PRODUCTION NOTE

University of Illinois at Urbana-Champaign Library
How Classifications Work: Problems and Challenges in an Electronic Age

Geoffrey C. Bowker
and
Susan Leigh Star
Issue Editors
Library Trends, a quarterly thematic journal, focuses on current trends in all areas of library practice. Each issue addresses a single theme in depth, exploring topics of interest primarily to practicing librarians and information scientists and secondarily to educators and students.

Editor:
F.W. Lancaster

Managing Editor:
James S. Dowling

Publications Committee: Leigh Estabrook, Janice Del Negro, Monica M. Walk, Betsy Hearne

Library Trends is published four times annually—in summer, fall, winter, and spring—by the Graduate School of Library and Information Science at the University of Illinois, Urbana-Champaign, 501 E. Daniel Street, Champaign, IL 61820-6211.

Subscriptions: Institutional rate is $75 per volume (plus $7 for overseas subscribers). Subscriptions for an individual are $50 (plus $7 for overseas subscribers). Registered students may subscribe for $25 (plus $7 for overseas subscribers). Individual issues are $18.50 (shipping included); back issues other than those from the present year are $10 (plus shipping). Claims for missing numbers should be made within six months following the date of publication. All foreign subscriptions and orders must be accompanied by payment.

Address orders to: University of Illinois Press, Journals Department, 1325 S. Oak Street, Champaign, IL 61820. For out-of-print issues, contact University Microfilms International, 300 North Zeeb Road, Ann Arbor, MI 48106. Postmaster: Send change of address to University of Illinois Press, 1325 S. Oak Street, Champaign, IL 61820-6903.

Copyright © 1998 by the Board of Trustees of The University of Illinois. All rights reserved. Printed in the U.S.A. ISSN 0024-2594.

Periodicals class postage paid at Champaign, Illinois.

Authorization to photocopy items beyond the number and frequency permitted by Sections 107 and 108 of the U.S. Copyright Law is granted by the Board of Trustees of the University of Illinois, provided that copies are for internal or personal use, or for the personal or internal use of specific clients and provided that the copier pay a fee of 10 cents per page directly to the Copyright Clearance Center (CCC), 222 Rosewood Dr., Danvers, MA 01923. The CCC code for Library Trends is 0024-2594/88 $0.10. To request permission for copies for advertising or promotional purposes, or for creating new works, please contact the Graduate School of Library and Information Science, Publications Office, 501 E. Daniel Street, Champaign, IL 61820-6903.

This journal is abstracted or indexed in Library and Information Science Abstracts, Current Contents, Current Index to Journals in Education, Information Science Abstracts, Library Literature, PAJS, and Social Sciences Citation Index.

Procedures for Proposing and Guest Editing an Issue of Library Trends

We encourage our readers to submit ideas for future Library Trends themes, issue topics are developed through recommendations from members of the Publications Committee and from reader suggestions. We also encourage readers to volunteer to be issue editors or to suggest others who may be willing to be issue editors.

The style and tone of the journal is formal rather than journalistic or popular. Library Trends reviews the literature, summarizes current practice and thinking, and evaluates new directions in library practice. Papers must represent original work. Extensive updates of previously published papers are acceptable, but revisions or adaptations of published work are not sought.

An issue editor proposes the theme and scope of a new issue, draws up a list of prospective authors and article topics, and provides short annotations of the article’s scope or else gives a statement of philosophy guiding the issue’s development. Please send your ideas or inquiries to F.W. Lancaster, Editor, Publications Office, 501 E. Daniel Street, Champaign, IL 61820-6211.
How Classifications Work: Problems and Challenges in an Electronic Age

Geoffrey C. Bowker
and
Susan Leigh Star

Issue Editors
This Page Intentionally Left Blank
How Classifications Work: Problems and Challenges in an Electronic Age

## CONTENTS

**Introduction**  
*Susan Leigh Star and Geoffrey C. Bowker*  
185

**Social Constructs and Disease: Implications for a Controlled Vocabulary for HIV/AIDS**  
*Jeffrey T. Huber and Mary L. Gillaspy*  
190

**Seeking the Subject**  
*Jennifer Tobias*  
209

**Grounded Classification: Grounded Theory and Faceted Classification**  
*Susan Leigh Star*  
218

**Mapping Beyond Dewey's Boundaries: Constructing Classification Space for Marginalized Knowledge Domains**  
*Hope A. Olson*  
233

**The Kindness of Strangers: Kinds and Politics in Classificatory Systems**  
*Geoffrey C. Bowker*  
255

**The Dynamics of Classification Systems as Boundary Objects for Cooperation in the Electronic Library**  
*Hanne Albrechtsen and Elin K. Jacob*  
293
Psychiatrists Make Diagnoses, but not in Circumstances of Their Own Choosing: Agency and Structure in the *DSM*  
Mark A. Spasser  

About the Contributors
Introduction

SUSAN LEIGH STAR AND GEOFFREY C. BOWKER

CLASSIFICATION IN THE WILD*

"CLASSIFICATION IS THE SLEEPING BEAUTY OF library and information science," said Hanne Albrechtsen (personal communication, November 1997). In some ways, at the most technical core of traditional library science, it also stands as a bridge builder between the past and the future of our field. Anthropologists have long seen classification as a tool for understanding culture. The distinctions that people make, as Lévi-Strauss (1969) argued in his famous report on The Raw and the Cooked, constitute cultural membership if not culture itself.

Classification involves the informal embedded in the formal and vice versa. If anthropological/cultural distinctions shape culture, the attempt formally to evaluate and improve those distinctions forms much of what we think of as information systems. We then take cultural cues from the systems so created.

In traditional library classification research, there have been two distinct challenges. The first is an ethnographic challenge: what distinctions does this specific group of clients use in forming their knowledge culture? How may we mirror in the thesauri, catalogs, and other search and retrieval tools we create for them? This challenge is one of *verstehen*—i.e., understanding the sense of vernacular terms. Here the information science researcher becomes an anthropologist—i.e., how to disambiguate

* These words are a play on Edwin Hutchin’s excellent Cognition in the Wild (Cambridge, MA, 1995), which explores aspects of planning, coordination, and cognition from a social/organizational and material viewpoint.
terms, decide what distinctions are really necessary, and which are reflections of ephemeral debates or fads?

The second challenge is a formal one in two parts concerning the structure of the information system or tool:

1. Occam's razor: how many terms can we afford given system capacity, user capability, and the means of distribution?
2. Structural soundness: are logical flows followed, are the branches of the tree carrying the right weight in proportion to the trunk, are the means of navigation easily grasped for accurate modeling?

There are no a priori solutions here: each scheme must be taken in its own context of use.

Classifications that work in the real world must meet both challenges simultaneously. For example, in studying the history of the International Classification of Diseases (Star & Bowker, 1994; Bowker & Star, in press), we noted that the designers of this global classification system must constantly make practical tradeoffs between the two challenges. In order to do justice to the range of subtle vernacular terms used by medical personnel around the world, a huge unwieldy list would have to be developed. In order for physicians and other users to actually employ the system, a much shorter key to filling out forms is the only possible alternative.

As the Internet, Web, and various digital libraries burst their boundaries and appear on desktops and in homes, the tension between these two challenges deepens. What do we understand about the interplay between vernacular classifications and the more formal structures underlying search engines, online catalogs, and other electronic guides? For groups of users that may be both global and unknown, what is the meaning of joining the two aspects of classification? What is usability in the context of both the Web and the intimate desktop?

The combination of the cultural and the formal in turn produces a third challenge—a moral and ethical one. For large-scale systems, whose voices will be heard and whose silenced? Whose culture will become the taken-for-granted and whose the exotic other? Where makers and users of classification systems do not address these questions, silent inequities prevail. The dominant voices may become the common sense of the designers or the loudest of the user voices (Forsythe, 1992).

The articles in this collection each address this set of issues from a variety of angles. Huber and Gillaspy's article tackles a core methodological issue surrounding the translation from vernacular systems of classification and vocabulary to the more formal controlled vocabulary systems such as those in LCSH, MESH, and ICD/DSM. Looking at AIDS and HIV vocabularies within the communities of gay men and IV drug users, they show how the political and historical situations of those affected help shape vocabularies and, in turn, the usability of more formal systems. They re-
port on their research-testing and refining vocabulary within a large-community AIDS provider organization via a series of face-to-face meetings with a range of users, from health care providers to patients and activists. They then move the findings to a large-scale vocabulary test sponsored by the National Library of Medicine. Grappling with the rapidly changing vocabulary, sensitivity to issues of stigma and race, and mapping the vernacular to the large-scale system, their research forms an exemplar of participant design and community-based research. As well, it points in some important directions for moral, political, and ethical aspects of applied classification work.

Jennifer Tobias's piece is similarly a manifesto for the examination of marginal, stigmatized, and experimental cultures under the rubric of "cyberspace." She continues in the tradition of Sanford Berman, challenging the worlds of cataloging and classification to keep up with cultural and social changes and to resist elitist or ethnocentric tendencies in large-scale systems. She argues that librarians need to become conversant in a range of specialized languages in order to provide the best possible services—and that this need is only heightened by the fluid nature of documents on the Web.

Another methodological piece, that by Star, provides a more abstract example of how "classification in the wild" could be joined with some foundational work in both library and information science on the one hand and sociology on the other. Star compares the faceted classification method developed for use in libraries by Ranganathan with the grounded theory method of qualitative analysis developed by Glaser and Strauss in sociology. Both systems struggle with the core dilemma posed above—i.e., how to braid the formal and the informal together in the study of classification. Future work in this area would, it is hoped, link some of the naturalistic findings of qualitative research with the developing technologies stemming from faceted classification and advanced networked information technologies for navigation (see Bradley & Sutton, 1993; McCombs & Maylone, 1998 for excellent overviews of research in the area).

Olson finds a theoretical mandate in the work of postmodern and feminist philosophy for some of the space mapped by the three discussions above. Perhaps instead of speaking in the "master voice" of the state-sanctioned list, she suggests that we might see the classification of marginalized domains as an exercise in cartography. Picking up on many of the same issues raised by Tobias and Huber and Gilaspy, she notes the implicit Western, and often sexist, constructions in the Dewey Decimal Classification. As an exercise in both imagination and method, she calls for a spatial imagery for neglected and stigmatized domains. Drawing on recent work in critical geography, she shows us both critique and a positive path for classification work and research.
Bowker's article adds a close reading of information-system as cultural artifact. He suggests that one can find traces of social and political debates in the classification of disease entities in the International Classification of Diseases (ICD). He argues that this encoding of social and political dimensions is a natural feature of such classification schemes, which thereby serve to encourage the development of some forms of knowledge while discouraging others. His article draws particular attention to the organization of time and space in the ICD—suggesting in turn that the scheme favors particular kinds of narrative of diseases, and that it best represents disease in the developed world.

Extending this theoretical contribution, Albrechtsen and Jacob conceptualize information in another inherently spatial fashion, that of information ecologies. They note that classification systems play a key role as boundary objects (Star & Griesemer, 1989) in the organizations in which they are used—i.e., they serve simultaneously as lingua franca and as specialized tools in particular domains. Drawing on public library examples from the Book House Project and Database 2001, they emphasize especially the continually reconstructed nature of classifications as organization and knowledge tools.

Finally, Mark Spasser examines issues of agency and structure in psychiatrists' use of the Diagnostic and Statistical Manual (DSM) of mental disorders. Drawing on Gidden's concept of structuration, he suggests that the DSM, through its propagation of a particular biomedical reading of mental disorders, severely constrains the kinds of research that can be done by psychiatrists. He argues that Gidden's work provides an analytic purchase on this constitution of psychiatric discourse through classification and maintains that the concept of structuration will also provide a useful tool for understanding the development and change of library classification systems.

CONCLUSION

Library and information science stands at a historical crossroads. New information tools appear daily and are used in more kinds of settings than before. They are part of not only desktops or kiosks but, increasingly, of living rooms, gyms, cars, banks, and hospitals inter alia. The formal tools of classification construction and evaluation, and decades of experience and research in working with client populations, give us a unique suite of tools for understanding this phenomenon. As sleeping beauty wakes up in this new world, there is a unique opportunity for her to build bridges across its rivers and canyons.

ACKNOWLEDGMENTS

The preparation of this special issue was supported in part by NSF Grant 9514744, from Societal Dimensions of Engineering, Science, and
Technology (formerly Ethics and Values Studies), and NSF/ARPA/NASA NSF 93-141 DLI. The authors gratefully acknowledge this support.

**NOTE**


**REFERENCES**


Social Constructs and Disease: Implications for a Controlled Vocabulary for HIV/AIDS

JEFFREY T. HUBER AND MARY L. GILLASPY

ABSTRACT

The body of knowledge associated with the human immunodeficiency virus (HIV) and the acquired immune deficiency syndrome (AIDS) represents complexity not present in any other disease. HIV infection is not only an extremely complicated disease process, but it also transcends the boundaries of biomedicine. Various domains shape the construction of HIV/AIDS as chronic disease with the societal construct circumscribing the body of knowledge concerning the pathological, mirroring the complexities of the malady itself. Disease, and the respective body of knowledge, co-exist within a social reality; consequently, a controlled vocabulary designed to facilitate knowledge organization and access relative to HIV/AIDS must reflect the complexities of this socially constructed reality.

INTRODUCTION

Social constructionists posit that reality is constructed through dynamic socialization and that the sociology of knowledge must examine the process in which this reality construction occurs (Berger & Luckmann, 1966, p. 1). Sociology of knowledge deals not only with empirical knowledge relative to various societies but also with the processes by which bodies of knowledge become established as social realities. In essence, reality evolves through continued socialization, yielding outcomes that result from

* This article is based, in part, on a presentation given at the 60th Annual Meeting of the American Society for Information Science, November 1-6, 1997, Washington, DC.
social interactions, negotiations, and power. Where the human immunodeficiency virus (HIV) and the acquired immune deficiency syndrome (AIDS) are concerned, social construction of reality is grounded in the spatialization and politicization of the pathological.

The body of knowledge associated with HIV/AIDS represents a complexity not present in any other disease. Furthermore, the epidemic has altered the model of information production and consumption and has spawned its own vernacular, one representative of a diverse population of information producers and consumers. Further compounding this complicated communication picture, the body of information surrounding HIV/AIDS continues to grow at an epidemic rate, often in tandem with the numbers of reported cases. Finally, HIV infection is not only an extremely complicated disease process, but it also transcends the boundaries of biomedicine. Various domains shape the construction of HIV/AIDS as chronic disease, including the political, social, economic, legal, philosophical, psychological, religious, and spiritual ramifications associated with the illness. The societal construct within which the body of knowledge concerning HIV/AIDS exists mirrors the complexities of the malady and the various controversies associated with it. This diseased body of knowledge—a body of knowledge that breathes life into the pathological by providing it visibility—exists because of HIV/AIDS. Disease, and the respective body of knowledge, co-exist within a social reality, a social reality that binds and circumscribes. Consequently, the organizational schema of a controlled vocabulary designed to facilitate knowledge organization relative to HIV/AIDS must be broad in coverage yet specific in terminology so that the multidisciplinary and interdisciplinary nature of the epidemic is reflected. In representing the dynamic nosological record of HIV/AIDS, the controlled vocabulary captures the societal construct circumscribing the pathological.

BACKGROUND

Complex Nature of HIV/AIDS

Infection with the human immunodeficiency virus results in a complex chronic disease process, complicated by various nonbiomedical factors. The disease itself is characterized by a constellation of signs and symptoms that culminate in a diagnosis of acquired immune deficiency syndrome. Ultimately, most individuals infected with HIV die of AIDS-related causes. From a biomedical perspective, what differentiates HIV from other chronic disease processes is the variety of opportunistic infections and cancers commonly associated with AIDS as well as HIV-related wasting and dementia and the wide variation in the dying trajectory. Although there have been numerous therapeutic advances where HIV is concerned, drug regimens, when available and accessible, have not proven uniformly effective. Combination therapy involving antiretrovirals and protease
inhibitors, while greeted with much fanfare, has proven to be a great disappointment to the many HIV-infected individuals who have failed to improve while taking the drug cocktails. Moreover, of the 612,078 AIDS cases reported to the Centers for Disease Control and Prevention (CDC) through June 1997 in the United States, 379,258 have died (Centers for Disease Control and Prevention, 1997, p. 14).

Further exacerbating the medical complexities of the illness, HIV is complicated by myriad factors outside the biomedical arena—economic, legal, political, psychological, religious, social, spiritual—that compound disease chronicity. These components of an individual’s psychosocial reality exist in tandem with the biophysical illness with stigma trajectory corresponding to disease course progression (Alonzo & Reynolds, 1995, p. 306). Although the spatialization of disease has been plotted along a historical continuum that forms the foundation for modern medicine (Foucault, 1975, pp. 3-20), the politics and stigma associated with HIV/AIDS prevents the illness from advancing to its ultimate position in the sequence (Huber, 1996, pp. 6-9). The pathological continues to exist within a socially defined set of spaces. As well as affecting the emotional, mental, and physical well-being of the HIV-infected individual, these nonbiomedical complications dramatically impact education and prevention efforts, treatment advances, and coping mechanisms. The repercussion of infection and disease manifestation is much more than an individual life event. In fact, given the multifarious nature of the pathological, HIV transcends the boundaries of life and death (Huber, 1993a, pp. 230-31).

HIV/AIDS Information and Communication

Since the beginning of the epidemic in the early 1980s, information has been viewed as a key resource in efforts to prevent HIV transmission, manage various disease complications, and ultimately prolong life. As HIV/AIDS-related information was initially limited, however, in size, scope, and availability (SantaVicca, 1987, p. 115), underground press networks, supported largely by affected individuals and community-based organizations, began emerging within the first few years of the epidemic. These networks subsequently have evolved into recognized legitimate information resources with many AIDS service organization newsletters now being indexed by the National Library of Medicine for inclusion in its HIV-specific bibliographic database, AIDSLINE.

The AIDS pandemic, in effect, has witnessed a confluence of roles regarding information creators, seekers, and providers (Ginn, 1987, p. 333). This paradigmatic shift has resulted in a nontraditional scientific communication model where traditional consumers of information are very often producers, and traditional producers are consumers. In a traditional scientific communication model, information is generated by researchers, disseminated, accumulated, distilled, and applied in the clinical arena (see
Figure 1). Practical information, when made available, is watered down into lay terms for public consumption (Patton, 1990, p. 5). However, in this nontraditional model, traditional consumers are active contributors to the discourse (see Figures 2 and 3). As a result, HIV-related information is currently produced and consumed at virtually every level—individual, institutional, organizational, community, local, regional, national, and international.
Producers

HEALTH CARE
PROVIDERS,
RESEARCHERS

PATIENTS,
LAY CAREGIVERS

EDUCATORS

AFFECTED
POPULATIONS
COMMUNITIES,
ET AL

Consumers

PATIENTS,
LAY CAREGIVERS

HEALTH CARE
PROVIDERS,
RESEARCHERS

EDUCATORS

AFFECTED
POPULATIONS
COMMUNITIES,
ACTIVISTS

DISCOURSE

MULTIPLE VERNACULARS

BIOMEDICAL
EDUCATIONAL
LEGAL
POLITICAL

ET AL.

PSYCHOLOGICAL
SOCIOLOGICAL
THEOLOGICAL
LAY PUBLIC

DISSEMINATION
& INTEGRATION

COMMUNITY RESOURCES

ELECTRONIC NEWSGROUPS
WORD OF MOUTH
HOTLINES
NEWSLETTERS
WORLD WIDE WEB
CONFERENCE PRESENTATIONS
JOURNAL ARTICLES

DISCIPLINES, APPROACHES

CONFERENCE PRESENTATIONS,
JOURNAL ARTICLES

Figure 3. Nontraditional Scientific Communication Model.
In addition to HIV-related information being produced and consumed at multiple levels, the epidemic has spawned its own vernacular, one representative of the diverse group of individuals infected with the virus and those working within the AIDS arena (Huber, 1993b, p. vii; Huber & Gillaspy, 1996b, p. 1). The language of the pandemic embodies cultural predispositions. This vocabulary consists not only of technical, scientific, and biomedical terminology but also includes verbiage germane to the lay population directly affected by the malady as well. The lexicon also reflects the various disciplines touched by the disease.

Further complicating access to HIV-related information, the body of knowledge concerning the epidemic is growing exponentially. The literature continues to increase in volume parallel to the rise in the number of documented cases of AIDS worldwide (Huber & Gillaspy, 1996a, p. 297). In addition, HIV-related information is currently produced in every conceivable format—audiovisual, electronic, print, realia—and is present in all discipline-specific bodies of knowledge affected by the pandemic.

Societal perceptions and individual perspectives fashion the pathology of the HIV/AIDS epidemic with the course of disease progression marred by politics and stigma. Complexities associated with both the pathological condition and the body of knowledge concerning HIV/AIDS exist within, and because of, social constructs circumscribing the pandemic. Illness, information, and intricacies are all entwined, evolving relative to both scientific advances and social interactions.

**Social Constructs**

*Reality Construction*

Although a person's conception of reality—fear and danger, abnormality and stigma, health and beauty—may be very individual, this perception, to a large extent, is culturally patterned (Ferrante, 1988, p. 224). Events, actions, attitudes, and beliefs are mediated by historical and cultural factors. Individual reality is a social construction and not necessarily an absolute truth. Personal conceptions are reflective of a much larger construct, one that transcends any individual. The relationship, however, between individual perception and social construct is an integral one in that dynamic socialization shapes the construction of reality.

Reality is constructed within three realms—social, physical, and individual—and is composed of societal definitions and interactions (Charon, 1992, pp. 37-38). Social, physical, and personal realities operate in conjunction to construct and define an individual's reality. Social reality is molded within the context of societal circumscription of the individual bounded by his or her culture. This spatial reality is grounded in social action and interaction. A physical reality also exists, independent of social reality, as a paradigmatic structure present at the instance of every situation. Physical objective reality consists of an individual's surroundings...
and current set of circumstances. Social reality responds, in part, to physical reality where physical reality is the existing situation. The situation, however, is defined by one's social reality. In addition, each individual possesses a personal reality based on that person's unique perspective. Personal reality, consciously and unconsciously, is shaped through socialization. Social reality, physical reality, and personal reality interact simultaneously to form one's cumulative perception of what is real. This cumulative perception then, when combined with other individuals' realities derived from a similar perspective, forms the basis for defining social constructs. Social constructs, however, being created from societal perceptions, may lack scientific foundation. In fact, political and economic elites very often generate media images and other forms of discourse to influence social construction of meaning and reality (Gamson, Croteau, Hoynes, & Sasson, 1992, p. 374). Social ideologies and political interests, in essence, shape the construction of reality. Where HIV and AIDS are concerned, the social construct within which the pathological exists is built upon the politics of bodies and disease.

**Bio-Politics Versus Bio-Power**

Within the historical development of civilization, bio-politics, or politics of the body, emerged as a result of the "proliferation of political technologies [that] ensued, investing the body, health, modes of subsistence and habitation, living conditions, the whole space of existence" (Foucault, 1978, pp. 143-44). Prior to, and in parallel with, the development of bio-politics, there was an "explosion of numerous and diverse techniques for achieving the subjugation of bodies and the control of populations, marking the beginning of an era of bio-power" (Foucault, 1978, p. 140). Both bio-politics and bio-power continue to mold reality construction today. Bio-politics relates to the empowerment of individuals, while bio-power may be thought of in terms of power over bodies by bodies—i.e., corporeal and social, individual and collective. Within the AIDS arena, politics of the body and the body politic are inextricably intertwined and often diametrically opposed. Nowhere perhaps is the strife between bio-politics and bio-power more obvious than in the debate between public health and individual rights, with regulation of individual sexual practices possibly being the best illustration (Gillaspy & Huber, in press). Societal normalization of sexuality is an instrument of power (Hewitt, 1991, p. 229). By defining what is normal, the body politic creates a tool that can be applied to control individual sexual behavior. Discourse—psychiatric, legal, moral, ethical—binds physical actions by rendering normative behavior. Social discourse is used and applied to define the rules of sexuality. The various standards, models, exclusions, limitations, and perversions of sexuality are derived from a particular discursive practice, based not on scientific discourse but on a
system of values and prohibitions (Foucault, 1972, p. 193). Organized religion's stance on homosexuality, the continued existence and enforcement of sodomy laws, the legal position concerning prostitution, the lack of inclusion of homosexuality as a legitimate sexual orientation in sex education curricula, and promotion of "just say no" campaigns regarding safe sex are a few examples of the establishment's efforts of regulating or administering sexual practices.

Given that gay men continue to constitute the largest affected population in the United States and a significant portion worldwide, along with the fact that engaging in unprotected sexual intercourse—heterosexual or homosexual—is the primary mode of transmission, issues involving sexuality cannot be cleaved from the complex discourse defining HIV/AIDS. In fact, where sexual practices are concerned, the struggle between politics of the body and the body politic extends well beyond the traditional establishment and into the gay community. Even within organized sexual communities, individual sex acts and identities vary widely (Vance, 1991, p. 878). For the gay community, this variation is often a source of conflict, given the difficult nature of striking a balance between maintaining sexual liberation gained in the wake of the Stonewall riots and seeking wider social acceptance among the heterosexual population. The relationship between sexuality and the AIDS pandemic has further exacerbated this debate. Prominent gay journalists advocating reduction in promiscuity as a means of HIV prevention have come under fire by gay theorists who counter that this position engenders "gay positive but sex negative" posturing (Crain, 1997, p. 28).

The contention that the struggle between politics of the body and the body politic occurs where HIV and AIDS are concerned is important because it exemplifies societal regulation—overt and covert—of the HIV-affected individual and community. Social policy and processes have been, and continue to be, used to shape the politics of the HIV-affected body, both individual and collective. These politics, supplemented by medical complexities and exacerbated further by the nonbiomedical complications of the disease, frame the social construct within which HIV exists as a chronic disease. Within the United States, this socially constructed platform is built largely upon individual perceptions and societal perspectives involving homosexuality, drug abuse, race, and gender.

Construction of Marginalized Populations

The complex nature of this chronic disease cannot be examined without considering the social construction of homosexuality given the close affinity of AIDS with the gay community in the United States. Homosexuality has been constructed socially in much of the world as a negative label, stigmatized largely because of perceived deviation from a broader societal norm. Moral entrepreneurs have toiled tirelessly in attempting to
persuade society that homosexuality is abnormal and immoral. "It is beliefs that homosexuality is evil, sick, or undesirable—and the corresponding efforts to punish, cure, or prevent it—that make homosexuality deviant" (Greenberg, 1988, p. 2). This conception of deviance has resulted in discrimination against, and repression of, individuals seeking to engage in same sex unions. There has been tremendous effort from within the gay community, however, to liberate gays and lesbians from the psychosocial stigma associated with their respective sexual orientations. "It was a historic step to have homosexuality changed from a medical anomaly to a psychological impairment in the early part of the century, and an equally significant step to have homosexuality removed from DSM-3 and ICD-9 in the early 1970s and later 1980s" (Patton, 1990, p. 3). The close association of HIV and AIDS with homosexuality, though, has threatened to unravel social tolerance extended toward gays and lesbians, with the stigmatization of one fueling the stigmatization of the other.

While the gay community has borne the brunt of the AIDS pandemic in the United States, other socially marginalized populations have been, and increasingly are being, woven into the disease-related web of devastation. Drug injection has been determined as the mode of exposure to HIV in 26 percent of reported AIDS cases among adolescents and adults in the United States with an additional 6 percent attributed to men who have sex with men and inject drugs (Centers for Disease Control and Prevention, 1997, p. 8). Moreover, of the adolescent and adult cases reported to the CDC from July 1996 to June 1997, 43 percent were black and 20 percent Hispanic (Centers for Disease Control and Prevention, 1997, p. 3). Further, incidence of AIDS among women in the United States now accounts for 15 percent of total reported cases (Centers for Disease Control and Prevention, 1997, p. 3). These groups live within social constructs that bind and circumscribe just as homosexuality is stigmatized by society.

Although men who have sex with men continue to constitute the greatest portion of the HIV-infected population in this country, injection drug users contribute significantly to the total number of AIDS cases. Like homosexuality, substance abuse has been modeled around issues involving morality and disease. Drug abuse has been constructed as a societal taboo and criminal problem, imposing a certain degree of forced invisibility upon members of that community. Moral panics and crusades relative to illicit drug use foster the perception that this is yet another disposable population devoured by demonic deviant behavior.

Perhaps one of the most visible forms of deviation, though, in a predominantly white culture, is that of race. Race is employed as a social concept to differentiate populations based on physical traits, blood types, genetic code patterns, and inherited characteristics. However, race also is applied to ascribe psychological and moral attributes, facilitating the justification of a discriminatory system exhibiting ethnocentric biases. In this
way, race categories support destructive social labeling, founded in societal perspectives rather than scientific fact. “Race categories are social constructs, that is, concepts created from prevailing social perceptions without scientific evidence” (Witzig, 1996, p. 675). Unfortunately, the continued use of race taxons, despite scientific evidence repudiating the validity of racial constructs, fosters the application of race as a negative descriptive social label. This is particularly poignant where HIV infection is concerned, given that the number of documented cases of AIDS is rising disproportionately among people of color. Blacks constitute approximately 13 percent of the U.S. population and roughly 35 percent of CDC documented AIDS cases, and Hispanics account for about 11 percent of the American populace but close to 18 percent of the documented AIDS cases (World Almanac and Book of Facts, 1997, p. 133; Centers for Disease Control and Prevention, 1997, pp. 9-11).

Tightly woven into the social fabric defining race, gender has been constructed in many cultures to portray female submissiveness and male domination as societal norm. Building on the conflict between politics of the body and the body politic, many feminists argue that the rules of sexuality have been delineated by men (Few, 1997, p. 619). In general, universalistic feminist theory views gender as being defined in terms of binary opposition—man/woman—and assumes that women are subject to gender subordination (Dugger, 1995, p. 139). Social order revolves around patriarchy, with women occupying secondary positions. While this view does not recognize the role of race, ethnicity, and nation in gender construction, it does serve to frame loosely the social construct within which HIV-infected women and those at risk for infection live in much of the world. This construct is even more binding, however, when applied specifically to women of color. Although white women are subject to societal circumscription and HIV-related prejudices, women of color are stigmatized further by individual perceptions, social expectations, cultural norms, and socio-historical development. For Hispanic women, male dominance is often typified by machismo attitudes. For black women, “womanhood was constructed not in terms of familial and domestic activities, but through black women’s role as laborer in slave, colonial, and market economies, and through their roles as domestics and surrogate mothers to white families” (Dugger, 1995, p. 140). HIV-infected women of all races are at the same time “innocent victims” and “immoral carriers,” further illuminating “prejudices which have long existed in medicine and law” (Van Vliet, 1993, p. 193).

Because of the biomedical complexities, disfiguring nature of the illness, and close association with death, HIV/AIDS would most likely have been stigmatized to some degree no matter who was initially infected (Herek & Glunt, 1988, p. 887). However, the American AIDS epidemic has been defined as a disease of marginalized populations with the resulting
social construct being shaped by this definition. The construct is not surprising, though, given that the construction has been based on a social response to a disease most prevalent among already stigmatized populations.

**Stigma**

Much of the construct circumscribing the HIV pandemic is plagued by stigma. Stigma may be thought of in terms of a language of relationships, resulting in the construction and application of deeply discrediting attributes (Goffman, 1963, p. 3). Stigma represents a deviation from some socially constructed ideal or expectation, such as adhering to an accepted sexual orientation or remaining free from a disfiguring disease (Alonzo & Reynolds, 1995, p. 304). Deviance is key to this broad multidimensional construct and acts as a negative discounting social label. However, deviation itself is not intrinsically immoral or pathological; rather, it is inferred from a culturally defined meaning. The social construct within which stigma exists allows stigmatized attributes to be discredited and tainted, resulting in prejudice and discrimination. In its extreme, stigmatization of disease withholds legitimate privileges afforded non-stigmatized maladies and imposes special obligations on those affected, thus resembling crime more than illness (Freidson, 1970, p. 236).

