On the Use of Research Categorizations as the Basis for Organizing Knowledge: A Test in the Domain of Information Behavior in Health Care

Abstract: This study considers the possibility of utilizing the categories employed by researchers within a subject domain as one approach for mapping the knowledge organization of that domain. The idea is tested using a sample of research reports dealing with the information behavior of people (e.g., physician, nurses, and patients) in the health care domain. It was relatively easy to identify category systems within the sample of texts. Consideration of the configuration of category systems employed across the analyzed texts shows a pattern of broad consistency. Many of the research reports employed some version of a diagnosis, treatment, and referral category system related to the information behavior of physicians, though there were many variations on this general category breakout. There were also many uses of alternative category systems (e.g., specialty area, information types, and dilemmas) that suggest the applicability and use of alternative views of the situations being studied. The key challenges for the future appear to be the specification of relations among category systems of different research reports and visualization of the identified rich relationships including the evolution of category systems over time.

1. Introduction

The purpose of this study is to consider the possibility of employing the categorizations adopted for or resulting from research as a vehicle for organizing knowledge in a specific domain. The idea for considering this possibility comes from several sources.

First, I have recently had the pleasure of serving on the examination/dissertation committees of two students, Lee Ann Paris and Keith Cogdill, who have interests in the area of information needs and seeking in the health care arena. Keith Cogdill, in particular, considered the issue of differing categorizations in research reports in his comprehensive examination materials. The experience of interacting with these students and their writings led to my use of the health care area in this study.

Second, Taylor's (1991) consideration of information use environments and the more recent work of Hjørland and Albrechtsen (1995) and Hjørland (1997) on domains and domain analysis suggest the importance of attention to disciplinary problems and disciplinary knowledge structures.

Thirdly, a long-standing interest in the structure and characteristics of texts, terminology and systems of categories drew my attention to the critical role of categories in research reports.

Fourthly, there seems to be an inherent tension in any system of categories between a need for stability, consistency, and control and a need for flexibility, variety, and adaptability. On the one hand, stability, consistency and control are needed so that the category structure can maintain itself and the possibility of communication is enhanced as users can rely on, for instance, the stability of the classification scheme and the consistency of its application over
time (Bowker, Timmermans, & Starr, 1995). On the other hand, flexibility, variety, and adaptability are needed so that a category structure can adapt to various situations, reflect differences in how things are called, be updated as the 'new' comes into play, and the possibility of communication is enhanced as users are able to find 'new' concepts represented (Albrechtsen & Jacob, 1997, Lin & Chan, 1997).

This dilemma of categories arises because communication requires both stability/consistency/control and flexibility/variety/adaptability. This communication role of categories seems particularly evident in Star's (1989) idea that category structures serve as boundary objects. These objects provide links between information systems and the variety of people, tasks, and purposes that the system serves. Thus, it seems possible that the body of categories (as boundary objects) reported in research reports might provide a means of reducing this tension in category systems by showing what is stable, what is evolving, and, more broadly, how a domain is represented through its research categories.

Finally, there is the place of categories in the research process. Research tends to be either extensive—drawing inferences relating to a small number of variables regarding a population on the basis of a sample of that population—or intensive—specifying the patterns of behavior, critical factors, or significant themes present in a small number of cases. Categories play an important role in both of these research modes. As Kwaśnik (1992) notes, «Classifications are really very much like theories» (p. 63). With intensive research methods, categorization is an essential part of the interactive process of data collection and analysis. The resulting categories ultimately play a major role in describing the behavior under study and setting apart the possibly general from the situational. With extensive methods, category choices orient the research and structure the responses of subjects to questions. For both types of research, the category inputs or outputs may hide or distort important aspects of a researcher's understanding of the research situation, problem or question (Dervin, 1989).

From this negative comes the possibility of a positive. By bringing together the diversity of category frameworks that have been reported in a research domain, it may be possible to provide a rich network of concepts and relations that can be utilized to ultimately advance research. This research advancement would take place by allowing researchers to learn about the research categorizations employed by others as well as to consider possible new views that may be evident in the connections and alternative visualizations of such a compendium of categories. Thus, different research traditions could be compared in juxtaposition and perhaps reconciled or lead to new views of the field. There would be other possible uses for such a product. In addition to providing a vehicle for expert learning, novices could visualize the research concepts of a research domain and perhaps ease their learning process. Such a tool could also provide a vehicle for text or concept retrieval.

