

Daddy, How do I Find a Book on Purple Frogs? : Representation Issues for Children and Youth

June Abbas

A key challenge to retrieval in any type of system is how to represent the resources appropriately so that the user(s) can find what they are looking for. In systems being used by children and youth, as well as those designed specifically for these younger users, there exist two fundamental representation problems: (1) the metadata or representation scheme of the system may not be designed with this specific user group in mind, and (2) few age-appropriate controlled vocabularies exist for use in creating metadata.

Children and youth have unique information needs and information seeking strategies (Walter, 1994). In today's increasingly digital world, children have global access to a wide variety of resources in many different formats. They access information by using a variety of information retrieval systems such as library online public access catalogs (OPACs), online database systems, and the Internet and the World Wide Web (WWW). Within the web environment children can now access specialized collections of resources in digital libraries, subject directories and web portals that are designed specifically for their use.

In order to maximize the success of children's information seeking and use of the information systems, the system design and the resources contained within the system must be represented at a level that is appropriate for this particular group of users. The metadata scheme and the metadata describing the system resources should reflect an understanding of how children access, organize, and use information, but it must also take into account the users' understanding of both how the system works and how the resources are represented within the system.

Representation of resources is a complex process and has been defined using many lenses such as library and information science, cognitive science, and linguistics, among others. Representation as "the set of means by which one thing stands for another [It is] a complex web of attributes of disparate objects and concepts, idiosyncratic and socially constructed codes and agreements, and neurological abilities" (O'Connor, 1996 p. 11). Marr (1982) defines the concept as "a system for extracting or highlighting some characteristics of concepts or things along with an explanation of the rules and reasons for the extraction" (p. 20). The representation process "seeks to establish [a] systematic correspondence between the

target domain and the modeling domain" (Jacob and Shaw, 1998 p. 146). Blair (1990) sees the problem of representation and information retrieval as linguistic in nature. How effectively we utilize language to represent an object, determines the success or failure of the information retrieval process. Blair's perspective shifts the focus of information retrieval away from interface or design issues and towards a socio-cognitive emphasis on "how language is used to communicate in various social activities" (Jacob and Shaw, 1998 p. 146).

Blair (1990) also posits that the language that we use to represent both our information needs or questions and that are used to index documents is learned in a social context or community. Blair explains the theory of "language games", as first developed by the early twenty century philosopher Ludwig Wittengenstein, and the process in which we learn language and meaning. We do not acquire language purely by learning the word and its definition, but instead learn its use and appropriateness within the context of our "forms of life" or everyday experiences. Children engage in "language games" as they go through their daily "forms of life" or experiences. Direct influences on their learning are their parents, teachers, the documents with which they engage (textbooks, web resources, etc.), and the information system they interact with. Learning and knowing the appropriate "language" or terminology to use within these contexts is vital to their success both in information retrieval and content understanding.

O'Connor (1996) also notes that the user's developmental and cognitive state, domain and system knowledge, and indexer's knowledge of the user's intended purpose(s) for the objects can affect representation and retrieval. His assertion is supported by the research community's exploration of children's information seeking, OPAC use, resource selection, and library use.

The cataloger/metadata creator's understanding of the domain knowledge of the users is also important to our understanding of how to create representations within a collection of resources for which children are the primary users. The terms a cataloger/metadata creator select impacts how successful an information retrieval session is for the user. The most prevalent controlled vocabulary in use in library systems, the Library of Congress Subject Headings (LCSH) contains words that are at mostly above a sixth grade level, thereby making their use inappropriate for younger users.

Representation for children and youth is therefore a complex socio-cognitive process in which many variables come into play. When creating representations for children the process is further complicated by: 1) our incomplete picture of this group of users, 2) metadata schemes designed for use by adults and not children, 3) the guidelines and tools to construct metadata, and 4) the differing cognitive abilities and system knowledge of children.

This paper will present an overview of the state of the research into representation for children and youth, examine the current controlled vocabularies being used in systems used by and developed for youth, and will present the preliminary findings of a current research project into the development of an age-appropriate controlled vocabulary for youth.

References

Blair, D. (1990) *Language and representation in information retrieval*. Amsterdam: The Netherlands: Elsevier Science.

Jacob, E. & Shaw, D. (1998). Sociocognitive perspectives on representation. In Williams, M.E. (Ed.), *Annual Review of Information Science and Technology*, 33, (pp. 131-185). Medford, NJ: Information Today.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman and Company: San Francisco.

O'Connor, B. (1996). *Explorations in indexing and abstracting: Pointing, virtue and power*. Englewood, CO: Libraries Unlimited, Inc.

Walter, V.A. (1994). The information needs of children.
Advances in Librarianship, 18, 111-129.