

Metadata and the Web

[Mehdi Safari](#)

Encyclopedia Islamica Foundation, Tehran, Iran. E-mail: mehsafar (at) yahoo.com

Received November 25, 2004; Accepted December 26, 2004

Abstract

The rapid increase in the number and variety of resources on the World Wide Web has made the problem of resource description and discovery central to discussions about the efficiency and evolution of this medium. The inappropriateness of traditional schemas of resource description for web resources has encouraged significant activities recently on defining web-compatible schemas named "metadata". While conceptually old for library and information professionals, metadata has taken more significant and paramount role than ever before and is considered as the golden key for the next evolution of the web in the form of semantic web. This article is intended to be a brief introduction to metadata and tries to present its overview in the web.

Keywords

World Wide Web, Metadata, Resource Description, Dublin Core, Semantic Web, Ontology

Introduction

The exponential growth of information on the World Wide Web and the characteristics of information on this medium, has posed many challenges in resource description and retrieval. The problem of information retrieval or resource discovery on the web, is one of the most important issues of the current web which has encouraged looking for the solutions.

Resource discovery is impossible without resource description and adequate resource description assures effective discovery (Dillon, 2001). Traditionally, the libraries have spent thousands of years developing systems for resource description and discovery. The Anglo-American Cataloging Rules, Library of Congress Subject Headings (LCSH), Library of Congress (LC) Classification System, Dewey Decimal Classification (DDC) System, and other procedures were developed with the aim of providing some access mechanisms to the information resources through

structured descriptive information. The descriptive information or "surrogates" made by traditional systems, had the main role in searching and accessibility of information stored in the libraries and traditional databases. Since these surrogates are information about information, Smith (1996) terms this characteristic the meta-information environment of a library. This meta-information environment provides a consistent and efficient infrastructure for the location, identification, retrieval and manipulation of information.

In the information environment of the web, resource description and discovery is the most challenging issue. What distinguishes current web resource discovery from the traditional library model is the sophisticated nature of the meta-information environment of the libraries that often involves many intermediaries between searching and finding information which needs the librarian's skills to match the information needs with the surrogates. But, surrogates have the limited, if almost non-existent, role in the process of web indexing tools. Search engines are the main tools for resource organization, through automated full-text indexing, and resource discovery, and the ease of use, where terms can be easily entered and searched, has made them as a first choice of the users seeking for information on the web. The way most of these tools operate is they run programs (called crawlers or spiders) that indiscriminately harvest whatever they can find and then do selective indexing on those contents. These tools make no use of surrogates in the process of full-text indexing. Therefore, because the content of the information resource has not the right and efficient information for it to be indexed effectively, some kind of descriptive information to impose pre-defined meaning on the web content is essential.

The unique characteristics of the Internet resources in terms of location, document versions, instability, redundant data and granularity (Heery, 1996) reveals some inappropriateness in using traditional schemas, such as cataloging rules, on the web. In addition, regarding the size of the web with millions of pages being added to it, it is not cost-effective (or really possible) to use the library cataloging rules to describe web resources and professionally catalogue each document. The challenges in the way of deploying cataloging rules for digital resources (Beacom, 2000; Lagoze, 2000; Huthwaite, 2001), have led to favoring "metadata" as a workable alternative to full library cataloging for web resources.

What is metadata?

Metadata, in general, is defined as 'data about data' or 'information about information'. In the other words, metadata is data that describe information resources. This broad definition covers various levels of descriptions and one can view this variety on a continuum from the simple to complex: A short descriptive note on a book, an informal description of search hits by search engines, a catalog and MARC (Machine Readable Cataloging) record, and a TEI (Text Encoding Initiative) header all are data that describe an information resource and hence

metadata. To refine this popular definition, metadata is considered "structured" data about data. Although excepts unstructured data such as descriptive note on a book and informal description of the search engines, this definition can actually interfere with comprehending the full scope of the metadata and needs to be more explained.

