A Sense of Wonder: Enhancing Access to Folktales through Task and Facet Analysis

Kathryn La Barre
University of Illinois
501 East Daniel, MC-493
Champaign, IL 61820
001.217.244.4449

klabarre@illinois.edu

Carol L. Tilley
University of Illinois
501 East Daniel, MC-493
Champaign, IL 61820
001.217.265.8105

ctilley@illinois.edu

ABSTRACT

Discusses the approach taken in Phase 1 of a three-phase project Folktales, Facets and FRBR [funded by a grant from OCLC/ALISE]. This project works with the special collection of folktales at the Center for Children's Books (CCB) at the University of Illinois at Urbana-Champaign, and the scholars who use this collection. The project aims to enhance the effectiveness and efficiency of folktale access through deep understanding of user needs. Phase 1 included facet analysis of the bibliographic records for a sample of 100 folktale books in the CCB, and task analysis of interviews with four CCB-affiliated faculty. Describes the information tasks, information seeking obstacles, and desired features for a discovery and access tool related to folktales for this initial group of scholarly users of folktales.

Categories and Subject Descriptors

H.3.1 [**Content Analysis and Indexing**]: Indexing Methods.
H.3.3 [**Information Search and Retrieval**]: Search Process.

General Terms

Performance, Design, Theory.

Keywords

Task analysis, Facet analysis, Search and Discovery.

1. INTRODUCTION

Folktales connect communities and people across time and space with each story evolving in its transformation from performance to text. Even as stories change, they continue to carry culturally unique values and ideals, anchored in shared human experience, with continuing relevance for its audience. Beyond their use in the specific communities where they are born, folktales find audiences among people of all ages and educational levels—from children hearing about stone soup for the first time

to an established scholar examining the transmission of treasure tales in the Dominican Republic.

A network of informants, adapters, compilers, storytellers, librarians, and scholars keeps stories alive through telling, collecting, and publishing, yet these efforts are frequently undermined by existing structures for representation and discovery in the bibliographic catalogs of libraries and similar institutions, potentially obscuring these stories from continued study and use. For instance, the records for single volume collections of tales seldom provide complete or searchable information about the titles and origins of each story in the volume, thus requiring a searcher to intuit a book's potential relevance and persevere to undertake an examination of the physical volume.

New strategies are needed in order to overcome the shortcomings of information retrieval systems that may hamper efficient information seeking for complex information resources such as folktales. The development of new strategies is complicated specifically for folktales because of the heterogeneity of users and tasks. For instance, scholarly users may want to undertake a comparative study of a particular tale type, while librarians designing a children's program may want to find multiple versions of a single tale in order to identify the most appropriate one for their needs (Goldberg, 2003). Even children, on their own or with the assistance of their caregivers, might want to explore different retellings of a favorite story such as East of the Sun and West of the Moon. Yet, each of these users must often rely on the brief descriptions in bibliographic records as they attempt to complete their information tasks.

Through an iterative combination of facet and task analysis that supports deep understanding of information tasks and allows the creation of new access models, this project aims to enhance discovery of and access to folktales and related resources. Kuhlthau's (2005) call for greater connection between the study of users' information-seeking behaviors and the design of information retrieval systems to better enable collaborative frameworks which would encourage and strengthen task-focused information seeking studies and user-centered system design motivates this project. The research design also draws from recommendations for more integrated and theoretically repositioned models for information seeking and use and

information retrieval (e.g. Hjørland, 1997; Ingwersen & Jarvelin, 2005).

2. METHOD AND FINDINGS

The scope of this preliminary work is limited in three important ways. First, the researchers interviewed only a small number of subjects, each of whom is engaged in scholarly activity related to folktales but none of whom would be identified primarily as folklorists. Second, the folktale collection that forms the basis for the facet analysis is comprised largely, but not wholly, of folktales that have been adapted for a juvenile audience; some folklorists (e.g. Goldberg, 2003) would consider a collection such as this one inadequate to support legitimate folklore scholarship. Third, as the informants and researchers are colleagues, the possibility for bias in both response and interpretation is amplified.