Blatant stigma-related rejection, prejudice, and discrimination are manifested—overtly and covertly—where HIV and AIDS are concerned. Attributes involving the disease are stigmatized because of a variety of biomedical and nonbiomedical factors, including modes of viral transmission, psycho-demographics of populations most affected, and obvious visibility of the disfiguring nature of the illness. Fear of contagion, homophobia, racism, sexuality, social perception of drug abuse, and the close association of AIDS with an unaesthetic form of death feed the stigma branding this pathological condition. The degree to which the disease is stigmatized affords the imposition of shame directly on those individuals who are HIV-infected as well as indirectly to family, friends, and partners in the form of a courtesy stigma. All too often, "existing societal fears and stereotypes quickly amalgamate with misrepresentations of medical and sociological facts" (Patton, 1992, p. 323). In fact, stigmatization of HIV and AIDS is so strong that the stigma trajectory has been conceptualized to mirror the course of disease progression, resulting in full manifestation being equated with passage to social and physical death (Alonzo & Reynolds, 1995, pp. 309-11).

The reality of the social trajectory of stigma is painfully apparent in the AIDS pandemic. Stigma and disease often outlive infected individuals through sheer irrationality and the continued politicization of the disease. "Cultural narratives of perversion and contagion seem endlessly capable of turning apparently interpretation-proof facts into ammunition for pan-
ics and discrimination" (Patton, 1992, p. 323). Social construction of the pathological condition ensures regulation of the HIV-affected, with the construct of the AIDS arena being described and recorded in the body of knowledge that has evolved since the inception of the epidemic.

Relationship between Disease and Body of Knowledge

Bodies of knowledge are bound by societal norms, policies, and processes, and scientific discoveries, social interactions, and personal beliefs are recorded within the containers of information that support those bodies of knowledge. Information and society are indelibly linked, as are pathology-specific bodies of knowledge and the diseases they represent. There is an integral relationship between a disease and the body of knowledge concerning that disease. In reality, one does not occur without the other. In order for a body of knowledge about a pathological condition to develop, the disease must exist and have been discovered. The body of knowledge concerning a disease, then, is generated to define and describe the malady, classify the pathological, and provide discourse regarding affected individuals. In return, this pool of knowledge breathes life into the pathological, providing it visibility. Without a representative body of knowledge, the disease remains invisible. Information concerning a disease, however, is bound by the life of that pathological condition and is circumscribed by any stigma associated with that illness. The body of knowledge is riddled with the same complexities as the malady itself, yielding a diseased body of knowledge where HIV and AIDS are concerned. Mirroring the complex nature of the epidemic and the controversies associated with the disease, the body of knowledge regarding the pandemic is circumscribed by the same societal construct as the pathological itself (Huber, 1996, p. 33).

HIV/AIDS Controlled Vocabulary

Given this societal circumscription, the controlled vocabulary and classification structure used to organize the body of knowledge associated with HIV/AIDS has been problematic. The very names first used to describe the disease reflected not only the community originally most affected but also society's stigmatization of it. These names included gay cancer, gay pneumonia, gay bowel syndrome, gay-related immune deficiency (GRID), acquired community immune deficiency syndrome (ACIDS), and community acquired immune deficiency syndrome (CAIDS).

The controlled vocabulary referenced here, HIV/AIDS and HIV/AIDS-Related Terminology: A Means of Organizing the Body of Knowledge (Huber & Gillaspy, 1996b), was developed in direct response to a need voiced within AIDS service organizations (ASOs) across the United States—i.e., a need for a system of organization and access to the ever growing and evolving data, information, and knowledge spawned by the epidemic. Organization
and access were complicated by the variety of formats in which information was appearing and the diversity of individuals needing to use it. The system needed to be comprehensive and yet simple enough for nonlibrarians to use, as many ASOs do not employ information professionals (Huber & Machin, 1995, p. 242).

The vocabulary was tested and refined at a large community-based AIDS organization that provided direct information services to physicians, educational staff, caregivers, patients, students, and allied health personnel. Resources included the variety typical across the board: videos, monographs, vertical files, brochures, serials, and even realia. The vocabulary evolved from efforts to organize this collection.

Central to the entire development process was the conviction that the vocabulary must be descriptive yet flexible to suit the needs of users and to portray and preserve the rapidly expanding, multidisciplinary body of knowledge. Using the Dewey Decimal Classification as a model, ten "umbrella" concepts, termed domains, were identified:

1. Generalities;
2. Epidemiology and Transmission;
3. Education and Prevention;
4. Clinical Manifestations of HIV and Complications, Malignancies, and Infections Associated with AIDS;
5. Treatments, Therapies, and Medical Management of HIV Disease;
6. Psychosocial and Religious Issues, Case Management;
7. Legal, Ethical, Economic, and Political Aspects;
8. Organizations, Funding Opportunities, and Health Policy;
9. Fine Arts; and

Having identified ten mutually exclusive domains consistent with the reality of the epidemic, the content within each domain was developed. Core medical references, an AIDS dictionary, and a curriculum from an HIV/AIDS Education and Training Center were key to ensuring inclusion of all relevant concepts and terms identified at the time. Newsletters, especially GMHC Treatment Issues and BETA: The Bulletin of Experimental Treatments for AIDS, were invaluable for identifying possible future research directions and assuring space in the scheme for the addition of terms resulting from such work.

Upon completion of the first draft of the vocabulary, it was compared to several existing works, including an early nonhierarchical HIV-specific arrangement developed by librarians at Philadelphia’s AIDS Information Network. Development of biomedical content using core medical texts was supplemented with the National Library of Medicine’s Medical Subject Headings (MeSH). The Thesaurus of Educational Descriptors, compiled by information specialists at the National AIDS Clearinghouse, was used to
determine completeness and accuracy of the third domain, Education and Prevention. Subject specialists reviewed the sections on religious aspects, United States government components, and virology and clarified the terminology.

Since some concepts were applicable across domains, standard subdivisions, called “universal subdivisions” in this work, were developed and may be appended to main concepts as necessary to facilitate access to the information. The categories of subheadings finally included were Age Ranges, Sexual Orientation, Gender, Stages of Infection, Ethnic Groups, Geographic Names, At-Risk Populations, Religious Faiths, Signs and Symptoms, and Special Populations.

The final tasks were the generation of an alphabetic listing of terms and cross-references, the tagging of MeSH and near-MeSH terms, and the composition of instructions for use. The alphabetic listing and cross-references are intended to be particularly valuable for users unfamiliar with hierarchical arrangements; the cross-references guide users to the preferred terms within the vocabulary but reflect the full spectrum of the vernaculars used within various communities to describe concepts applicable to the pandemic.

Participation in the Large-Scale Vocabulary Test

In 1986, the National Library of Medicine embarked on a complex research and development project designed to link various biomedical vocabularies to a single system, today known as the Unified Medical Language System (UMLS). One of the four knowledge sources for the UMLS is the UMLS Metathesaurus, a “uniform, integrated distribution format for more than 30 biomedical vocabularies and classifications” (National Library of Medicine, 1997). Between July and December 1996, individuals were invited to participate in a test of the UMLS to determine to what extent existing biomedical language schemes fulfill the needs of health information systems.

HIV/AIDS and HIV/AIDS-Related Terminology: A Means of Organizing the Body of Knowledge (Huber & Gillaspy, 1996b) was one of the test vocabularies. The work as it stands contains 1,457 terms, excluding the Universal Subdivisions. Of these, 537 were identified at the time as MeSH headings while 98 were considered near-MeSH headings. (Since publication, some of the terms have been added to MeSH.) There were 822 terms that were considered unique from MeSH and approximately 336 of these (23 percent) have no equivalent in the UMLS. Such a significant number reflects not only the multidisciplinary nature of the disease but also the distinctive flavor of the lexicon itself, one that has added new terms (e.g., safe sex) to the language and increased both visibility and notoriety among the marginalized groups of people associated with the epidemic in the mind of the body politic.
Examples from the Controlled Vocabulary

The instructions for use state clearly that the evolution of the epidemic and progress on the scientific front guarantee that the lexicon will grow. Therefore, users are encouraged to add terms to the scheme as they are developed. For example, this work went to press shortly before HIV protease inhibitors, integrase inhibitors, and non-nucleoside reverse transcriptase inhibitors burst onto the scene; these are perhaps only the most obvious examples of terms that must be added for the vocabulary to remain current and useful.

Biomedical terms, however, are eventually accessible from other sources. The strength of a controlled vocabulary devoted solely to all aspects of the HIV/AIDS epidemic is that it includes terms commonly used among those affected by HIV but not reflected in other schemes. Some examples of vernacular terms that have no equivalent form or have a different meaning in other vocabularies but are, nonetheless, commonly employed within the discourse of the pandemic may serve to illustrate the usefulness of the arrangement. One series of terms, for example, describes unique individual responses to HIV infection: rapid progressors, nonprogressors, and long-term survivors.

Rapid progressors appear to be infected with a particularly virulent strain of HIV. Their blood counts tend to fall precipitously over a short span of time, perhaps just within two to three years, and their overall health fails rapidly (Khanlou, Salmon-Ceron, & Sicard, 1997, p. 163). Nonprogressors are patients who have tested positive for HIV antibodies but who retain normal blood counts and good health over a period of seven to twelve, or even more, years. Long-term survivors are patients at any stage of HIV disease whose conditions remain stable over several years. Even if their disease has progressed to AIDS, they tend not to be stricken with opportunistic infections but to remain relatively healthy (“AIDS Medical Glossary,” 1997, pp. 20-21). Of these three terms, only the concept of long-term survivors has roughly equivalent MeSH headings: Survivors, Survival Analysis, and Remission Induction/Methods, where methods is a subheading attached to the main heading. None of these terms, however, is common to the literature and discourse surrounding HIV/AIDS, though to experienced searchers, a relationship is evident.

With purely vernacular terms, however, no such relationship that would be meaningful in community settings exists. For example, a primary means of HIV transmission is shared needles among injection drug users. In international urban areas, a common venue for this activity is shooting galleries, often abandoned buildings or similar areas where users assemble for the specific purpose of injecting various substances into their bodies in a social setting. MeSH headings used to describe documents where shooting galleries is a text word include Substance Abuse, Intravenous, or Needles and Risk-Taking. While information professionals find such indexing logi-
cal, populations needing access at the community level to information about these behaviors are unlikely to use such terminology. Moreover, this controlled vocabulary is meant to be a record of the lexicon of the epidemic, mandating the inclusion of "street language," especially when it records a place with a social ecology that facilitates the transmission of the virus.

While the term "shooting galleries" or its equivalent does not exist in the biomedical literature, some other terms do but are defined differently in the community. An example of such a term is frottage. Dorland's Illustrated Medical Dictionary (1994) defines frottage as "paraphilia in which sexual arousal or orgasm is achieved by rubbing up against another person, who is unaware of the activity, as when pressed close to others in a crowd, usually without specific genital contact. Called also frotteurism" (p. 665). The Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV), used in constructing the UMLS Metathesaurus, employs the term frottage in defining Frotteurism. In the section on Sexual and Gender Identity Disorders, DSM-IV states that "the paraphiliac focus of Frotteurism involves touching and rubbing against a nonconsenting person" (Diagnostic and Statistical Manual of Mental Disorders, 1994, p. 527). Using computer systems, such as the National Library of Medicine's Internet Grateful Med (IGM) that incorporate the UMLS Metathesaurus in their search functions, entering the term frottage automatically maps to Frotteurism. However, among both heterosexual and same-sex partners, frottage has long been considered a consensual act, one that can lead to sexual gratification with little chance of either disease transmission or unplanned pregnancy. Educational units of some AIDS service organizations teach frottage as a safe-sex alternative to penetration. Researchers studying contraception effectiveness among married couples in Ireland noted that "about half reported using oral sex and/or frottage (body rubbing) [to achieve sexual release]" (Bonnar, Lamprecht, & O'Connor, 1997, p. 173). The equating of frottage in authoritative biomedical sources with a defined, deviant, nonconsensual act is a prime example of the body politic's continuing efforts to regulate sexual behavior by declaring deviance of both heterosexual and homosexual community norms. Such declarations further isolate, burden, and assign stigma to already marginalized populations.

CONCLUSION

As an evolving narrative, the discourse surrounding the HIV/AIDS pandemic represents the dynamic nosological record of the disease. This record, as captured in an HIV/AIDS controlled vocabulary, reflects the social construct within which the pathological condition and respective body of knowledge exists. For HIV/AIDS, examining this social construct is imperative because it sheds light on the direction in which disease and socio-scientific response have developed.
Although HIV is a complex chronic disease process, it has, to a large extent, been defined in the United States by the body politic. Discourse originating from the empowered elite fosters idealized conceptions, conjured through rhetoric, that are ultimately disseminated to the public. This discourse then becomes a tool central to constructing reality and building social constructs. However, social constructs, deriving from individual perceptions and societal perspectives, may be destructive in nature. Given the complexities associated with HIV/AIDS, the importance of this potentially negative effect cannot be overlooked. By representing HIV/AIDS as being reflective of particular socio-sexual categories and marginalized populations in public discourse, the body politic is provided the opportunity to promote the normalcy of "traditional" behavior and the abnormality of "deviant" conduct (Nzioka, 1996, p. 567). Discourse facilitates the shared construction of meaning, positive or negative, but only with socialization does the discourse yield consequences. In this way, HIV, through public discourse, becomes synonymous with promiscuity, permissiveness, and moral decadence, thus facilitating stigmatization of the disease and fostering prejudice, discrimination, and blame. By politicizing and stigmatizing the pathological, the biomedical complications of HIV/AIDS are further exacerbated.

The organizational schema of a controlled vocabulary intended to facilitate knowledge organization relative to HIV/AIDS must be reflective of the various biomedical and nonbiomedical complexities connected with the disease. Similarly, the structure needs to be flexible enough to accommodate evolution of the discourse, and the controlled vocabulary itself should be representative of the multifarious intricacies defining the body of knowledge associated with the pathological. Both disease and respective body of knowledge exist within the societal construct circumscribing HIV/AIDS, with social interactions and scientific advances delineating this construct being recorded in the controlled vocabulary. However, a lexical representation devoted to HIV/AIDS does not exist irrespective of the influence of the construct; rather, the vocabulary is affected by, just as it is reflective of, social ideologies and scientific realities framing the pandemic.

REFERENCES


Seeking the Subject*

JENNIFER TOBIAS

ABSTRACT

As an exercise in pondering cataloging in the networked environment ("networked environment" meaning electronic information sources interconnected through the Internet), this article compares traditional cataloging and Web-based description of two topics—the concept of "green cards" and a recent nonfiction work. This involves, first, outlining intellectual access issues as they apply to reference services today ("intellectual access" meaning the formal or informal description of a work for purposes of its discovery by others). Following this is an outline of key cataloging issues, per Sanford Berman, and corresponding issues in Web-based intellectual access. Ways that catalogers and public service librarians can address these issues conclude the article.

THE REFERENCE SCENE

Several key issues in intellectual access apply to reference services today. Perhaps the most crucial is an increasing demand for what this author calls "naïve" access—i.e., access to specialized subject knowledge by nonspecialists in that subject. Two major trends contribute to this demand. The first is the sheer volume of scholarly, professional, and popular publication. The second is a general intellectual trend toward interdisciplinarity (for one perspective on the implications of interdisciplinarity, see Messer-Davidow, Shumway, & Sylvan, 1993).

* Excerpted from "Issues of Intellectual Access in Our Electronic Age," with Elliott Shore (Director of Libraries, Bryn Mawr College) and Sanford Berman (Chief Cataloger, Hennepin County Library, Minnesota) a presentation of the Rutgers University SCILS Professional Development Program, April 1996.

Jennifer Tobias, Reference Department, Library, Museum of Modern Art, 11 West 53rd Street, New York, NY 10019
© 1998 The Board of Trustees, University of Illinois
As a result, librarians must be fluent in many subject vocabularies. Consider, for example, a historian seeking technical information about medical effects of lead use in ancient material culture. What vocabularies apply (historical, medical, chemical, sociological, or material)? The librarian must be able to communicate across these disciplinary vocabularies. This is a crucial skill for reference librarians today.

A third key issue is that increasing sophistication of hypermedia creates higher expectations by information seekers. In concrete terms, if home-Internet and twenty-four hour news channels (not to mention cars, phones, and the occasional coffee maker) appear to respond so readily to our everyday information needs, why is it so hard to pursue a question at the library?

This sophistication is also redefining what constitutes a scholarly work. Contemporary scholarship in the humanities and social sciences, for example, more often looks to the medium of messages—i.e., the way ideas are conveyed as much as the ideas themselves (see this line of thinking applied to hypertext navigation in Aarseth, 1997, chap. 8). The rise of media studies and the methodology of deconstruction are but two examples of this.

Further, the products of such scholarship are increasingly likely to be expressed in multiple media. In the humanities, a good example of this is the Perseus Project (http://www.perseus.tufts.edu), a hypermedia work thoughtfully integrating history, geography, literature, cultural studies, material culture, and mythology. In the “hard” sciences, as computing becomes increasingly integrated into methodology, the results of research increasingly integrate computing. Think of the Human Genome Project (http://www.nhgri.nih.gov/HGP), the international gene-mapping collaboration.

Finally, fast-paced changes in information technology are having obvious effects upon reference services. The task of integrating reference resources in diverse formats is one. The breakdown of distinctions between reference services and computing/information services is another. Ways to address these issues are discussed in the concluding section.

The Cataloging Scene

To outline some key issues in subject cataloging, this discussion will now turn to the indefatigable Sanford Berman (1993) and summarize his longstanding critique of traditional cataloging—i.e., deficiencies in traditional (AACR2) cataloging conceals works. These deficiencies are illustrated by a search in a traditional catalog (the Library of Congress catalog is used here) for information about “green cards” (representing “resident alien” immigration status in the United States). These deficiencies include:
• Anachronistic subject headings (A previous heading in LCSH was *Alien registration receipt cards (United States)*)

• Table-of-contents absent

• Lack of added titles

• Lack of notes

• Poor cross-referencing (for comparison, see the Hennepin County Library's treatment of the topic, in particular the scope notes visible in the catalog).

Compare this to indexing of this topic with the Yahoo! Internet search engine (http://www.yahoo.com) (for purposes of argument, the dominant advertisements for immigration lawyers have been ignored). The result includes the site illustrated in Figure 1, which leads to an authoritative site about U.S. immigration (http://travel.state.gov/visa_services.html) (see Figure 2).

![Figure 1. One Result from “Green Card” Search in Yahoo!](image)

Compared to traditional cataloging, what do we observe about the two pages in Figures 1 and 2?

• Traditional cataloging data are absent, nonstandardized, or nonapplicable. There is no reliable information about authorship, title, date, and place of publication.

• A fuzzier idea of discrete work—where are its boundaries, and are boundaries a useful way to think about the content?

• Many more contextual cues about the “spin” or point-of-view of the work conveyed through its organization and graphic packaging (for a thoughtful approach to “spin,” see Crowe, 1986).

• Heterogeneous, unstructured, popular subject vocabulary.

• Commodification of subject terms.

• Cross-referencing through hypertext.
SUBJECT ACCESS SHORTCOMINGS

The "green card" question exemplifies Berman's longstanding critique of much subject cataloging, starting with the problem of applying a nineteenth-century idea of indexing to the twentieth-century scene. The limits of the Dewey Decimal System and the plethora of specialized indexes indicate that the world cannot be organized into a single coherent vocabulary. And even if it could, it can be easily argued that data structure should be different for different disciplines, professions, and populations. Attempts to organize diverse information such as news broadcasts, chemistry literature, visual materials, and fiction in one truly useful vocabulary have been less than successful.

Berman also criticizes the lack of organic holistic cross-references in vocabularies such as LCSH. Extensive synonym relationships and inclusion of popular vocabulary, Berman argues, is a crucial element of any subject vocabulary. It is worth noting that his HCL system of subject headings is far more associative than hierarchical with many more "see also" references than broader-term and narrower-term relationships.

SUBJECT ACCESS ALTERNATIVES

Networked information such as that found on the Web offers some alternatives to these dilemmas. This is enabled, most obviously, by the interlinked nature of the Web and relational databases: hypertext is literally a cross-reference (of course, hypertext also enables not-so-useful linking of information, the most prevalent being the lists of links so prevalent on the Web). Richer search results are also enabled by the presence of more searchable content. Simply, there is more data in a given work to search—one is not searching simply a cataloging record (a description of a work) but more of the work itself. This can be used for indexing in new ways, as we see in
search engines based upon prevalence and proximity of terms (for an overview of mechanisms at work in search engines, see Steinberg, 1996).

However, full-text searching and its multimedia equivalents should not be the last word in indexing. Computing presents an opportunity to rigorously interlink diverse vocabularies—to create a thesaurus of thesauri. The Getty Institute's *a.k.a.* project (http://www.gii.getty.edu/vocabulary/aka.html) attempts to put this idea into action. The *a.k.a.* initiative attempts to cross-reference search terms across several subject authority files. The governing idea is this: it does not matter what you call something as long as it is linked.

Perhaps the most interesting alternative to traditional cataloging presented by networked information is the opportunity for self-determination in indexing and retrieval. The Web presents the clearest evidence of this; self-publishing is the norm, a byproduct of keyword indexing of these works' titles and major subdivisions.

Traditional cataloging rules such as LC's emphasize the assignment of subject terms based on title keywords and the “rule of specificity,” but this filtering process has a perverse tendency to remove the author's voice from searching for works by subject, sometimes so much as to render the work invisible. While the intent of subject cataloging is enhanced retrieval through normalization of terms, often this is not the effect (quantifying the effectiveness of subject cataloging is debated by Mann, 1997).

This is not to argue for abandoning subject access in favor of keyword searching, however. Even with sophisticated query languages, keyword searching of full-text databases, citation indexes, and Internet search engines reveal definite shortcomings.

To examine these strengths and weaknesses more closely, we will examine subject access to *Escape Velocity* (Dery, 1996), a nonfiction work, in two contexts.

**Seeking Cyberculture**

The first context is Library of Congress subject access. How might we get to *Escape Velocity*? Searching the Library of Congress catalog (http://lcweb.loc.gov/catalog) by title reveals the results illustrated in Figure 3.

In terms of subject access, we observe the familiar bibliographic standard with its reliable provision of title, author, publisher, date, and physical description. We also find subject access: “Computers and Civilization” and “Internet (computer network)—Social Aspects.” The subject headings describe the book's content to a degree—a limited degree.

One of the limits is the number of subject headings assigned, a longstanding criticism by Berman and one that LC has tried to address in recent years. Within the LCSH vocabulary (20th ed., 1997), subject access would be enhanced by the addition of: *Computer sex, Cyborgs, Fantasy games, Internet* (Computer network), and *Self-organizing systems.*
Another major limitation is the lack of access to subtopics in the book. In other words, what else is the work about? Additional headings could be assigned for body marking, cyberpunk fiction, and perhaps the social effects of the millennium.

A still better approach would be adding the table of contents to the record. A particular advantage to adding table-of-contents information returns us to the issue of self-determination in subject access, as will be seen later in this discussion.

In light of these limitations, we next observe how Escape Velocity identifies itself on the Web. Searching for some of the suggested headings above (again in the Yahoo! search engine) leads to the self-promotional site illustrated in Figure 4.
Exploring the site, we see these major supplements: author information, excerpts of (selected) reviews, and a table of contents (see Figure 5).

First, note the evocative words from the table of contents: “synth-rockers,” “cyberdelia,” and “mechanical spectacle.” The words are unsystematized, jargonistic, and perhaps ephemeral, but they express the work in a way that standardized vocabulary would obscure. (Interesting as well that these idiosyncratic terms might escape search-engine stop-word lists, unlike generic terms like “cyberspace,” “computer,” and “Internet”—the number and nature of stop words in Internet indexing is a rich topic in itself.)

For comparison, consider grassroots cataloging. These descriptors (with underlined terms hyperlinked to a definition) are assigned to Escape Velocity in a site about Cyberpunk authors (http://euro.net/mark-space/bkEscapeVelocity.html): nonfiction, cyberculture, cyberpunk, identity, culture, posthuman, future, Pat Cadigan, William Gibson, Mark Pauline, Stelarc, social history, edge, and 1990s. Besides explicit descriptors, we also observe indirect cues to content. These are conveyed through, first, the URL (the sub-subdirectory of a small commercial site conveys a different impression than a Federal agency, for example). Other cues include editorial style (Cyberpunk-speak), site organization (blending excerpts with press kit), graphic design (techno-chic, tending toward the ominous), and related links (to the presumed milieu of the book).

SEEKING PROGRESS

These comparisons are intended to show some cataloging alternatives presented by networked information, perhaps for adoption into
standard cataloging practice. To develop these alternatives further, I conclude with recommendations for developers, catalogers, library educators, and reference librarians.

For developers and catalogers, first, continue to develop subject and keyword indexing systems; both are useful. Second, enable interlinking of existing vocabularies through construction of thesauri. Getty's *a.k.a.* project, described earlier, is an important step in the right direction. Another experiment to follow is the graphically oriented "hypertextual searcher's thesaurus" of Johnson and Cochrane (1995). Third, adapt metadata standards to reap the indexing benefits of traditional and new media. Specifically, adapt MARC to accommodate new media. The use of the MARC 856 field for URLs and the use of Web-based catalogs using the Z39.50 standard are positive steps. Fourth, work toward standards for metadata (standardized descriptive information embedded into electronic works). In the short term, seek the integration of metadata into HTML (HyperText Markup Language) (the HTML "meta" tags move toward this goal. See http://www.w3.org/TR/WD-htm140-970708/struct/global.html#edef-META) and SGML (Standard Generalized Markup Language).

Metadata initiatives are moving along slowly; track the progress in an IETF (Internet Engineering Task Force) draft (http://www.ietf.org/ID.html) and in the report of RLG's January 1997 Metadata Summit (http://www.rlg.org/meta97.html).

Finally, to address the issues of transience and fluid boundaries in hypermedia, work toward development of persistent identifiers for networked information. Current initiatives include OCLC's Persistent Uniform Resource Locator (PURL) service (http://purl.oclc.org/OCLC/PURL/SUMMARY) and the proposal of a Digital Object Identifier (DOI) (http://www.doi.org) (for a useful evaluation of DOI by Lynch, see http://www.arl.org/newsltr/194/identifier.html).

For educators and reference librarians, be conversant in the languages of different disciplines. When teaching, encourage users of electronic resources to expect changes of interface, syntax, or dates of coverage. Concentrate instead on technology-independent methods for seeking and evaluating information. After all, media are transient; thinking critically about media content is not.

**NOTE**

Commodification through *sponsoring* search terms in popular search engines. When a searcher enters one of these terms, the sponsor's ad displays as a sidebar. For example, a search for *cars* or *Phillies* or *Planned Parenthood* could result in links to Honda or ESPN—or to Operation Rescue. A given query is thus linked, behind the scenes, to a particular kind of "related term."

**REFERENCES**


Grounded Classification: Grounded Theory and Faceted Classification

Susan Leigh Star

ABSTRACT
This article compares the qualitative method of grounded theory (GT) with Ranganathan's construction of faceted classifications (FC) in library and information science. Both struggle with a core problem—i.e., the representation of vernacular words and processes, empirically discovered, which will, although ethnographically faithful, be powerful beyond the single instance or case study. The article compares Glaser and Strauss's (1967) work with that of Ranganathan (1950).

INTRODUCTION

There are some striking similarities...between field work and library research. When someone stands in the library stacks, he is, metaphorically, surrounded by voices begging to be heard. (Glaser & Strauss, 1967, p. 163)

Classification is an uncovering of the thought-content of a written or expressed unit of thought. . . . The reference librarian . . . applies the classification scheme in the ultimate stage of library service which is effecting contact between the right reader and the right unit of thought in a personal way. (Ranganathan, 1951, p. 116)

The landscape of information retrieval is shifting rapidly (with networked distributed computing, large-scale digital libraries, and enormously powerful search engines). As the introduction to this issue notes, formerly firm boundaries between library and office, catalog and desktop are transmogrifying. The change means that a wider range of human
activities come under the purview of library and information science. When the library and the desktop become seamless, then practices of work organization become part of the cataloging and indexing process. This merger calls for methodological creativity and cross fertilization between previously disparate methodological domains.

One fruitful direction for this creativity is in blending the methods of library and information science (LIS) with those of sociology and anthropology. LIS brings the strengths of order and sensitivity to domains and documents and a long tradition of struggling to find representations that are both useful and elegant. Sociology and anthropology bring strengths based in the empirical chaotic process of analyzing work, perspectives, conflict, and representations that are themselves the site of struggles.

Some of the tough challenges faced by classification in environments such as the World Wide Web or large digital libraries include: how work settings and the flow of real-life tasks give rise to information needs and strategies; how different vernaculars and representational schemes may work together heterogeneously; and how informal and formal classifications interact in information retrieval and use (Cochrane, 1993; Svenonius, 1986). In parallel fashion, some of the cutting edge challenges faced by grounded theorists include: assessing the quality and completeness of analysis; managing large amounts of unstructured textual data; and accounting for a basis for theoretical sampling. The two endeavors offer each other some aid in meeting their respective challenges.

Both faceted classification (FC) and grounded theory (GT) began as reform movements against powerfully entrenched a priori schemes with claims on universality. Grounded theory offers a way to include processes and actions in the analysis of vernacular representations (a question introduced as a core theoretical problem by Ranganathan). It is at the same time a source of theoretical richness for the understanding of intermingled types of work. Faceted classification offers a way to assess the structural integrity and architecture of a particular theory via facet analysis and other analytical tools used in thesaurus construction and assessment; with automated thesauri tools, FC is an aid for managing large bodies of text that will augment current qualitative methods software.

After writing the first draft of this article, a colleague brought Clare Beghtol’s (1995) superb paper, “‘Facets’ as Interdisciplinary Undiscovered Public Knowledge: S.R. Ranganathan in India and L. Guttman in Israel,” to my attention. Recursively, of course, our papers are an example of undiscovered public knowledge converging. Beghtol draws parallels between the work of Ranganathan and that of Louis Guttman, a sociologist who developed a faceted theory for the analysis of qualitative data, principally as an aide to the analysis of survey research data.

Though Beghtol (1995) notes that we will never know if proximate or remote contact transpired between Guttman and Ranganathan, she maps
out ways in which the two systems might profitably cooperate. They are, she notes, solving analogous problems of data analysis and management (p. 237). For all the structural reasons noted in the introduction above, there are now unique opportunities to exploit these previously unlinked bodies of research.

**Classification au Naturel**

The notion that classification schemes are neither innocent nor arbitrary is core to several disciplines. Anthropologists map the complex taxonomic schemes of a culture as a way of understanding worldview and norms. Library researchers, going back to Ranganathan's original foundational work, see classification as core to mapping, in Ranganathan's words, "the universe of knowledge." Social critics of classification systems argue that the choice of categories reflects political choice and (the often silent) wielding of bureaucratic exercises of power (Berman, 1984; Kirk & Kutchins, 1992; Bowker & Star, 1994, In press; Bowker, Timmermans, & Star, 1995). Others have argued for the historical specificity of schemes of classification (Hacking, 1995; Young, 1995).

It has, however, been uncommon for two things to converge: (1) the idea that a qualitative social scientist might use the structures of formal classification systems as a proactive tool for generating and assessing theory; or (2) the idea that the theories and tools of qualitative social science might actively guide classification and indexing activities in library and information science. In other words, it is uncommon to see information systems classification as an ethnographic or theoretical enterprise, even where it has sometimes been seen as political. However, there are potential benefits to seeking this convergence. This article proposes a comparison of one of the more common qualitative methods—i.e., grounded theory (Glaser & Strauss, 1967; Strauss, 1987; Glaser, 1978) with the construction of faceted classifications in library and information science (compare Aitchison, Gilchrist, & Bowden, 1997; Vickery, 1960, 1966).

**Basic Definitions**

*Grounded Theory*

Grounded theory is a method for analyzing data; it is most commonly used on naturalistic field data but has also been used to analyze historical and documentary data (compare Clarke, 1990; Star, 1989). Barney Glaser and Anselm Strauss, who trained several generations of graduate students in sociology and nursing, developed grounded theory in the 1960s. The method has its roots in symbolic interactionist sociology and American Pragmatism, as well as, to some extent, Lazarsfeld's analysis of variables and their valences. (GT is enormously popular as a method in social science analysis. Its use—and some might argue misuse—extends from simply "empirical and inductive" to much more formal and thorough appli-
cations of the method. A recent volume by students and colleagues of Strauss provides a good overview of more thoroughgoing developments [Strauss & Corbin, 1997].