The first part of the above definition is "structured data". Structured data as Greenberg (2002, p. 245) says, implies the systematic ordering of data according to a metadata schema specification. This specification is an official representation of a unified and structured set of rules developed for object documentation and functional activities (ibid, p. 247). It seems that in the web jargon the structured data implicitly means machine readability and understandability. As Day (1999) points out, in today's jargon, this data is considered to "[be] structured so that it can become machine-understandable as well as machine-readable [. . .] and has largely been identified with issues of Internet resource discovery".

Defining "Aboutness" is not straightforward and is something controversial especially in the context of information retrieval. In the context of metadata definition, it means the data that metadata capture. Burnett et al (1999, p.1212) define this term from two main approaches contributing to metadata discussion: bibliographic control and data management. From the bibliographic control perspective, the focus of the aboutness is on the characterization of the source data for identifying the location of information objects and facilitating the collocation of subject content. In this sense, metadata is a set of data elements that can be used to describe and represent information objects. The focus of data management approach to aboutness is to enhance the use of the source data. In this sense, metadata is any data that supports the effective use of data, including information that can facilitate data management, data access, and data analysis. The data that metadata capture to describe an information resource can be divided into two categories that Burnett et al (1999, p.1215) discuss them as intrinsic and extrinsic data. Aboutness, with this categorization in view, may imply to intrinsic and extrinsic data. Intrinsic data, as Weibel et al (1995) state, refer to the properties of the work that could be discovered by having the work in hand, such as its intellectual content and physical form. This is distinguished from extrinsic data, which describe the context in which the work is used. In the other words, intrinsic data are some salient and inherent features or characteristics extracted directly from the information resource such as title, author, and subject, while the extrinsic data are those related to the administration and other non-bibliographic data such as author email, author department, password or digital signature. The first is useful for management and administrative purposes while the second facilitates resource description, identification, and discovery.

The last part of the definition is the data being described by metadata. The information being described by metadata, may be considered at the first look as corporal and digital information resources such as books, newspapers, journals, photographs and so on. Greenberg (2002) refers to this data as "object" and states

that this object "is any entity, form or mode for which contextual data can be recorded. The universe of objects to which metadata can be applied is radically diverse and seemingly endless, ranging from corporeal and digital information resources, such as a monograph, newspaper or photograph, to activities, events, persons, places structures, transactions, relationships, execution directions and programmatic applications" (p. 245).

Metadata, therefore, captures the wide range of intrinsic or extrinsic information about a variety of objects. These intrinsic or extrinsic characteristics and features are described in the individually structured data elements that facilitate object use, identification and discovery.

Taking the metadata definition as structured data about data reveals that metadata is not new, but a new coinage. Standard bibliographic information, indexing and cataloging information and classifications are all structured data that describe the characteristics and contents of information resources to facilitate their discovery and use. But what is new, is a new information environment with new challenges and problems that have made metadata more eminent than before, expanding the metadata efforts beyond the traditional library environment.

Metadata development activities and problem of interoperability

Today's metadata activities are unprecedented. The increasing number of metadata schemas with the various levels of richness and complexity originated from the different communities (Heery, 1996; Dempsey & Heery, 1997; Burnett et al., 1999) is an indicative of the wide spread interest in metadata. There has been an exponential growth in the literature of library and information science and computer science on the topic of metadata and considerable decrease in the topic of cataloging (Ercegovac, 1999, p. 1165).

As the number of metadata standards increases, the problem of interoperability among metadata schemas becomes more crucial. Various individual communities have developed different metadata standards with different levels of complexity that address their particular needs. These metadata standards are different in terms of their structure, syntax and semantics.

The literature on the information retrieval has substantiated this proposition that searching for information is likely to be effective and efficient when the searcher is familiar with the classification, structure, content and purpose of the information being sought (Cortez, 1999, p. 1218). In the traditional libraries, the searcher can consult with the trained librarian, as an intermediary, to interpret the metadata used for resource description; but, in the web the story is different. The information is provided by a wide range of resource description communities, each with his own metadata, and accessed through one portal. The need to search all of this

heterogeneous and distributed metadata in a systematic way and simultaneously from the web entails the metadata standards being interoperable.

