The Intersection of Words and Meaning:
Creating Metadata for Children's Materials for Effective Information Retrieval
Diamond Alexander
LIS 4010 - Spring 2018

Abstract: While children are often thought of as "little adults," science has proven time and again this simply is not the case. One area this is often seen is in library materials and resources. Materials aren't described in terms that children would naturally search for and are also missing elements that children would use to search for items. This paper looks at the information seeking behavior of children as primary indicators of how materials should be described for them, showcases an example that is describing materials well for children, and survey literature that can inform metadata creation for children's materials in a way that can positively impact their information seeking.

"It's hard because you have to find the right words to put into the box"

"Writing words is hard for me because I'm not really good at the writing"

"It doesn't do all the words you say" (Druin et al., 2009).

These quotes are from children surveyed by the Human Computer Interaction lab at UMD College Park when asked about the frustration they were experiencing when using search engines. All under the age of twelve, each child expressed their frustration with choosing words and the difficulty of having to match the right words with what they imagined this abstract search engine want them to use. It is as if there was a secret code the children knew they must crack in order to get the information they were looking for but did not know how to solve the code. The conflict here is that children search for information differently than adults do and children's materials are often lacking descriptive metadata that is relevant to their inquiries. This paper features research done on the information seeking behaviors of children, an example of an organization who embodies proper metadata and search tool creation for children, and future positive implications of proper materials descriptions for children.

At its core, metadata is considered "structured information that describes the attributes of information resources for the purposes of identification, discovery, selection, use, access, and management" (Taylor & Joudrey, 2018). NISO writes in an installment of its "Primer Series" that "It is this notion of structure that turns raw information into actionable metadata" (Riley, 2017). Information resources and materials require this "data about data" in order to be discoverable and provide value to users. Proper metadata is also important for the staff that works with these materials as better searching and linking of materials is possible when the most accurate terms and data are applied. Without appropriate metadata, objects (and digital or

primarily visual objects, in particular) are basically invisible. Thankfully, structured and standardized metadata is what allows users to find what they are looking for. Authority control, thesauri, local lists, etc. all provide the structure needed to properly and consistently describe materials for discovery and management. Online Public Access Catalogs (OPACs) allow for both the broad searching and narrow faceting of surrogate records before a user ever steps into a building or accesses the materials online. This is made possible through, primarily, typed searches that intersect with metadata elements. Users learn to articulate their desires through typed words and alter their search terms to fit the retrieval system being used. When it comes to children and their searching behaviors, this process is not as simple.

Often referred to as "digital natives," children in our current world have grown up with the ubiquitous Internet and use portable devices that can connect to it at all times. In spite of this, researchers find that this early exposure does not necessarily make them better keyword searchers. Druin et al. (2009) found that "although today's children have been exposed to computers for most of their lives, spelling, typing, query formulation, and deciphering results are all still potential barriers to finding the information they need" (pg. 89). Prior research found that younger children would often search using natural language keywords or even phrases and often preferred to browse rather than explicitly search for materials. An interesting finding on typing and spelling in particular is that younger children who are experiencing mechanical difficulties when typing have compounded difficulties with spelling whereas children that have improved typing skills have less trouble with spelling (Solomon, 1993).

This is important background information because how children search for information has to be considered when developing metadata for their materials. The ways in which children think information resources would be described informs the way they search. Although librarians

may not be able to change the way children choose to search for materials, they do have control

over the ways materials are described to better match the way children search for them.

Fundamentally, this is the purpose of descriptive metadata - to be tailored to a user group in a

way that makes the information "actionable." Abbas (2005), in "Creating Metadata for

Children's Resources: Issues, Research, and Current Developments," writes that

> "In order to maximize the success of children's information seeking and use of the
>
> information systems, the resources contained within the system should 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" (pg. 304).

Information resources are only as valuable to children as the resource's surrogates are accurately

described in a way that makes sense to them.

The author continues by writing, "How effectively we utilize language to represent an

object determines the success or failure of the information-retrieval process" (Abbas, 2005, pg.