Grounded theory relies on several components:

1. An empirical iterative approach to the collection and analysis of data—i.e., data are collected, analyzed, and revised cyclically as checked against empirical findings.

2. A constant comparative approach to the development of theory. Similarities across disparate domains are sought in order to reveal the dimensions present in a situation. Their discovery lends a kind of anthropological strangeness to the analysis of situations otherwise taken for granted. In this, less emphasis is placed on the degree to which a given variable presents itself in a situation. The best example of constant comparison comes from Everett Hughes (1970), a long-time colleague of Strauss's, who asked: "why is a priest like a prostitute?" (Answer: They both hear confessions in private, outsiders find their work somewhat mysterious, etc.) (p. 316). The point was to find the common dimensions, thus illuminating something about their work conditions—not to level the obvious disparities between the cases.

3. An approach to sampling which is theoretical rather than site or population driven—i.e., emphasis is put on making theories as richly complex as possible rather than on proving instantiations of hypotheses or applications of previous theories.

4. Theory development that works from substantive (close to descriptive) through to formal (abstract) levels as constant comparison proceeds over time. For example, early grounded theory studies looked substantively at dying patients in hospitals (Glaser & Strauss, 1965), detailing the many dimensions of the nursing, medical, and family situations. One of the important substantive focuses was: who was aware of the status of the dying patient as terminal, and what conditions gave rise to these differences in "awareness context?" Years later, Strauss (1978) took the awareness context concept and applied it to a variety of other circumstances in which awareness of conditions might be important, viz., being a spy, coming out as gay, being on either side of a bargaining table. The formal theory was developed as the comparisons ranged across substantive cases.

Faceted Classification

Faceted classification 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). Suggested in the 1930s by Ranganathan and codified in his system of classification, it has become an important tool in library and information science for constructing thesauri, building retrieval schemes for particular
groups of users, and in many circumstances for cataloging information. Important points are:

1. the division of fields of knowledge into categories that may express different aspects (facets) of the knowledge (especially from the point of view of information retrieval). This stands in contrast to schemes that would assign each document (book, article, and so on) to a single rigid value in a universal hierarchical classification scheme;

2. the combination of a system of notation, of analysis of knowledge classes, and the physical storage and retrieval of documents and parts of documents into an integrated system;

3. an iterative and evolving set of classifications which may flexibly serve the needs of particular groups of users;

4. the importance of comparing and synthesizing analytic facets in order to reflect changing knowledge and changing user needs;

5. a movement away from a flat proliferation of particular (phenomenological) aspects of a field of knowledge, toward a synthetic representation that includes basic (both abstract and concrete) categories. These latter, crucially, remain open to revision. (This article focuses on similarities between GT and FC. It does not do justice to all the important developments in FC, such as those proposed by the Classification Research Group in the 1960s on integrative levels, or the work in medical classification. A fuller history of classification research would examine these contributions.)

COMMON GROUND

Both grounded theorists and designers of faceted classifications struggle with a common core problem. This is the question of how to represent vernacular words and processes. In both cases, the categories are empirically discovered in an almost self-contradictory fashion. The contradiction comes with the attempt simultaneously to represent, on the one hand, the local, specific, and empirical and on the other, abstractions and generalizations. The difficulty lies in making this representation both ethnographically faithful (faithful to the needs of users and particular populations), yet simultaneously powerful beyond the single instance or case study. Both grounded theory and faceted classification began as reform movements against powerfully entrenched a priori schemes with claims on universality (compare Vickery, 1960). These are unusual in that this reform did not consist of abandoning the attempt to formalize and systematize.

This set of common core methodological problems has been present since well before the advent of the Internet and the World Wide Web. Ranganathan struggled against the rigidities of the dominant universalistic library classification schemes, many of which originated in the nine-

However, the landscape of information retrieval is shifting rapidly (with networked distributed computing, digital libraries, and large-scale and enormously powerful search engines). Although from the beginning Ranganathan argued for classification of documents by both physical existence and the ideas they contain (down to a very fine degree of analysis), today the nature of documents is in extreme flux and more than ever demands such analysis. The boundaries of documents are unclear as people modify and distribute them electronically; authorship is changing as multiple versions and annotations proliferate (Brown & Duguid, 1996). The ability to fracture and use pieces of documents as well means that library classification is now linked not only with traditional genres but also with work processes, communication, and writing (Levy & Marshall, 1994; Levy, 1994; Bishop & Star, 1996).

The landscape of qualitative research is similarly in flux due to the challenges posed by networked information technology. What does it mean to “observe” someone’s writing on the Internet or World Wide Web? How do we “do fieldwork” when actions are taking place in such a geographically distributed fashion? How do we understand the links between local mixes of online/offline activities and those that appear on the Web? It is a two-edged sword—on the one hand, it seems that infinite ready-typed field notes lurk out there waiting for the analyst; on the other, little in traditional qualitative social science methodology can manage this volume of data and geographic dispersion.

As noted in the introduction to this volume, there has developed over the past several years a lively strand of qualitative inquiry in library and information science as well as in management information science. Dervin’s (1992) sense-making methodology, for example, has been adapted in a number of empirical investigations of information use. A Web page for qualitative research in information systems is maintained at http://www.auckland.ac.nz/MSIS/isWorld/index.html.

A comparison of grounded theory and faceted classification offers some important cross-fertilization in addressing these situations and lines of research. Some of the challenges faced by classification in environments such as the World Wide Web or large digital libraries include: how work settings and the flow of real-life tasks give rise to information needs and strategies; how different vernaculars and representational schemes may work together heterogeneously; and how informal and formal classifications interact in information retrieval and use.

At the same time, some of the challenges faced by grounded theorists include assessing the quality and completeness of analysis, managing large amounts of unstructured textual data, and accounting for a basis for
theoretical sampling. The two endeavors offer each other some aid in meeting this challenge. Grounded theory offers a way to include processes and actions in the analysis of vernacular representations (a question introduced as a core theoretical problem by Ranganathan) and a source of theoretical richness for the understanding of intermingled types of work (Strauss, 1994). Faceted classification offers a way to assess the structural integrity and architecture of a particular theory, via facet analysis and other analytical tools used in thesaurus construction and assessment with automated thesauri tools, a means for managing large bodies of text that will augment current qualitative methods software (Schatz, Johnson, Cochrane, & Chen, 1996).

**Classifications as Theory Developing Tools**

In an important article, Kwasnik (1992) places the theoretical aspect of classification schemes center stage. She states that:

> Classifications are really very much like theories. Like theories, classification schemes can provide an explanatory shell for looking at the world from a contextually determined perspective. Classification schemes not only reflect knowledge by being based on theory and displaying it in a useful way...but also classifications in themselves function as theories do and serve a similar role in inquiry. (p. 63)

She notes that, in the attempt to impose order and specify relations, classification schemes are inherently theoretical, just in the way that scientific theories are. Kwasnik goes on to use Ranganathan's faceted classification scheme to assess the structure of three scientific classificatory enterprises: the periodic table, psychiatric classification as it appears in the DSM, and classification in software re-use. This novel evaluative use of facet analysis hints at a valuable tool for assessing theory construction. In Kwasnik's words: "Classifications have structural properties that lend themselves to representing knowledge in a given situation" (p. 80).

It follows from this that the construction of classification schemes is also an inherently methodological enterprise—i.e., one must make choices about analytic tools guided both by theoretical concerns, as Kwasnik suggests, and by questions of reliability, validity, doability, audience, and even the ripeness of particular scientific questions.

An article by Solomon (1991) also indicates the possibility of using classification schemes in research, this time specifically from a qualitative perspective. He argues that the construction of classification schemes is a form of technology development and one that must be closely linked to user semantics. Naturalistic methods of inquiry meet the requirements of handling what is often found in the field—i.e., ambiguity, multiple meanings, context dependence, and a gap between what users say and what they do (p. 164). In discussing his case study material, Solomon also notes that:
The experience of the case study suggests that a multiple dimensional classification is needed to satisfy the diverse interests and information needs of the users involved. By factoring the interests and concerns of the managers involved, the unidimensional classification becomes less fuzzy and highlights key concerns in the resource allocations process: management requirements, scope of effort, resource requirements, and resource characteristics. (p. 169)

It is important to note that the evaluative component can be both used in theory construction (e.g., evaluating the usefulness of the classification scheme in process); in theory deconstruction (showing the theory-ladenness of all classification schemes); and in post-hoc and participatory user studies of extant classification schemes such as LCSH (Library of Congress Subject Headings) (Rosenberg & Borgman, 1992).

Some parallels between the early mandates of grounded theory and of Ranganathan's vision will now be discussed.

THE BIG PICTURES

Critiques

Both Glaser and Strauss's (1967) work, *The Discovery of Grounded Theory*, and Ranganathan's (1950) foundational classification work (especially on the Colon Classification) read like manifestos. The enemy in both cases is reified rigid attempts at universal descriptions of knowledge that are not grounded in people's needs or experiences. From the grounded theory perspective, this meant taking on much of institutionalized American sociology, at that time (as now) largely quantitative, survey-oriented, and (then) functionalist:

The qualitative research is generally labeled "unsystematic," "impressionistic," or "exploratory". . . . These critics, in their zeal for careful verification and for a degree of accuracy they never achieve, have forgotten both the generation of theory and the need for carefully appraising the different degrees of plausibility necessary for sociology's diverse tasks. (Glaser & Strauss, 1967, p. 223)

Glaser and Strauss go on vehemently to denounce Robert K. Merton for his attacks on qualitative methods, basically calling him an "armchair theorist." "His reasoning necessarily leads to the position that data should fit the theory, in contrast to our position that the theory should fit the data" (p. 261). "Verification" in the grounded theory vocabulary becomes a dirty word (later Glaser will extend this even to the word "scholarship," which he says is no substitute for getting out there and seeing for yourself).

Similarly, the impetus for Ranganathan's reform movement within library classification was first given as the explosion of knowledge following World War I and the attendant inability of older rigid classification systems to adapt and accommodate new and divergent viewpoints. Parallel with the grounded theory denunciation of verification above,
Ranganathan (1950) states that: "Hundreds have seen the attempts to represent specific subjects by arbitrary symbols without any organic relation to the ideas represented. Practically in all such cases a breakdown has come sooner or later" (p. 47).

Later, Reese (in Vickery, 1966) notes that faceted classification schemes are "mission-oriented rather than discipline oriented... designed for user groups whose interests cut across the traditional fields" (p. 14). Faceted classifications do not follow pre-set categorization schemes deriving from disciplinary status-quo; rather, they demand semantic sensitivity and are designed to incorporate novel—that is to say, grounded—user needs. Vickery (1966) notes: "A faceted classification differs from the traditional in that the facets so distinguished are not locked into rigid, enumerative schedules, but are left to combine with each other in the fullest freedom, so that every type of relation between terms and between subjects may be expressed" (p. 13).

**AN OPEN UNIVERSE OF KNOWLEDGE**

Both grounded theory and faceted classification see the universe(s) of knowledge as potentially infinite, open, and evolving. Ranganathan (1965) says:

> For in the true Tree of Knowledge, one branch is grafted to another at many points. Twigs too get grafted in a similar way among themselves. Any branch and any twig are grafted similarly with one another. The trunks too become grafted among themselves. Even then the picture of the Tree of Knowledge is not complete. For the Tree of Knowledge grows into more than three dimensions. A two dimensional picture of it is not easily produced. There are classes studded all along all the twigs, all the branches, and all the trunks. (pp. 32-33)

A similar complexity is clear in all the grounded theory work, both in terms of interconnectedness and openness: "One of our deepest convictions is that social phenomena are complex phenomena... this is why grounded theory methodology emphasizes the need for developing many concepts and their linkages in order to capture a great deal of the variation that characterizes the central phenomena studied during any particular research project" (Strauss, 1987, p. 6).

Glaser and Strauss (1967) note that:

> The theorist's task is to make the most of his insights by developing them into systematic theory. His sociologist's perspective is never finished, not even when he writes the last line of his monograph—not even after he publishes it, since thereafter he often finds himself elaborating and amending his theory, knowing more now than when the research was formally concluded. (p. 256)

Some of the practical problems posed by both these approaches include developing schema for management of notation, managing the proliferation of codes (classes), responsible abstraction, and ongoing revisions.
PARALLELS IN APPROACHES: SOME KEY TECHNICAL DETAILS

The openness and centrality of complexity to both grounded theory and Ranganathan's faceted classifications approaches have made these both attractive and often difficult to learn. There is a constant tension between faithfulness to empirical detail and a desire to make the complexity usable via abstraction. Both FC and GT are techniques with long histories, schools of practice, and subtleties of interpretation far beyond the expository capabilities of this discussion. The following details are not exhaustive but are suggestive of key parallels in technical approach between the two systems. The GT examples rely heavily on Glaser (1978), perhaps providing the most formal statement of GT problems of coding and classification.

Constant Comparison and Analytic Synthesis

Both grounded theory and faceted classification have strong components of comparison and synthesis. From the GT point of view, as with the Hughes example of the priest and the prostitute, the comparison of even seemingly discrepant phenomena may illuminate valuable dimensions. Glaser (1978) notes: “Actually apparent non-comparability is irrelevant, if the variable to be compared has a value in each group. Comparing on the basis of properties of groups has the purpose of generating theory. . . Comparing the apparently non-comparable increases the broad range of groups and ideas available” (p. 42) (emphasis in original).

Compare this with Vickery’s (1960) discussion of FC: “[F]rom the theoretical point of view, faceted classification breaks free from the restriction of traditional classification to the hierarchical, genus-species relation: by combining terms in compound subjects it introduces new logical relations between them, thus better reflecting the complexity of knowledge” (p. 13). As Aitchison, Gilchrist, and Bowden (1997) note, FCs are designed so that new concepts may be built by combining existing class marks rather than by exhaustive enumeration (p. 55).

Levels of Formality

Both grounded theory and faceted classification (especially in Ranganathan’s original formulation) emphasize orthogonal, but simultaneous, operations of coding categories. In GT, “[s]ubstantive codes conceptualize the empirical substance of the area of research. Theoretical codes conceptualize how the substantive codes may relate to each other as hypotheses to be integrated into the theory” (Glaser, 1978, p. 55). Substantive codes are arrived at by asking questions of the data that will result in classes, such as “of what is this an example?” Often constant comparison (or simply lateral thinking) will act to generate a class in this fashion (Strauss, 1987, p. 272).

In grounded theory, the substantive gives rise to the theoretical by asking questions of relationships between substantive categories. This is
exactly the interrogation made in analyzing a faceted classification scheme: Is this category broader or narrower than that? Which is the more basic? (Aitchison, Gilchrist, & Bowden, 1997). These questions are necessary to support the analytico-synthetic approach of Ranganathan.

"Coring Out" and Basic Classifications

As humans demand both simplicity of representation and the ability to combine and recombine, the problem noted above of simultaneous specificity and abstraction appears for both grounded theory and faceted classification as a very tricky mapping problem. Ranganathan (1965) says, on mapping these relations:

The multi-dimensional universe of knowledge has to be transformed into a one-dimensional universe. Here arises an insoluble problem. It is well known that in the transformation of an n-dimensional space into a space of smaller number of dimensions and into a one-dimensional space or line in particular—or its equivalence, in the mapping of an n-dimensional space on a space of small number of dimensions and on a line in particular—many of the Immediate-Neighborhood-Relations among the classes are necessarily lost. (p. 33)

A similar mapping problem in grounded theory is called the core category problem, arrived at through open coding of field data. In this process, rapidly generated classes are related to each other, then recurring instances become core categories. As the data are coded and recoded and relationships specified, they are said to become saturated. This means that the mapping problem is solved through specifying a series of relationships, with the result of eventual convergence. In faceted classification, this is phrased as moving from the phenomenal to the seminal level. In grounded theory, more than one core category can originate from the same data over time, resulting in different focuses or emphases; it is not a matter of one underlying truth or form but rather the fashion in which relationships are specified. This is also true in the construction of FCs in the sense that multiple special thesauri may rely on the same data sources.

Both Ranganathan (1965) and Glaser (1978) argue that moving down to very fine points in the data helps the discovery of these classes. "In the view of the Postulate of Fundamental Categories, we should descend down and down, and allow the various subjects and ideas to become absorbed and reassembled, reabsorbed and again reassembled, and so on; until we find only five ultimate generic ideas—standing out" (Ranganathan, 1965, p. 198). These categories, often cited in library science, are personality, matter, energy, space, and time (PMEST)—basic attributes of all knowledge.

Similarly, Glaser (1978) recommends "fracturing the data"—i.e., looking at data line by line (pp. 57-58). Simultaneously, he notes that there are also several core (he lists eighteen) theoretical codes which can be
used, similar to the PMEST categories, to "maintain [a] conceptual level in writing about concepts and their interrelationships" (p. 73). Although more fine-grained than the PMEST category system, there are again interesting resonances as these families also cover space, time, and character. Glaser (1978) lists eighteen families of theoretical codes including: Process—stages, phases, transitions, ranks, etc.; Degree—limits, ranges, amounts, etc.; Dimensions—elements, pieces of, properties of, slices, segments, etc. (part-whole relations); and Ordering (including temporal ordering) (for a complete discussion of the families, see pp. 74-82).

CROSS-FERTILIZATION

Why are the parallels between grounded theory and faceted classification of interest? Earlier in this discussion it was mentioned that changes in the nature of information retrieval, networked computing, and thus of qualitative research all make the search for ways of ordering classes and categories more urgent. Automated thesauri and retrieval systems have made important advances in the direction of recognizing deep semantic similarities, often explicitly addressing problems in faceted classification (e.g., Schatz et al., 1996; Pollitt, 1997; Pollitt, Smith, & Brackevelt, 1996). At the same time, badly needed are the theoretical developments that will both help model complex data and be useful in naturalistic settings. This author suggests that FC may provide a helpful tool to analyze and construct grounded theories.

There are several software packages that support the analysis of qualitative data. Two were specifically targeted at grounded theory analysis—i.e., NUDIST and Atlas/ti. Without going into extensive comparison, both support flexible document coding. Atlas/ti captures many features of the discussion above in supporting flexible coding structures and bundles of codes for data collected using grounded theory (for a demo see http://www.cs.tu-berlin.de/~muhr/atlasti.html). It is thus possible to use Atlas/ti to build a thesaurus from one's own field notes and interrogate its structure as one would a faceted classification. As we understand the theory-ladenness of classification schemes, we may also come to understand more about the classification schemes embedded in our qualitative theories and methods.

ACKNOWLEDGMENTS

The author would like to thank the following people for rich discussions and very helpful comments: Marcia Bates, Marc Berg, Ann P. Bishop, Geoffrey C. Bowker, Pauline Cochrane, Mike Hales, the late Anselm Strauss, and Jürg Strübing. All errors are my own—this exploration is at an early stage of bridge building between library and information science and qualitative social science research methodology. The writing of this article was
supported in part by grants from the National Science Foundation No. SBR 9514744 (Ethics and Values Studies) and by the NSF/ARPA/NASA Digital Library Initiative NSF93-141. Their support is gratefully acknowledged.

NOTES
1 Beghtol notes a possible indirect or remote connection between the work of George Kelly, a psychological methodologist, and both Guttman and the classification community (p. 214). Another parallel exists here in my own work. As an undergraduate psychology major, I wrote my honors thesis using a combination of Kelly's Repertory Grid Method and Kuhn's notion of paradigm. This work reminded one of my advisors of the work of Glaser and Strauss, and they directed me to The Discovery of Grounded Theory. Some years later I became Strauss' student. On reading a draft of this paper, a colleague in Britain sent me a helpful message pointing out the similarities between grounded theory and Repertory Grids (personal construct theory, and directing me to a Web site at the University of Calgary dedicated to the latter [http://ksi.cpsc.ucalgary.ca:80/PCP/] (Mike Hales, e-mail communication to the author, 29 October 1996).

2 An important exception, of course, is the work of anthropological linguists and taxonomists which falls into this category, and I do not mean to exclude their important contributions. However, much of their work does not develop theories in the sense that I am using the term in this paper, which should be clear from the context below. Rather, the findings are used as primary data from which theories are developed, as in structuralist anthropology. In any event, it would also be important for future work to compare the process of anthropological taxonomy with some of the LIS tools discussed.

3 The Diagnostic and Statistical Manual, the major classification for psychiatry.

REFERENCES


Mapping Beyond Dewey's Boundaries: Constructing Classificatory Space for Marginalized Knowledge Domains

Hope A. Olson

ABSTRACT

Classifications are bounded systems that marginalize some groups and topics by locating them in ghettos, diasporized across the system. Other marginalized groups and topics are totally excluded from these systems, being outside of their territorial limits. Because classifications are locational systems, spatial analyses borrowed from various disciplines have potential to identify and address their problems. The philosophical basis for the analysis in this article is Lorraine Code's (1995) conception of "rhetorical spaces" as sites where topics can be taken seriously as legitimate subjects for open discussion. In existing classifications, there is rhetorical space for most mainstream social and scholarly knowledge domains but not for marginalized knowledge domains. Geography offers concepts for building a theoretical framework to ameliorate the biases of classification. This article describes such a framework and how it is applied using techniques such as Gillian Rose's (1993) "paradoxical spaces," which are simultaneously or alternately in the center and at the margin, same and other, inside and outside to develop a more complex and meaningful classification for women and other marginalized groups. The project described here operationalizes these theoretical openings by applying them to the Dewey Decimal Classification as both critique and as techniques for change.

INTRODUCTION

The problem of bias in classification can be linked to the nature of classification as a social construct. It reflects the same biases as the culture.
that creates it. Existing literature has critiqued the most widely used classification in the world, the *Dewey Decimal Classification (DDC)*, for its treatment of women, Puerto Ricans, Chinese and Japanese Americans, Mexican Americans, Jews, Native Americans, the developing world (including Africa, the Middle East, and Melanesia), gays, teenagers, senior citizens, people with disabilities, and alternative lifestyles.¹ To look at these biases with a fresh eye, a theoretical construct capable of revealing the complexities of classification and its social construction was sought. The theoretical framework that subsequently evolved draws on the spatial metaphors that have become so prevalent in cultural criticism in recent years.

As Lorraine Code (1995) points out:

> [use of] spatial metaphors picks up a late-twentieth-century concern with location: with territories, mappings, positionings where resources are variously available, subjectivities are variously enacted, and identities are constructed and continually reconstructed in the enactings; and where hierarchies of power and privilege always contribute to shaping these processes . . . . (p. ix)

In this spirit, this discussion will move from a description of the construction of classification to the development of spatial imagery as a metaphorical mechanism with the ability to discover the processes by which powerful and privileged discourses shape information and with the potential to inform change. What will then evolve will be a multidisciplinary theoretical framework based on spatial conceptions in the context of a specific project, concluding with suggestions for further research.

**THEORETICAL MODELS FOR THE
SOCIAL CONSTRUCTION OF CLASSIFICATION**

The idea of classification as a social construct is not new. A. C. Foskett (1971) suggests that classificationists are the products of their times. Therefore, since classifications are the products of classificationists, classifications also reflect the biases of their times. Examining the ideological construction and present needs for reconstruction of the former Soviet classification (Sukiasian, 1993) or the Confucian, and later Maoist, classification in China (Studwell, Wu, & Wang, 1994) makes it easy to see that classifications reflect philosophical and ideological presumptions of their cultures and not only the times but also the places. Classifications arrange concepts according to accepted cultural discourses whether those discourses are Leninist or Maoist communisms, the Seven Epitomes of Confucian doctrine, or Dewey’s apparent reversal of Francis Bacon’s classification scheme.

Allocation of 80 percent of *DDC*’s religion section (the 200s) exclusively to Christianity and the existence of a separate section for American literature (the 810s) when all other literatures are arranged by language
is not surprising given the origins of this classification. Finding the topic “concubinage” under customs in 392.6 where it is gathered with topics such as chaperonage and dating or “suttee” in 393.9 and all combined with funerals and wakes has a certain ethnocentric logic. The other major North American classification, the *Library of Congress Classification (LCC)*, exhibits similar biases. For example, the allocation of space and the sequence of development of Class K for law, with separate volumes for individual North American and European countries, was published in the 1960s and 1970s with only one volume appearing in 1993 covering Asia, Eurasia, Africa, Pacific Area, and Antarctica. In each of these cases, there is a tendency to simply accept that these powerful discourses operate, and that change is too expensive and impractical.

**IN SEARCH OF A THEORETICAL MODEL**

Building a theoretical framework or model to analyze and address the biases of classification in a practical way requires a model capable of revealing the complexities of classification and the discourses that construct it. Developing this framework or model requires examination of the characteristics of classification, testing various conceptions against those characteristics, and reflexively reworking the model.

Two major characteristics of classification are that it gathers similar information together and places it in proximity to related information. If there is to be only one ordering of information, then it is useful for classification to reflect the relationships perceived in the wider society. Because the relationships between concepts can be drawn in a variety of ways, classifications will give more advantageous space in the overall structure to some concepts than to others. As the literature cited earlier suggests, classification tends to reflect the most mainstream version of these relationships. Classificatory structures are developed by the most powerful discourses in a society. The result is the marginalization of concepts outside the mainstream.

Classifications are also closed systems in that they represent some concepts and not others. No classification will ever be all inclusive. Since classifications are notationally controlled vocabularies, these inevitably have limits. Legal scholar Drucilla Cornell (1992) has suggested that any system or structure has limits, and that replacing one system with another will simply define different limits rather than being all inclusive. A system of any kind is defined by what it is not and, because systems tend to be dynamic, like classifications, the definition of what the system's limits are is always deferred (p. 2). It is an instance of Jacques Derrida's deconstructive concept of *differance* that limits are constructed by their exclusions and are in a state of constant flux because they are socially constructed. The question for classification then becomes, What is left
beyond the limit? What is excluded? Given Foskett's observations on classificationists, it comes as no surprise that the limits of classifications are also constructed by the powerful discourses within a society, and that what is excluded is what is further from the mainstream.

Further, library classifications have responded to the needs of libraries to classify published works into a browsable collection. Therefore, what exists in published form will dictate, to a greater or lesser degree, what is included in a classification. Even a classification that does not limit itself to literary warrant will be irresponsible if it ignores the published record. Since what gets published is also limited by powerful social discourses, it too tends to produce a corpus largely representing mainstream thought.

The result of these factors is classification, which might be seen as a dense mainstream core of aptly juxtaposed concepts with marginal concepts scattered around the edges or not represented at all. This conjures up concentric circles of degrees of representation quality forming a distribution—similar to Zipf's, Lotka's, or Bradford's—of: a few core concepts best represented, a middle ground adequately represented, and a large periphery of poorly represented marginal concepts with some concepts outside of the limits not represented at all (see Figure 1). This image of Zipf's core of word occurrences or Lotka's core of published authors or Bradford's core of journals effectively documents the distribution of what currently exists. It also suggests the effectiveness of a spatial conception of classification. However, it does not provide a metaphor for analyzing the historical discourses that have shaped the present nor does it offer suggestions for changing the status quo.

---

Figure 1. Degrees of Representation Quality.
This distribution does, however, probably correlate with how representative mainstream views actually are—i.e., they show only what is concentrated at the perceived center and how that puts others in the margins. The core does not consist of many different entities. Putting this distribution into social terms creates a small core surrounded by margins. If one takes all so-called “special interest groups” out of the social equation, there is little left in the mainstream. In North American society, taking away women, African Americans, Hispanic Americans, French Canadians, Native peoples, Asian Americans, lesbians and gay men, people with disabilities, anyone who is not Christian, working class and poor people, and so forth, one is left with a very small “core.” An image that shows the complexity of these overlapping categories is that of a huge Venn diagram with many sets limited by Boolean ANDs. The white AND male AND straight AND European AND Christian AND middle-class AND able-bodied AND Anglo mainstream becomes a very small minority (Figure 2), and each set implies what it is not. The implication of this image is that not every person, not every discourse, not every concept, has equal weight. Some discourses simply wield more power than others. Different discourses have different levels of power to construct our realities. The Venn diagram helps to represent the discourses constructing classification. However, Venn diagrams operate on the basis of dualities. Something either is or is not in the circle (fuzzy sets could help somewhat in this respect). Further, as the sets overlap, these estimate only quantity and not the shape or relative location of the dominant.

Mapping as Metaphor

Zipf-like distributions and Venn diagrams are descriptive representations of a phenomenon. They do not contain the potential to inform change. A more powerful device that does have this potential is metaphor. Michael A. Arbib and Mary B. Hesse (1986) suggest that:

Metaphor is potentially revolutionary. . . . Scientific revolutions are, in fact, metaphoric revolutions, and theoretical explanation should be seen as metaphoric redescription of the domain of phenomena . . . . Metaphor causes us to “see” the phenomena differently and causes the meanings of terms that are relatively observational and literal in the original system to shift toward the metaphoric meaning . . . . Meaning is constituted by a network, and metaphor forces us to look at the intersections and interaction of different parts of the network. (p. 156)

The diagrams in Figures 1 and 2 are spatial representations but not spatial metaphors. To redescribe classification through theoretical explanation, to explore its network of interactions and intersections, requires a more complex device.
Figure 2. Representation of Classification Complexity in Overlapping Categories.

To see classification differently, spatial metaphors were chosen beginning with maps. Whether maps of the earth's surface or maps of knowledge, which we call classifications, these are socially constructed representations. Werner Bies (1996) notes that cartographical and architectural images are particularly prolific as metaphors for classification (p. 4). He suggests that the way we organize knowledge is "an essential part of the cultural memory" just as is the knowledge itself (p. 7). Therefore, analyzing the metaphor as well as the system will reveal the construction with greater clarity. Use of a metaphor is itself part of the social construction. To take this a step further, it is suggested that understanding this cultural metaphor can help us understand the construction of classification.

Mapping is not a new metaphor for classification. To look at the metaphor of mapping and the way it is used in knowledge organization, the discussion will now turn to two of the fathers of knowledge or information organization, Berwick Sayers and B. C. Brookes. In his classic work, Manual of Classification, Berwick Sayers (1926) states categorically that: "A classification scheme is really a map of knowledge... A general classifica-
tion is, then, a map of the universe within and without the mind of man; it covers all things we may have known, know or can know. In the language of metaphysics it covers all *being* (pp. 65-66). Sayers is using the metaphor of the map to suggest the vastness that classification can cover. He is making a presumption that a boundless system is actually possible and that a universal classification can exist. This presumption is at odds with Drucilla Cornell's idea, described earlier—i.e., that systems are defined by their limits. The contrast between Sayers and Cornell is that of empiricist versus poststructuralist. Sayers implies that there is a single knowable reality that classification can represent in its entirety. Cornell suggests that there are multiple realities such that no system can represent reality in its totality.

In the fourth edition of his work, Sayers (1967) goes further to say that:

> [A classification's] task is to provide for the field of knowledge or part of it, as comprehensive and clear a statement as the cartographer is able to make of a territory of the earth. For just as a map makes clear the relationship between place and place so a classification strives to show the relationship of each branch of knowledge to the other branches. (p. 32)

That is, the map as a metaphor works because it is comprehensive (including everything), clear, and shows relationships. In this sense, the map works from a poststructural view as well since the map is just as constructed as the classification.

B.C. Brookes's (1980) perception, expressed in one of his articles on the foundations of information science, adds the characteristic of objectivity. Brookes suggests that a map is an objective representation of a landscape, independent of the perspective one would have in the view from a window. One "objective map accommodates all possible subjective views of the same scene" (p. 270).

Together, then, Sayers and Brookes suggest that classifications can be all-encompassing, accurate, and objective relational representations of knowledge or information—just as maps are. However, there is no reason to think that maps have any of these characteristics. Maps are no more objective or free of perspective than classifications, in spite of their basis in accurate measurement as indicated by Brookes. Maps are just as culturally bound as classifications and classificationists' reliance on cartographic imagery as being neutral and has allowed the continued existence of the illusion that classification can also be neutral. Examining the limitations of maps can help to define what questions we should ask ourselves about classification.

Maps have always determined the limits of our worlds. In 500 B.C., Hecataeus created a map of the earth as a disk representing the world known to him, with the Mediterranean (from the Latin; middle + land or
earth) at the center. It is like the Zipf-like diagram in Figure 1. What is at
the center is what he knew best. In addition, Hecataeus's map defines the
outer limits of the earth's disk as being the ocean beyond Europe, Ethio-
pia, and India. Present classification schemes are somewhat like maps of
the earth as a disk—i.e., limited to a cultural perspective.