Interoperability, as Miller (2000) discusses, is multidimensional and all-pervasive concept that covers a vast variety of features. The ability of the two systems and their applications to work together effectively to exchange information in a useful and meaningful manner is a basic definition of interoperability (Moen, 2001, p.163). This condition achieved when two or more technical systems can exchange information directly in a way that is satisfactory to the users of the systems (Mooney, 2001). In the metadata speaking, interoperability can be defined as the ability of metadata systems to work together, providing the systems with an effective and efficient information inter-exchange both semantically and syntactically, and the users with more effective and satisfactory search through simultaneous search of the heterogeneous metadata systems.

Interoperability among metadata standards requires common conventions about semantics, syntax and structure. The semantics or meaning of metadata addresses the particular needs of the various fields. Syntax is the systematic arrangement of data elements for machine-processing, which facilitates the exchange and use of metadata among multiple applications. Structure can be thought of as a formal constraint on the syntax for the consistent representation of semantics (Miller, 1998).

One of the well-known approaches to making metadata standards interoperable is Metadata mapping, also called crosswalks. In the mapping mechanism, it is tried to identify what elements of one metadata set corresponds to elements of another. The projects of mapping between metadata formats (Day, 2002) show the wide spread use of mapping for metadata interoperability. Container architecture is another way to metadata interoperability. Warwick Framework (Lagoze, 1996), developed in 1996 in an invitational metadata workshop, is a container architecture for diverse sets of metadata. This framework is a mechanism for aggregating logically, and perhaps physically, distinct packages of metadata. In 1998, the World Wide Web Consortium (W3C) specified a new architecture for metadata on the web known as Resources Description Framework (RDF) (Miller, 1998; Lassila & Swick, 1999). It is an infrastructure that enables the encoding, exchange, and reuse of metadata. This infrastructure is a foundation for processing metadata and enables metadata interoperability through the design of mechanisms that support conventions of semantics, syntax and structure. RDF provides a foundation and ability for transforming the current web into a more useful and powerful information resource in the form of semantic web.

Dublin Core metadata: a simple metadata for web resources

Among the various metadata standards, it seems that Dublin Core Metadata Initiative (DC for short) has gained the special importance among the resource

description communities. Within the diverse resource discovery activities of the mid 90, ranging from unstructured indexing of full-text resources by search engines to richly-structured data like MARC and TEI records, DC arose as a means to mediate these extremes. DC developed in the *March 1995 Metadata Workshop* (Weibel et al., 1995), sponsored by the Online Computer Library Center (OCLC) and the National Center for Supercomputing Applications (NCSA), to advance the state of the art in the development of metadata records for networked information resources. One of the main goals of the workshop was to reach a consensus on a simple and core set of metadata elements to describe networked resources. The result of the workshop was a set of 13 metadata element which was called Dublin Core Metadata Element Set (DCMES) for describing what called Document-Like Objects. By the third workshop (Weibel & Miller, 1997) the elements set was developed to 15 elements.

It was believed that resource discovery is the most pressing need that metadata can satisfy (Weibel et al., 1995); therefore, only descriptive data elements required to support resource discovery were considered and data elements covering other characteristics of the resources such as terms and conditions, archival status, and other types of metadata were not included (Dempsey & Weibel, 1996). The Dublin Core Metadata Element Set includes: Title, Creator, Subject, Description, Publisher, Contributor, Date, Type, Format, Identifier, Source, Language, Relation, Coverage, and Rights (Dublin Core, 1999).