304). A child's language choice when searching for a resource must match the language chosen

to represent that object in the retrieval system or it will not be returned from the child's query.

What follows then is an understanding that the language used to describe and represent

information resources must be tailored to the words and language skills of the children using the

system, not just adults. Even if a child were to use the same language used by the system to

represent an object, as Abbas cites a 2002 Theimer article, sometimes a "lack of congruence may

occur between a user's meaning for the search term and the meaning or definition of the term by

the controlled vocabulary's creators" (pg. 306). Referring back to the children surveyed in the

2009 Druin et al. conference paper, "[a] child explained, 'It doesn't do all the words you say,'

suggesting that even though he entered several keywords, he still could not find the information he wanted" (pg. 91).

A premier example of collaborating with children to create a search environment and descriptive metadata that better matches their behavior is the International Children's Digital Library (ICDL). The ICDL was created through the collaborative efforts of the College of Information Studies and the Human Computer Interaction lab at the University of Maryland to create a multilingual, multicultural digital books resource for children and families all over the world. The project also received invaluable contributions from six children who participated in a program called "College Park KidSteam" in Maryland where they worked with adults in the lab. It is through the expertise of this combined group's skills and experience that the ICDL was created "for children by children." ICDL's Mission Statement echoes this sentiment, stating that it was created to "[build] a digital library of outstanding children's books from around the world and [support] communities of children and adults in exploring and using this literature" ("Mission Statement", International Children's Digital Library). What many library systems may be lacking in their approach to children's materials, ICDL excels in.

As Abbas (2005) writes in her "Creating Metadata for Children" article and is evident on the ICDL Simple Search page (Figure 1), the digital library allows children "to access the resources by several innovative means…browsing with three different interactive screens and searching with traditional and nontraditional access points (pg. 309). For example, children can search for books that have yellow covers, are short in length, and feature animal characters. These categories are displayed textually and visually with colors and drawings.

Figure 1 – Source: International Children's Digital Library

Each facet is shown to progressively narrow the search with actual "+" signs and an "=" sign to designate the final results of their query (see Figure 2). Facets that can no longer apply to the current query are "greyed out," indicating their unavailability and preventing an empty return of results due to conflicting parameters.

Figure 2 – Source: International Children's Digital Library

Searching by nontraditional access points such as color, emotion, ratings, book length, etc. is made possible because the ICDL created unique metadata elements for these different categories, thus allowing their retrieval system to return very specific results. By creating metadata elements specifically for different categories such as age level, genre, emotions, and colors, materials can be found without keywords. As mentioned earlier, younger children often

searched using natural language or phrases, not controlled vocabulary. In the Druin et. Al (2009)

paper, the authors found that "One of the most consistent results in this study across all ages was

in how the children's typing abilities impacted their interaction with and use of the screen (pg.

92). How this relates to children finding materials they need becomes clearer as we consider

most search engines for materials are prominently keyword focused. Additionally, they found

that in "children ranging in age from 4 to 8 ...the keyboard was the slowest method of input

across all age groups when compared with the mouse and the joystick" (pg. 90). When one

considers the difficulty with which children have mechanically and mentally using keyboards

and words to find what they're looking for, the innovation of the University of Maryland team in

creating the ICDL is that much more impressive. They designed a system for children with

children in mind. However, the focus should not rest solely on the design and search behavior.

The focus of this project was the metadata.

ICDL's website features an extensive webpage dedicated just to its metadata standards

and usage ("Metadata Specification," ICDL). Those familiar with Dublin Core (DC) will find

many of those elements present in a typical ICDL entry, such as "contributor," "publisher," and

other common ones in the full specification list. ICDL writes that they chose DC "because of its

international background, its ability to be understood by non-specialists, and the possibilities to

extend its basic 'elements' to meet [their] specific needs" ("Bibliographic Conventions," ICDL).