A major characteristic of classification is that it is meant to place re-
lated concepts in proximity to each other. This factor suggests that classi-
fication is a spatial ordering. Certainly, thinking of its use for shelving
books in a library, one can recognize its existence in what amounts to one-
dimensional linear space. A book on a shelf can be to the left of one book
and to the right of another but cannot sit next to a third. Maps also
determine the perceived dimensions of spaces. Like the two-dimensional
projections of the three-dimensional earth classifications, these must dis-
tort all knowledge in its infinite multidimensionality into a linear arrange-
ment suitable for creating a browsable list or locations on shelves. For
example, traditional map projections, like the Mercator projection, skewed
the size of different parts of the world relative to each other. Therefore,
the European and North American colonial powers in the temperate lati-
tudes appear much larger than countries of the South that are closer to
the equator (for further explanation and an interesting alternative, see
Map of the World, n. d.), in the same way that traditional classifications
have allocated more space to mainstream topics and less to marginal topics.

Maps of the same area may be differently constructed depending on
cultural discourses. Maps by Native North Americans assisting European
explorers included rivers, mountains, and other physical features relative
to each other and to the settlements of different tribes (see, for example,
three maps on plate 59 in the Historical Atlas of Canada, 1987). These
maps did not divide up the entire space into discrete units like euro-settler
maps. The latter tradition, which we still follow, divides the "pie" into
separate pieces that take up the whole space. There is no common space
left. Each piece of the terrain "belongs" to someone, reflecting a particu-
lar cultural concept of property. Every inch is part of a jurisdiction. In
looking at classification as a sort of mapping, we see how similar discrete
domains with boundaries are created. Each concept is limited by its defi-
nition, and the definition is the boundary of what is or is not a given
concept. The definition marks the territory. Since the different territo-
ries or categories in this type of map are ideally mutually exclusive, defini-
tions that criss-cross each other cause problems that need to be dealt with
by breaking them into facets and creating hierarchical arrangements—
i.e., cities within states, states within countries, and countries within conti-
nents. However, going back to the contrast between Native American
maps and conventional political maps, it is apparent that even this ap-
proach of mapping by dividing up territories is not a universal concept—
it is culturally bound.
This idea of the social construction or mapping of information is very interesting in and of itself, but of what use is it? We know that classifications are not perfect, and that they reflect social biases. Can spatial metaphors of mapping help address these biases? Can these be theoretically revolutionary as Arbib and Hesse (1986) suggest above?

To test the power of the metaphor, an explanation of the development of a theoretical reason drawn from spatial constructions of a range of poststructural and feminist theorists will be given. The specific project that has been a vehicle for evolving these ideas is an effort to map the terminology of a marginalized knowledge domain to a mainstream classification, the *Dewey Decimal Classification*. The project takes the terms of a widely used feminist vocabulary, *A Women's Thesaurus* (Capek, 1987), and links these to numbers in *DDC*. This approach is a way of linking the margins and the center to create a sort of network or web instead of concentric circles with no overlaps. According to Arbib and Hesse (1986), the network creates the meaning. This project constructs a network of intersections different from those in the original *DDC* so that it creates meaning differently. To enable this network of links, Dennis Ward, a colleague in the School of Library and Information Studies, University of Alberta, is developing Windows and World Wide Web interfaces to reflect the theoretical framework.

*DDC* is a good representation of mainstream thought. It has not been left to the peculiarities of its nineteenth century origins. For example, it is a long time since it located the status of women in 396 between etiquette and outcast races as it did in earlier editions. However, it does still show the basic structure it inherited from Melvil Dewey's milieu. It is not limited by literary warrant, but revisions are often based on how literature is used in the disciplines represented in *DDC* such as recent revisions in public administration and the life sciences. Therefore, *DDC* continues to represent mainstream arrangements effectively and is constructed by various mainstream voices and not just one dominant discourse.

In comparison, the concepts represented in *A Women's Thesaurus* and the relationships between those concepts cross different disciplines, setting up an alternative structure that is fundamentally different. It is a structure developed in the margins—i.e., in the marginalized knowledge domain of women's studies and feminist thought.

By linking these two modes of representation, some might consider this as trying to fit round pegs into square holes or comparing apples and oranges or some other similar metaphor. However, continuing with these spatial metaphors, it is suggested that what is actually being done is constructing paradoxical spaces. Paradoxical space is a concept developed by feminist geographer Gillian Rose (1993). It is simply a practice that allows...
existence on both sides of a limit simultaneously or alternately. It is both inside and outside, center and margins. In this way, it does not put a new structure in place of the old but puts a different spin on existing concepts that come to coexist with concepts from the margins (p. 140).

An everyday example of paradoxical space is the concept of separate spheres, public and private, associated with men’s roles and women’s roles. The private sphere represented by the white bourgeois concept of “home” and the public sphere represented by the paid workplace have been sites of paradoxical space in a variety of ways (Rose, 1993, pp. 52-56; Haraway, 1991, p. 170). Female-intensive professions like nursing, home economics, teaching and, of course, librarianship brought the ethic of care from the private women’s sphere to the public male sphere in the nineteenth century. Later in this century, the necessary revisions of white middle-class feminism came to recognize that the private sphere, the home, is a workplace not only for the women who live there, but also for the women who leave their own homes to work in the homes of others. These women are mostly women of color who bring a very different perspective to the idea of the private sphere as a place for women’s work. Recognition of the widespread existence of wife abuse also upsets the idea of the private sphere as the place where women are in control. Technology is now reviving the old cottage model of exploitation in the home. The electronic cottage and telecommuting bring the public sphere and its values into the private sphere (see, for example, Fulton, 1997). These examples of the fuzzy boundaries between public and private make both into paradoxical space. It is no longer possible to define the limits between public and private. “Home” is not a simple concept—it never was except in our naïve constructions of it. However, we can still understand concepts like “home” because paradoxical spaces can exist.

QUALITY OF CONSTRUCTED SPACE

We can also purposely create paradoxical spaces. In this project, I have worked with research assistants to link the concepts from A Women’s Thesaurus to DDC, creating paradoxical space. As we began, the idea worked reasonably well and seemed to have potential but, as we progressed, it became apparent that some concepts mapped to positions qualitatively better than others. In seeking some way to analyze the qualities of the links, what was first considered was their coextensiveness. Coextensiveness is considered here in spatial terms: the “shape” of the topic and the “shape” of the representation are the same, or, as Jessica Milstead (1984) puts it, coextensiveness is “the extent to which the index term reflects the precise content of the item of information . . . ” (p. 143). Milstead suggests the limitations of coextensiveness for classification when she opposes the predetermined pigeonholes of classification to the potential coextensiveness of thesauri. Most classification constructs pigeonholes, which are pre-
formed without reference to the subjects of particular documents. Therefore, documents are put into the pigeonhole "closest in size to the subject" (p. 144). Coextensiveness is based on the subjects of individual documents while pigeonholing is based on the structure of the system.

Coextensiveness became a useful measure for this project in a sort of Goldilocks-and-the-three-bears way. Some matches between feminist topics and DDC numbers were too broad, others were too narrow, and others were just right. Of course the analogy did not hold up because some were overlapping in an associative, rather than a hierarchical, manner and other concepts simply had no number to represent them at all. The coextensiveness problems were especially acute in the treatment of topics from the gendered perspective implied by a feminist thesaurus that required our assessment of coextensiveness to be split into general and gendered forms (the basic approach of the research and the variables are discussed by Olson & Ward, 1997a, 1997b).

However, even more problematic, and far more theoretically interesting, were problems of gathering and proximity. Classification gathers works on a particular topic or group of topics and places them in close proximity to related topics. What became interesting as we progressed with mapping A Women's Thesaurus to DDC was that the gathering and proximity sometimes created odd, and even unfriendly, environments. To address this idea, feminist philosopher Lorraine Code's (1995) concept of rhetorical space was used:

Rhetorical spaces . . . are fictive but not fanciful or fixed locations, whose (tacit, rarely spoken) territorial imperatives structure and limit the kinds of utterances that can be voiced within them with a reasonable expectation of uptake and "choral support": an expectation of being heard, understood, taken seriously. They are the sites where the very possibility of an utterance counting as "true-or-false" or of a discussion yielding insight is made manifest. Some simple examples will indicate what I mean the term to achieve. . . . Imagine trying to make a true statement about whether it is more convenient to fly into Newark or La Guardia airport in the year 1600. The statement would not be false but meaningless: it could neither be true nor false within the available discursive possibilities. Or imagine trying to have a productive public debate about abortion in the Vatican in 1995, where there is no available rhetorical space, not because the actual speech acts involved would be overtly prohibited, but because the available rhetorical space is not one where ideas on such a topic can be heard and debated openly, responsibly. . . . what I want this terminology [rhetorical space] to do [is], namely to deflect the focus of philosophical analysis away from single and presumably self-contained propositional utterances pronounced by no one in particular and as though into a neutral space; and to move it into textured locations where it matters who is speaking and where and why, and where such mattering bears directly upon the possibility of knowledge claims, moral pronouncements, descriptions of "reality" achieving
acknowledgment, going through. Often in such spaces discourse becomes a *poiesis*, a way of representing experience, reality, that re-makes and alters it in the process. And the making is ordinarily a communal process, dependent for its continuance on receptive conditions, on engaged responses both favourable and critical. (p. x)

Code proposes a new concept in spatial imagery that helps explain what goes wrong for marginalized topics in classification. Even more than coextensiveness, rhetorical space helps explain why the mapping of classification supports mainstream biases. To demonstrate, key phrases will be taken from Code's defining quotation and will be elaborated on in relation to classification and in light of other feminist and poststructuralist theorists' work.

1. "Rhetorical spaces . . . are fictive but not fanciful or fixed. . . ." That is, they are constructed, made-up (fictive), but not arbitrary (not fanciful) and dynamic (not fixed). This is also true of classifications. Classifications are not innate or natural but are constructed. Their construction has some logical basis, and they change as discourses continue to act upon them.

2. "[T]erritorial imperatives structure and limit . . .". The spaces have boundaries. They are limited by the way they are constructed and by the imperatives of the discourses that construct them. Again they are like classifications which are limited systems, including some concepts and excluding others. As Cornell suggests, limits define systems and spaces.

3. "[B]eing heard, understood, taken seriously . . .". What is limited when positive rhetorical space is lacking is voice. Voice is a given for mainstream discourses, but for marginalized discourses it is something more vital. As bell hooks (1989) puts it:

Moving from silence into speech is for the oppressed, the colonized, the exploited, and those who stand and struggle side by side. . . a gesture of defiance that heals, that makes new life and new growth possible. It is that act of speech, of "talking back," that is no mere gesture of empty words, that is the expression of our movement from object to subject "the liberated voice." (p. 9)

So a positive rhetorical space allows marginalized discourses to be heard as legitimate statements—i.e., to be acknowledged as worth listening to. Cornell (1992) proposes a responsibility for those who control a system to make its limits permeable so that they can approach an ethical relationship with those who are excluded (p. 62). The permeability of the limits allows the voices of the excluded—the oppressed, the colonized, the exploited—to be heard in the system. Those of us with authority must constantly throw the system off balance to maintain this permeability (p. 80).
4. Code intends the idea of rhetorical space "to deflect the focus of philosophical analysis from single and presumably self-contained utterances pronounced by no one in particular as though into a neutral space..." (p. 10). This apparently neutral space is like postmodern theorist Henri Lefebvre's (1991) "transparent space" which denies the existence of anything excluded from its mapping because it appears to be all there is. "Transparent space" is the illusion that location is neutral—i.e., that mapping territory can be a true representation of some essence of reality as B.C. Brookes (1980) suggests. However, there is an implied exclusion defining the transparent space that is hidden by it. Because space has boundaries and always includes and excludes something, it cannot be neutral. Making the exclusions visible means identifying the space's boundaries to allow recognition of what is outside those boundaries. It is identifying the implied opposite of a deconstructive binary opposition. It shows the constructed nature of the space. It moves the discourse "into textured locations where it matters who is speaking and where and why, and where such mattering bears directly upon the possibility of knowledge claims" (p. x). Transparent space is difficult to discern because of its apparent neutrality. The efforts at neutrality that are made in classification can mask exclusions.

5. "[P]oiesis, a way of representing experience, reality, that remakes and alters it in the process." Poiesis is a creative and creating production. In poiesis, we understand that the representation of reality is the construction of reality. The representation of information, through classification, is part of the construction of information. Classification remakes and alters information by constructing a particular context for it—gathering, scattering, and juxtaposing topics in relation to each other. How broadly or narrowly topics are represented will enhance or mask their visibility. In these ways, classification produces information in a creative process. This process of poiesis is a locational one. Feminist sociologist Elspeth Probyn (1990) proposes that:

Through location knowledges are ordered into sequences which are congruent with previously established categories of knowledge. Location, then, delineates what we may hold as knowable and, following Foucault, renders certain experiences "true" and "scientific" while excluding others. (p. 178)

She continues on to point out that this act of creation or construction determines not only what is knowable, but whose voices are heard. So the creation of classification creates the space in which some knowledges are central and others are peripheral.

6. This creation of reality is "a communal process, dependent for its continuance on receptive conditions, on engaged responses both
favourable and critical." That is, both the context and the process affect the construction of reality. It could be called reflexive or holistic. It is akin to the death of the author and the ascendancy of the reader in literary criticism. The author does not create the text. It is created in the process of reading and depends upon the "receptive" and "engaged" reader for its meaning and existence. It involves interpretation, in the case of classification, by classifiers and users. In this sense it places responsibility for the construction of information not just on classificationists who write classifications but also on the individuals and institutions who use classifications.

**Operationalizing the Theory**

To operationalize the concept of rhetorical space, our research project examines the DDC context of individual concepts from *A Women's Thesaurus* by looking at:

1. What other topics share the number?
2. How is the number described?
3. What is the hierarchical context?
4. What topics sit on either side?

Two examples from our pilot study illustrate how the variable of rhetorical space works to reveal whether or not feminist topics can be taken seriously in DDC.

The first topic is colonialism. The following entry from *A Women's Thesaurus* implies the scope of this term:

Entry from *A Women's Thesaurus*:

**colonialism**
- UF imperialism
- NT neocolonialism
- RT apartheid
- cultural imperialism
decolonization
developing nations

Colonialism appears in the index to DDC and points unequivocally to the number 325.3. This concept is not excluded from DDC. However, its rhetorical space is not as neutral as it at first appears. The following entry from DDC shows how 325.3 is defined and what it includes:

Entry from *DDC*:

325.3 **Colonization**
Class here exercise of political dominion over distant territories
The location of colonialism with colonization is an example of transparent space. It seems neutral but is actually one-sided, showing colonization from the point of view of the colonizing power as opposed to the people and culture being colonized. Colonialism is linked to this number not in its caption but only as a reference from the DDC relative index. The entry describes the colonized territories as “distant” from the colonizing powers, not the other way around. Colonies are not distant in the view of colonized people.

The following summary shows the hierarchical and sequential contexts of 325.3 that reinforce the perspective of the imperial power observing the colony:

Summary from DDC:

300 Social sciences
320 Political science (Politics and government)
325 International migration and colonization

325.1 Immigration
325.2 Emigration
325.3 Colonialization
325.4-325.9 International migration to and colonization in specific continents, countries, localities in modern world

In categories 325.4-325.9, the geographic subdivisions that combine “migration” and “colonization” move from colonizing country “to” colony in its subdivisions and define colonization as “in” the colonized locale and not “by” the colonizers. As with the description of 325.3, the geographic subdivision is entirely from the perspective of the imperial power. Colonization involves movement—being colonized does not. The location of colonialism in the established category of colonization fulfills Probyn’s (1990) prediction that the voices of hooks’ colonized become peripheral.

A further indication of this marginalization is found by looking higher up the classificatory hierarchy. The DDC principle of hierarchical force dictates that what is true for 325 is also true for its subdivisions including 325.3:

Entry from DDC:

325 International migration and colonization
Including involuntary population transfer, population exchange
Class movement of people associated with a specific event in history with the event in 909 or 930-990; class interdisciplinary works on international movement of people in 304.82
While colonialism involves people from the imperial power going to the colonized territory, its disruptive nature often causes other movements of people—most notably the scattering of colonized peoples that form diaspora. For people moving away from the site of colonization as part of the diasporas that often result from colonialism, the instructions under category 325, *International migration and colonization*, suggest two options. First, movement of people associated with a specific event in history is to be classed with the event in the 900s. This option has at least two problems for representing postcolonial diasporas: first, it is difficult to pin down colonization as a specific event because it tended to happen over a diffuse period of time not conducive to classification and, second, by putting these movements into history, we would take them out of the present day where their results, typically including racism, must be addressed.

The second option is to use the number for interdisciplinary works on international movement of people in 304.82:

Entry from *DDC*:

304.82  International movement

Class international emigration in 304.809; class international immigration in 304.83-304.89

The number 304.82 is a subdivision of movement of people under the broader concept of social behaviors, thus taking this issue out of the political realm of colonialism. It does allow for geographic subdivision in either direction—the country to which people went or the country which they left. The latter applies to the postcolonial diaspora, but it uses the less preferred number. Preference will be given, according to *DDC*’s rule of zero, to international immigration rather than emigration. Therefore, people leaving India for other places in general will be classified with movement from India, but people leaving India for specific destinations will be classified with movement toward each of those places (such as movement toward England) thus diasporizing the diaspora.

These details about the siting of colonialism in *DDC* help to reveal that the apparently neutral transparent space is actually skewed toward a mainstream interpretation because it has been constructed by a mainstream discourse.

A second example is the poor representation in mainstream schemes of the concept of unpaid labor. The heading in *A Women’s Thesaurus* is:

Entry from *A Women’s Thesaurus*:

unpaid employment

UF  nonwage labor
BT  employment
NT  unpaid household labor
RT  economic value of women’s work
homemaking
unpaid labor force
valuing children
volunteer work

The term *unpaid employment* in *A Women's Thesaurus* is a subdivision of *employment*. However, when we try to classify it in the same way in *DDC*, we find that the relative index sends us to 331.125, *Labor actively employed*:

Entry from *DDC*:

331.125  Labor actively employed
That portion of the total available supply of labor employed at any given time
Including types of *employment*
Class here utilization of human resources, employment, comprehensive works on employment and compensation

DDC Index Terms:
- Employment
- Human resources—utilization—economics
- Occupations—active employment

The “class here” note indicates that the number includes “comprehensive works on employment and compensation” implying only paid labor. This implication is confirmed by hierarchical force since 331.125 is a subdivision of 331.12:

Entry from *DDC*:

331.12  Labor market
The activities of and opportunities for *buying and selling labor*

The phrase “buying and selling” confirms that only paid labor is included. The same is true if we look at the labor force:

Entry from *DDC*:

331.11  Labor force
All who are employed or available for employment
Class here human resources, manpower and womanpower, labor supply, size of labor force

The scope note indicates that this definition of *labor force* includes “all who are employed or available for employment” which is not likely meant to include people who are available for becoming housewives and househusbands. If the latter are part of the labor force under this definition, it is more likely as people available for paid employment.

Some conventional types of unpaid employment are represented in other places in the classification with the result that they are not treated
as labor. Homemaking, one of the related terms to unpaid employment in A Women's Thesaurus, is located in home economics, which is a subdivision of technology:

Entry from DDC:

640 Home economics and family living
Class here management of home and personal life, domestic arts and sciences
DDC Index Terms:
Domestic arts
Domestic sciences
Home economics
Homemaking
Homes
Homes—home economics
Household management

Summary from DDC:

600 Technology (Applied sciences)
640 Home economics and family living
640.4 Specific aspects of household management
640.41 Helpful hints and miscellaneous recipes
640.42 Management of money
640.43 Management of time
640.46 Household employees
640.49 Survival housekeeping

While this is an excellent place to put information about the processes and production of homemaking, acknowledging it as appropriate to be adjacent to agriculture or engineering, it does not include the aspect of the people who do this labor except in a subdivision for household employees, the people who are paid to do housework.

Another example of unpaid labor is voluntarism, which is defined as one aspect of social participation along with encounter groups and sensitivity training:

Entry from DDC:

302.14 Social participation
Including communalism, competition, cooperation, encounter groups, sensitivity training, voluntarism

While anyone who has done volunteer work may appreciate its links to encounters and sensitivity, this location treats voluntarism as "social participation" and not as labor. While it is important not to belittle the value of social participation in a world in which governments are cutting back
essential social services which then fall to voluntary agencies, the labor of voluntarism also becomes a key to economic well-being.

A way of addressing this problem is to create a paradoxical space by locating the general concept of unpaid employment alongside paid employment. While these are two different topics, they may sit adjacent to each other to create viable rhetorical space and give legitimate voice to the unpaid labor performed everyday around the world. One option would be to tuck it into other generalizations of the labor force, say in the gap between “qualifications and personal characteristics” and “systems of labor”:

Summary from DDC with potential addition:

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>Social sciences</td>
</tr>
<tr>
<td>330</td>
<td>Economics</td>
</tr>
<tr>
<td>331</td>
<td>Labor economics</td>
</tr>
<tr>
<td>331.1</td>
<td>Labor force and market</td>
</tr>
<tr>
<td>331.11</td>
<td>Labor force</td>
</tr>
<tr>
<td>331.110*</td>
<td>&lt; Zero Subdivisions &gt;</td>
</tr>
<tr>
<td>331.111</td>
<td>Geographic distribution</td>
</tr>
<tr>
<td>331.114</td>
<td>Qualifications and personal characteristics</td>
</tr>
<tr>
<td>$\Rightarrow$</td>
<td>Economic basis of labor</td>
</tr>
<tr>
<td>$\Rightarrow$</td>
<td>Unpaid labor</td>
</tr>
<tr>
<td>$\Rightarrow$</td>
<td>Paid labor</td>
</tr>
<tr>
<td>331.117</td>
<td>Systems of labor</td>
</tr>
<tr>
<td>331.1172</td>
<td>Free labor</td>
</tr>
<tr>
<td>331.1173</td>
<td>Compulsory labor</td>
</tr>
<tr>
<td>331.118</td>
<td>Labor productivity</td>
</tr>
<tr>
<td>331.119</td>
<td>Labor force by industry and occupation</td>
</tr>
</tbody>
</table>

Here a section with a title something like “economic basis of labor” with the subdivisions “unpaid labor” and “paid labor” (with the mainstream interpretation coming second to upset the hierarchy a bit) could set this topic alongside basic concepts of the labor force. Because this number would be hierarchically encompassed by “labor force,” the definition of the latter would have to be adjusted to include unpaid labor as would 331.1, Labor economics, in a reversal of hierarchical force—the subdivision driving the definition of the dominant concept. Mapping a marginal concept in the midst of a mainstream concept will not alone create positive rhetorical space. However, careful placement of such interpolations will make transparent space visible and will create paradoxical spaces where discussion of issues can continue openly.

CONCLUSION

Following Drucilla Cornell’s concept of systems’ limits, there would be no point in abandoning our existing classifications in the hope of achieving that objective map described by B.C. Brookes. All systems will exclude
and marginalize in some way. However, it is possible to shift between mainstream and margin in our mapping, creating paradoxical spaces and defining the limits differently. With a new theoretical framework, it is possible to make changes in our mapping akin to the changes made when Pythagoras determined that the earth is round and not flat.

The new theoretical framework developed throughout this article is offered as a new way of mapping knowledge in classification. It has potential for both analysis and amelioration. The categories of classification—because they typically reflect a cultural mainstream—appear neutral, objective, and transparent. This makes marginalizations and exclusions difficult to identify. Therefore, to analyze the problems of classification in relation to marginalized knowledge domains, the framework poses three assumptions drawn from feminist and poststructural literature examined in this discussion. First, classification, like any map, is constructed by dominant cultural discourses. Second, classification, like any system, has constructed boundaries or limits that result in exclusions. Third, the construction of classification is a form of location that defines and sequences what is accepted as knowledge, thus marginalizing as well as excluding. Regarding classification as a text and reading it with these three assumptions in mind will make what was transparent and invisible opaque and visible, elucidating the biases and the discourses that construct and enforce them.

To ameliorate the biases of classification, this framework proposes that the limits of a classification be made unstable and permeable to allow the voices of those who have been excluded to be heard. In this way, the classification approaches an ethical relationship with previously silenced voices. Further, to address the marginalizations within classification, this theoretical stance advocates the creation of paradoxical spaces that are neither mainstream nor marginal but are both simultaneously or alternately. By mapping A Women's Thesaurus to DDC, this project creates such spaces. The same concept can offer other ways of deconstructing and reconstructing not only the limits but also the structures of classification.

The creation of paradoxical spaces can become a poiesis—i.e., that alters representation in ways that make boundaries permeable. In the next stage of this project, suggestions will be made for revision, supplements will be devised, and optional practices offered to further develop paradoxical spaces for women’s studies and feminist thought in DDC. It will allow for more dimensions and, thus, more creative connections between places/spaces/concepts than have hitherto been available. Further, we hope that it will be a prototype for a poiesis applicable to other marginalized knowledge domains. With care, paradoxical spaces will appear throughout classifications, thereby keeping them from stagnating and keeping them vital and exciting.
NOTES

1See, for example, Afolabi (1992) on Africana; Hamdy (1980) on Arabic materials; Iwuji (1989) on Africa; Lochhead (1985) on women; McConnell (1984, 1985a, 1985b) on Melanesia; Milstead Harris & Clack (1979) on nearly all of these groups; Mowery (1989) on Mexican Americans; Pacey (1989) on Africa; Steinberg (1974) on women; and Wolf (1972) on gays and lesbians.

2Entries and summaries from DDC are from the electronic version, Dewey for Windows. They have been edited in format. The relative index terms are included only when they are relevant to the discussion and related Library of Congress Subject Headings have been omitted. Other omissions are shown by ellipses. Emphasis is mine to facilitate interpretation.

3Hierarchical force (DDC Glossary). The principle that the attributes of a class as defined in the heading and in certain basic notes apply to all the subdivisions of the class, and to all other classes to which reference is made" (Dewey for Windows).

4Rule of zero (DDC Glossary). The rule instructing that subdivisions beginning with zero should be avoided if there is a choice between 0 and subdivisions beginning with 1-9 in the same position in the notation. Similarly, subdivisions beginning with 00 should be avoided when there is a choice between 00 and 00" (Dewey for Windows, 1996).

REFERENCES


The Kindness of Strangers: Kinds and Politics in Classification Systems

GEOFFREY C. BOWKER

ABSTRACT
This article offers a formal reading of a classification scheme of international scope and long duration: the *International Classification of Diseases (ICD)*. The argument is made that this classification scheme retains many traces of its own administrative and organizational past in its current form. Further, it is argued that such traces operate normatively to favor certain kinds of narrative of medical treatment while denying others. It is suggested that the *ICD*, like other large-scale classification systems, is able to do its work so effectively precisely because these traces permit a coupling of classification scheme and organizational form.

INTRODUCTION
In so far as the coding scheme establishes an orientation toward the world, it constitutes a structure of intentionality whose proper locus is not the isolated, Cartesian mind, but a much larger organizational system, one that is characteristically mediated through mundane bureaucratic documents such as forms. (Goodwin, 1996, p. 65)

In the digital libraries that are being constructed today, a burgeoning number of formal classification systems are being inscribed deep into the infrastructure of the information system.

In this discussion, some medical classification systems with a long history will be examined—notably the *International Classification of Diseases (ICD-9-CM, 1996; ICD-10, 1992)*, in operation since the 1890s—in order to discern the relationship between the use of the classification as an information storage and retrieval mechanism and its use to encode multiple political and ethical agendas.

Geoffrey C. Bowker, Graduate School of Library and Information Science, 501 E. Daniel Street, University of Illinois, Champaign, IL 61820

© 1998 The Board of Trustees, University of Illinois
One classic division between kinds of classification system is that drawn by Taylor (1995), who distinguishes between Aristotelian classification and prototype classification. The prototype classification was defined by experimental psychologist Eleanor Rosch (1978). This distinction is going to be an important one throughout this discussion and will be explored in some detail. An Aristotelian classification works according to a set of binary characteristics, which the object being classified either presents or does not present. At each level of classification, enough binary features are adduced to place any member of a given population into one, and only one, class. So we might say that a pen is an object for writing within a population consisting of pens, balls, and bottles (Taylor, 1995). We would have to add in one more feature in order to adequately distinguish pens, for example, from pencils, balls, or bottles. A technical classification system operating by binary characteristics is called monothetic if a single set of necessary and sufficient conditions is adduced ("in the universe of polygons, the class of triangles consists of figures that have three sides"), polythetic if a number of shared characteristics are used (in our example, the pen could be described as thin, cylindrical, used for writing, has a ball point, and so forth) (Blois, 1984). Desrosières (1993) indicates a typical breakdown between monothetic and polythetic classifications in the work of statisticians. He associates the former with Linnaeus and the latter with Buffon (who engaged in local classification practices, just using the set of traits needed to make a determination in a specific instance) and writes: "These local practices are often carried out by those working in statistical centers, according to a division of labor whereby the chiefs are inspired by Linnaean precepts but the working statisticians apply, without realizing it, Buffon's method" (p. 296 [authors' translation]). Aristotelian models—monothetic or polythetic—have traditionally informed formal classification theory in a broad range of sciences, including biological systematics, geology, and physics.

Rosch's (1978) prototype theory argues that, in daily life, our classifications tend to be much fuzzier than we might at first think. We do not deal with a set of binary characteristics when we decide that this thing we are sitting on is a chair. Indeed, it is possible to name a population of objects that people would in general agree to call chairs that have no two binary features in common.

According to prototype theory, there is a broad picture in our minds of what a chair is, and this picture is extended by metaphor and analogy when trying to decide if any given thing that we are sitting on counts. We call up a best example and then see if there is a reasonable direct or metaphorical thread that takes us from the example to the object under consideration. Prototype theory has been powerfully developed within the field of sociolinguistics by George Lakoff (1987) and John Taylor (1995). One finding of the theory is that different social groups tend to have quite
different prototypes in mind when classifying something—e.g., a piece of furniture. Thus, when surveyed, a group of Germans came up consistently with a different set of best examples than a group of Americans (Taylor, 1995, pp. 44-57). For the Americans, chair and sofa are best fits for furniture, for the Germans, asked about möbel, it was bed and table. An important implication of the theory is that there are levels at which we most easily and naturally distinguish between objects in the world, and that supervenient or subvenient levels tend to be more technically defined. Looking at a picture of a Manx coon cat, a nonexpert will say that this is a picture of a cat. An expert might call it either a Manx coon cat or a vertebrate.

This distinction between two main types of classification is a very useful one. However, there are a number of reasons for saying that it is not an absolute distinction—indeed, one could say that we all probably have a personal prototype of the ideal Aristotelian classification system, but that no one system in practice fully meets a single set of Aristotelian requirements. We stress “in practice” here, since it is practice that this discussion is largely about. Turning to an example from the workplace, it is possible to begin to see how practice and location mediates such divisions. In the medical arena, it emerged from a survey of physicians in 1979 in the United Kingdom that general practitioners “had a constant tendency to regard a wider range of phenomena as disease” than the hospital physicians, who in turn were more inclusive than the lay public—the perceived need for medical intervention being the determining axis (Prins, 1981, p. 176; Campbell, Scadding, & Roberts, 1979). An influential factor, Prins notes, seems to have been whether or not medical intervention was required. For the lay public, “measles” and “mumps” might be prototypical diseases, but “arthritis,” a card-carrying *ICD-10* (1992) disease, might be seen rather as a condition.

So why do we seem in practice prototypical even if in principle Aristotelian? For two main reasons: (1) because each classification system is tied to a particular set of coding practices, and (2) because classification systems in general (we are not making this as an *ex cathedra* pronouncement) reflect the conflicting contradictory motives of the sociotechnical situations that gave rise to them.

**PRACTICES**

Consider the *International Classification of Diseases* (*ICD-9-CM*, 1996; *ICD-10*, 1992). When originally drawn up, it had a maximum of 200 categories, not because this was the number of diseases in the world but because this had been the number of lines on Austrian census forms. If too many diseases got identified, then there would be no way of maintaining and analyzing registers of causes of death as the technology would not hold more information.
In addition to this inheritance, there is a practical Occam’s razor. When doctors come to code causes of death, they are frequently faced with a set of difficult judgments (that may require an autopsy and further diagnostic work). They can simply go for the easiest solution—i.e., by using a generalized “other” category. They can then get back to dealing with their live patients (Fagot-Largeault, 1989). So the classical beauty of the Aristotelian classification gives way to a fuzzier classification system that shares in practice key features with commonsense prototype classifications—i.e., heterogeneous objects linked by metaphor or analogy.