The characteristics of Dublin Core that distinguish it as prominent candidate for description of electronic resources fall into several categories: simplicity, semantic interoperability, international consensus, and flexibility (Weibel, 1997). The Dublin Core is intended to be of a maximum simplicity and flexibility. Simplicity of Dublin Core is its main feature and originates from this fact that the elements of Dublin Core designed to be used by the creators of the resources, who are not trained catalogers or have no knowledge of cataloging, to describe the resources. Since Dublin Core provides "core" and internationally agreed upon elements that are commonly understood among various communities and fields, it promotes some level of semantic interoperability. The interoperability of Dublin Core enables resource description records created using it to be implemented across disciplines and in different fields. On the other hand, Dublin Core can be qualified and extended to meet the requirements of a wide variety of communities. Though it concentrates on describing intrinsic properties of the information object, the extension mechanism will allow the inclusion of extrinsic data for objects that cannot be adequately described by the small and simple set of Dublin Core elements. It is possible to encode many controlled vocabularies and description standards such as LCSH, MeSH (Medical Subject Headings), DDC, and LC in Dublin Core elements through Dublin Core qualifiers (Dublin Core, 2000).

Among the current metadata standards, Dublin Core has the potential of being adapted as an international standard for resource description and discovery on the

web and as a *lingua franca* for metadata, partly because of its simplicity. Its simplicity promotes general applicability but also suggests an important problem that is lack of consistency and trust. Regarding the importance of and the need to trust and provenance of data and metadata in the web (see Lynch, 2001), the greatest problem of the author-generated metadata is the inability to rely on its accuracy.

Ontology: a new form of semantic metadata for the new form of the web

The information in the current web is designed to be accessed, extracted and interpreted by human users not machines. The next generation of web, called semantic web, is based on the machine-processable semantics of the information, stored in the machine processable metadata. This web is not a separate web but an extension of the current web in which the information is given well-defined meaning, better enabling computers and humans to work in cooperation (Berners-Lee et al., 2001). The prerequisite of this web, as its definition implies, is metadata that explicitly represent semantics of data, which called ontology.

Ontology is a concept borrowed from philosophy where ontology is a systematic account of existence. This term has a different meaning in the context of knowledge representation: an explicit specification of a conceptualization (Gruber, 1993). Conceptualization simply can be thought of as objects and concepts in a domain and the relationships that hold them. In a more precise word, conceptualization refers to an abstract model of the phenomena in the world by having identified the relevant concepts of those phenomena and the relationships among them. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined (Ding & Foo, 2002a, p.123).

Identifying the relevant concepts of a domain and the relationships among them, and specifying those concepts and links explicitly in a systematic way resembles the conventional tools, such as classification schema and thesauri, traditionally used in library and information communities to represent the subject content of the documents. But, the prevalence of digital information raised issues regarding the suitability of these semantic metadata systems. The new information environment requires a more versatile and flexible semantic metadata in the machine understandable form. So, ontology as an important emerging discipline in Artificial Intelligence (AI) is used as a solution to such issues mainly due to the unique ability explicitly to specify the semantics and relations and to express them in a machine understandable language. It has the crucial role to play in enabling knowledge-based access, syntactic and semantic interoperability, and acting as the backbone of the next generation of the web transformation in the form of semantic web. Conventional semantic metadata systems such as classification schema and thesauri may be reminiscent of ontologies in a way that they define concepts and relationships systematically, but they are less expressive than ontologies when they

come to machine language. Semantic relationships among different concepts are reflected through broader terms, narrower terms and related terms in thesauri, and a hierarchical structure in classification schemes. But, the semantic relationships in ontologies are much richer, based on the various needs, than what exist in these traditional tools. While some (Soergel, 1999) believes that the ontology is a classification with another name and that the use of different terms is symptomatic of the lack of communication between scientific communities, but as Jacob (2003, p.23) states, it is a unique system that integrates within a single structure the characteristics of more traditional approaches such as hierarchies and thesauri. To equate ontology with any one type of traditional semantic metadata is to diminish both its function and its potential in the evolution of the semantic web.