2. Study Approach

To test the possibility of research categories being employed for knowledge organization, I selected a sample of 16 texts that reported research relating to information behavior in health care. I began by searching MEDLINE and Library and Information Science Abstracts (LISA) for texts related to this subject domain. I, then, purposively selected 20 texts to cover the time range represented by the retrieved items and to represent the orientation of the total collection of retrieved texts towards physicians, medical students, nurses, patients, or combinations thereof. Overall, I was trying to preserve the variety of the collection of the retrieved documents. Four of the 20 texts were dropped—one because it was on methodology and not a research report and three others because I could not get access to the texts.

I read each of the 16 texts to be reasonably sure that the all categorizations were
identified. In most cases it was quick and easy to identify the categorizations as they were usually set off in some way (e.g., in tables, lists, and/or ita). I arrayed the identified categorizations in a table format (see Table 1), which includes a citation, the gist of the study—subjects, methods and goal, and a list of the categorizations with the name/purpose of the categorization in italics and a list of associated categories. If the category name did not seem self-evident, an example is provided. These entries use the researcher’s words.

3. Discussion

My suspicion that categorizations are inherent in research reports was supported by the analysis of the selected texts. All in the sample contained at least one categorization; most contained multiple categorizations. The one initially selected text that was dropped because it was a methodological case study did not contain any categorization. It may be that categorization is a robust characteristic of certain kinds of texts and not of others.

The predominant set of categories employed in studies of the information behavior of health professionals in this sample was diagnosis, treatment, and referral. This result is not surprising as these categories reflect the primary process aspects in the provision of health care. Yet, this is an analytical result, as these terms were not employed exclusively. That is, there is some variety in terminology (e.g., treatment and therapy). Perhaps most interesting is the extension of the categories by some researchers to include, for instance, etiology (e.g., causes) along with diagnosis. Knowledge of etiology and diagnosis might, accordingly, come together in treatment. The variety within even related categories across the 17 studies provides a basis for investigating terminological and classificatory relationships across time and space for a variety of researchers.

Also, some studies of the information behavior of health care professionals have provided different cuts or views of information requirements, seeking, and use. Distinctions between medical procedures and facts (procedural and declarative knowledge); identification, explanation, and description in information use; and medical records and the literature as sources of information provide a basis for enriching the analytical conceptualizations employed in research.

A variety of other categorizations beyond information behavior come into play and address aspects that reflect on the situation of health care. These include medical specialties, problems, dilemmas, and barriers. Categorizations that reflect on the characteristics of information sought or employed are evident also. These include information sources, type of information, time frame, and information assessment criteria. In addition, there was even an attempt to express the meaning of information in the form of a typology formed by formal vs. informal and general vs. specific dimensions.

4. Future Directions

Among possible next steps, three seem particularly critical: 1) Mapping relationships among categorizations to reflect possible relationships of potential interest to those who might utilize the knowledge content of Table 1. This activity seems to be necessary if the information potential of such texts is to be realized through a knowledge organization approach of this sort (Hjørland, 1997). 2) Investigating the flexible visualization of such relationships along with the possibility of allowing users to explore relationships according to their own definitions. This seems a strategy for reducing the tension between stability and change in category systems. Here users would be able to change what they see by focusing on mapped relationships within and across categories or defining their own. A capability of
viewing the change in categories over time would also seem useful as well as other exploratory visualization mechanisms that would allow a change in focus to highlight different views of the subject matter. 3) Testing the resulting product with possible users.

References
Kwaśnik, B. (1992). The role of classification structures in reflecting and building theory. Advances in Classification Research, 3, 6

Analyzed Texts

Advances in Knowledge Organization, Vol.6(1998)
Biomedical Research, 25, 181-200.