General International Standard Bibliographic Description and Anglo-American Cataloging

Rules, 2nd Edition (AACR2) are used to better describe elements as they relate to other

important communities' standards, such as librarians and international community members.

ICDL is even more unique in this regard as it is a multilingual library, both in metadata and in

display languages. Children are able to search for books in a specific language as well as find

books in a specific language (for example, searching for a Russian book using Korean language keywords or interface).

In a solo paper, Druin (2005) writes that "Developmental psychologists have shown that children think quite differently from adults. Therefore, the way that information is categorized for children might not be the same as for adults. Yet, information systems typically in place are designed for adults', not children's, worldviews" (pg. 22). This means that traditional cataloging standards and practices, while useful for adults and for creating a standard vocabulary for describing materials, is not helpful for children. Druin (2005) includes findings from a Busey and Doerr study from the Kid's Catalog project where they not only used Dewey Decimal Classification to organize materials but also included terms such as 'scary stuff,' 'castles,' and 'jokes and cartoons' based on their research (pg. 23). She also shares revelations from a 2002 Cooper study of seven-year-olds in a school setting where they found that "children looked primarily at the book's cover, pictures, and title when selecting books from a shelf and used these as selection criteria" (pg. 23). Based on a Pejtersen paper, Druin (2005) writes that "it is quite typical for a child to ask for a book in this way: 'On the front cover there is a boy and a rhinoceros, and it is red'" (pg. 24). These examples are shared because they each illuminate how differently children search for and seek information than adults, and also how the ICDL project addresses many of these differences in its design. Children can search the ICDL either by clicking on images or keyword searching; they can facet their options by color, human vs animal characters, and genre; and the search results all feature book covers prominently. This is not surprising as Druin was a part of the University of Maryland team that began the ICDL project and has undoubtedly provided much of the direction for the digital library.

Looking more deeply at how ICDL's metadata compares to more traditional forms of description, Figures 3 and 4 are tables taken from a Beak and Olson (2011) paper where they attribute research done on children's book choices to those unique characteristics, as well as compare AACR2 and ICDL's Metadata Schema side by side across the various characteristics unique to children's book choices, respectively. It is interesting to note, particularly from Figure 4, that ICDL is successful because it has created actual metadata elements to allow for unique descriptions of materials, whereas an older standard such as AACR2 relegates these unique descriptions to the "Note" field, forcing a child to use keywords to find the material they seek.

Unique characteristics of children's book choices		Origins
Physical Characteristics	Book cover's illustration or character including its objects and colors Size of the print Hardback / Paperback	Kragler & Nolley (1996) Wendelin & Zinck (1983) Moore (1988) Pejtersen (1986)
Intellectual difficulty	Difficulty of words The number of pages or words Age appropriateness	Moore (1988) Robinson, Larsen, and Haupt (1997) Pejtersen (1986)
Prior knowledge (particular character, series, titles, etc.)		Kragler & Nolley (1996) Robinson, Larsen, and Haupt (1997)
Recommendation (by peers, teachers, or family) Awards		Kragler & Nolley (1996) Wendelin & Zinck (1983)
Topics (learning, social activities, or interests)		Kragler & Nolley (1996) Pejtersen(1986)
Media connection (such as TV show, movies)		Kragler & Nolley (1996) Wendelin & Zinck (1983)
Emotions (happy, sad, funny, adventurous, etc.)		Pejtersen (1986)
Frame (The setting in time and place of the subject)		Pejtersen (1986)
Genre		Robinson, Larsen, and Haupt (1997)

