scientists build hierarchies with an eye toward inheritance, one of the most powerful concepts in software development. Machines can correctly understand a number of relationships among entities by assigning properties to top classes and then assuming subclasses inherit these properties. For example, if Ricky Martin is a type of "Pop Star" in a hierarchy marked "Singers," then a software program can make assumptions about Mr. Martin even if the details of his biography are not explicitly known. An ontology may express the rule, "If an entertainer has an agent or a business manager and released an album last year, then assume he or she has a fan club." A program could then readily deduce, for example, that Ricky Martin has a fan club and process information accordingly. Inference rules give
ontologies a lot of power. Software doesn't truly understand the meaning of any of this information, but inference rules allow computers to effectively use language in ways that are significant to the human users.

By contrast, librarians think of inheritance in terms of hierarchical relationships and information retrieval for patrons. Taking the example above, the importance of the taxonomy rests in its ability to educate patrons. Someone who's been tuned out of popular culture might use the Pop Star hierarchy to learn the identities of singers who are currently in vogue. A searcher could also uncover the various types of Pop Stars that exist in mass culture: Singers, Movie Stars, Television Stars, Weight-Loss Gurus, Talk Show Hosts, etc. Finally, a patron could hop from one synonym to another--from "Singer" to "Warbler" to "Vocalist"--and discover associative relationships that exist within this category.

**TOPIC MAPS AS NEW WEB INFRASTRUCTURE**

Topic maps are closely related to the Semantic Web and point the way to the next stage of the Web's development. Topic maps hold out the promise of extending nimble-fingered distinctions to large collections of data. Topic maps are navigational aids that stand apart from the documents themselves. While topic maps do not include intelligent agents, other aspects of this technology--metadata, vocabularies, and hierarchies--fit well within the Semantic Web framework. According to Steve Pepper, senior information architect for Infostream in Oslo, Norway, in "The TAO of Topic Maps: Find the Way in the Age of Infoglut" [www.gca.org/papers/xml-europe2000/papers/sll-01.html], his presentation at IDEAlliance's XML Europe 2000 conference, topic maps are important because they represent a new international
standard (ISO 13250). Topic maps function as a super-sophisticated system of taxonomies, defining a group of subjects and then providing hypertext links to texts about these topics. Topic maps lay out a structured vocabulary and then point to documents about those topics. Even OCLC is looking to topic maps to help its project of organizing the Web by subject.

An important advantage of topic maps is that Web documents do not have to be amended with metadata. While HTML metatags are embedded in the documents described, topic maps are information structures that stand apart from information resources. Topic maps can, therefore, be reused and shared between various organizations or user groups and hold great promise for digital libraries and enhanced knowledge navigation among diverse electronic information sources.

Katherine Adams [kadams@lapl.org] is a reference librarian at the Los Angeles Public Library and a freelance writer.