The powerful habits of practice with respect to the humble tasks of filling out forms are often neglected in studies of classifying. Goodwin (1996) provides an elegant description of working student archaeologists matching patches of earth against a standard set of color patches—the Munsell color charts. He notes that earlier cognitive anthropological work on color assumed a universal genetic origin for color recognition but failed to examine the kinds of practices that informed the ways in which color tests were designed and carried out in the course of this research. Goodwin (1996) notes:

Rather than standing alone as self-explicating textual objects, forms are embedded within webs of socially organized situated practices. In order to make an entry in the slot provided for color an archaeologist must make use of another tool, the set of standard color samples provided by a Munsell chart. This chart incorporates into a portable physical object the results of a long history of scientific investigation of the properties of color. The version of this chart that archaeologists bring into the field has been tailored to the distinctive requirements of their work situation. (p. 66)

The archaeologists constantly compare the pieces of earth against the chart, negotiate with each other, and transform their everyday terms for the earth into the formal numbered categories on the chart. The uncertainties they face along the way are removed once the numbers are selected and reported: “The definitiveness provided by a coding scheme typically erases from subsequent documentation the cognitive and perceptual uncertainties that these students are grappling with, as well as the work practices within which they are embedded” (Goodwin, 1996, p. 78).

**Contradictory Requirements of Classification Systems in General**

Classification systems in general inherit contradictory motives in the circumstances of their creation. This is very clearly illustrated by items in the *ICD* covering such charged ethical or religious issues as abortion or stillbirth. Over the years, defining the moment of birth differed radically from Protestant to Catholic countries and with technological changes. The final definitions given in the *ICD* directly reflect the charged political and
ethical atmosphere of the subject, distinguishing, for example, legal and illegal abortion as separate categories. In this sense, the ICD can also be read as a kind of treaty, a bloodless set of numbers obscuring the behind-the-scenes battles informing its creation. This dryness itself contains an implicit authority, seeming to rise above uncertainty, power struggles, and the impermanence of the compromises.

Indeed, one might observe that technical classification schemes are constructed in such a way as to fit our commonsense prototypical picture of what a technical classification is. Thus when the International Committee for the Nomenclature of Viruses, to which we shall return, floated the idea of using “siglas”—a series of code letters attached to the virus name to indicate its characteristics—Matthews (1983) describes the response as follows: “Leading virology journals were only lukewarm to try out cryptogram ideas. Among comments from this period: ‘Why should they be given funny names? Are we not exposing ourselves to the laughter of the general public? Do we want to join the ranks of old-fashioned botanists and zoologists so soon?’” (pp. 13-14). A good technical classification should not only be correct in Aristotelian terms, it should, in good prototypical fashion, look and feel scientific. This is not an isolated case. The developers of the Nursing Interventions Classification (NIC) have made similar observations—e.g., they initially did not classify “leech therapy” not because it was not a scientific intervention but because it did not look and feel like one. With respect to the ICD, there has been a long debate within the patient community about naming chronic fatigue syndrome, for example (as there was for AIDS). Consider this discussion among patients suffering from chronic fatigue syndrome:

Many patients feel that one of the greatest burdens of having chronic fatigue syndrome is the name of the illness. The word “fatigue” (which many patients refer to as the “F word) indicates everyday tiredness. It reinforces negative perceptions that remain with the public and most medical doctors, despite a decade of steady, gradual research advances. (Chronic Fatigue Syndrome Electronic Newsletter, 20 February, 1997)

One option was to name it after Darwin, but it was felt that, although he had the scientific cachet, he did not necessarily have the disease. Inversely, Florence Nightingale’s diagnosis is more certain but less prestigious:

Nightingale’s. (A general note: no historical figure has been definitively diagnosed with CFS/M.E. Purists may object to choosing any person in history, who may not have actually had the disease, as the basis for an eponym.) Florence Nightingale is a widely respected and world-renowned figure who founded the International Red Cross and the first formal school for nursing. For decades she had an undiagnosed, severely debilitating illness with symptoms similar to CFS. Despite Nightingale’s considerable talents and her personal character, many doubted that she had a physical illness. Her illness was
quite controversial. A 1996 paper by D.A.B. Young that appeared in the British Medical Journal indicates that Nightingale’s illness was likely to have been chronic brucellosis (a disease with symptoms similar but not identical to CFS). Patient groups have promoted Nightingale’s birthday, May 12, as International CFIDS/M.E. Awareness Day, and Nightingale is a familiar symbol to those who know this disease. However, some argue that women’s diseases often have difficulty in getting recognized and accepted. Choosing Nightingale’s name as an eponym might add to, rather than offer relief from, current name-associated problems. (Chronic Fatigue Syndrome Electronic Newsletter, 20 February 1997)

More generally, Taylor, from a linguistic perspective, and Durkheim and Mauss (1968) (for whom primitive social classifications “seem to link, without any discontinuity, with the first scientific classifications” [p. 82]) from an anthropological one have observed that our technical classifications grow out of, and have to answer to, commonsense socially comfortable classifications. It just would not be socially feasible to call a donkey a fish no matter how good your scientific grounds.

There is no great divide between folk and scientific classifications. Below, we discuss one particular fault line between the two: a fracture that is constantly being redefined and changing its nature as the plate of lived experience is subducted under the crust of scientific knowledge. This fault line is the ways in which temporal experience—i.e., history, experience, development, memory, evolution—is registered in, and expressed by, two formal classification systems—the ICD and the INV. The crack comes when the messy flow of bodily and natural experience must be ordered against a formal neat set of categories. We will trace this particular faultline across the two classification schemes. It is the case that all complex classification schemes in fact have multiple sets of faults and fractures arising from similar tensions. On a meta level, the system of faults and tensions forms a kind of texture of any given organizational terrain; mapping this texture is a major research challenge for the field of social informatics.

**The International Classification of Diseases is a Pragmatic Classification**

In order to communicate information in the aggregate, it must first be classified. At any time over the past 100 years, one can find complaints about the Tower of Babel that afflicts the storage and communication of medical knowledge.1 David Rothwell (1985) notes that:

More than two hundred statistical systems are being used by the United States government to monitor health, occupational and environmental conditions through the country. Despite the incredible amount of information accumulated, there is no method of coordinating these data into a single coherent database, a national health information system. (p. 169)
Mark Musen (1992) complains:

The medical-informatics community suffers from a failure to communicate. The terms that WMR uses to describe patient findings generally are not recognized by Medline. The manner in which Iliad stores descriptions of diseases is different from that of Dxplain. Therapy plans generated by ONCOCIN are meaningless to the HELP system. . . . Each time another developer describes yet another formalism for encoding medical knowledge, the number of incompatibilities among these different systems increases exponentially. (p. 435)

Musen indicates that there is no clear relationship between “the Unified Medical Language System [UMLS] advanced by the National Library of Medicine and the Arden syntax proposed by the American Society for Testing and Materials as a standard for representing medical knowledge” (p. 436). The ICD, he points out, originated as a means for describing causes of death; a trace of its heritage is its continued difficulty with describing chronic, as opposed to acute, forms of disease. This is one basis for the temporal faultlines that emerge in its usage. The UMLS originated as a means of information retrieval (the MeSH scheme) and is not as sensitive to clinical conditions as it might be (p. 440).

The two basic problems for any overarching classification scheme in a rapidly changing and complex field can be described as follows: first, any classificatory decision made now might, by its nature, block valuable future developments. If we decide that all instances of Sudden Infant Death Syndrome are to be placed into a single box (ICD-10, 1992, vol. 1, R 95, p. 890), then we are not recording information that might be used by future researchers to distinguish possible multiple social or environmental causes of the syndrome. We are not making it impossible to carry out such studies, but we are making it difficult to retrieve information. Second, inversely, if every possible relevant piece of information was stored in the scheme it would be entirely unwieldy.

The decision not to collect is the most difficult for any classification on these grounds, whether it be the acquisition department of a library, the curator of an art museum, or the collector of information for vital statistics. There are always practical budget and storage issues. These are balanced against two other factors: (1) the need for a well ordered and, in some sense, parsimonious repository that can be used, and (2) the side bets that are made about what material will be useful in the future. This latter is particularly difficult.

Collectors and curators of all sorts must become future forecasters and decide the boundaries of what will be useful for the future. There is no perfect answer, only a set of practical tradeoffs. This is a problem that has plagued museums of natural history, for example. Fossils found in the nineteenth century might come along with general information about
the depth at which they were discovered and the surrounding geological features (though they often did not). Even if this information was included, it was never as precisely noted as would be useful for geologists and paleontologists today since there was just no conception at that stage of the kinds of dating techniques that are used today. The museum is then faced with the choice between recording as much as possible now (which is very expensive and possibly not useful anyway) and having the collection perhaps last longer into the future or recording a judicious amount now (which will keep the administrative costs down) and having the collection possibly be not so useful in the future. The latter has generally been the de facto choice and is generally a reasonable one to have made since new criteria of relevance just cannot be predicted.

Different designers of the classification system have different needs—and the shifting ecology of relationships between the disciplines using the classification will necessarily be reflected in the scheme itself. As with the insurance company example above, these relationships must be resolved in order to make a usable form, often obscuring power relationships in the process. As Goodwin (1996) notes: “A quite different kind of multivocality, one organized by the craft requirements of a work task rather than the genres of the literary academy, can be found in mundane bureaucratic forms" (p. 66). But one must dig to find the voices. The process of filling out the forms may further obscure them. For example, the designers of the *ICD* recommend that its classification scheme be interpreted economically:

> The condition to be used for single-condition morbidity analysis is the main condition treated or investigated during the relevant episode of health care. The main condition is defined as the condition, diagnosed at the end of the episode of health care, primarily responsible for the patient’s need for treatment or investigation. If there is more than one such condition, the one held most responsible for the greatest use of resources should be selected. . . . (*ICD-10*, 1992, vol. 2, p. 96)

This reflects a constant condition of the use of the ICD; it has been recommended throughout its history that priority should be given to coding diseases that represent a threat to public health. This goal is clearly a good one; equally clearly it can discriminate selectively against the reporting of rare noncontagious conditions.

Faced with these problems, the WHO has been consistently pragmatic in its aims and clear in its explanations of the ICD. From the time of the ninth revision on, it has been recognized explicitly that “the ICD alone could not cover all the information required and that only a ‘family’ of disease and health related classifications would meet the different requirements in public health” (*IDC-10*, 1992, vol. 2, p. 20). This “family” is pictured in ICD-10 (see Figure 1).
The family itself is a diverse one: there are various standard modifications of the *ICD*. The most significant is the *ICD-9-CM* (1996) where CM stands for "clinical modification." This has a complex history, originating in the development of modifications of the *ICD* for use in hospital information systems. It is now the classification of record in a wide variety of medical settings and is used for billing, insurance, and administration as well as in-patient medical records. This institutional entrenchment has made it very difficult for *ICD-10* (1992) to be fully adopted in the United States with the clinical modification necessarily lagging behind the production of the classification itself.
When we observe the ways in which culture and practice interweave in the text of the ICD, we are not unmasking a false pretender to the crown of science. We are drawing attention to an explicit positive feature of ICD design: "The ICD has developed as a practical, rather than a purely theoretical classification. . . . There have . . . been adjustments to meet the variety of statistical applications for which the ICD is designed, such as mortality, morbidity, social security and other types of health statistics and surveys" (ICD-10, 1992, vol. 2, p. 12). The preamble to the classification defines a classification of diseases as "a system of categories to which morbid entities are assigned according to established criteria" (ICD-10, 1992, vol. 1, p. 1). A statistical classification, such as the ICD, "must encompass the entire range of morbid conditions within a manageable number of categories" (ICD-10, vol. 2, p. 1). It is not meant to be a net to capture all knowledge but a workable epidemiological tool. This practical goal does not make it less scientific, of course; all classification systems are developed within a context of organizational practice. The goal of the ICD's designers is to create what Latour (1988) has called immutable mobiles—inscriptions that may travel unchanged and be combinable and comparable. Indeed, the term "immutable mobile" might almost have been in the designers' minds when they wrote:

The purpose of the ICD is to permit the systematic recording, analysis, interpretation, and comparison of mortality and morbidity data collected in different countries or areas and at different times. The ICD is used to translate diagnoses of diseases and other health problems from words into an alphanumeric code, which permits easy storage, retrieval, and analysis of the data. (ICD-10, 1992, vol. 2, p. 2)

The ICD has become the international tool for "standard diagnostic classification for all general epidemiological and many health management purposes" (p. 2).

The world has changed since the ICD was first introduced, and the classification scheme has evolved to try to encompass these changes. The ICD is thus both highly responsive and tightly constrained. A large-scale change in the way that people die (Israel, Rosenberg, & Curtin, 1986, p. 161) has led to an alteration in one line in the internationally recommended Death Certificate. This is, of course, one of the main bureaucratic uses of the ICD—i.e., the recording and compiling of causes of death from bureaus of vital statistics via coroners, hospitals, doctors, or priests:

In considering the international form of medical certificate of cause of death, the Expert Committee had recognized that the situation of an aging population with a greater proportion of deaths involving multiple disease processes, and the effects of associated therapeutic interventions, tended to increase the number of possible statements between the underlying cause and the direct cause of death: this meant that an increasing number of conditions were being entered
on death certificates in many countries. This led the committee to recommend the inclusion of an additional line (d) in Part 1 of the certificate. *(ICD-10, 1996, vol. 1, p. 18)*

Thus there is now one more blank line on the form to indicate multiple causation (see Figure 2).

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Approximate interval between onset and death</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>I</strong></td>
<td></td>
</tr>
<tr>
<td>Disease or condition directly leading to death*</td>
<td>(a)........................................................................</td>
</tr>
<tr>
<td>Antecedent causes</td>
<td></td>
</tr>
<tr>
<td>Morbid conditions, if any, giving rise to the above cause, stating the underlying condition last</td>
<td>(b)........................................................................</td>
</tr>
<tr>
<td></td>
<td>due to (or as a consequence of)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(c)........................................................................</td>
</tr>
<tr>
<td></td>
<td>due to (or as a consequence of)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(d)........................................................................</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>II</strong></td>
<td></td>
</tr>
<tr>
<td>Other significant conditions</td>
<td></td>
</tr>
<tr>
<td>Contributing to the death, but not related to the disease or condition causing it</td>
<td></td>
</tr>
</tbody>
</table>

* This does not mean the mode of dying—e.g., heart failure, respiratory failure. It
(otherwise there would have been too many cases). The National Tuberculosis Association's (1955) edition of *Diagnostic Standards and Classification of Tuberculosis* notes that new laboratory tests had made it more difficult to decide whether a particular case of TB was active or inactive—activity could now be seen at sites previously considered inactive, and yet one would not necessarily want to call the "new" active sites cases of TB since they very well may not progress to the point of needing treatment. The committee cites the 1955 version of the book:

> The Committee, however, recognizes the fact that all classifications are ephemeral. They are useful only as long as they serve their purpose. The purpose of a clinical classification of tuberculosis is, however, a most important one. On it depend such matters as legal requirements for isolation, medico-legal considerations with respect to compensation for disability, standards for the return of patients to work, and similar matters. (p. 6)

For another example, the discovery of the lentiviruses led to the description of a new set of disease entities—i.e., slow acting viruses from which one could suffer asymptomatically for extended periods.

In the interests of creating a working infrastructure, Aristotelian principles are deliberately violated:

> C15 Malignant neoplasms of oesophagus

*Note:* Two alternative subclassifications are given:

.0 - .2 by anatomical description

.3 - .5 by thirds

*This departure from the principle that categories should be mutually exclusive is deliberate since both forms of terminology are in use, but the resulting anatomical divisions are not analogous (ICD-10, 1992, vol. 1, p. 190).*

Where the state of the art is unclear, so is the scheme itself:

*Note:* The terms used in categories C82-C85 for non-Hodgkin's lymphomas are those of the Working Formulation, which attempted to find common ground among several major classification schemes. The terms used in these schemes are not given in the Tabular List but appear in the Alphabetical Index; exact equivalence with the terms appearing in the Tabular List is not always possible.

*Includes:* morphology codes M959-M994 with behaviour code /3.

*Excludes:* secondary and unspecified neoplasm of lymph nodes (C77-).


There are several specialty-based adaptations of the *ICD* originating in different national or international bodies (dermatology, stemming from the British Association of Dermatologists, and, under development, rheumatology and orthopaedics from the International League against Rheumatism) *(ICD-10, 1992, vol. 2, pp. 5-6).*

The *ICD* is also directly responsive to changes in the world. Diseases
themselves die (smallpox), are superseded (Gay-Related Immune Disorder becomes AIDS), are newly born (radiation sickness with the discovery of radium), or fall into disrepute (hysteria or neurasthenia). Since this is a statistical classification, a disease with no incidence is of no interest. Thus smallpox was still well defined within *ICD-9-CM* (1996):

\[050\] **Smallpox**

Excludes: arthropod-borne viral diseases (060.0-066.9)

- Boston exanthem (048)
- Hemorragic (pustular) smallpox Malignant smallpox Purpura variolosa
- Alastrim
- Variola minor
- Modified smallpox
- Varioloid


By the time *ICD-10* was developed, this had collapsed into “B03 Smallpox” with a footnote: “In 1980 the 33rd world Health Assembly declared that smallpox had been eradicated. The classification is maintained for surveillance purposes” (*ICD-10*, 1992, vol. 1, p. 150). Or again, malnutrition is defined in relativistic fashion—as the population changes so does the definition:

The degree of malnutrition is usually measured in terms of weight, expressed in standard deviations from the mean of the relevant reference population. When one or more previous measurements are available, lack of weight gain in children, or evidence of weight loss in children or adults, is usually indicative of malnutrition. When only one measurement is available, the diagnosis is based on probabilities and is not definitive without other clinical or laboratory tests. In the exceptional circumstances that no measurement of weight is available, reliance should be placed on clinical evidence. (*ICD-10*, 1992, vol. 1, p. 290)

In these cases, then, the fact that the world is changing is reflected directly in the classification scheme. Another source for this recognition is of course the development of accident categories that also display in their explanations a historical cultural specificity. For example, this set of accident categories describes a series of tumbles more common in the industrial world than for a nomadic tribe:

- **E884** Other fall from one level to another
- **E884.0** Fall from playground equipment
  Excludes: recreational machinery (**E919.8**)
- **E884.1** Fall from cliff
- **E884.2** Fall from chair
E884.3  Fall from wheelchair
E884.4  Fall from bed
E884.5  Fall from other furniture
E884.6  Fall from commode
       Toilet
E884.9  Other fall from one level to another
       Fall from: Fall from:
       embankment  stationary vehicle

There is a relatively impoverished vocabulary for talking about natural accidents—the ICD is richest in its description of ways of dying in developed countries at this moment in history. It is not that other accidents and diseases cannot be described, but these cannot be described as well. Differentiating insect and snake bites, for example, is very important for those living in the rural tropics. However, while arthropods, centipedes, and chiggers are singled out under “bites” in the ICD index, snakes are only divided into venomous and nonvenomous, as are spiders. Clearly this makes sense to some extent, given that this is a pragmatic classification. There is only a point in making fine distinctions between types of accident if those distinctions might make a difference in practice to some agency—medical or other. Simultaneously, those agencies have traditionally been more accountable to Western allopathic medicine and to the industrial world than to traditional systems.

So the ICD bears traces of its history as a tool used by public health officials in developed countries. It also reflects changes in the world at large—either the eradication of diseases or culturally charged changing understandings of certain conditions. Further, it is very much an entrenched scheme. There is a natural reluctance to make changes since each renders a previous set of statistics incomparable and hence less useful.

The first and last entries in the ICD describe a sociotechnical trajectory. The first disease in the ICD over the years has been cholera, unsurprising since cholera was the issue that, in the 1850s, brought participants to the table in an attempt to deal with this international threat. As we noted in the introduction, this threat was exacerbated by the development of steamship technology, which allowed cholera sufferers to carry the disease further before dying. The last condition given in the book takes us to the other end of the sociotechnical arc—i.e., the creation of cyborgs. The last condition listed in the ICD is Z99, “Dependence on enabling machines and devices, not elsewhere classified,” with the very last entry, Z99.9, being “Dependence on unspecified enabling machine and device” (ICD-10, 1992, vol. 1, p. 1175). By some standard, we all now qualify for the Z99.9 condition.
The original sequence produced by William Farr (1885) is reproduced in the latest ICD:

The ICD is a variable-axis classification. The structure has developed out of that proposed by William Farr in the early days of international discussions on classification structures. His scheme was that, for practical epidemiological purposes, statistical data on diseases should be grouped in the following way:

- epidemic diseases
- constitutional or general diseases
- local diseases arranged by site
- developmental diseases
- injuries. (p. 232)

This pattern can be identified in the chapters of *ICD-10* (1992). It has stood the test of time and, though in some ways arbitrary, is still regarded as a more useful structure for general epidemiological purposes than any of the alternatives tested (*ICD-10*, 1992, vol. 1, p. 13).

This classification scheme, then, makes no exaggerated claims to timeless truth. On the contrary, its designers have attempted to paint a fluid picture of the world of disease—one which is sensitive to changes in the world, to sociotechnical conditions, and to the work practices of statisticians.

**There are Many Aids to Storytelling in the ICD**

The classification system that is the *ICD* does more than provide a series of boxes in which to place diseases; it also encapsulates a series of stories, which are the preferred narratives of the *ICD*’s designers. Certain attributions of intentionality are easy to make, others are rather difficult. Some ways of life are clearly considered to be well led, others are called into question. Sometimes context is important, sometimes it can be ignored. Stories also come and go, narratives fade in importance (viz. the example of AIDS moving, in medical terms, from a specifically gay-linked disease to a more general one). If one should have doubts about how to encode a given story, one can turn to volume 2 (*ICD-10*, 1992) of the classification, which gives an extensive set of rules for the interpretation of causes of death. In this section, we will observe the various aids to storytelling to be found within the ICD.

*The Setting*

Frequently, when diseases were first named, they took on the name of their first scientific describer, of a famous victim, or of the place where they occur. Each of these kinds of naming strategy tells a simple story to accompany the classification. Throughout the history of classification systems over the past 200 years, such specifications have progressively been winnowed away to make way for new kinds of context and new kinds of description now considered more interesting and relevant.
What is known by many sufferers as Amyotrophic Lateral Sclerosis (Lou Gehrig's Disease) is coded by the *ICD-10* (1992) as G12.2: Motor Neuron disease (p. 398). (With the famous physicist Stephen Hawking now suffering from the disease, it may in future be more well known to the lay public as Hawking's Disease, as baseball player Lou Gehrig brought it to public awareness the first time.) In the index to the *ICD*, the Parisian neurologist Charcot can lay claim to an arthropathy (tabetic), a cirrhosis, a disease (tabetic arthropathy), and a syndrome. In the body of the text, Charcot's name tends to slip away—i.e., Charcot's syndrome becomes "173.9 Peripheral vascular disease, unspecified"; there is no mention of Charcot (p. 504). The 173s (Other peripheral vascular diseases) are an interesting category. They show the various forms of modality (173.0 is still proudly "Raynaud's syndrome," 173.1 is "thromboangiitis obliterans [Buerger]" (p. 503), 173.8 is "Other specified peripheral vascular diseases" and includes "Acroparaesthesia—i.e., simple (Schultze's type) or vasomotor (Nothnagel's type)" (p. 504). In general, as the modalities get deleted, the name of the person goes from being the name of the disease to a bracket after the name, to an entry in the index, until finally it slides gracefully out of the index onto the scrap heap of history. A similar process occurs with deletion of detail and the uncertainties of discovery in any scientific publication, as Latour and Woolgar (1979) noted in their classic *Laboratory Life*.

Places follow a similar path to abstraction and formal representation. The ideal *ICD* disease is not tied to a particular spot. It is rather identified with a particular causal agent. However, up to and including *ICD-9-CM* (1996), leishmaniasis was a classification that told a travelers' tale—i.e., not only do we know what you got sick of but where you got sick:

```
085  Leishmaniasis
085.0 Visceral [kalaazar]
     Dundum fever Leishmaniasis:
     Infection by Leishmania:  dermal, post-kalaazar
donovani  Mediterranean
     infantum  visceral (Indian)
085.1 Cutaneous, urban
     Aleppo boil                   Leishmaniasis,
     Baghdad boil                  cutaneous:
     Delhi boil                    dry form
     Infection by Leishmania       late
     tropica (minor)               recurrent
     Ulcerating
     Oriental sore

085.2 Cutaneous, Asian desert
     Infection by Leishmania tropica major
```
Leishmaniasis, cutaneous:
  Acute necrotizing
  Rural
  Wet form
  Zoonotic form

085.3 Cutaneous, Ethiopian
  Infection by Leishmania ethiopica

Leishmaniasis, cutaneous:
  Diffuse
  Lepromatous

085.4 Cutaneous, American
  Chiclero ulcer
  Infection by Leishmania mexicana

Leishmaniasis tegumentaria diffusa

085.5 Mucocutaneous (American)
  Espundia
  Infection by Leishmania braziliensis
  Uta

Leishmaniasis, unspecified (ICD-9-CM, 1996, p. 16)

Similarly, for ICD-10 (1992), we can still find Delhi boil in the index, but the main entry itself is a svelte:

B55   Leishmaniasis
B55.0 Visceral leishmaniasis
  Kala-azar
  Post-kala-azar dermal leishmaniasis

B55.1 Cutaneous leishmaniasis
B55.2 Mucocutaneous leishmaniasis
B55.9 Leishmaniasis, unspecified (ICD-10, 1992, vol. 1, p. 166)

So we go from primacy being given to a place (Baghdad boil) to primacy being given to a kind of place (urban cutaneous) to primacy given to a universal agent. Gradually the narrative of travel inscribed in the patient’s code, present earlier, is deleted.

The loss of eponymy and place markers can, of course, be read as a story of the advance of science—the replacement of the local and specific with the general; the thing with the kind; the mutable immobile with the immutable mobile; and the concrete instance with the formal abstraction. However, another line of argument also deserves attention. As we have already seen, the ICD also reflects historical states of the world. The world has changed. With the huge increase in international travel over the past century and a half, it is more rare for a disease to be tied to a particular location—diseases themselves tend to spread to kinds of location. The malaria map of the world hanging on the wall at the WHO headquarters
in Geneva shows the expected tropical venues—and small red circles around major airports—as mosquitoes are transported from the tropics. We are, as a world, becoming more “abstract” in this way.

Similarly, research now is not attributed to single great figures who can claim sole responsibility for a discovery. Medical work was always done in teams, but these have become larger, involving complex social and institutional relationships of attribution as Gallo and Montaignier would be the first to remind us (in Grmek, 1990). A typical scientific article has so many authors that the death of the individual scientific author seems certain. In general, the ICD has gone from being the holder of a set of stories about places visited, heroic sufferers, and great doctors to holding another set of stories.

The Context of Disease

As people and places have moved out of eponymous and loconymous classification, they are replaced by a general set of categories—what we are calling the kindness of strangers. By this we mean that the classification system indicates a shift from our being individuals experiencing the world to our being kinds of people experiencing kinds of places. The constructions of social and natural science and of the legal world have moved in. Broken legs and ski resort locations co-evolve as do cancer rates and toxic waste dumps. The classification system, as we shall see in this section, has become a site which holds these constructions together and, through excluding other kinds of story, makes them more real. With the ICD providing the main legitimate means for describing illness, the social, economic, and political stories woven into its fabric become, by extension, the main legitimate narrative threads for the science of medicine.

Although particular places have moved out, two places have come to play a more significant role in the classification system—i.e., the laboratory and the “sociological home.” The latter appears in the extra categories developed for ICD-9 as supplemental codes, which in ICD-10 (1992) have become fully integrated—what we might call the context codes. Thus housing is one of the conditions that can be broken down and described as part of the classification. In ICD-9-CM (1996) it is described as follows:

- **V60** Housing, household and economic circumstances
  - **V60.0** Lack of housing
    - Hobos
    - Social migrants
    - Tramps
  - **V60.1** Inadequate housing
    - Lack of heating
    - Restriction of space
    - Technical defects in home preventing adequate care
V60.2 Inadequate material resources
   Economic problem  Poverty NOS
V60.3 Person living alone
V60.6 Person living in residential institution
   Boarding school resident
V60.8 Other specified housing or economic circumstances
V60.9 Unspecified housing or economic circumstances. (vol. 1, p. 267).

The related code in *ICD-10* (1992) is expanded to include discord with neighbors and lack of adequate food (vol. 1, p. 1152). In both, the name of the city gives way to the name of the social category and social condition.

These context codes define what is considered to be medically relevant in one's material surroundings. They make it easy to structure studies in these terms (e.g., what effect does poor housing have on the incidence of tuberculosis). Simultaneously, they do make it much more difficult to deal with unrecognized contexts (what effect does conspicuous consumption have on cholesterol levels?). It is not impossible to do these latter studies, but the information is not "to hand" in the way that it is for medically sanctioned contexts. The reason for stressing this point is that it can be taken as a sign of the correctness of allopathic medicine that it has isolated the basic variables that must be taken into account in the development of public health policy or medical science. However, it is important to note that, although the *ICD* is a powerful tool, in this sense it also, as infrastructure, enforces a certain understanding of context, place, and time; it makes a certain set of discoveries (which validate its own framework) much more likely than an alternative set outside of the framework (since the economic cost of producing a study outside of the framework of normal data collection is necessarily much higher).

This sort of convergence is an important feature of large-scale networked information systems. Star, Bowker, and Neumann (In press) define convergence as:

*Convergence.* . . is the double process by which information artifacts and social worlds are fitted to each other and come together. On the one hand, a given information artifact (a classification system, a database, an interface, and so forth) is partially constitutive of some social world. That is to say, the sharing of information resources and tools is a dimension of any coherent social world—be it the world of homeless people in Los Angeles sharing survival knowledge via street gossip, or the world of high-energy physicists sharing electronic preprints via the Los Alamos archive. On the other hand, any given social world itself generates many interlinked information artifacts. The social world creates through bricolage, a (loosely coupled but relatively coherent) set of information resources and tools. People without houses also log into the Internet, and physicists indulge in street gossip at conferences—as well as engage in a whole set of other information practices. Put briefly, information artifacts undergird
social worlds, and social worlds undergird these same information resources. We will use the concept of convergence to describe this process of mutual constitution.

With these processes of convergence, the site of the medical work itself has gained in importance. The classification of tuberculosis, canonically difficult to diagnose accurately (compare Latour, In press), retains the story of what has been done in the laboratory as well as what has occurred in the body.

A15  *Respiratory tuberculosis, bacteriologically and histologically confirmed*

A15.0  *Tuberculosis of lung, confirmed by sputum microscopy with or without culture*

  *Tuberculous:*

  bronchiectasis /
  fibrosis of lung /
  pneumonia confirmed by sputum microscopy with or without culture
  pneumothorax /

A15.1  *Tuberculosis of lung, confirmed by culture only*

  Conditions listed in A15.0, confirmed by culture only

A15.2  *Tuberculosis of lung, confirmed histologically*

  Conditions listed in A15.0, confirmed histologically

A15.3  *Tuberculosis of lung, confirmed by unspecified means*

  Conditions listed in A15.0, confirmed but unspecified whether bacteriologically or histologically. (*ICD-10, 1992, vol. 1, p. 113*)

In this case, the disease itself is always classified in terms of the work that has been done in the medical laboratory. Again, as new technologies are invented, historical shifts occur, as with the relationship between epilepsy and the EEG machine as diagnostic many decades ago.