Ontology as a new emerging form of metadata is revolutionizing the current classificatory approaches towards semantic metadata. Constructing the traditional metadata systems with ontological view, such as thesauri (Bechhofer & Goble, 2001) and card catalog systems (Welty & Jenkins, 1999), as well as converting the controlled vocabularies into the ontology (Qin & Paling, 2001) indicates this change. Now, what is the added value of ontologies compared with classificatory approaches used in library and information communities? The experience of Qin and Paling (2001) at the GEM (Gateway to Educational Material) showed that the major differences between two models lies in the values added through deeper semantics in describing digital objects, both conceptually and relationally. They recognized that the ontologies have the following added values:

- Higher levels of conception of descriptive vocabulary;
- Deeper semantics for class/subclass and cross-class relationships;
- Ability to express such concepts and relationships in a description language; and
- Reusability and share-ability of the ontological constructs in heterogeneous systems.

Ontology research and development has gained substantial interest recently and researchers are diverse and come from the different fields, mainly from library and information science, computer science, artificial intelligence, e-commerce and knowledge management (Ding & Foo, 2002a; 2002b). But, the study of methodologies of ontology building (Gruninger & Fox, 1995; Uschold & Gruninger, 1996; Jones et al., 1998; Fernandez-Lopez, 1999; Ding & Foo, 2002a; Corcho et al., 2003) shows that there is not a unique standard methodology for building and developing ontologies. It seems that the best known guidelines have been offered by Gruber (1993). Fernandez-Lopez (1999) believes that an important difference between a technical field that is in its infancy and another that has reached adulthood is that the mature field has widely accepted methodologies. In his comparative study of methodologies for building ontologies, he concludes that none of the methodologies are fully mature comparing them with the IEEE Standard for Developing Software Life Cycle Processes.

The deployment of the semantic metadata on the web, in the form of ontologies, needs to develop standards for specifying and exchanging this metadata. RDF, OIL (Ontology Inference Layer) and DAML+OIL (DARPA Agent Mark up Language) can be mentioned as some well-known standards for representing ontologies on the web. The comparison of the ontology representation languages (Gomez-Perez & Corcho, 2002; Corcho et al., 2003) indicates that these languages are in the development phase and varies due to their abilities and expressiveness.

Conclusion

The web of today is a mass of unstructured information. To structure its contents and, consequently, to enhance its effectiveness, the metadata is a critical component and "the great web hope". The web of future, envisioned in the form of semantic web, is hoped to be more manageable and far more useful. The key enabler of this knowledgeable web is nothing but metadata. Moving towards semantic features of the information beyond the syntactic forms, concentrating more on machine-understandability than machine-readability of information, and, therefore, providing for a high level of semantic and syntactic interoperability among heterogeneous systems are what semantic web is looking for and exactly what the semantic metadata in the form of ontologies are to meet. Though the methodologies for building and developing this metadata are under development and cannot be considered as fully matured, certainly metadata are the integral part of the web in the future. As Iannella (1999) states, "the future of metadata is the Internet and the future of the Internet is metadata".

References

- Beacom, M. (2000). [Crossing a digital divide: AACR2 and unaddressed problems of networked resources](http://www.loc.gov/catdir/bibcontrol/beacom_paper.html). Retrieved November 20, 2004, from http://www.loc.gov/catdir/bibcontrol/beacom_paper.html
- Bechhofer, S., & Goble, C. (2001). Thesaurus Construction through Knowledge Representation. *Data & Knowledge Engineering*, 37, 25-45.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, May, 284(5), 345-355.
- Burnett, K., Ng, K. B., & Park, S. (1999). A comparison of the two traditions of metadata developments. *Journal of the American Society for Information Science*, 50(13), 1209-1217.
- Corcho, O., Fernandez-Lopez, M., & Gomez-Perez, A. (2003). Methodologies, tools and languages for building ontologies: where is their meeting point? *Data & Knowledge Engineering*, 46, 41-64.
- Cortez, E. M. (1999). Use of metadata vocabularies in data retrieval. *Journal of the American Society for Information Science*, 50(13), 1218-1223.
- Day, M. (1999). [Metadata and electronic information](http://www.ukoln.ac.uk/metadata/presentations/circe/birmingham.html). Retrieved November 20, 2004, from <http://www.ukoln.ac.uk/metadata/presentations/circe/birmingham.html>