2.1 Task analysis

Task analysis is a repertoire of techniques commonly used in the field of human-computer interaction to support the development of systems and interfaces from a user-based perspective. In the past decade, task analysis has increasingly been used to understand people's information seeking processes (Vakkari, 2003). No universal definition of tasks exists and it can be difficult to disambiguate task and goal, but for the purposes of this study, task is best defined as any information-seeking activity necessary to complete some scholarly goal (cf. Xie, 2008). A first step in conducting any task analysis, then, is to understand users' goals.

As part of this preliminary study, we conducted one-hour semi-structured interviews with four of the five faculty members who use this collection, in order to determine the manner in which they conduct research in this area as well as their use of the CCB collection. Although not conventionally folklorists, the subjects are engaged in scholarly activity related to folklore, including editing collections of folktales, reviewing folktales adapted for children, studying audience engagement in storytelling performance, and documenting the history of literary transmission of folktales. Each of the subjects also teaches in the area of youth services librarianship, which has a strong tradition of oral storytelling (cf. Hearne, 1998), so folklore permeates their discussions and work with students. Finally each of the subjects has performed folktales orally as part of professional work experiences outside academe.

The purpose for these interviews was to ascertain 1) folktale-related scholarly practices (i.e. goals); 2) obstacles the informants have encountered in information seeking; and, 3) suggestions for an ideal tool that would help them in their information-related activities. Although we asked direct questions to elicit relevant insights, we also asked each subject to talk more broadly about other areas including their experiences working with folktales and their educational experiences related to folktales in order to capture information relevant to our interests that may not have

been revealed through direct questioning. The interviews were recorded and transcribed for coding, after which we developed the coding framework on an emergent and iterative basis in order to identify scholarly practices.

2.1.1. Scholarly practices in folklore

Six categories of scholarly practices surfaced in the interviews:

(1) *Exploring* (e.g. Reading tale collections for possible future uses; monitoring websites or journals to stay current on scholarly issues pertaining to folktales)

(2) *Creating* (e.g. Adapting a folktale for performance; designing a library program based on a folktale)

(3) *Synthesizing* (e.g. Critiquing a published adaptation of a folktale for a juvenile audience; documenting the published variants of a particular tale; preparing lecture notes and other instructional materials)

(4) *Studying* (e.g. Conducting research on audiences' responses to oral performance; examining the relationship between women's personal narratives and folktales)

(5) *Collecting* (e.g. Building a personal folktale library to support scholarship; keeping notes about folktale variants to support scholarship)

(6) *Searching* (e.g. Using a bibliographic tool to identify a variant; following cited references to identify relevant information)

Some of the goals overlap with Palmer, et al's (2009) synthetic model of scholarly information practices. For instance, she and her co-authors identified "collecting" as a core scholarly activity. Searching appears in their model as well as ours, but we have a related category—"exploring"—as well that represents non-directed searching activities that we identified; in contrast, Palmer, et al subsume a similar activity—"browsing"—beneath "searching." The activity Palmer, et al termed "writing" is similar to "synthesizing" that emerged from our data. Both "studying" and "creating" are unique to our framework with the latter category representing an activity similar to "synthesizing" but with a greater emphasis on creative transformation.

We anticipate conducting further interviews and observation of practice with these scholars and other folklorists, as part of the next phase of our research. From this added observation and interview data, we will be able to refine the list of tasks and to identify the tasks essential to supporting successful goal completion for these users.

2.1.2. Obstacles to information seeking

Regarding obstacles to information seeking, the interview data clustered in two categories: disciplinary-related and discovery and access-specific. Examples from the latter category are not especially unique in that they relate to lack of awareness of useful bibliographic tools or problems with the tools themselves (e.g. quickly outdated).

More interesting are the disciplinary-related obstacles, several of which touch on the variable nature of folktales. For instance, the names given to tales may vary from one collection or one

community to another; similarly, tale variants may share motifs, although the variants have quite different effects or themes. A disciplinary-related obstacle such as this one, however, also speaks to problems with existing tools (e.g. limited or no cross-references).