## Table 1: Summaries of the Analyzed Texts

<table>
<thead>
<tr>
<th>Text Reference</th>
<th>Gist of Study (Subjects/Methods/Goal)</th>
<th>Categories Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northup, Moore-West, Skipper, &amp; Teal (1983)</td>
<td>Physicians, residents, medical students, Critical incident technique, Everyday information needs and sources</td>
<td>Purpose—diagnostic, therapeutic, epidemiological, etiologic, Problems—disease related, procedure, drug related, basic science, other, Motivation—direct patient care, curiosity, share with others, research, patient education, referral decision, other, Specialty—internal medicine, surgery, pharmacology, primary care, basic science, other clinical sciences, other, Disease—single-body system, multi-body system, vector-caused/injuries/poisoning/occupational, female genital/pregnancy/endocrine, high prevalence (cardiovascular/neoplastic) (from MESH), Type of Information—standard medical, new, rare, controversial, other, overview/background, Time Frame—today, 1-2 days, week or more, Type of Resource Used—colleagues, journal article, book, lab tests, audiovisuals, other, Sources of Resources—personal library, reprint file, library, other</td>
</tr>
<tr>
<td>Covell, Uman, &amp; Manning (1985)</td>
<td>Physicians (internists), Interviews, Self-reported information needs relating to patient management</td>
<td>Use of information in Patient Care: diagnosis, drug information, obtaining records or information from other health care professionals, patient services, use or interpretation of test results, general review of disease topic, Type of Information—Medical Fact, Medical Opinion, Non-medical information, Subspecialty area—Cardiology, Hematology/Oncology, Endocrinology, Neurology/Psychiatry, Rheumatology, Pulmonary Disease, Gastroenterology, Infectious disease, Nephrology, Not Internal Medicine, Information Sources—Print (general texts, specialty texts, pharmaceutical texts, journals, drug company information, self-made compendia), Human (specialist, generalist, office partner, pharmacist, nurse, other health professional), Other (laboratory data, patient response, other), Barriers—Awaiting laboratory reports or patient response, source unknown, human source not available, print source not available, source not reliable, too time-consuming, other</td>
</tr>
<tr>
<td>Text Reference</td>
<td>Gist of Study (Subjects/Methods/Goal)</td>
<td>Categories Employed</td>
</tr>
<tr>
<td>----------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Timpka & Arborelius (1990)                         | Advice nurses  
Videos, stimulated recall, consultation maps  
Dilemmas during phone advice | Dilemma situations—medical knowledge (history taking, physical examination, conclusion), social knowledge (communication breakdown—did not understand patient, patient took control of consultation, bizarre behavior, disagreements over medical facts, patient not telling whole story, interruptions; organizational knowledge—information missing from medical record, unavailable medical record, scheduling, prescription routines, coordination with nurse practitioners), personal competence (faux pas)  
Dilemma specialty areas—orthopedics, internal medicine, psychiatry, neurology, pharmacology, infectious diseases, urology, otolaryngology, social medicine |
| Forsythe, Buchanan, Osheroff, Bankowitz, Blumenfeld, & Miller (1991) | Physicians, residents, interns, medical students  
Observation  
Clinical requests for information | Motivation—patient care, evaluating trainee knowledge, teaching, combinations of the above, other  
Subject—specific patient (history, exam findings, lab studies), therapy (drugs, other), diagnosis, other  
Sources—patient record, hospital information system, library (textbook journal, MEDLINE), patient, experience physician, trainee, automated information system, other  
Type of Response Required—fact, synthesis of patient and medical knowledge, speculation, other  
Generality—patient-specific, class-specific  
Urgency—same day, within a few hours, now, ambiguous  
Comprehensive information needs—currently satisfied, consciously recognized, unrecognized |
| Forsythe, Buchanan, Osheroff, & Miller (1992)      | Physicians  
Ethnographic observation/conversation  
Clinical information needs | Information needs—Formal, globally-applicable, informal, local; Broad issues in medical practice (ability to pay);  
Form of expression—direct questions, indirect questions (identifiable as questions via context), nonverbal (gestures); tacit (not expressed, but underlays the conversation)  
Meaning of information—1) formal vs. informal; 2) general vs specific (global vs. local; temporal); typology formed by 1) and 2) |
<table>
<thead>
<tr>
<th>Text Reference</th>
<th>Gist of Study (Subjects/Methods/Goal)</th>
<th>Categories Employed</th>
</tr>
</thead>
</table>
Request information on a case, evaluated impact  
Impact of library services on clinical decision making | Information impact on patient care—handled situation differently, diagnosis, test choice, drug choice, other treatment choice, hospital stay, patient advice, post-hospital care or treatment  
Information impact on adverse event avoidance—hospital admission, patient mortality, hospital-acquired infection, surgery, additional tests or procedures, additional outpatient visits  