- Day, M. (2002). [Metadata: mapping between metadata formats](#). Retrieved November 20, 2004, from <http://www.ukoln.ac.uk/metadata/interoperability>
- Dempsey, L., & Heery, R. (1997). [A review of metadata: a survey of content resource description formats](#). Retrieved November 20, 2004, from <http://www.ukoln.ac.uk/metadata/desire/overview/>
- Dempsey, L., & Weibel, S. L. (1996). [The Warwick Metadata Workshop: a framework for the development of resource description](#). *D-Lib Magazine*, July/August. Retrieved November 20, 2004, from <http://www.dlib.org/dlib/july96/07weibel.html>
- Dillon, M. (2001, January 23). [Metadata for web resources: how metadata works on the Web](#). Retrieved November 20, 2004, from http://www.loc.gov/catdir/bibcontrol/dillon_paper.html
- Ding, Y., & Foo, S. (2002a). Ontology research and development: Part 1- a review of ontology generation. *Journal of Information Science*, 28(2), 123-136.
- Ding, Y., & Foo, S. (2002b). Ontology research and development: Part 2- a review of ontology mapping and evolving. *Journal of Information Science*, 28(5), 383-396.
- Dublin Core Metadata Initiative (1999). [Dublin Core Metadata Element Set, Version 1, 1: Reference Description](#). Retrieved November 20, 2004, from <http://dublincore.org/documents/1999/07/02/dces>
- Dublin Core Metadata Initiative (2000). [Dublin Core Qualifiers](#). Retrieved November 20, 2004, from <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers>
- Ercegovac, Z. (1999). Introduction. *Journal of American Society for Information Science*, 50(13), 1165-1168.
- Fernandez-Lopez, M. (1999). [Overview of methodologies for building ontologies](#). In: *Proceedings of the IJCAI-99 workshop on ontologies and problem-solving methods*, Stockholm, Sweden. Retrieved November 20, 2004, from <http://www.lsi.upc.es/~bejar/aia/aia-web/4-fernandez.pdf>
- Greenberg, J. (2002). Metadata and the World Wide Web. In: *Encyclopedia of Library and Information Science* (Vol.72, Supplement 35, pp.244-261). New York: Basel.
- Gruber, T. R. (1993). [Toward principles for the design of ontologies used for knowledge sharing](#). In: Roberto Poli Nicola Guarino (Ed.), *International Workshop on Formal Ontology*, Padova, Italy, 1993. Available as technical report KSL-93-04, Knowledge Systems Laboratory, Stanford University. Retrieved November 20, 2004, from <http://ra.crema.unimi.it/lenya/turing/live/know/knowpubs/gruber93toward.pdf>
- Gruninger, M., & Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. *IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing*. Montreal, Quebec, Canada.
- Gumez-Perez, A., & Corcho, O. (2002). Ontology languages for the semantic web. *IEEE Intelligent Systems*, January/February, 54-60.