Another intriguing set of disciplinary-related obstacles pertains to "translating," or working across boundaries (cf. Palmer, et al, 2009). The subjects identified translation problems as they sought and accessed information from a variety of scholarly (e.g. literary criticism, psychoanalysis, anthropology) and disciplinary perspectives (e.g. structuralist, historical-geographic). Translation problems also occurred as subjects moved from understandings of tales informed by personal experiences (e.g. recalling stories told by family members, reading tales in childhood) to understandings constructed through scholarly practice.

2.1.3. Desired features for search and discovery tools

The features these subjects identified as essential for an ideal discovery and access tool for folktale scholarship reflected both their work as scholars and their professional experiences in storytelling and youth services librarianship. For instance, the scholarly focus is evident in requests for searchable fields for source notes and cultural attributions, as well as descriptor fields for motifs such as characters. The professional focus is clearly visible in proposing the inclusion of programming ideas, ties to learning standards, and suggested audience ages for performance. All subjects indicated preferences for a tool that would permit both directed searching and serendipitous discovery and that offered extended synopses or the full text of tales for searching.

Worth noting is that some existing bibliographic record structures such as MARC already partially support informants' requests (e.g. for a searchable field for cultural attribution), although these structures are seldom leveraged fully to these ends. In a separate paper (Tilley and La Barre, 2010) we offer a provisional model for bibliographic records that shows where existing MARC fields overlay those arising from this study.

2.2 Facet analysis

This study proposes that facet analysis is a necessary and useful first step towards the creation of user-oriented search and discovery systems. The facet-analytic dimension of this study builds on a traditional understanding of facets, as articulated by Ranganathan, who viewed them as basic concepts that are inherent in a given subject. A facet may be a concept, characteristic, attribute or aspect that may assist in the identification of a set of distinct entities. Facets are uncovered through a technique known as facet analysis, which requires the conceptual analysis of a subject area into a set of fundamental categories. The essence of facet analysis is the sorting of terms in a given field of knowledge into homogeneous, mutually exclusive facets, each derived from the parent universe by a single characteristic of division (Vickery, 1960 p. 12). The entire process of facet analysis is governed by a canon composed of principles (specific rules), postulates (guidelines) and devices (Vickery,

1960). Additional guidance for the facet-analytical approach used in this study comes from Cochrane (1965).

After establishing a complete shelflist of folktale books in the CCB's collection, we created a stratified (i.e. by decade of publication) random sample of 100 folktale books to form the core collection of materials that were subjected to facet analysis in Phase 1 of this project. The CCB is one of the world's premier reviewing and examination centers for children's books and related materials. It houses more than 15,000 English-language print and non-print resources in its non-circulating collection. Folktales published in single-tale volumes and multiple-tale collections, and scholarly resources related to folklore and storytelling comprise approximately ten percent (or 1500 items) of the collection. Publication dates for the print materials span the 20th and 21st centuries, but a majority of the items were published after 1960; this distribution is reflected in our sample.

For each item in our sample, we examined several artifacts for the facet-analysis portion of the protocol. First, we examined the books themselves. Second, we inspected the local bibliographic records as well as the most complete bibliographic records for each item we were able to obtain through WorldCat. Finally, we scanned reviews—primarily those published in the *Bulletin of the Center for Children's Books* but also from other sources—for items in the sample. The hope is that this will present an opportunity to uncover a variety of facets that might be useful beyond those typically represented by the fields now being leveraged in library catalogs.

Our analysis of facets is still ongoing at this preliminary phase. For instance, we have yet to engage in a rigorous facet analysis of users' information tasks, or deep facet analysis of the indices and controlled vocabularies used by folklore scholars to assist them in locating relevant stories. The names given to each facet may not be entirely reflective of the terms used by scholars. Both the terms used, and the facet groupings will be subject to further refinement as more interviews and observations are conducted and subject to facet analysis. Refinements are also expected as a result of further facet analysis of the subject access tools, such as folklore-specific controlled vocabularies and classifications used by folklorists.

2.2.1 Preliminary facets derived from the collection

Based on the preliminary analysis, we have identified the following facets (*in italics*) and areas where the focus of each facet may be sharpened or refined [in parentheses]:

Agent [may include: author/narrator, translator, adapter, editor/compiler, illustrator, etc.]