The doctors themselves enter the story at the moment of classification, the patient rarely does. This comes out clearly if we compare migraine and epilepsy in *ICD-9-CM* (1996). Epilepsy is a condition that is defined by the doctor in the context of laboratory and so is a real condition:

345  *Epilepsy*

The following fifth-digit subclassification is for use with categories 3450, 1..4.9:

0  without mention of intractable epilepsy
1  with intractable epilepsy (*ICD-9-CM, 1996, vol. 1, p. 80*)

So here the question is whether or not the patient objectively has intractable epilepsy in the opinion of the doctor. The determination of
intractable migraine relies on the voice of the patient and so is marked as a suspicious designation:

346 Migraine

The following fifth-digit subclassification is for use with category 346:

- 0 without mention of intractable migraine
- 1 with intractable migraine, so stated (ICD-9-CM, vol. 1, p. 80)

The laboratory context then is the "real" context of the disease; the classification serves to reinforce the separation of the patient from ownership of their condition. We should note at this point that we are not arguing that this makes the ICD a tool for evil and oppression. On the contrary, what we are trying to do is work out what kind of a tool it is—i.e., what work it does and whose voice appears in the unfolding narrative.

The legal context is often enfolded into the classification system. Thus the classification of blindness considers the American system of medical benefits:

369 Blindness and low vision

Note: visual impairment refers to a functional limitation of the eye (e.g., limited visual acuity or visual field). It should be distinguished from visual disability, indicating a limitation of the abilities of the individual (e.g., limited reading skills, vocational skills), and from visual handicap, indicating a limitation of personal and socioeconomic independence (e.g., limited ability, limited employment).


Note that definitions of blindness vary in different settings. For international reporting WHO defines blindness as profound impairment. This definition can be applied to blindness of one eye (369.1, 369.6) and to blindness of the individual (369.0).

For determination of benefits in the USA, the definition of legal blindness as severe impairment is often used. This definition applies to blindness of the individual only.

369.0 Profound impairment, both eyes

369.00 Impairment level not further specified

Blindness:

- NOS according to WHO definition
  - both eyes

369.3 Unqualified visual loss, both eyes

Excludes: blindness NOS:

- legal [USA definition] (369.4)
WHO definition (369.00)

369.4 Legal blindness, as defined in USA

Blindness NOS according to USA definition


Note in the above example that "blindness of the individual" might be psychogenic—i.e., due to brain damage or other organic cause outside the eye itself. The problem of localized versus "whole organism" conditions forms a serious source of coding problems. For example, depending on one's theory of cancer, it would be an immune disorder affecting the whole person or a localized phenomenon to be surgically removed with many gray areas in between for the different types of cancer.

In the example above, the legal definition can take precedence over the cultural and social. Thus cannabis dependence has its own category, while the culturally profoundly different absinthe and glue addictions are combined:

304.3 Cannabis dependence

Hashish Marihuana

Hemp

304.6 Other specified drug dependence

Absinthe addiction Glue sniffing


Few would argue that glue sniffing and absinthe addiction are similar phenomena. The former leads to more serious physical conditions than "cannabis dependence" (a category many would challenge) and yet does not rate its own category. Absinthe addiction is, one suspects, a hangover from earlier days. Because the origins of the ICD were French and absinthe abuse an important problem in Paris in the nineteenth century, it persists. These accidents of history, practice, and crime contain many clues to re-narrativizing the ICD. E970 to E979 in ICD-9-CM (1996) is an interesting set; it covers injuries caused by legal interventions:

Legal Intervention:

Includes: injuries inflicted by the police or other law-enforcing agents, including military on duty, in the course of arresting or attempting to arrest lawbreakers, suppressing disturbance, maintaining order and other legal action

Excludes: injuries caused by civil insurrections (E990.0-E999) (ICD-9-CM, vol. 1, p. 304)

This set includes state executions. Note that civil insurrections, where the definition of legal intervention is on the table, are classified together with
war. The definition of legal, of course, may be subject to its own retrospective reconstruction, as in the case of Rodney King.

Abortions, which may be to all intents and purposes the same medically, are marked differently in the ICD according to their legality:

635 Legally induced abortion
- Includes: abortion or termination of pregnancy:
  - elective
  - legal
  - therapeutic
- Excludes: menstrual extraction or regulation (V25.3)

636 Illegally induced abortion
- Includes: abortion:
  - criminal
  - illegal

Each type of abortion (spontaneous or 634, legally induced, illegally induced, unspecified, failed attempted, or 638) has the same set of complications attached—i.e., nine difficulties, each accorded a digit (one of the most closely coded category sets in the ICD). When the issue arises, then, the ICD privileges the voice of the doctor and the laboratory over the voice of the patient and legal discourse over cultural and social discourse. One can read another order of social history from the nature of the silences in the story as well.

In general, the ICD carries with it its own context. This is a common feature of classification systems. One way of reading these is that they provide a stabilizing force between the natural and the social worlds. They hold in place sets of arrangements that allow one to read the natural as stable and objective and the social as tightly linked to it. For the ICD, this means describing disease in a way that folds the socially and legally contingent into the classification system itself and so naturalizes it. Inversely, the disease entity out there in the world is brought into the laboratory where the social and organizational work of its stabilization can best be guaranteed.

**Cutting Up the World**

In order to tell stories of the sort with which we are most familiar, one needs objects in the world that can be cut up spatially (compare Berg & Bowker, 1997) and temporally into recognizable units. Narrative structures are typically formed with a moving timeline, protagonists, and a dramatic structure unfolding over time. The ICD does in fact operate this kind of dissection, which we will discuss later. In the last section, we saw the constitution of a context within the ICD; in this section we will see the constitution of actants to populate that context and those stories.
Time Story 1: The Life Cycle

Temporally, the classification system provides a picture of acute (temporally bound) episodes within an otherwise well-ordered life. It is notoriously bad for describing chronic diseases; the interest is on the episode of treatment (Musen, 1992). Let us go through some temporal units presented by the ICD. Birth is extremely important and is very closely defined:

Live birth is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered liveborn. (ICD-10, 1992, vol. 2, p. 129).

Time flows very quickly for the newborn, and so temporal units vary accordingly: The neonatal period commences at birth and ends 28 completed days after birth. Neonatal deaths (deaths among live births during the first 28 completed days of life) may be subdivided into early neonatal deaths, occurring during the first seven days of life, and late neonatal deaths, occurring after the seventh day but before 28 completed days of life.

Age at death during the first day of life (day zero) should be recorded in units of completed minutes or hours of life. For the second (day 1), third (day 2) and through 27 completed days of life, age at death should be recorded in days. (ICD-10, 1992, vol. 2, p. 131)

Given the bump in mortality that occurs around birth, this is not surprising. When adult life begins, things start to slow down. Adults are defined in ICD-9-CM (1996, p. xiii) as people between 15 and 124 years old. If you make it to 125, you are "hors de categorie."

In the middle period, adulthood, there are some indications of what constitutes a good life. It should be well-ordered and rhythmic. Things should happen at the right time. Thus sexual development has its own timing:

259 Other Endocrine disorders
259.0 Delay in sexual development and puberty, not elsewhere classified
Delayed puberty
259.1 Precocious sexual development and puberty, not elsewhere classified PED
Sexual precocity:
NOS
constitutional
cryptogenic

Similarly, problems with temporal regulation of menstruation are well-
defined—i.e., too early, too late, too frequent, not frequent enough. Natural rhythms should not be upset.

A relatively recent temporal problem addition here is jet lag:

307.45 Phase-shift disruption of 24-hour sleep-wake cycle
Irregular sleep-wake rhythm, nonorganic origin
Jet lag syndrome
Rapid time-zone change

The reference to the "nonorganic origin" highlights that this is a situation-bound condition: the context (jet travel or night shift work) is directly folded into the disease.

To an outside observer, there is remarkably little reference to the process of aging. An adult is a timeless being who should be healthy; disease is not, in general, indexed by age. Further, the body is not present as something that gets used up and worn out; such stories have to be superadded (indeed the category of being "worn out" was in earlier additions of the ICD but has since been removed).

If you rent a house, your agreement with the landlord includes a "fair use" or "normal wear and tear" category; it is expected that the house depreciates over time and this is written into the legal code. There are only two references to normal wear and tear in the whole ICD. First, one can, as an adult, step out of the well-ordered life and suffer from premature or delayed senility, puberty, birth, and aging. Among the conditions under "delay" are delayed birth, development (including intellectual, learning, reading, sexual, speech, and spelling), menstruation, and puberty. In this case, the cycle structure is the same, but the patient is taking the steps too early or too late. Second—and there is only one example of this—you could use your body badly. The only specific instance of this, however, is that you can grind or otherwise mismanage your teeth:

521 Diseases of hard tissues of teeth

521.1 Excessive attrition
Approximal wear. Occlusal wear (ICD-9, vol. 1, p. 125)

In ICD-10, abrasion of teeth carries with it an illuminating set of contexts: dentifrice, habitual, occupational, ritual, and traditional. Occupational abrasion in earlier times included the hazard "tailor's tooth," for example, where the teeth were abraded due to biting off the thread for hand sewing. In principle, the timeless adult could do many things excessively. There are categories for excessive thirst, secretion, salivation, sex drive, sweating, and binocular convergence among others. However, that superfluity is, only in this one case, indexed against an aging body. Note
that there are, of course, diseases associated more broadly and often implicitly with excessive wear and tear—e.g., cirrhosis of the liver associated with alcoholic excess. But here we are concerned directly with the representation in the classification system.

This curious invisibility of aging as wear and tear is one way in which the ICD stabilizes context and disease entity—the human body as the substrate of both is outside the flow of time. The human adult body becomes the unmarked category—i.e., the cipher against which laboratory, social, and natural time can be coordinated. Indeed we could go one step further and see the adult male body as the unmarked category—since there are many more diseases restricted to women than restricted to men; there are sixteen categories or clusters of categories that apply only to males and forty-two that apply only to females (ICD-10, 1992, vol. 2, p. 26). Feminist critics of medicine have long remarked on the relative pathologizing of the female body (Ehrenreich & English, 1973).

Nobody Dies of Old Age. To finish with the life cycle before moving on to other temporal features, we should note that death itself is remarkably poorly defined in comparison to life. One can scarcely die of old age (Fagot-Largeault, 1989). The closest that one may get comes under a banner disclaimer:

*ILL-DEFINED AND UNKNOWN CAUSE OF MORBIDITY AND MORTALITY
(797-799)

797  S en i l i ty without mention of psychosis

Old age

Senile:
Senescent debility
Senile asthenia exhaustion


The ICD's life cycle for humans, then, is as follows: a spurt of intense activity at birth; timeless adulthood, when one is afflicted with a range of woes that carry their own temporalities (more on this anon); and an inglorious ill-defined end. The effect of this is, paradoxically, to make the individual an undefined tabula rasa onto which various diseases are inscribed. From this blank sheet, one can read various stories (with the aid of the ICD), restoring first context and then interpretation (which we shall deal with in the next section).

Time Story 2: The Virus

Diseases themselves change over time. HIV, for example, mutates rapidly in the individual so that no two people suffer from the "same" disease nor may the disease be the same even within a person. This extreme variability of the object world is a problem for any classification
system. The case of virus classification illuminates many features of categorizing difficulties and the strategies used to control them. We look here at some of the work of the International Committee on Taxonomy of Viruses (ICTV) (Murphy et al., 1995) so as to see how diseases that are present differently in every individual, and often vertiginously mutate, can be usefully classified.

Throughout the history of virology, there have been acerbic debates over just what are viruses. The great virologist Lwoff, echoing Gertrude Stein no doubt, declaimed in 1953 that "viruses should be considered as viruses because viruses are viruses" (Matthews, 1983, p. 7). Viruses themselves have moved from scientific category to scientific category. In the early twentieth century, the central definition of a virus was entirely negative. As Waterson and Wilkinson note, a virus was any disease organism which could be filtered through one of the "filter candles" developed for the purpose. This was a useful definition in that it excluded all other known disease agents; however, it did not guarantee the homogeneity of the category itself as Andrewes noted in 1930 when describing animal viruses:

judgment must be suspended . . . in the case of the invisible viruses or so-called "filter-passing" organisms. Here our ignorance is almost complete; they are possibly a heterogeneous group but in the case of creatures which we cannot see and whose very existence is, in many cases, a matter of inference only, it is idle to talk of classification in the usual sense. (Matthews, 1983, p. 4)

So there was no one definition, or rather, the ultimate encompassing residual category. Here be dragons.

Equally, there was no one discipline studying the matter of virus classification. There was no study of virology per se until the 1980s. There was an a priori assumption, entrenched in disciplinary specialties, that animal and plant viruses were not the same. This was disproved in the 1940s when it was shown that some plant viruses could also affect insects (Matthews, 1983, p. 7). Groups that were not used to working together were forced to cooperate, and they did not necessarily like it. As with the numerous and passionate battles between cladistics and numerical taxonomy in biology (Duncan & Stuessy, 1984), there was a series of strong virological arguments that have left their traces in the literature. The arguments can be read in two ways. They are simultaneously about a struggle for professional authority on the part of the various disciplines involved and an attempt to find a single language with which to talk about the complex temporal and spatial properties of viruses.

The role of the classification systems in knitting together (or not) the specialties is clear in all accounts of virus taxonomy. Matthews (1983) notes, "in the period 1966 to 1970 there was considerable controversy regarding some of the rules, which developed into a serious rift between
most of the plant virologists, and some animal virologists" (p. 13). He comments on Fenner's presidency of the ICTV in the period 1970-1976:

In retrospect perhaps the major contribution made by Fenner during his Presidency was to keep the plant virologists working within the ICTV organization. This really meant stopping the insistence of Lwoff's supporters on a hierarchical classification and Latinized binomials, and also, as noted earlier, deleting the rule regarding new sigla. In addition, Fenner exerted pressure to ensure that following two vertebrate virologists, a plant virologist should be the next President of the ICTV. (p. 20)

Murphy et al. (1995) note that even today: "Virus taxonomy is a polarizing subject when it comes up in hallway conversations." They go on to praise the ICTV for its work of

true international consensus building, and true pragmatism—and it has been successful. The work of the Committee has been published in a series of reports, the Reports of the International Committee on Taxonomy of Viruses, The Classification and Nomenclature of Viruses. These Reports have become part of the history and infrastructure of modern virology. (p. v)

We see then that the development of the classification system is an occasion for the construction of the community for which that system will act as information infrastructure. The system is built as a political compromise between specialties. The kinds of truth and the kinds of stories that it can contain by their nature recognize this. As Murphy et al. (1995) state, the resulting classification system is in some sense arbitrary:

Today, there is a sense that a significant fraction of all existing viruses of humans, domestic animals and economically important plants have already been isolated and entered into the taxonomic system. . . . [The] present universal system of virus taxonomy is useful and usable. It is set arbitrarily at hierarchical levels of order, family, subfamily, genus, and species. Lower hierarchical levels, such as subspecies, strain, variant, etc., are established by international specialty groups and by culture collections. (Murphy et al., 1995, p. 2)

The apposition of specialty groups (professionalization work) and culture collections (naturalization work) is unsurprising; Murphy et al. (1995) offer it in a different form later in the same work: "Unambiguous virus identification is a major virtue of the universal system of taxonomy . . . and of particular value when the editor of a journal requires precise naming of viruses cited in a publication" (p. 7).

Thus a first temporality associated with viruses is that the field itself has formed and changed rapidly, much like the organisms that it studies. This is an unsurprising echo, as the fact that the viruses transgress spatial boundaries and mutate enormously rapidly has contributed to the change.

So what is the problem with correlating virus time with laboratory time? The overwhelming difficulty has been that it is extremely difficult
for viruses to produce the kind of "genetic classification" whose genealogy Patrick Tort (1989) has so brilliantly traced across the social and natural sciences of the nineteenth century. A genetic classification is one that classifies things according to their origins—rocks might be metamorphic or sedimentary; languages might be Indo-European or Nilotic. Viruses have multiple possible origins—i.e., they look and feel the same since they pass the filter test and make you sick, but they got that way along multiple paths. This is an old problem in medical philosophy and diagnosis—a cure does not necessarily reflect a cause, and there may be many paths to a single symptom.

Ward (1993) gives four theories for viral origins. First, it is possible that some viruses "evolved from autonomous, self-replicating host cell molecules such as plasmids or transposons, by acquiring appropriate genes that code for packaging proteins" (p. 433). In this picture, they are simple chemical combinations that have acquired the replication habit of their material substrate. Second, "some viruses arose by degeneration from primitive cells in a manner similar to that proposed for the evolution of cellular organelles such as mitochondria and chloroplasts from bacteria" (p. 434). Here they are complex organisms that devolved. Third, "some RNA viruses are descendants of prebiotic RNA polymers" (p. 433). According to this theory, viruses might have co-evolved with life itself. Finally, there is the possibility that "some viruses evolved from viroids or virusoids, although it is equally possible that these small RNA, rather than being progenitors of viruses, are recent degenerative products of the more complex self-replicating systems" (p. 434). Where you do not have a single origin story, you cannot have a single biological classification system. Viruses have been classed into families and then into increasingly controversial supervenient categories (only one "order"—the Mononegavirales—has been approved to date by the ICTV). The supervenient categories frequently have the inconvenience of separating viruses that had been considered grouped together. With the lack of a single origin, the central class of virus "species" has been defined: "A virus species is a polythetic class of viruses constituting a replicating lineage and occupying a particular ecological niche" (Van Regenmortel, 1990, p. 49).

A "polythetic" class is a class that is defined by the congruence of multiple characteristics no one of which is essential. The attribution of occupation of a particular niche is essential for dealing with obligate parasites. This relatively loose definition opens up a space for the professionalization work that needs to be done in conjunction with the alignment of competing temporalities (of the virus and of the laboratory). There has, in recent years, developed a line of argument that with genome sequencing it will be possible to produce a coherent history of viruses that will make the species concept more historically accurate. This reflects a wider trend across many social and natural sciences to recover
the origin—in geology, the tide has turned against uniformitarianism (Allegre, 1992); in philosophy, Foucault’s archeology has grown up in opposition to the postmodern denial of origins. However, even today a strictly genetic classification of viruses is possibly leading to category death:

if mammalian viruses are descended from mammals, snake viruses from snakes, and honeybee viruses from honeybees, the group “virus” would cease to have any formal classificatory validity. It could be retained as a nonclassificatory group, analogous to the group of “animals with wings,” but if it is not a monophyletic group, there is no doubt how cladism would deal with it; it presents no philosophical difficulty: the taxonomic category “virus” should be exploded. (Ridley, 1986, p. 51)

The demotion to a nonclassificatory group would again have professional consequences.

We see with the history of virus classification, then, that there has been a deliberate effort to create something that looks and feels like other biological classifications, even though the virus itself transgresses basic categories (it jumps across hosts of different kinds, steals from its host, mutates rapidly, and so forth). This has been somewhat of a deliberate political decision on the part of the international virus community: you need such classification systems in order to write scientific papers, provide keywords for indexing and abstracting, and compare results. Even in this most phenomenologically difficult of cases, the world must still be dissected into recognizable temporal and spatial units—partly because that is the way the world is and partly because that is the only way that science as we know it can work.

STORIES OF CARVING UP THE BODY: THE VERMILION BORDER OF THE LIP

In Regions of the Mind, Leigh Star (1989) examined the ways in which researchers, seeking to localize cerebral functions, cut up the brain into meaningful units. The process is a messy one since brains themselves come in many shapes and sizes. During the early days of research, a diagram of a “typical” monkey brain, with minutely localized and labeled regions, is transposed onto a representation of a human brain in an attempt to produce a standardized diagram. Human brains are of a much different size than monkey brains. Nevertheless, the need for standardized representations was so urgent that the physiologists overlooked this source of uncertainty, among others (Star, 1985). A similar problem occurs with the dissecting of bodies for medical purposes. Stefan Hirschauer (1991) has noted this for the practice of the surgeon’s trade; Berg and Bowker (1997) have discussed the same phenomenon with respect to the development of medical records.

The ICD bears traces of this sort of uncertainty, most notably at liminal sites (those with borders that are unclear or are used in several different
categories), and with respect to roving categories like neoplasms (the cancer may overlap the ICD categories). We can use the vermilion border of the lip, also known as the "lipstick area," as a tracer for this. An early appearance in ICD-9-CM is as follows:

4. Malignant neoplasms overlapping site boundaries
Categories 140-195 are for the classification of primary malignant neoplasms according to their point of origin. A malignant neoplasm that overlaps two or more subcategories within a three-digit rubric and whose point of origin cannot be determined should be classified to the subcategory .8 "Other." For example, "carcinoma involving tip and ventral surface of tongue" should be assigned to 141.8. On the other hand, "carcinoma of tip of tongue, extending to involve the ventral surface" should be coded to 141.2, as the point of origin, the tip, is known. Three subcategories (149.8, 159.8, 165.8) have been provided for malignant neoplasms that overlap the boundaries of three-digit rubrics within certain systems. Overlapping malignant neoplasms that cannot be classified as indicated above should be assigned to the appropriate subdivision of category 195 (Malignant neoplasm of other and ill-defined sites).

140.0 Upper lip, vermilion border
Upper lip:
NOS
external

The "NOS" in this classification stands for "not otherwise specified"—a protean modifier throughout the classification.

If we consider ICD as a prototype classification system, we can see the way of treating the vermilion border as part of a general strategy of distinguishing central members of certain categories from outliers. The vermilion border is strictu sensu part of the skin of the lip, but it is not a good member of that category:

173.0 Skin of lip
Excludes: vermilion border of lip (140.0-140.1, 140.9) (ICD-9-CM, 1996, vol. 1, p. 32)

Equally, it is definitely skin but is a special subcategory:

238.2 Skin
Excludes: anus NOS (235.5)

skin of genital organs (236.3, 236.6)


Or again, it is definitely soft tissue but is an outlier:

239.2 Bone, soft tissue, and skin
Excludes: ...


In *ICD-10*, its marginality is explicit:

**D00.0 Lip, oral cavity and pharynx**

*Aryepiglottic fold:*

**NOS**

*Hypopharyngeal aspect*

*Marginal zone*


This multiple reference to the vermilion border of the lip is a typical *ICD* naming strategy. If a region of the body might fall under several categories, its membership in a special category is explicitly marked.

In principle at least, the world itself—that messy, sprawling, sociotechnical system—should be divided into regions of relevant causal occurrence. The *ICD*’s work here is necessarily far from complete. Here, however, is one typically precise definition of a liminal zone in the outside world:

*A public highway (trafficway) or street is the entire width between property lines (or other boundary lines) of every way or place, of which any part is open to the use of the public for purposes of vehicular traffic as a matter of right or custom. A roadway is that part of the public highway designed, improved, and ordinarily used, for vehicular travel.* (*ICD-9*, 1996, vol. 1, p. 274)

The *ICD* records accident statistics, including the place or mode. Such precision is needed for the compilation of, for example, effective safety statistics. This drive for precision is in principle unending—how much of the social and natural worlds would have to be described within the *ICD* in order to produce an exhaustive system?

The point here is not that these are bad definitions of lipstick areas and streets. It is that they are ineluctably arbitrary ways of cutting up the world. The goal with a classification system is to produce homogeneous causal regions. Homogeneous causal regions are zones without effective subdivision. For the vermilion border, there is no real distinction between upper and lower; for streets, there is no real distinction between tarred and gravel. There is no possibility, in principal, that this can be other than a bootstrapping operation. All research work that explores medical causality has the *ICD* or a similar system as its base referent and so necessarily assumes the *ICD*’s set of homogeneous regions in order to design its tests, experiments, or projects. It is analytically always possible to act otherwise and carve the world up differently into other kinds of causal regions. Latour (1987) reminds us of this in *Science in Action*. He posits the thought experiment: How would someone challenge the basic premises of quantum mechanics? No one would deny that it is possible that these premises are wrong nor that an experiment might be designed to prove this. However, the economic and administrative cost of doing so would be huge. Who would fund the proposal? Who would referee the papers? How, in short,
would the inertia of the networks involved be overcome? In the same way, it is always possible (and somewhat more common than in the quantum mechanics case) to challenge basic ICD categories. However, it is in practice much easier to hypostatize and duplicate them for local usage. Exceptions occur when particular categories are linked with social movements and social problems; an outstanding example of this occurred with the de-medicalization of homosexuality in the DSM-3 after challenges from the gay community (Kirk & Kutchins, 1992).

We have seen in this section that medical classifications split up the world into useful categories. They do not describe the world as it is in any simple sense. They necessarily model it. This modeling within classification systems of all sorts is where the real work gets done in terms of the enfolding of social, political, and organizational agendas into the scientific work of describing nature—in this case, in the form of disease entities.

**INTERPRETATION IS ALSO ENFOLDED INTO THE ICD**

We saw in the last section how the ICD divides the world into standard Aristotelian units of time and place and, in doing so, how it produces favored readings of the body and of the world at large. The WHO goes one step further. It not only provides, through the ICD, a set of possible stories, it also provides, bundled up in the classification system, explicit rules for the interpretation of those stories.

In order to follow this through, we need to look at the form of the standard international death certificate (see Figure 2 above). Ann Fagot-Largeault (1989) has done a wonderful philosophical analysis of this document; our own description will not attempt to be as complete. It is the death certificate that constitutes the archetypical use of the ICD—indeed, until ICD-5, the classification only covered causes of mortality and did not seek to represent morbidity. The death certificate itself has as a major heading, “Cause of Death.” It is split into sections, “Cause of Death,” “Approximate interval between onset and death,” and other contributing factors or significant conditions.

It is a difficult task to summarize a complex series of conditions to a single cause of death, and the work of interpretation begins on the form itself. A single cause is favored for very practical reasons. In the first place, it is hard enough to compile statistics at all, and the task could get overwhelming if multiple causes were allowed. Further, a single cause of death provides the lowest common denominator over multiple collection systems—from medical examiners in a large hospital to medical paraprofessionals in the underdeveloped rural areas. Finally, as the ICD’s developers point out, the goal of the classification system is not to describe complex phenomenologies but to prevent death:

> From the standpoint of prevention of death, it is necessary to break the chain of events or to effect a cure at some point. The most effective
public health objective is to prevent the precipitating cause from operating. For the purpose, the underlying cause has been defined as "(a) the disease or injury which initiated the train of morbid events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury." (ICD-10, 1992, vol. 2, p. 31)

This statement revealingly indicates a recognition by the system’s developers that reality is indeed more complex than their registration system can describe. All the analytic points made to date in this discussion can be read into this one statement: the ICD is a pragmatic classification ("the most effective public health objective" [p. 31]), and it divides the world spatially and temporally into causal zones that underwrite preferred stories ("it is necessary to break the chain of events...at some point" [p. 31]).

The cause of death as given on the death certificate by the attending physician is frequently not, as Fagot-Largeault (1989) points out, the cause of death that is entered into the statistical record. The classifications entered on the certificate are themselves systematically re-coded so as to constrain the kinds of story that the statistics tell.

One standard algorithm is that precision always beats no precision (this is an echo of John King’s [personal communication] wonderful observation about technical arguments in the policy domain: “[N]umbers beat no numbers every time”). On a deeper epistemological level, the substitution of precision for validity is often a needed expedient in getting work done (Star, 1989; Kirk & Kutchins, 1992). It may also become a kind of gatekeeping tool in theoretically defining a ground of knowledge. It functions as follows in the ICD:

Where the selected cause describes a condition in general terms and a term that provides more precise information about the site or nature of this condition is reported on the certificate, prefer the more informative term. This rule will often apply when the general term becomes an adjective, qualifying the more precise term.

“Example 57: I (a) Meningitis
Tuberculosis
Code to tuberculous meningitis (A17.0). The conditions are stated in the correct causal relationship.” (ICD-10, 1992, vol. 2, p. 48)

This is doubtless a very reasonable rule. However, it is significant that it sets in motion a process that begins placing mediating layers between what the doctor says and what gets reported.

In general, these mediating layers refashion the story that the act of classification permits. The records clerk is given the license to change the doctor’s classification in such a way that it will reflect the best current medical theories:

Rule 3. If the condition selected by the General Principle or by Rule 1 or Rule 2 is obviously a direct consequence of another reported condition, whether in Part I or Part II, select this primary condition. (ICD-10, 1992, vol. 2, p. 34)
Thus, for example,

[w]here the selected cause is a trivial condition unlikely to cause death and a more serious condition is reported, reselect the underlying cause as if the trivial condition had not been reported. If the death was the result of an adverse reaction to treatment of the trivial condition, select the adverse reaction. (ICD-10, 1992, vol. 2, p. 45)

Derrida (1980) reminds us that it is through what is excluded as trivial that we can frequently understand systems of thought by pointing directly at what is important. Similarly, this opening of the door to an undetermined attribution of triviality is one significant moment, hidden in the third volume of a massive classification system, where the work of reifying current categories is done. Only certain causal chains will be permitted at the moment of classification. This in turn naturally impacts the interpretation at the other end of “raw data” in the form of epidemiological statistics:

The expression “highly improbable” has been used since the Sixth Revision of the ICD to indicate an unacceptable causal relationship. As a guide to the acceptability of sequences in the application of the General Principle and the selection rules, the following relationships should be regarded as “highly improbable.” (ICD-10, 1992, vol. 2, p. 67)

After this passage, there follows a series of unacceptable chains. For example, a malignant neoplasm cannot be reported as “due to” any other disease than HIV; haemophilia cannot be “due to” anything, and no accident can be reported as due to any other cause—except epilepsy (ICD-10, 1992, vol. 2, p. 68).

An acceptable string of classifications in a death certificate is one which fits into an internally consistent chain that reflects current medical knowledge. In the process of arriving at such a chain, all qualifiers should be removed: “Qualifying expressions indicating some doubt as to the accuracy of the diagnosis, such as ‘apparently,’ ‘presumably,’ ‘possibly,’ etc., should be ignored, since entries without such qualification differ only in the degree of certainty of the diagnosis” (ICD-10, 1992, vol. 2, p. 88).

In the process of achieving this certainty, multiple causality often has to be arbitrarily collapsed into unicausality—here by a principle of first come first served:

If several conditions that cannot be coded together are recorded as the “main condition,” and other details on the record point to one of them as the “main condition” for which the patient received care, select that condition. Otherwise select the condition first mentioned. (ICD-10, 1992, vol. 2, p. 106)

Any working classification system will have such standard rules attached. Such rules are theoretically interesting for several reasons. First, the ICD developers have explicitly recognized that it is not enough to control the classification (the name of the disease). They also have to attempt to exercise control over the language game in which the classification is
inserted—this indeed is the purpose of the rules contained in volume 2. This attention to both the base level and its meta-level is a bureaucratic necessity that simultaneously conjures the wild world of the patient's body into the ordered world of medical knowledge. Second, the rules themselves serve to systematically reduce ambiguity and uncertainty, even where these are integral to the attendant physician's depiction of the situation of the patient. Those who see the patients are aware of this uncertainty; those who apply the rules also know of it; those who read the final statistics are shielded from it. The patients live it. Finally, there is of course a potential infinite regress in the control of, first, the name of the disease, then on rules for using these names and so forth. The final level at which regress occurs is in the presentation of results. The WHO recognizes that, when dealing with small populations, you can get wild fluctuations of information on mortality or morbidity from year to year. In order to achieve stability and certainty at this level, one needs to sacrifice precision: to go up to broader ICD rubrics, aggregate data over a longer period, use the broadest of the recommended age groupings and aggregate areas (ICD-10, 1992, vol. 2, p. 137). The regress itself to ever higher levels of control marks the fact that the world is always slightly out of reach—it cannot be contained in the classification system, or the system plus set of rules, or the system plus set of rules for interpretation plus set of rules for change, or the system plus set of rules for interpretation plus set of rules for change plus set of rules for presentation.

**CONCLUSION**

At the start of this discussion, we looked at two basic kinds of classification system—i.e., Aristotelian and prototype. We have seen in the course of our analysis that medical classification systems are "naturally" prototypical and that they nevertheless have to appear Aristotelian in order to bear the bureaucratic burden that is put on them. This burden is to act as a gateway between the world of the laboratory and the hospital (with precisely defined closed environments) and the workaday world. As we consider the stories embedded in the system, from the point of view of work and practice, we understand that both the "intuitive" and the "technical" are always present in systems such as the ICD.

The way in which this gateway function is provided is twofold. First, the Aristotelian classification embeds within itself a set of implicit narratives that align the artificial categories of the ICD with the real world. Second, the rules for interpretation and presentation sit on top of the ICD and nudge its categories along prepared legitimate pathways. This combination of embedded and supervenient narrative provides the "give" through which the prototypical classification can be made to look and feel Aristotelian.
Increasingly, we will see work classifications and formal library classifications merging in the digital library of the future—the UMLS, which includes both the ICD and classifications of nursing work, among others, is a good example. In this discussion, the argument has been made that, as this happens, we need to pay due attention to the political and ethical undergirding of classification systems before they become so deeply inscribed in our information infrastructure that they are lost to sight (while their consequences propagate).