Information assessment criteria—quality (relevant, accurate/current), cognitive value (refreshed memory, substantiated prior knowledge, new knowledge), patient care (information was of clinical value, better informed clinical decisions, contributed to higher quality care), saved time  
Information sources—library, diagnostic imaging, lab tests, colleagues |
| Dee & Blazek (1993)          | Physicians (rural)  
face-to-face interviews, observation, patient charts  
information needs and behavior | Information needs—Treatment, diagnosis, etiology, psychological aspects  
Information sources—colleagues, medical meetings, textbooks, hospital library |
| Cameron, Corbett, Duncan, Hegyi, Maxwell, & Burton (1994) | Hospital patients  
survey  
patient information needs & satisfaction levels | Gaps—hospital life (ward routine), map, long term effects of illness, care once released from hospital, effects of drugs, clarification on staffing, prevention of illness, social effects of illness  
Time dimension—Prior to admission, during stay, just prior to release  
Type of information—verbal, leaflet, booklet, book, video, tape  
Source of information—doctor, nurse, other medical staff, relative/friend |
| Wildemuth, de Biek, Friedman, & Miya (1994) | Medical Students, first year  
Clinical Scenarios, toxicology  
Questions for librarians vs. physicians | Form/purpose—identification—personal object (toxins), impersonal object (references, chemical composition), action (effects, tests); definition; description; placement—location of toxin; explanation—cause, effect, environment; process—treatment, associations, follow-up, environmental effects, other |
interviews  
clinical questions and motivation for information seeking | Information needs—diagnosis, patient management  
Factors influencing pursuit of new information: knowledge, uneasiness, urgency, potential help, potential harm, generalizability, edification, liability, patient expectation, peer knowledge, existence of an answer, difficulty  
Information sources—textbooks, drug information, consultant (other specialty), colleague (same specialty, non-physician (e.g., pharmacist), personal reprints, other source, computer search (e.g., MEDLINE) |
<table>
<thead>
<tr>
<th>Text Reference</th>
<th>Gist of Study (Subjects/Methods/Goal)</th>
<th>Categories Employed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pound, Bury, Gompertz, &amp; Ebrahim (1995)</td>
<td>Stroke patients in depth interviews components of care valued by patients in retrospect</td>
<td>Components of care—being cared about (kindness, warmth, individual attention, encouragement, genuine indications), clinical care (decisive actions, administration of tests, rehabilitation), nursing care (relief that patient or family did not have to perform care functions, inflexibility of ward regime (e.g., visiting hours), information or advice (about condition, prognosis)</td>
</tr>
<tr>
<td>Stavri (1995)</td>
<td>Residents, internal medicine &amp; family practice case vignettes under experimental conditions relationship between diagnostic problem (urgency, etiology) and information seeking question</td>
<td>Information seeking question—verification (yes or no), disjunctive (one or the other), concept completion (where, what), feature specification (value), quantification (how much, how many), causal antecedent (history), causal consequence, goal (motives for action), enablement (when), instrumental/procedural (perform action) Medical Problems—base, etiology, urgency, and etiology/urgency</td>
</tr>
<tr>
<td>Florance (1996)</td>
<td>Physicians relevance (clinical applicability) scoring exercises, semi-structured interviews specifications for a prototype clinical extract</td>
<td>Information sought from articles: patient (age, sex, condition, history, diagnosis), Study (# of patients, location, treatment, length, outcome), Article (prognostic information, natural history, survival data, publication type, scope of content) Relevance factors—positive (matches case, contextual data, applicable data, scientifically sound, respected source, good presentation format, support statistical inferences, up-to-date); negative (wrong therapy, wrong subject aspect, not enough cases, found better article, unfamiliar source, wrong type of disease, wrong aspect of care, too old, wrong article type, presentation problem, no abstract</td>
</tr>
<tr>
<td>Baker (1997).</td>
<td>Multiple sclerosis patients, women questionnaires information needs</td>
<td>Information needs—physiological (e.g., fatigue, muscle spasticity), psychological (e.g., dealing with a hidden disability, feelings of uncertainty), social (e.g., support groups), economic (e.g., insurance, career, discrimination) Information orientation—monitoring (active information seeking), Blunting (avoidance distraction from information seeking)</td>
</tr>
<tr>
<td>Cogdill &amp; Moore (1997)</td>
<td>Medical students, first year clinical scenario information needs</td>
<td>Information needs—general disease information, diagnosis, treatment/therapy, general management, patient specific information, procedures, psycho-social issues, other Type of resources—diagnostic test, general medical literature, medical textbooks, reference material, bibliographic retrieval, journal articles, human experts, patient chart or interview, educational software</td>
</tr>
</tbody>
</table>