Area [of source] [of story]

Association [award] [aggregations of multiple stories] [related materials] [stylistic dependencies] [source] [work]

Content [characters] [illustrations] [language] [mood] [moral] [motif] [narrative structure] [story type]

Context [age of story] [audience] [function of story] [language of source] [manner of dissemination] [style] [type of variant]

Documentation [external sources like bibliographies or indexes]

Genre [type of story]

Origin [cultural] [ethnic] [geographical] [theoretical] [of source]

Time [of source] [of story]

Transmission [oral] [print] [function]

Viewpoint [theoretical] [cultural] [ethnic]

As the list indicates, a variety of facets emerged through our analysis. For instance, folklore publications are typically careful to articulate authorial responsibility, or *agent*, by clearly distinguishing among authors, editors, adapters, translators, illustrators, and retellers. Another important facet pertains to *genre*, which acknowledges differences among story types such as: folktales, fairy tales, fables, legends, and myths. The *origin* facet indicates cultural attribution—whether according to geographic region or by reference to a particular ethnic or cultural group. The *documentation* facet supports cross references or direct linkage to external sources such as notes or bibliographies. Such linkage is especially important for recently published works which may be available in digital, full text format. [Motif], here shown as a focus of the *content* facet, emerged as another important characteristic of folklore material. Within folkloric analyses of folktales, motifs refer to small persistent elements of individual stories such as actors (e.g. a princess, Baba Yaga), items (e.g. a magical stick, a curse), and plot elements (e.g. a contest, burial alive)(cf. Thompson, 1946).

These preliminary facets echo several aspects of Uther's recommendations to guide the creation of new motif indexes and related tools to assist folklorists. "In establishing concepts for new indexes and integrating the narrative material for a region or ethnic group, the following should be required:"

(1) clearly defined time and area,

(2) theme-oriented presentation,

(3) indication of structural elements

(4) chronological and structural listings of variants,

(5) suggestions of related items,

(6) year of publication,

(7) references to external sources and literature,

(8) indexing by subject, names, places, narrators (Uther, 1997, p. 215).

2.2.2 Facets derived from bibliographic tools

In addition to examining the sample of books from the CCB, we also examined selected bibliographic tools to aid in the discovery of and access to folktales (e.g. Ashliman, 1987; MacDonald and Sturm, 2003; American Folklore Society, n.d.) along with some core scholarly and overview works related to folktales (e.g. Dorson, 1972; Thompson, 1946; Toelken, 1996). Suggestions for the works we examined came both from our interview subjects and from bibliographies such as the one provided by the Folk Narrative Section of the American Folklore

Society. Our analysis of these materials provided further support for the validity of the facets derived from the book sample.

Many, but not all, of the preliminary facets already have underlying bibliographic record fields that may support facet display, but are not fully leveraged by library catalogs. For instance, La Barre (2010) queried the term "folktale" in 200 library catalogs using one of six next-generation integrated library systems (ILS)(e.g. AquaBrowser, Koha) in order to determine which facets are currently used and supported. She found the following facets (number in parentheses refers to the number of ILS systems using each facet):

- *subject/topic (6)*
- *author (5)*
- *date of publication (5)*
- *format (4)*
- *genre (4)*
- *location (4)*
- *availability (3)*
- *language (3)*
- *series (3)*
- *call number (2)*
- *subject: geography (2)*
- *subject: time (2)*

3. IMPLICATIONS AND CONCLUSION

Full-text resources have become ubiquitous, whether through subscription databases or digitization projects or the Internet. Add to this reality the perception held by many laypersons and scholars (and even some librarians) that a few keyword searches performed in Google will retrieve a universe of information. The result is that too often the value of providing systematic, reliable, and meaningful access to the intellectual contents of texts is negated. Libraries themselves play a role in this negation when, in an effort to save human and financial resources, they increasingly rely on copy cataloging, purchased records, and other similarly conceived records to provide access to the resources in their collections, believing that these frequently all too minimal descriptions will provide adequate access.

Yet, the ability to search full-text sources is not the *automagical* tool some scholars and laypersons would have us believe. In a study still relevant today, Blair and Maron (1985) demonstrated that the recall rate for relevant documents when users used free-text searching in a large data set not constructed for the purpose of testing retrieval was on average below 20%. As Blair and Maron argued, "it is impossibly difficult for users to predict the exact words and combinations, and phrases that are used by *all* (or most) relevant documents and *only* (or primarily) by those documents" (295). Even in an era with improved natural language processing algorithms to facilitate searching, the results are unsatisfactory (e.g. Tomlinson et. al 2007).