- Heery, R. (1996). [Review of metadata formats](#). *Program*, 30(4), 345-373. Prepublication draft. Retrieved November 20, 2004, from <http://www.ukoln.ac.uk/metadata/review.html>
- Huthwaite, A. (2001). [AACR2 and its place in the digital world: near-term solutions and long-term direction](#). Retrieved November 20, 2004, from http://www.loc.gov/catdir/bibcontrol/huthwaite_paper.html
- Iannella, R. (1999). [The future of Internet metadata](#). Keynote Paper from the *5th Hong Kong Web Symposium*, 4-6 October 1999. Retrieved November 20, 2004, from <http://archive.dstc.edu.au/RDU/reports/HKSYM99>
- Jacob, E. K. (2003). Ontologies and the semantic web. *Bulletin of the American Society for Information Science and Technology*, April & May, 29(4), 19-22.
- Jones, D., Bench-Copon, T., & Visser, P. (1998). [Methodologies for ontology development](#). In: J. Cuenca (Ed.), *Proceedings of ITi and KNOWS Conference of the 15th IFIP World Computer Congress*, Budapest, August 1998, pp. 62-75. Retrieved November 20, 2004, from <http://www.iet.com/Projects/RKF/SME/methodologies-for-ontology-development.pdf>
- Lagoze, C. (1996). [The Warwick Framework: a container architecture for diverse sets of metadata](#). *D-Lib Magazine*, July/August. Retrieved November 20, 2004, from <http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>
- Lagoze, C. (2000). [Business unusual: how "event-awareness" may breathe life into the catalog?](#) *Bicentennial Conference on Bibliographic Control for the New Millennium*, Library of Congress, November 15-17. Retrieved November 20, 2004, from <http://www.cs.cornell.edu/lagoze/papers/lagozelc.pdf>
- Lassila, O., & Swick, R. R. (1999). [Resource Description Framework: \(RDF\) model and Syntax Specification](#). Retrieved November 20, 2004, from <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>
- Lynch, C.A. (2001). When documents deceive: trust and provenance as new factors for information retrieval in a tangled web. *Journal of the American Society for Information Science and Technology*, 52(1), 12-17.
- Miller, E. (1998). [An introduction to the Resource Description Framework](#). *D-Lib Magazine*, May. Retrieved November 20, 2004, from <http://www.dlib.org/dlib/may98/miller/05miller.html>
- Miller, P. (2000). [Interoperability: what is it and why should I want it?](#) Retrieved November 20, 2004, from <http://www.ariadne.ac.uk/issue24/interoperability/>
- Moen, W. E. (2001). The metadata approach to accessing government information. *Government Information Quarterly*, 18, 155-165.
- Mooney, S. (2001). [Interoperability: digital rights management and the emerging e-book environment](#). *D-Lib Magazine*, January, 7(1). Retrieved November 20, 2004, from <http://www.dlib.org/dlib/january01/mooney/01mooney.html>

- Qin, J., & Paling, S. (2001). [Converting a controlled vocabulary into an ontology: the case of GEM](#). *Information Research*, 6(2). Retrieved November 20, 2004, from <http://InformationR.net/ir/6-2/paper94.html>
- Soergel, D. (1999). The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science*, 50(12), 1119-1120.
- Smith, T. R. (1996). [The meta-information environment of digital libraries](#). *D-Lib Magazine*, July/August. Retrieved November 20, 2004, from <http://www.dlib.org/dlib/july96/new/07smith.html>
- Uschold, M., & Gruninger, M. (1996). Ontologies: principles, methods and applications. *Knowledge Engineering Review*, 11(2), 93-155.
- Weibel, S. (1997). [The evolving metadata architecture for the World Wide Web: bringing together the semantics, structure and syntax of resource discovery](#). Retrieved November 20, 2004, from <http://www.dl.slis.tsukuba.ac.jp/ISDL97/proceedings/weibe.html>
- Weibel, S., Godby, J., Miller, E., & Daniel, R. (1995). [OCLC/NCSA Metadata Workshop Report](#). Retrieved November 20, 2004, from <http://www.ifla.org/documents/libraries/cataloging/oclcmeta.htm>
- Weibel, S., & Miller, E. (1997). [Image description on the Internet: a summary of the CNI/OCLC image metadata workshop](#). *D-Lib Magazine*, January. Retrieved November 20, 2004, from <http://www.dlib.org/dlib/january97/oclc/01weibel.html>
- Welty, C. A., & Jenkins, J. (1999). Formal ontology for subject. *Data & Knowledge Engineering*, 31, 155-181.

Bibliographic information of this paper for citing:

Safari, M. (2004). "Metadata and the Web". *Webology*, 1(2), Article 7. Available at: <http://www.webology.ir/2004/v1n2/a7.html>

This article has been cited by other articles:

- Gladney, H.M. (2005). [Information Retrieval](#). *Digital Document Quarterly*, 4 (1). Available at: http://home.pacbell.net/hgladney/ddq_4_1.htm
- Lee, C.S., Kiu, C.C. (2005). A concept-based hybrid graphical-neural approach for ontological interoperability. *WSEAS Transactions on Information Science and Applications*, 2(6), 761-770.

Copyright © 2004, Mehdi Safari.