NOTES


2. Ironically, the slogan, "nobody dies of old age" was an anti-ageist aphorism first popular in the 1980s and used by groups such as the Grey Panthers. It was meant to imply that the social invisibility of old people led to them being medically invisible or overlooked as well. It is an interesting example of the inversion of the prototypical and Aristotelian aspects of death.

3. The general Principle is: "when more than one condition is entered on the certificate, the condition entered alone on the lowest used line of Part I should be selected only if it could have given rise to all the conditions entered above it." (p. 34)

4. Recommended age groupings and regional groupings are:
   - 1, 1-4, 5 year groups from 5-84, 85+
   - 1, 1-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+
   - 1, 1-14, 15-44, 45-64, 65+ (128)
   "Classification by area should, as appropriate, be in accordance with:
   each major civil division;
   each town or conurbation of 1,000,000 population and over, otherwise the largest town with a population of at least 100,000;
   a national aggregate of urban areas of 100,000 population and over;
   a national aggregate of urban areas of less than 100,000 population;
   a national aggregate of rural areas." (p. 128)

REFERENCES


The Dynamics of Classification Systems as Boundary Objects for Cooperation in the Electronic Library

HANNE ALBRECHTSEN AND ELIN K. JACOB

ABSTRACT

The notion of the classification scheme as a transitional element or "boundary object" (Star, 1989) offers an alternative to the more traditional approach that views classification as an organizational structure imposed upon a body of knowledge to facilitate access within a universal and frequently static framework. Recognition of the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production indicates the dynamic role of classification in supporting coherence and articulation across heterogeneous contexts. To this end, it is argued that the library should be an active participant in the production of knowledge, and that this role can be effected by the development of classificatory structures that can support the needs of a diverse information ecology consisting of a complex web of interacting agents, users, and technologies. Within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm but must evolve in cooperative interaction between librarians and their user groups.

INTRODUCTION

A bibliographic classification system is intended to provide both an overall structure for a document collection and a set of concepts that will guide the information searcher into the knowledge domains encompassed by the collection. Traditionally, classification research has approached these objectives by developing schemes based on a one-size-fits-all-search-
ers paradigm—i.e., We have created a standard system, because, deep down, all users are the same. Such classificatory tools often fail to fulfill their function of supporting the searcher's access to, and navigation through, the domain structure. In most databases, including catalogs on the Web, the searcher may find it difficult to comprehend the organizational structure that has been imposed upon the materials. This is not due simply to the often exotic notations of a scheme or to the surface characteristics of the classificatory data. Rather, the problem is often a product of a lack of match between the structure imposed upon the retrieval system by the classification scheme and the user's individual knowledge structures and search strategies.

Classification research has responded to this problem by collecting the terminology of individual users and compiling the results to generate larger, broader, and, it is hoped, more successful sets of access points for users—i.e., If we design an end-user thesaurus, that should do the trick. In his recent book on information seeking and subject representation, Hjørland (1997) argues that such endeavors to compile end-user vocabularies are generally conducted without recourse to an underlying theory of knowledge. Because failure of the classificatory structure to support user access is generally interpreted as a mechanical question of matching between different individual knowledge structures—i.e., among those of the searcher, the author, and the librarian as mediator (compare, for example, Ingwersen, 1992)—the underlying relationship between user access and the collective knowledge structures that are the basis for knowledge production has not been widely recognized.

From the perspective of the sociology of science, Star (1989) has argued that the Turing test, which is intended to measure the degree to which an expert system is able to perform as a human expert in its interaction with individual users, should be replaced by a "Durkheim test," where the system is evaluated on its ability to support the goals of a specific community of users. Star points out that scientific work is not all one piece but is distributed and heterogeneous, with differing viewpoints emerging only to be reconciled within the existing knowledge base. In her view, information systems should not be designed simply to represent consensus but to accommodate the dissent that can be expected to appear among the various communities participating in their use. To this end, she brings forward the concept of boundary objects as a method for resolving problems of heterogeneity in knowledge production and use or, in terms of library and information science (LIS), problems of variation or inconsistency in the representations by information producers, information mediators, and information users.

In this article, we will investigate how classificatory structures can act as transitional elements or boundary objects (Star, 1989) to support coherence and articulation in the heterogeneous and sometimes distributed...
contexts where knowledge is produced and mediated. In particular, we will review, within the context of the library, two perspectives put forward by Hjørland (1997) and by Star (1989) that analyze information systems as dynamic social constructs. We will build an analogy between a scientific enterprise and the library as an active participant in the general production of knowledge and use this analogy to develop a view of modern classification research that engages the library directly in the development of classificatory structures that can accommodate information searching by heterogeneous user groups. Following Nardi and O'Day (1996), we regard the library as a diverse information ecology, comprising a complex web of interacting human agents, users, and technologies. And we will argue that, within such an information ecology, a classificatory structure cannot follow a one-size-fits-all paradigm but must evolve in cooperative interaction between librarians (and other information intermediaries) and their user groups. In this context, we draw on examples of information systems in Danish public libraries—i.e., the Book House (Pejtersen, 1980) and Database 2001 (Albrechtsen, 1997).

Classification Systems: From Rationalism and Empiricism to Social Constructivism

Hjørland (1997) argues for a philosophical and sociological orientation for classification research. In his view, the problem of the searcher's uncertainty is a function of relative task uncertainty in the user's problem domain. Because information searching takes place within a particular social framework—e.g., an academic discipline—task uncertainty in searching is often the result of the relative task uncertainty within the discipline itself. Albrechtsen and Hjørland (1994) have earlier shown how such task uncertainty within knowledge domains may be a function of various social factors involved in the production of knowledge, such as the degree of interdisciplinarity or maturity within a domain. Such uncertainties will not only be manifest in the searchers' difficulty in formulating queries for IR-systems but will also be inscribed in the relative plasticity and variety of the concepts and terminology applied within the domains.

Classification research has too often neglected such broader social backgrounds that inform information searching and knowledge organization and has relied, more or less implicitly, on either a one-size-fits-all paradigm (rationalism) or on the accumulation of data about user behavior (empiricism). While the rationalist approach argues that we just need to get everyone to understand this, the empiricist counters that we just need to get more data about users and proceeds to collect more or less meaningful sets of "facts" on the individual user's relative success measured as the number of "hits" resulting from a series of search queries.

Figure 1 divides the different approaches to classification research and practice into two broad epistemological categories: Rationalism/
Empiricism on the one side and Historicism/Social Constructivism on the other. Both rationalism and empiricism are based on assumptions regarding the nature of truth and the objectivity of knowledge. From the empiricist approach, knowledge is reduced to sensory observations or facts. In classification research, empiricism is the prevalent epistemology in bottom-up thesaurus construction based either on user warrant or on terminology warrant, particularly when the process lacks grounding in a theory of knowledge. In contrast, rationalism strives to reduce knowledge to an all-embracing structure of concepts that is intended to be universally comprehensive. It is, for example, the epistemological foundation for Ranganathan's notion of universal facets. Rationalism is also closely related to more sociopolitical actions undertaken by a particular agency or from a specific disciplinary viewpoint—i.e., actions which are intended to impose one view of knowledge on all research and practice within that domain. In a paper discussing the role of dialogue in the development of classificatory structures, Jacob and Albrechtsen (1997) have shown how the American Psychiatric Association's construction of DSM-IV (American Psychiatric Association, 1994), the international classification for mental disorders, used dialogue to create a device for marginalizing and eliminating the viewpoints of competing professions such as psychology (see also Kirk & Kutchins, 1992). In short, both empiricist and rationalist approaches to classification are primarily looking for invariant structures that can be imposed on encyclopedic knowledge (universalist approaches) or data compiled from local observations (e.g., grounded theory approaches).

In contrast to these more formalized structure-seeking approaches to classification, social constructivism, or historicism, offers a view of knowledge as a product of historical, cultural, and social factors, where the fundamental divisions and the fundamental concepts are products of the divisions of scientific/cultural/social labor in knowledge domains. According to a social constructivist epistemology, the concepts and the structures are inseparable in a classification system, and hence the schemes must reflect the development, variety, plasticity, and use of both within a particular knowledge domain. This implies that scheme designers are not primarily looking for ways to impose one single structure on knowledge, including one set of all-embracing facets. Rather, the designers should operate as "epistemic engineers," attempting to articulate and represent the dynamics of knowledge in such a way that the searcher can proceed from the topic of his initial query to other related perspectives on the same topic or to related materials within the same knowledge domain. In this manner, epistemic engineering of classificatory schemes can provide for multidimensional classification schemes where the concepts are represented in a variety of different conceptual structures, functioning to articulate the multiple discourses performed in different domains. In the
<table>
<thead>
<tr>
<th>View of knowledge in information systems:</th>
<th>Rationalism/Empiricism: Knowledge is infallible and objective.</th>
<th>Historicism/Social Constructivism: Knowledge is historically, culturally, and socially determined.</th>
</tr>
</thead>
<tbody>
<tr>
<td>View of concepts in information systems:</td>
<td>Concepts are objective and exist as modules of knowledge or universal facets.</td>
<td>Concepts are culturally determined, domain-dependent, and evolve from experience and use.</td>
</tr>
<tr>
<td>View of language and dialogue:</td>
<td>Dialogue is secondary to objective knowledge and can be controlled through standard classifications.</td>
<td>Dialogue is central to knowledge production and mediation and should be facilitated, not controlled.</td>
</tr>
<tr>
<td>Example: DSM-IV</td>
<td>Example: HIV/AIDS vocabulary</td>
<td></td>
</tr>
<tr>
<td>View of information systems and their designers:</td>
<td>Information systems are value-free gateways to knowledge. Designers are engineers whose primary function is to exert control in support of performance.</td>
<td>Information systems are meaningful historical products—social and cultural constructs. Designers are epistemic engineers and knowledge catalysts whose primary function is facilitation.</td>
</tr>
<tr>
<td>View of mediating tools, such as classifications systems:</td>
<td>What is a classification?</td>
<td>When is a classification?</td>
</tr>
</tbody>
</table>

Figure 1. Division of the Approaches to Classification and Research.

role of epistemic engineer, then, the scheme designer operates as an active participant in the process of knowledge production and mediation.

Such involvement on the part of the classificationist is particularly evident in areas of interdisciplinary research that engage participation from many different professions. The HIV/AIDS vocabulary, developed by Huber and Gillaspy (1996), provides an illustrative example of such involvement on the part of the scheme designers. This system, which was not intended as a classification per se but as a mediating vocabulary, was developed to support dialogue between the different communities involved with the HIV/AIDS epidemic, including clinical and medical researchers,
practitioners of alternative medicine, nutritionists, psychotherapists and other professionals, as well as those individuals who are either living with the disorder themselves or are caring for someone who has contracted the disease. The HIV/AIDS vocabulary is built on a theory of knowledge generation that explicitly eschews the standard life cycle for knowledge production in medicine—a knowledge cycle that proceeds in a top-down fashion from theory developed at universities and other research institutions, to applied clinical research, to daily clinical application. Rather, according to the epistemological view driving the HIV/AIDS vocabulary, research in lived experience must necessarily feed into basic clinical research. Accordingly, this scheme was not developed solely as a tool for retrieval of information in the database of the local community, but as a tool for facilitating communication both within and across diverse interest groups, from the so-called layman to the cloistered scientist. In its role as communicative facilitator, the scheme is also hospitable to adaptations and extensions as an indexing language in local contexts. For instance, specific drug names are not articulated in the scheme but are left to local instantiations of the indexing language. In Star’s (1989) terms, the HIV/AIDS scheme serves as a boundary object precisely because it supports cooperation and common understandings among the various interest groups touched by this epidemic.

**Classification and Boundary Objects**

The notion of “boundary objects” was developed by Star (1989) as a structure for coordinating distributed work, such as may occur with a scientific enterprise that not only involves heterogeneous actors, elements, and goals but also incorporates different research methods, values, and languages. From her field work with scientific communities, Star has found that scientists are able to cooperate without consensus or shared goals. They can work together successfully because they create objects that function in the same way as a blackboard in a distributed artificial intelligence system:

I call these *boundary objects*, and they are a major method of solving heterogeneous problems. Boundary objects are objects that are both plastic enough to adapt to local needs and constraints of the several parties employing them, yet robust enough to maintain a common identity across sites. They are weakly structured in common use, and become strongly structured in individual-site use.

Like a blackboard, a boundary object “sits in the middle” of a group of actors with divergent viewpoints. Crucially, however, there are different types of boundary objects depending on the characteristics of the heterogeneous information being joined to create them. (Star, 1989, pp. 46-47. Emphasis in original)

Accordingly, Star (1989; Star & Griesemer, 1989) has identified different types of boundary objects in her various case studies, including:

- *repositories*—databases, libraries, or museums;
ideal types or platonic objects—diagrams, atlases, or abstract concepts such as, for example, the concept of "species" used by both the creators of a zoological museum and other interested parties involved in its construction;

coincident boundaries—common objects with the same boundaries but having different internal contents, such as maps of a geographical area that cover the same terrain but are outlined according to different knowledge interests such as, for example, the life zones identified by biologists contrasted with the trails and collection sites identified by museum conservationists;

standardized forms—forms created as methods of common communication across distributed work groups such as, for example, the forms completed during field work or the cataloging formats used for cooperation and networking between libraries where the content fields may or may not be part of each repository's database.

Unlike the model of the ideal universal computing machine whose goal, as proposed by Turing, is to emulate individual human mental capacities in all domains, boundary objects are advanced as an ecological concept—i.e., a concept that respects local contingencies and the viewpoints of different knowledge interests. In a case study on the formation of Berkeley's Museum of Vertebrate Zoology (Star & Griesemer, 1989), a classificatory structure of the species and subspecies of mammals and birds in California constituted an important boundary object. Thus the scientific classification scheme served as a shared conceptual structure and provided a shared vocabulary that facilitated communicative exchanges and cooperation across the different social and intellectual worlds represented by the scientists and the groups of amateurs who were involved in building the museum's collection.

Although they approach the problem of classificatory structures and knowledge access from two different angles, Star's exposition of the communicative and integrative functions of classificatory structures in the general knowledge production of the sciences is closely related to Hjørland's (1997) discussion of the epistemological positions adopted in classification research and his argument for following a more pragmatic philosophy of classification. Star builds on case studies and theoretical work in scientific communication and knowledge production, while Hjørland builds on case studies and theoretical work in the area of information searching for knowledge production. Both argue that classifications always serve pragmatic purposes in the same way that science serves human action. According to Hjørland's theory, scientific classifications reflect a highly abstract and generalized method of knowledge organization, in contrast to classifications with more local contingencies, such as categorizing fruit and vegetables in a supermarket or the amateur horticulturist's
categorization of plants by use or cultural preferences. Such variations in
taxonomic structure could be argued to reflect different levels of ambi-
tion among the interested parties and thus to function as boundary ob-
jects, created and negotiated by different social worlds, with the scientific
structure functioning as a more specific taxonomy dictated by the needs
of the scientific community itself. However, with respect to its specific
role within the praxis of a formal disciplinary community, the scientific
taxonomy is just as concrete as the pragmatic systems of classification that
reflect local contingencies. Indeed, when viewed from a broader socio-
logical perspective, these latter systems may actually be interpreted as more
abstract or generalized.

THE ROLE OF CLASSIFICATIONS IN DIVERSE INFORMATION
ECOLOGIES

American anthropologists Nardi and O'Day (1996) have introduced
the concept of "diverse information ecology" to describe the sociotechnical
network of heterogeneous materials, people, and practices that consti-
tutes a modern library:

What we learned in the library suggests the possibility of a socio-tech-
nical synthesis, an opportunity to design an information ecology that
integrates and interconnects clients, human agents and software
agents in intelligent ways congenial to extending information access
to, potentially, all of humanity. As we design the global information
infrastructure, the ultimate goal should be to design an ecology, not
to design technology. (p. 83)

Because information ecologies are situated within human practice,
they are dynamic and constantly changing. An information ecology can-
not be controlled by any one single agency but evolves through the col-
laboration of heterogeneous socio-technical networks, whose elements
strive constantly to achieve coherence and wholeness. The notion of an
information ecology also implies a collective view of information systems
as striving to meet heterogeneous community goals rather than the goals
of a single agency or individual. In their study of two research libraries in
software companies in the United States, Nardi and O'Day (1996) explored
how the work practices and expertise of librarians can serve as a model for
the design of computerized information services. They found that librar-
ians are exemplary agents who evince particular expertise not only in com-
municating with users but also in searching for information. These two
skills are closely interrelated in that the librarian's search strategy tends to
evolve in collaboration with the user's project. Nardi and O'Day propose
to extend this working relationship between the librarian and the user to
the collaborative design of information ecologies.

In an information ecology, a classification system should function as a
boundary object, supporting coherence and a common identity across
the different actors involved. In its role as boundary object, a classification would be weakly structured in common use, while remaining open to adaptation in individual communities. Across diverse information ecologies, classification schemes would function as discursive arenas or public domains for communication and production of knowledge by all communities involved. This approach to the development of classification schemes also implies that the task of constructing such a scheme would no longer be invisible work. This view of classification systems is in line with the concept of "coordination mechanisms" in distributed collaborative work, as put forward by Schmidt and Simone (1996). More importantly, the understanding and appreciation of classification schemes as boundary objects and discursive arenas, in cooperation with heterogeneous user groups and technology, engages the library as a facilitator of connections and ensures its continuing participation as an active contributor in the general process of knowledge production.

In the following discussion, we will illustrate how the role of classification systems is changing within the information system that is the library, shifting from reliance on a single standardized or mainstream view of order, where classification is the invisible precursor to the organization of a collection, toward the creation of more diverse information ecologies, where the development of a classification scheme takes place within an arena of discourse to create a shared order across heterogeneous social worlds.

**Something Old, Something New, Something Universal, Something Local**

As indicated in Figure 2, classification systems have served different pragmatic purposes in the history of libraries and information retrieval systems. In a recent European study of public libraries in the information society (Thorhauge & Segbert, 1997), it was demonstrated that public libraries have progressed through three distinct stages, evolving from manual paper-based services, via the automated library, to the current phenomenon of the electronic multimedia library. This progression should not be understood to imply that the current status of libraries has been driven entirely by technology. Rather, the electronic multimedia library must be understood from a more integrated socio-technical point of view, where the various actors, including librarians, computer suppliers, and researchers in computing and information science, constitute a heterogeneous network of agencies that bring certain technologies to the foreground while marginalizing others. In the recent development and use of communication technology, for example, there is a convergence of hitherto separate, even disparate, media and activities. This is apparent in the development and application of Web technology, which integrates text-based materials, graphic illustrations, and audio materials with interactive
features such as online conferences and e-mail. It is characteristic of this
development that the technology is not only plastic and customizable to
almost any context of use, rather like a boundary object, but is constantly
renegotiated and redeveloped through such use.

<table>
<thead>
<tr>
<th></th>
<th>Manual Paper-Based Services</th>
<th>The Automated Library</th>
<th>The Electronic Library</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary means of access</strong></td>
<td><strong>Collection building.</strong></td>
<td>Circulation, acquisition, stock control, etc.</td>
<td>Local access to global information. Networking.</td>
</tr>
<tr>
<td><strong>to knowledge:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Technology:</strong></td>
<td>Cards, phone, fax.</td>
<td>IT for housekeeping functions. OPACs.</td>
<td>Internet multimedia Web catalogs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Organizational culture:</strong></td>
<td>Introvert and bureaucratic.</td>
<td>Some change in attitude toward users.</td>
<td>Project oriented culture.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Role of classification systems:</strong></td>
<td>Order and control of collections.</td>
<td>Order and control. Subject access via OPACs.</td>
<td>Support of dialog in information services.</td>
</tr>
<tr>
<td></td>
<td>Invisible work.</td>
<td>Some experiments with thesauri.</td>
<td>Integration and infrastructure in diverse information ecologies.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Examples from Denmark:</strong></td>
<td>DDC is adapted and maintained in Denmark by central agency.</td>
<td>Verbal indexing in Danish national catalog.</td>
<td>Local experiments with classification and indexing in Danish public libraries.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Classification Research and Use in Different Stages of Public Library Development.
In the recent past, manual paper-based libraries focused on collection building. Intermediaries, or librarians, served both as collection builders and as agents controlling and interpreting the order of the libraries. Classification systems were frequently standardized in order to support interlibrary cooperation with the result that classification research was itself dominated by the development of universal schemes which could be adopted by central agencies to control the organization of knowledge across libraries. As a result of such standardization, classification became invisible work performed without regard to the needs of the local community of users. And, because maintenance and development of these classification schemes was often based on literary warrant, reflecting only those subjects represented in large national collections, they can be interpreted as imposing an implicitly empiricist view of knowledge. There was, then, at this stage in the library evolution, a mix of rationalist and empiricist epistemologies underlying classification research and development.

The role of the librarian as intermediary was challenged during the 1980s by the development of online retrieval systems and, in particular, by the introduction of online public access catalogs (OPACs) for end-user searching. During this decade, classification research was dominated by work on thesauri and indexing systems. There were numerous experiments with automated indexing, including the application of text analysis techniques developed in computational linguistics. OPAC development was often based on studying users, sometimes in naturalistic settings, but generally without prior analysis of their different social worlds or the functional role of libraries in knowledge production and mediation. Research in information retrieval systems was very much oriented by a mechanistic conception of human competence in information searching, indexing, and classification, thereby neglecting the variety and heterogeneity with which human agents (both librarians and users), information sources, and technology interact in different settings. Furthermore, as technological fixes were thrust to the foreground, displacing the search competence of the librarians, the librarian's role as intermediary between the searcher and the collection was gradually becoming marginalized as invisible work—the preliminary work of representing and organizing the collection that occurs in isolation from, and therefore without recognition by, the users.

During the 1990s, the library has increasingly switched its service emphasis from building and guarding the collection or offering users access to the collection through the local OPAC to providing local access to global information resources available on the World Wide Web. This represents a shift from a closed to an open system. In some European public libraries, for example, traditionally introverted and bureaucratic organizations have migrated toward a project-oriented culture, where librarians and users cooperate on the development of new services, using the interactive affordances of Web technology and the Internet. In general, such
projects have not involved the library schools in Europe, the traditional research communities in the library and information sciences. Close cooperation between libraries and the community of LIS researchers in Europe has yet to be manifested (Albrechtsen & Kajberg, 1997). In the United States, communities of LIS researchers have come together in workshops and research projects related to the social informatics of what are called “digital libraries” but could equally well be termed “electronic libraries” (Bishop & Star, 1996). In this research area, major topics include how knowledge is structured in digital libraries, including cataloging and classification, and how digital libraries are used—i.e., how knowledge is produced, communicated, applied, and recycled in distributed social worlds. Research methods comprise ethnographic studies of communication and knowledge production in (digital) libraries as well as comprehensive sociological studies of professional classification schemes in medicine (Bowker & Star, 1994) and nursing (Bowker, 1996). Thus it seems apparent that classification research is gradually evincing a more sociological and historical orientation.

**Classifications as Boundary Objects in Libraries: Librarians and Users in Mutual Design Activity**

Ballerup Public Library is a medium-sized Danish library on the outskirts of Copenhagen. There is, in this library, a tradition of direct collaboration between the librarians and their users. Until recently, a majority of the librarians regarded themselves as cultural workers—as intermediaries between collection and user, very much in line with the traditional perspective described above for libraries in the manual stage. In 1995, the library started a new project called Database 2001. This project, which was evaluated by Albrechtsen (1997), involved the development of an enriched multimedia catalog on the Web. In addition to the evaluation researcher, the project group for Database 2001 included six librarians with different areas of expertise: several in the group were experienced intermediaries and online searchers, while others were specialists in catalog design and in the management of the library’s technological resources. However, none of the librarians had experience with Web design or Internet browsing.

During the development of Database 2001, the project group collaborated with user groups and colleagues in the library to identify different kinds of materials, including books, musical recordings on CD, CD-ROMs, and audiotapes of books. Text, pictures, and sound were selected as enrichment for the database, the idea being to emulate a kind of virtual library on the Web. The menus were designed as graphical layers of icons representing both user groups and the kinds of materials available. The subject icons in Database 2001, which represent the subject content of materials in the database, went through several iterations. In addition,
the interface designed for browsing the menus was customized for both children and adults. The librarians arranged evaluation sessions with users who represented different user communities and their evaluations were very positive; users with different interests were able to use the icon-based interface for browsing in the database even though they had very different interests and different goals for searching.

In the database, documents were indexed using standard call numbers from the Danish variant of the Dewey Decimal Classification (DDC). Even though indexing by class number would take advantage of the hierarchical structure of DDC and thus would be potentially useful for browsing by users, the librarians knew from their practice as intermediaries that users found it very difficult to understand the standard classification. They experimented with a more pragmatic and much more weakly structured classification which could reflect the kinds of questions actually posed to library staff by the different user groups. For example, for subject browsing by children, they worked with the seven categories listed below and designed a unique icon to be used on the Web site:

1. computers;
2. astronomy, nature, animals, environment;
3. first love, star signs, being young today;
4. horses;
5. excitement, humor;
6. fantasy, science fiction; and
7. books that are easy to read.

From a semantic or disciplinary point of view, the separation of subjects like animals and horses would appear to be "incorrect" or "illogical." For the children, however, this classification worked very well. Category 2 (astronomy, nature, animals, environment) was intended for a broad group of interests, including fact literature, whereas category 4 (horses) was intended, in particular, for girls interested in novels about horses. There is, in Denmark, a special research tradition within children's librarianship, based on Wanting's (1984) research on how children ask questions in libraries, that advocates mediating literature according to the different user interests of children. Pejtersen (1994) has also studied children's use of libraries in Denmark and their communication with librarians. In her development of the Book House system in the 1980s, Pejtersen used a collaborative prototyping approach, engaging librarians, information scientists, and users in Danish public and school libraries, and subsequently designed a special interface of subject icons for browsing of the Book House system by children. Database 2001 took advantage of both of these research approaches to children's information searching.

The Book House (Pejtersen, 1994) is a retrieval system for fiction and is based on a general conceptual model that seeks to surround users with
an adequate resource space within which to situate their own search spaces. The design involves multidimensional representations of different kinds of user needs, search strategies, and literary paradigms as well as authorial intentions. This multidimensional structure for subject access is intended to match the different levels of user interest. The system interface is constructed around the metaphor of a "house of books," guiding the users through the rooms of a library where they can browse the collection. Users can also switch between different search strategies, including analytical search in the multidimensional database structure, visualized as icons for each dimension, and browsing of subjects, visualized as icons in a picture gallery. The design of these icons involved classification experiments using both word association experiments and evaluations of suggested icons in Danish public libraries.

The icons for browsing subjects in the Book House and in Database 2001 serve similar functions—to provide the users with an overview of the subjects included in the databases. Because the Book House system builds on the central design metaphor of rooms in a library, it provides a single uniform interface. Database 2001, in contrast, is realized as a mixture of interfaces that include the Web layer of icons, designed by the librarians; a more or less standard search client offering conventional text-based searching; and a database structured according to a standard cataloging format that uses traditional call numbers to represent the subject content of the documents. While the Book House is a general system for fiction retrieval, which in its present form cannot be customized by individual libraries to support the idiosyncratic needs of specific user communities, Database 2001 is a localized experiment with system design and classification drawing upon a range of technologies that reflect the heterogeneity of tools used in today's libraries, from conventional customizable applications such as the closed systems of the database and the search client to the open systems supported by interactive Web technologies.

Collaborative Development and the Agency of Libraries

Both the Database 2001 project and the Book House system were realized using a collaborative approach among librarians, users, researchers, and technicians. In this way, users were involved in negotiating classificatory structures and the design of subject icons in the interfaces of the two systems. Because the Book House was a new approach for interface and database design in the 1980s, it had to be developed technically from scratch. Database 2001, on the other hand, was able to take advantage both of the design ideas generated during development of the Book House system and of the possibilities for integrating modern Web capabilities within existing technology. The process of designing an interface adapted for local needs quite naturally involved local experiments with
classification. In Database 2001, the graphic Web layer and its icons were intended to represent both the users' needs and the existing technology. Decisions regarding the subject icons, as well as those pertaining to the search client and the database, were determined as much by the users as by the demands of the Web technology itself. Thus the icons employed in the graphic interface constitute an integrated system of boundary objects that mediate among the library, the users, and the technology. In this way, Database 2001 exists as an open system in that it makes the library available not only to local users but to other users as well through the medium of the technology. Without the interface of icons, the system would have been technically open but conceptually closed.

Design of the Book House and Database 2001 involved heterogeneous human actors, elements, and goals, which are also found in Star's (1989) description of a scientific enterprise. Star draws upon the example of a scientific enterprise to put forward a more collective concept of design than the psychological approach generally employed for the design of AI systems. Traditionally, design of library systems is based on a consensus model, or a one-size-fits-all approach. Multidimensional classifications providing different views of concepts in IR systems are still the exception (Albrechtsen & Hjørland, 1994; Jacob, 1994). But in the Book House system and in Database 2001, classificatory structures can perform as boundary objects by accommodating both the heterogeneous information needs and the various search strategies of different user interests as well as different knowledge communities.

Figure 3 juxtaposes some important boundary objects developed in the Book House and Database 2001 with Star's typology in order to illustrate the analogy between boundary objects in a scientific enterprise and the creation of a library system. Obviously, this analogy between the library and a scientific enterprise, even when supported by parallel structures, does not establish that what goes on in a library is isomorphic to what goes on in a scientific enterprise. Hjørland (1997) has proposed a theory of classification at multiple levels, from specific classifications developed in accordance with local contingencies to those general classifications developed by the so-called "hard" sciences, such as biology and medicine. However, analysis of the role of dialogue in the creation of classificatory structures indicates that traditional classification schemes frequently function as control structures that forestall an interpretive approach to scheme design through the imposition of controlled vocabulary that limits the impact of dissonant viewpoints (Jacob & Albrechtsen, 1997). In this manner, current developers of classification systems do not function as epistemic engineers, creating a discursive arena or forum for multiple views of knowledge, but rather as engineers of one episteme or worldview seeking to control the flow of knowledge production within a given domain by systematically legitimizing a single universal classificatory scheme,
thereby disenfranchising those researchers and practitioners who do not participate in the resulting structure.

<table>
<thead>
<tr>
<th>Star's Types</th>
<th>Book House</th>
<th>Database 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repositories:</td>
<td>Database with multiple orderings of information.</td>
<td>Enriched database.</td>
</tr>
<tr>
<td></td>
<td>Multiple kinds of materials.</td>
<td></td>
</tr>
<tr>
<td>Ideal type:</td>
<td>Icon for subjects and ordering dimensions.</td>
<td>Icons for subjects and target groups.</td>
</tr>
<tr>
<td></td>
<td>Multimedia enrichments.</td>
<td></td>
</tr>
<tr>
<td>Coincident</td>
<td>Design metaphor.</td>
<td>Web pages.</td>
</tr>
<tr>
<td>boundaries:</td>
<td>Rooms in library.</td>
<td>Interactive features (e-mail etc.).</td>
</tr>
<tr>
<td>Standardized</td>
<td>Database structure.</td>
<td>Database structure.</td>
</tr>
<tr>
<td>forms:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In general, however, the library and its organizational structures must be viewed as important actors in the general process of knowledge production because their primary goal is to mediate between knowledge producers and users. This role is generally realized through the provision of information services to users and producers who are very often members of the same knowledge communities. Although the scenario sketched here is traditionally understood as a closed world—i.e., where librarians mediate between documents and users—it could equally well be described as an information ecology—i.e., as a practice that builds environments by bringing together heterogeneous materials and actors.

The librarians' practice of building information ecologies is based on both explicit and tacit knowledge. The explicit knowledge is typically based on principles and formalisms for presenting classificatory structures in the form of universal classifications or faceted thesauri. The tacit knowledge includes knowledge of the interests of their user communities, the users' levels of computer and information literacy, and preferred tactics for "mediated" versus "unmediated" information services. In mediated services, the librarians communicate with the users, either directly or by e-mail, and guide them to relevant information sources. In unmediated services, such as the Book House system or Database 2001, the users may search a card catalog, a database, or a contingent local classification scheme prior to browsing the conceptual space within a domain. Such
“unmediated” services are, in fact, “silently” mediated by librarians or other information professionals who designed or customized a conceptual space for end-user searching. The librarian’s service to the users has been “translated” or formalized through the classification scheme. It has been abstracted or “disembedded” from the work context of a human intermediary interacting with a user.