Figure 3

Unique characteristics of children's book choice		Metadata elements from crosswalk			
		AACR2+		ICDL's metadata schema	
		Main Elements	Sub Elements	Sub Elements	Main Elements
Physical characteristics	Size of the print	X	X	X	X
	Hard / Paper cover	X	X	X	X
		X	X	Format*/ Shape*	Others*
	Book cover's color	Note	Physical description	Cover colors*	
Physical characteristics / Prior knowledge	Book cover's characters or objects/ Particular characters	X	X	Characters*	
Prior knowledge	Series	Series	Title proper of series	Series title	Title Information
Intellectual difficulty		X	X	Length*	Others*
	# of pages or words	Physical description	Pagination	Page count	Physical characteristics
	Difficulty of words	Note	Summary	Abstract	Abstract, keywords, etc.
	Age appropriateness	Note	Audience	Age range	
Topics		[MARC21: 6XX]		Subject & Keywords	
		[MARC21: LitF]		Type	
Genres		[MARC21: 655]		Genre	
Frame		[MARC21: 648, 650 |y |z, 651]		Setting: When & Where	
Recommendation / Awards		[MARC21: 586]		Award	
		X	X	Rating*	Others*
Emotions		X	X	Feeling*	
Media Connection		Uniform title [MARC21: X30, 6XX, 700 |t]		X	X

Figure 4

Metadata created specifically for children will afford them the pleasure of searching in a way that feels natural to them and of finding materials well-suited to what they are specifically looking for. Looking to the future, there can be research done into how speech or voice searching, such as the technology Google or smartphones use, could positively impact searching for children where spelling and typing are barriers but where vocabulary words can be spoken. The future of metadata specifically for children also has implications on its impact even other marginalized user groups. How might this help English-speaking children's librarians find resources for children in a language other than English. Visual cues such as emotions, size, animals, and colors often span would-be language barriers and this could be helpful in these

library interactions to help children find the resources they are looking for in their own language. Abbas (2005), in "Out of the mouths of middle school children: I. Developing user-defined controlled vocabularies for subject access in a digital library," articulates that "systems designed for children should adapt to children and their unique information seeking needs, instead of expecting them to use…systems designed with adults as primary users (pg. 1512). Designing systems and describing materials in a way that is tailored to this young population will mean making shifts and adjustments that may not seem important when considered in the context of adult users. However, these changes will be necessary if institutions value children and desire to support their information needs just as much as they support adult information needs.

References

Abbas, J. (2005). Out of the mouths of middle school children: I. Developing user-defined
    controlled vocabularies for subject access in a digital library. *Journal of the American
    Society for Information Science and Technology, 56*(14), 1512-1524.
    doi:10.1002/asi.20245

Abbas, J. (2005). Creating Metadata for Children's Resources: Issues, Research, and Current
    Developments. *Library Trends, 54*(2), 303-317. doi:10.1353/lib.2006.0012

Beak, J., & Olson, H. A. (2011). Analysis of Metadata Schemas for Children's Libraries (Vol. 3,
    pp. 1-12, Rep.). Toronto: Proceedings from North American Symposium on Knowledge
    Organization. http://dx.doi.org/10.7152/nasko.v3i1.12785

Druin, A. (2005). What Children Can Teach Us: Developing Digital Libraries for Children with
    Children. *The Library Quarterly, 75*(1), 20-41. doi:10.1086/428691

Druin, A., Foss, E., Hatley, L., Golub, E., Guha, M. L., Fails, J., & Hutchinson, H. (2009). How
    children search the internet with keyword interfaces. *Proceedings of the 8th International
    Conference on Interaction Design and Children - IDC 09,* 89-96.
    doi:10.1145/1551788.1551804

International Children's Digital Library. (n.d.). Metadata Specification. Retrieved May 29, 2018,
    from http://en.childrenslibrary.org/about/policies/metadata.shtml

Riley, J. (2017). *Understanding Metadata: What is Metadata, and What is it For?: A Primer*
    (Rep.). Retrieved
    https://groups.niso.org/apps/group_public/download.php/17443/understanding-metadata

Solomon, P. (1993). Childrens information retrieval behavior: A case analysis of an OPAC.
    Journal of the American Society for Information Science, 44(5), 245-264.
    doi:10.1002/(sici)1097-4571(199306)44:53.0.co;2-#

Taylor, A. G., & Joudrey, D. N. (2018). *The organization of information*. Santa Barbara, CA:
    Libraries Unlimited, an imprint of ABC-CLIO, LLC.