Folktales are but one example of resources that are often obscured by the movement to full-text searching and the subsequent reduction in the provision of rich bibliographic records. Oral histories, archival materials, museum artifacts, musical scores, and many other types of texts are similarly difficult to for users to locate. By seeking to understand how users of these resources integrate them into their work tasks, and by systematically analyzing the domains in which these resources are situated, we look to design alternative models for bibliographic records that highlight, rather than obscure, these resources for the people who turn to them most frequently.

4. ACKNOWLEDGMENTS

5. REFERENCES

[1] American Folklore Society. Ethnographic Thesaurus Online. <et.afsnet.org>.

[2] Ashilman, D. 1987. A guide to folktales in the English language: Based on the Aarne-Thompson Classification System. Greenwood Press.

[3] Blair, D. C., Maron, M.E. 1985. An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System. Communications of the ACM 28 (3): 289-299.

[4] Cochrane, P. 1965. American Institute of Physics, Documentation Research Project: A review of work completed and in progress, 1961-1965. American Institute of Physics.

[5] Dorson, R. 1972. Folklore and folklife: An Introduction. U of Chicago.

[6] Goldberg, C. 2003, Folktale research and the Pantheon Fairy Tale and Folklore Library. J Am Folklore. 116 (Spring 2003) 217 – 218.

[7] Hearne, B., 1998, Midwife, witch, and woman-child: Metaphor for a matriarchal profession. In Story, from fireplace to cyberspace : connecting children and narrative (Papers presented at the Allerton Park Institute held October 26-28, 1997), ed. Betsy Hearne, et al. 37 – 51

[8] Hjørland, B., 1997, Information seeking and subject representation: An Activity-Theoretical Approach to Information Science, Greenwood.

[9] Ingwersen, P., Jarvelin, K. 2005. The Turn: Integration of information seeking and retrieval in context, Springer-Verlag.

[10] Kuhlthau, C. C., 2005, Towards collaboration between information seeking and information retrieval, *Inform Res* 10 (2) paper 225.

[11] La Barre, K. 2010. Facets, Search, and Discovery in Next Generation Catalogs: Informing the Future by Revisiting Past Understanding. In *Paradigms and conceptual systems in Knowledge Organization, Proceedings of the eleventh international conference of the International Society of Knowledge Organization,* ed. Claudio Gnoli.

[12] MacDonald, M.R., Sturm, B. 2003. The Storyteller's sourcebook: A subject, title, and motif index to folklore collections for children, 1983 – 1999. Gale.

[13] Palmer, C., Teffeau, L., Pirmann, C., 2009. Scholarly information practices in the online environment: Themes from the literature and implications for library service development, www.oclc.org/programs/publications/reports/2009-02.pdf

[14] Thompson, S. 1946. *The Folktale*. Dryden Press.

[15] Tilley, C., La Barre, K. 2010. New Models from Old Tools: Leveraging an Understanding of Information Tasks and Subject Domain to Support Enhanced Discovery and Access to Folktales. In *Paradigms and conceptual systems in Knowledge Organization, Proceedings of the eleventh international conference of the International Society of Knowledge Organization,* ed. Claudio Gnoli.

[16] Toelken, B. 1996. The dynamics of folklore (Rev. and Exp. Edition). Utah State U.

[17] Tomlinson, S., Oard, D.W., Baron, J.R., Thompson, P. 2007. Overview of the TREC 2007 Legal Trac. In *The Sixteenth Text Retrieval Conference (TREC 2007) Proceedings* (NIST Special Publication SP 500-274)

[18] Uther, H. 1997. Indexing folktales: A critical survey. J Folklore Res 34 (3): 209-220.

[19] Vakkari, P. 2003. Task-based information searching. ARIST. 37. 413-464.

[20] Vickery, B. C. 1960. Faceted classification. A guide to the construction and use of special schemes. ASLIB.

[21] Xie, I. 2008. Interactive information retrieval in digital environments. IGI.