Following Star and Strauss (1999), much of the mediating practice of librarianship may be considered “invisible work.” Even though the librarian as human intermediary is visible within the traditional library setting, his or her work is frequently considered to be “background work” involving the identification and delivery of books or other materials in support of the “real thing”—i.e., the user’s immediate work task or particular interest. When the work of the intermediary is abstracted from the work setting, this “invisible work” may become “visible” in that the task now falls to the user, but the dialogue between the user and the intermediary is effectively silenced. No longer is there a human intermediary to inform the user and ensure equality of services.

Gross and Borgmann (1995, cited in Bishop & Star, 1996) point out that: “Even home shopping requires informed consumers” (p. 904). When the librarian’s mediation work is silenced in the high-tech home shopping environment of electronic libraries—when the task of the user is no longer supported by, or facilitated through, dialogue with the human intermediary—some users will not be informed but will be reduced to mere consumers of standardized information services. Obviously, then, the information ecology of the electronic library cannot be responsive to the needs of the individual user without achieving a balance between visible and invisible work. As Star and Strauss (1999) point out: “Making visible can incur invisibilities; obscuring may itself become a visible activity.” In “unmediated” information services, cooperation between librarians and users in the design and maintenance of classificatory structures may be one method for achieving this balance between the visible and the invisible and for ensuring the evolution of an information ecology that is contingent upon the needs of an informed public.

**CONCLUSION**

Classification systems and indexing languages have been constructed as organizational tools in order to provide structure to a body of knowledge, but they frequently have the effect of limiting or restructuring individual conceptual structures during a process of information searching (Tang & Solomon, 1998). Established approaches to classification research and development appear to suffer from a fear of touching the real thing—the social worlds constituting an information system and the collective conditions for knowledge production. However, in LIS and the sociology of science, new approaches to classification research are emerging,
approaches that build on the idea of information systems as open and collaborative systems. A similar trend toward development of open systems has been identified in the public libraries in Europe which are evolving from manual, paper-based services to the electronic multimedia library. In the modern electronic library, classification is similarly transformed from a tool for establishment of order and control over the collection to a boundary object functioning to create cohesion across diverse information ecologies. The modern information ecology is a socio-technical network comprised of heterogeneous materials, people, and practices. Within this emerging network, the classification scheme constitutes a discursive arena facilitated by the library and functions as a boundary object for the various interests that exist among users and librarians. Such an information ecology is at the same time a situated network and an open system wherein the classification scheme supports coherence and articulation across the domains encompassed by the network both locally and globally.

The practice of classification is changing from invisible work carried out in centralized agencies to articulation work emerging within socio-technical networks. As the role of the library evolves from collection guardian to facilitator of connections, the role of classification is similarly transformed from control of collections to facilitation of communication, maintenance of coherence, and establishment of a shared conceptual context. From this perspective, then, the intelligent intermediaries of today are the human agents in diverse information ecologies who facilitate the process of knowledge production by collaborating with communities of users in the creation and use of boundary objects such as classification schemes.

ACKNOWLEDGMENTS

We gratefully acknowledge the inspiration and comments by Susan Leigh Star for this article. We also want to acknowledge the Danish National Library Authority for support of the project Database 2001 in Ballerup Public Library. An earlier version of this article was presented and discussed at the ASIS SIG/CR workshop in Washington, DC in 1997.

REFERENCES


ADDITIONAL REFERENCE

Psychiatrists make Diagnoses, but not in Circumstances of Their Own Choosing: Agency and Structure in the DSM

MARK A. SPASSER

ABSTRACT

Psychiatric classification is a profoundly important activity that directs subsequent treatment decisions, assumptions about etiology, and prognostic considerations. While the ideal classification scheme would be clear, concise, comprehensively inclusive of, and hospitable to, the entities under consideration, in practice, all classification systems reflect tradeoffs and embody flawed structures. Accordingly, it is essential to be fully cognizant of the shortcomings, biases, and tacit assumptions of extant systems so that classifications can be improved and so that misrepresentations will not be blindly repeated or reproduced. Modern psychiatric classification and diagnosis are almost exclusively defined within the context of the nomenclature and diagnostic categories of the American Psychiatric Association’s Diagnostic and Statistical Manual of Mental Disorders (DSM). This article adapts Giddens’s (1984) theory of “structuration” to explain how at least some of the consequences of relying on the DSM for classification result in unexamined conditions of its use and unintentionally reproduced its underlying assumptions. This article uses the DSM to explicate agency in structuration theory and structuration theory to illuminate the structure and use of the DSM. The discussion suggests that Mouzelis’s (1995) four-fold duality-dualism typology, by empowering the agent not only virtually but in actuality, is a necessary and salutary modification of structuration theory. Finally, it will be suggested that several prominent issues and concerns in psychiatric nosology resonate profoundly with those that have concerned, and continue to interest, library classificationists.
INTRODUCTION

Classifying (understood most broadly as arranging or grouping phenomena on the basis of some system or principle) is fundamental to, and underlies, all human thought (see, for example, Svenonius, 1983) and is thus an indispensable tool for understanding contemporary knowledge structures as well as their era-specific historical place and development. Classification is both the scientific origin and an arena of ongoing evolution, evaluation, and contention for the development of systematic knowledge. All classifications of knowledge, including library classifications and psychiatric nosologies, have been objects of contention because they ineluctably harbor tacit presuppositions of all kinds—e.g., scientific, sociocultural, practical, and ideological, to name but a few.

The ideal classification scheme would be clear, concise, and comprehensively inclusive of, or hospitable to, the entities and to the approaches to classifying the entities under consideration. Obviously, realizing such a classification is impossible. Thus, all attempts to classify reflect, to some degree, trade-offs, compromises, biases of omission and/or commission, possibilities, impossibilities, successes, and failures. Because an ideal classification is impossible, it becomes essential to be fully cognizant of the shortcomings, biases, and tacit assumptions of extant systems so that classifications can be improved and so that problems will not be repeated blindly or reproduced. This discussion is an attempt to contribute to that project.

This article elaborates on a theoretical framework for analyzing the operation of the official diagnostic classification system within the mental health professions. However, it is perhaps unremarkable that many of the most prominent issues and concerns of psychiatric nosologists resonate with those that have concerned, and continue to interest, library classificationists. The birth of psychiatry (as well as of library and information science [LIS]) was characterized by the introduction of classifications with a three-fold purpose (Pichot, 1986)—i.e., social, scientific, and pragmatic: “The early psychiatric nosology tried simultaneously to attain these three goals. Basic to this position was the conviction that, if the classification was ‘natural,’ i.e., scientific, it was at the same time the most pragmatic” (p. 63).

Compare the above assertion to Bliss’s resolutely held conviction that the library classification that best mirrors the scientific and educational consensus would also be the most useful to library patrons (see, for example, Bliss, 1929). Clearly, the status and validity of the warrants, if any, that underwrite classification schemes are a source of ongoing controversy.

Critical, recurring, and LIS-relevant issues in the history of psychiatric classification include the following (see, for example, Freedman, Silverman, Brotman, & Hutson, 1986). First, what is classified in a
psychiatric nosology—disease, disorder, syndrome, individual patients, or patient/client groups? This problem of identifying the unit or object of classification has its parallel in librarianship with the problem of distinguishing, descriptively, among the work, book, or manifestation.

Second, for whom is the classification scheme created? Is it for the researcher, mental health practitioner, or the courts? In other words, there are ongoing concerns with the audience for, and purpose of, the psychiatric classification. In LIS, classifications have traditionally been constructed for use by librarians but recently, with the advent especially of online public access catalogs, it has become increasingly clear that classification schemes need to be useful to the patron or end-user as well.

Third, there are concerns about the social inputs and consequences of classifications in terms of which fashions, societal trends, or pressures influence (or bias) the categories of a classification and of how types of knowledge (and people) get represented (and obliterated). Because of the relatively compelling economic and political implications of psychiatric classification (and, conversely, of the seeming absence of such ramifications with classifications of library materials), such sociocultural and ethical concerns have received far less attention in library science than in psychiatry.

Fourth, two related, common, and recurring themes in the history of psychiatric nosology arise directly from its ineluctable subjectivity: lumping versus splitting and the categorical/hierarchical versus dimensional (or, in library and information science terms, faceted) approach to classification construction (Mack, Forman, Brown, & Frances, 1994). The number and granularity of categories, and whether they can be considered discrete isolatable entities, are ongoing and potentially insoluble problems for both psychiatry and library science. These issues concern the epistemic status of our classifications and the distinctions they make and have equal applicability both to the classification of diseases/disorders in patients and to that of subjects/topics in LIS materials.

Finally, perhaps most symptomatically, both disciplines have been deeply concerned with the consistently inconsistent manner in which their classifications have been applied. In psychiatry this concern goes under the name of inter-rater unreliability, while in LIS it has been referred to as inter-indexer inconsistency. The intractability of this vexatious problem in both professions suggests their foundational relevance to each other.

Thus, it can now be readily appreciated that the study of psychiatric classification has much to offer library classification in terms of the relevance of, and overlap among, common and recurring themes. This article will therefore exploit the sociological-sensitive research about the former to frame and illuminate the latter.

This article will focus on the Diagnostic and Statistical Manual of Mental Disorders (hereafter the DSM-III, DSM-III-R, and DSM-IV will be collectively
referred to either as the DSMs or as the collective singular, the manual, unless otherwise indicated). The DSMs comprise the official nomenclature and classification system of the American Psychiatric Association and as such delineate the boundaries within which psychiatry claims epistemic and professional authority (Kirk & Kutchins, 1992; McCarthey & Gerring, 1994). However, the DSMs reflect a compromise of interests. While their primary goal is the pragmatic one of clinical utility, their underlying structures reflect not only (or even primarily) researchers and clinicians, but also the interests of lawyers, statisticians, epidemiologists, insurers, and disability claims personnel, among others. Accordingly, many decisions were made on the extra-clinical and non-empirical basis of expert consensus (Blashfield, 1984; Kirk & Kutchins, 1992) in the absence of empirical data. The DSMs are documents of mixed origins and conflicting purposes, based partly on scientific interests but also reflecting other clearly political and social (including professional) concerns.

While the DSMs have been the object of intense scrutiny, especially scientific, philosophical, and linguistic (see, for example, the contributions in Sadler, Wiggins, & Schwartz, 1994), they have not as yet been read from a “structurational” perspective. The purpose of this article is to analyze the DSMs by employing Giddens’s theory of “structuration.”

This discussion will begin with an outline of the major tenets of structuration theory, highlighting those principles especially applicable to classification in general and to diagnostic identification in particular. This will be followed by a close structured reading of two situated activities related to the development and use of the DSMs. To illustrate the analysis of strategic conduct, I will reread, from a “structurational” perspective, McCarthey’s (1991) review of the use made of DSM-III by one child psychiatrist in her hospital-based clinical practice. This will be compared to the article with McCarthey and Gerring (1994) in which the child psychiatrist of the 1991 article, as a co-author, rhetorically analyzes the sociopolitically motivated revision process leading to DSM-IV.

This comparative analysis will illustrate an important weakness of Giddens’s duality of structural theoretical framework and the utility, at least in terms of the analysis of strategic conduct, of maintaining the duality-dualism distinction. Briefly, Giddens’s structuration theory simply does not exhaust the types of relationships that actors have toward rules, resources, and social objects, such as classifications. In fact, opting, as Giddens does, for subject/object duality conflates agency and structure so that the possibility for actors to distance themselves from social resources to view, and orient toward, them strategically is severely curtailed, contradicting his useful distinction between institutional and strategic conduct analysis. In effect, this limitation in Giddens’s duality-of-structure notion limits the ability to distinguish the effects of classifications on classifiers and classificationists from those of the latter on the former.
STRUCTURATION THEORY

Giddens's structuration theory is especially useful in attempting to understand the social context and consequences of classification. Structuration theory is concerned with the conditions governing the continuity or transmutation of structures and therefore the (re)production of social systems (Giddens, 1979, 1984, 1993, 1995).

Basic concepts of structuration theory especially relevant to this discussion can be adumbrated as follows:

1. The duality of structure refers to the fact that social structures are simultaneously produced and modified by human agents and are used as resources; structures are dual in the sense that they are both the medium and outcome of the interactions and institutions they recursively organize.

2. Structure is a virtual order of rules and resources that exists only when instantiated in interaction and simultaneously both constrains and enables knowledgeable and skilled human agency. Conversely, systems are reproduced relations between actors, organized as regular/routine social practices; systems are the observable patterns of social interaction and can be said to exhibit, rather than have, structures or structural properties.

3. Modalities of structuration are rules that guide action (normative and interpretative) and facilities that empower action (authoritative/political and allocative/economic resources). Modalities (i.e., interpretative schemes, norms, and resources) are understood to be drawn upon by actors in the production of meaningful interaction: communication, sanctions, and power while, simultaneously, they are the reproductive media of the structural components of interaction systems: structures of signification, legitimation, and domination. The analytic significance of the modalities is that they provide the coupling elements whereby the analysis of interaction is linked to the (re)production of the structural components of social systems.

Critics of Giddens's explication of structure (see, for example, Archer, 1982; Layder, 1987, 1990) accuse him of obscuring the ontological status of structures. Because Giddens claims that structures exist only when instantiated in human activity, they reason that structures must be recreated anew each time. In other words, structures are created by human agency but must pre-exist any given actor's appropriation of them as resources in activity. However, supporters of Giddens have suggested that such criticisms exaggerate this difficulty by focusing almost exclusively on structures as necessarily instantiated in action, neglecting their continuity as "memory traces" (Giddens, 1984, p. 17).

However, there is a more serious problem with Giddens's concept of duality of structure. It conflates agency with structure and, in doing so,
simply does not address all the important relationships between agents and the rules and resources that comprise social objects (Mouzelis, 1995).\(^5\) As we will see when examining a psychiatrist’s use of the manual, by maintaining the duality-dualism distinction, agency can be better theorized both sociologically and critically. It is essential for agents to be able to distance themselves from rules so that we can account for their strategic use (and possible transformation).\(^6\)

**THE Diagnostic and Statistical Manual of Mental Disorders**

The *Diagnostic and Statistical Manual of Mental Disorders* is the official classification system of mental disorders published by the American Psychiatric Association. The third edition of the *DSM (DSM-III)* was published in 1980, was revised (*DSM-III-R*) in 1987, and the fourth edition (*DSM-IV*) was published in 1994. It can be considered a charter document in that it “establishes an organizing framework that specifies what is significant and draws people’s attention to certain rules and relationships. . . . defines as authoritative certain ways of seeing and deflects attention away from other ways . . . stabilizes a particular reality and sets the terms for future discussions” (McCarthey, 1991, p. 359).

Since 1980, the *DSMs* profoundly influenced the way in which the mental health field defines itself, the way in which it conducts its clinical and research work, the way it educates and socializes new professionals entering the field, and they have shaped legal and financial arrangements, including which treatments are eligible for insurance reimbursement (McCarthey & Gerring, 1994). Moreover, through the manual, a relatively small group of closely knit psychiatrists, known as the neo-Kraepelinians (Blashfield, 1984; Kirk & Kutchins, 1992),\(^7\) has attempted to accomplish three other things.

The primary goal of the neo-Kraepelinians is to assert the primacy and dominance of the biomedical model in the mental health field. There are at least two competing and contradictory models that have dominated psychiatry. The first, the biomedical-empirical model, comprises two primary assumptions. First, there are real discrete entities to which disease labels such as “dysthymia,” “schizophrenia,” or “attention deficit disorder” ought legitimately to be applied. These disorders are seen as generic and applicable across cultures, and there is the related assumption of underlying behavioral, psychological, or biological dysfunction: the disturbance is not to be located in the relationship between the individual and society. Second, the model employs the assumption of specific etiology, which in medicine states that diseases are caused by a single biological factor. That assumption has been modified in psychiatry to include patterns of multiple, discrete, and interacting etiological factors: biological, psychological, genetic, environmental, and/or social. Because these etiological factors are not well understood, the *DSMs* have adopted a fully or purely
descriptive approach: they attempt to describe comprehensively the manifestations of disease—i.e., they are intended to be atheoretical as regards the etiology of mental disorder.

The second model is hermeneutic-intuitive and fundamentally evaluative. The mentally ill patient is seen as an individual whose symptoms have meaning particular to him or her. In this model, the focus is less on distinguishing, describing, and classifying symptoms as manifestations of some unknown (and heretofore unknowable) underlying disease process, than on the meanings that those symptoms have for the individual. Mental health professionals, working within the hermeneutic model (many with psychodynamic and psychoanalytic orientations), understand the patient as an individual with a "story to tell" that must be understood and explained, while those working within the biomedical model see the individual as a member of a group with impairments to be explained. A psychiatrist's choice of perspective, which is often taken without awareness but which has profound ramifications for how the patient is conceptualized, is a result of personality, education, interests, and situational and professional pressures, a point that has important implications for this argument.

By imposing the biomedical model on mental health classification, this small group of psychiatric researchers (who are, most importantly, not clinicians) has attempted to accomplish two other more exclusively professional goals: (1) to achieve superiority over neighboring non-medical disciplines within the mental health field; and (2) to strengthen their affiliation and to achieve parity with other medical specialties (Kirk & Kutchins, 1992; McCarthey, 1991; McCarthey & Gerring, 1994).

The publication of DSM-III has often been referred to as a landmark event and a major scientific achievement (Kirk & Kutchins, 1992). According to Blashfield (1984), four major changes were made between DSM-I, DSM-II, and DSM-III and beyond: "(1) the use of diagnostic criteria; (2) a multiaxial approach to patient evaluation; (3) expanded descriptive information; and (4) a reorganization of the diagnostic categories" (p. 112). However, these diagnostic systems were, and are, controversial. Criticism has come from a wide variety of perspectives, some focusing on specific diagnostic entities and categories and others on broader conceptual issues, such as diagnostic boundary problems and the implications of a categorical classification for the measurement of comorbidity (Clark, Watson, & Reynolds, 1995). Without undertaking the impossible task of reviewing all critiques of the DSMs, three recurrent and important ones will be highlighted below.

One of the major criticisms of these diagnostic manuals is the focus on the individual, its individualistic metaphysics: "minds reside in brains, which in turn reside in individual persons. Minds, and subsequently mental disorders, do not reside in the social world" (Sadler & Hulgus, 1994,
The underlying assumption is unrealistic in that all psychiatric disorders (in fact, all human experience) are deeply embedded in social, community, or family networks (see, for example, the essays in Sadler, Wiggins, & Schwartz, 1994). This underlying structural principle undermines the usefulness of the DSM's multiaxial structure (especially with regard to axes IV and V) at least as it is currently constructed.

The second major area of criticism has been most strongly voiced by psychoanalysts and other dynamically oriented psychodiagnosticians. It focuses on the historical emptiness in the DSMs (McHugh & Slavney, 1983), claiming that they largely ignore the life story of the person: "The etiological, clinical, and practical significance of these [historical events such as job loss, catastrophic loss of loved ones, marital discord, and other stressful life events] and other life events in the patient's past are pushed into the nosological background" (Sadler & Hulgus, 1994, p. 262).

This fundamental disregard for the temporal and contextual dimensions of lived experience tends to reify or naturalize diagnostic categories. Instead of seeing DSM nosological entities as potentially useful abstractions, clinicians are encouraged to see their patients in terms of—and as being coextensive with—concrete diseases. Giddens (1984) talks about reification in a manner particularly apposite to classifications:

The concept [reification] should not be understood simply to refer to properties of social systems which are "objectively given" so far as specific, situated actors are concerned. Rather, it should be seen as referring to forms of discourse which treat such properties as "objectively given" in the same way as are natural phenomena. That is to say, reified discourse refers to the "facticity" with which social phenomena confront individual actors in such a way as to ignore how they are produced and reproduced through human agency. Reification thus should not be interpreted to mean "thing-like" in such a connotation; it concerns, rather, the consequences of thinking in this kind of fashion . . . . The "reified mode" should be considered a form or style of discourse, in which the properties of social systems are regarded as having the same fixity as that presumed in laws of nature. (p. 180)

As a result, a vast literature exemplifying the vital relevance of recent and remote historical life events to psychiatric problems, as well as an equally vast literature on human development and its pertinence to such problems, are excluded from consideration in the DSMs.

The third, and perhaps most celebrated, problem area in the DSMs has to do with their alleged atheoretical stance toward etiology. However, while no overt declaration is made in the manuals, they describe or structure diagnostic reality so that some etiological theories are more applicable or relevant than others (Faust & Miner, 1986). The diagnostic approach selects operationalized individualistic signs and symptoms as the relevant clinical data, whereas other kinds of contextual and temporally
sensitizing data are ignored as classifierically irrelevant. As Sadler and Hulgus (1994) observe:

This descriptive, syndrome-bound approach to diagnosis fits the needs of a biological psychiatry much better than other etiological models as, for instance, a family interactional model . . . or a developmental, life story approach . . . . Because DSM-III-R [as well as DSM-III and DSM-IV] fit biological psychiatry's theory base better than other psychosocially oriented therapies, the DSM-III-R diagnosis tends to make biological conceptualizations of the patient primary and the psychosocial secondary. In summary, DSM-III-R may not state a theory, but the metaphysical structure of its classification prefers the theoretical bases of descriptive/biological psychiatry. (p. 263)

STRUCTURATION THEORY, CLASSIFICATION, AND THE DSMs' MODALITIES OF STRUCTURATION

As has been already stressed, all human action and interaction is inextricably and simultaneously composed of structures of meaning, morality, and power. In terms of the modalities of structuration, social practice links the realm of human agency with that of social structure. Interpretative schemes are standardized shared stocks of knowledge that humans draw upon to interpret behavior and events, thereby achieving meaningful interaction. They are the cognitive means by which each actor makes sense of what others say and do. Resources are the means through which intentions are realized, goals are accomplished, and power is exercised. Norms are the rules governing sanctioned or appropriate conduct, and they define the legitimacy of interaction within a locale's moral order. As Orlikowski and Robey (1991) state: "Those three modalities determine how the institutional properties of social systems mediate . . . human action and how human action constitutes social structure" (p. 148).

Interpretative Schemes

From the point of view of strategic conduct, human interaction involves the communication of meaning which is achieved via interpretative schemes—i.e., stocks of mutual knowledge that agents draw upon in the production and reproduction of interaction. "These form the core of the mutual knowledge whereby an accountable universe of meaning is sustained through and in processes of interaction" (Giddens, 1979, p. 83). Interpretative schemes do more than merely enable the communication of shared meaning; they also serve as media for the imposition of structural constraints and affordances.

From the viewpoint of institutional analysis, interpretative schemes comprise structures of signification that represent the social rules that enable, inform, and constrain the communication process itself. Thus, in any interaction, mutual knowledge does not merely provide background for the communication process but is constitutive of it, in part organizing it and in part being constituted by the process itself.
As such, a diagnostic nosology like the DSM is an interpretative scheme that mediates between signification structure and social interaction in the form of meaningful communication among researchers, clinicians, patients, and such other organizational actors as insurers and government agencies. The signification structure in those cases comprises the shared rules, concepts, and theories that are drawn upon to make sense and organize communication about etiology, diagnosis (including reliability and validity issues and concerns), treatment plans, efficacy, and of course reimbursability.

Facilities (Resources)

From the point of view of strategic action, power enters into human interaction by providing the facilities and capabilities to accomplish outcomes. For example, the DSMs provide clinicians and researchers with categories that determine the applicability of various types of treatments. Power is understood here in both its broader meaning as transformative capacity—that is, the ability to transform or to affect the social and material world—and in the narrower sense of “power over”—that is, power as the domination of some individuals over others. Its use in organizations is mediated by the resources that agents appropriate within interaction.

All social systems and institutions are characterized by an irreducible asymmetry of resources (involving relations of both autonomy and dependence). the existing structure of domination is reinforced through the use of those resources, and it is when the existing asymmetry of resources is explicitly challenged or resisted, via what Giddens calls the dialectic of control, that the existing structure of domination may be creatively transformed.

This is especially the case with psychiatric diagnoses. For many, if not most, of the reasons mentioned above, both those diagnosing or applying the classification and those diagnosed may use a diagnosis (or assignment) for purposes of their own, purposes for which the nosology was not intended. Kirk and Kutchins (1992) explain in some detail the use of the manual to misdiagnose (both to under- and over-diagnose patients for purposes of stigma avoidance or to ensure reimbursability, respectively). Several authors (see, for example, Starr 1992; Hacking 1992) have called attention to the fact that, while classifications of the natural world are one-way relationships in that only people categorize natural objects, “[p]eople, however, have their own ideas about group membership—not only ideas but strong sentiments. When institutions classify, therefore, they often confront the self-conceptions of the subjects” (Starr, 1992, p. 158). Nowhere do those concerns, essentially with power, apply more problematically than in psychiatric classification.

Norms

From the viewpoint of strategic action, norms are organizational rules
or more or less binding conventions legitimating appropriate conduct. Such moral codes for legitimate conduct are created out of the continuous use of sanctions by agents in interaction. Norms play an active role in the shaping of institutional notions of "correct" behavior, that is, in what is to be regarded as vice or virtue, as important or trivial, and as obligatory or merely contingent. In this way, the practice of psychiatric diagnosis ineluctably involves the communication of a definite set of values (see Fulford, 1994, concerning the repolarization of illness and disease in terms of a value-based perspective on classification). The practice of diagnosis and of consequent classification can then be seen as involving the communication of notions of what should be, and it is primarily on the basis of those notions that sense is made of (or meaning constructed and imposed on) what psychiatrically exists.

From the viewpoint of institutional analysis, norms articulate, conventionally reproduce, or creatively (usually incrementally) transform established structures of legitimation. The legitimation structure institutionalizes the reciprocal rights and obligations of social actors and mediates, through norms and moral codes, the sanctioning of particular actions and interactions. Systems of psychiatric classification provide and legitimize the categories to which people can be assigned. For example, a DSM diagnosis is necessary for reimbursement from insurance companies or other third parties for treatment costs. The classification system embodies norms (such as reliability, validity, and conceptual operationalism) that legitimize diagnoses made from within its descriptive biomedical framework.

However, the modalities, either from the perspective of strategic conduct or from that of institutional analysis, are only isolatable for analytic convenience; in the flow of conduct and institutional life these are inextricably intertwined in each action and interaction. From an institutional perspective, modes of signification, domination (and subordination), and legitimation are intersecting dimensions of the wholeness of institutional social practice. From the point of view of strategic conduct, any interaction simultaneously exemplifies "three fundamental elements: its constitution as meaningful, its constitution as a moral order, and its constitution as the operation of relations of power" (Giddens, 1993, p. 110).

We can now explore, in detail, the use of the DSMs from a structurational perspective. A diagnostic classification system, as does all classification, exists in and as language (Hodge & Kress, 1993). Once inscribed in language, and after legitimizing sociocultural resources are mobilized and aligned (i.e., after much necessary, but often invisible, social and political work is transacted), a classification scheme becomes capable of coordinating and controlling action across long durations of time and large tracts of space. Moreover, as language, a classification scheme can be seen, structurally, as a set of generative rules and resources
which are drawn upon (and, often, in the process, reproduced) in its application, in this case classifying.

However, to understand the actual operation of such systems, it is helpful to go beyond descriptive and conceptual accounts and examine the conditions and consequences of its use in actual situated practices. In practice, different people will perceive a particular system in a variety of ways, and their appropriation of systemic resources will both reflect and reproduce their various interests. Moreover, the use of a classification system will inevitably reflect its unacknowledged conditions and generate unintended consequences because, according to structuration theory, actors, while inherently knowledgeable, may be unaware of the conditions of at least some of their actions and certainly of all the consequences that feed forward from previous—and feed back to subsequent—action.

To illustrate the application of structuration theory to the study of diagnostic classification, I will analyze one child psychiatrist's experience with the *DSM* in her clinical practice, suggesting both that we pay too high an analytic price by eliding the subject-object distinction and that to re-energize agency in structuration theory, we must allow for dualism, as well as for duality subject-object relations. Actors must be permitted to stand back and distance themselves from rules, resources, and interactive situations for the vitally important, and commonly observed, purposes of strategy or monitoring (Mouzelis, 1995).

**Strategic Conduct and the *DSMs***

McCarthey (1991) and McCarthey and Gerring (1994) have provided a detailed picture of the use of the manual by the child psychiatrist Gerring, who coauthored the 1994 article. In particular, this comparative analysis, while certainly not parallel, illustrates how Gerring's conception of (and, by implication, use of) the manual obviously changed over several years. The material presented in this section relies heavily on McCarthey's (1991) and McCarthey and Gerring's (1994) papers, which should be read in the original for a detailed and more complete accounting of their research. On the one hand, McCarthey (1991) will afford the analysis, from a structurational perspective, of the strategic conduct of one child psychiatrist to better understand and illustrate the ways in which structures (that are virtually present in the classification system) are appropriated and drawn upon to constitute social action as meaningful, legitimate, and enabling (and simultaneously constraining), while also being unintentionally reproduced through their appropriation and use. On the other hand, McCarthey and Gerring (1994) present the same psychiatrist interacting with the manual strategically and critically (however, not clinically). In terms of Mouzelis's duality/dualism typology, the former article illustrates paradigmatic duality and syntagmatic dualism, while the latter illustrates paradigmatic dualism and syntagmatic duality.
McCarthey (1991) used a multi-methodological approach to study the epistemological and textual consequences of *DSM-III* for the diagnostic work of Gerring, who was a child psychiatrist on the staff of a university hospital-based rehabilitation team that ministered to children who had suffered brain injury. McCarthey's detailed analysis follows Gerring through one case as she conducts interviews and draws her diagnostic conclusions. As might be expected, a biological psychiatric model, rather than a hermeneutic perspective, is more likely to be adopted (and reinforced) by psychiatrists working in this setting. Moreover, Gerring admitted to McCarthey that her own training emphasized the biomedical approach to studying psychiatric disorders as a result of studying pediatrics for years before undertaking her psychiatric training.

McCarthey presents her analysis in terms of how the manual structured and determined the gathering of data, the presentation of data, and finally, and most importantly, the analysis of the data that were gathered and presented in the psychiatrist's clinical diagnostic evaluations of one brain-damaged and comatose patient. McCarthey breaks those down into two general areas for analysis.

The first area concerns "*DSM-III* selectivity" (McCarthey, 1991, p. 365). The diagnostic classification determined the type and amount of data that were gathered about patients. Thus, while the categories of the manual were enabling in that they facilitated the collection of detailed information about some aspects of the patient's condition, they more problematically constrained Gerring from seeing other important data about the patient. In terms of the aforementioned structural principles that tacitly underlie the manual, the data required to make a *DSM* diagnosis do not include contextual and, for the most part, historical data about people (Sadler & Hulgus, 1994).

The manual assumes that mental disorders are real discrete entities that can be identified in patients by their clinical symptoms. Not only did the psychiatrist neither speculate as to the underlying meaning of the symptoms nor attempt to specify their etiological significance (unknowable in terms of the *DSM*), she used a highly structured interview schedule based on, and derived from, the *DSM* itself. In fact, not surprisingly, the schedule is designed to lead specifically and rigidly to a *DSM* diagnosis. We can see, then, that the conditions of its use reproduce the structural properties of the *DSM*. For example, as reported by McCarthey (1991): "If . . . [she] found no symptoms for a particular disorder, she moved on quickly. However, when her questioning revealed the presence of some [*DSM*-validated] diagnostic criteria for a disorder, she questioned . . . further" (p. 366). While Gerring reports feeling frustrated by what the manual and the interview schedule leave out, it is not surprising that the time constraints she feels—"[t]ime is the problem" (McCarthey, 1991, p. 368)—are the logical result of using a *DSM*-based interview schedule.
that merely reproduces the manual's lack of concern with temporal issues. Speeding through the interview looking for "legal" diagnostic criteria ensures that these, and only these, will be found.

Moreover, in terms of how the data are presented, the psychiatrist evinces her commitment to the biomedical model that tacitly underlies the DSM. As McCarthey (1991) points out, the headings that are used closely follow the hidden logics of the manual (Fulford, 1994) and the manual-based interview schedule. Three pages of the five page report that she completes on the patient are devoted to the data elicited from using the manual-based interview schedule. Two pages are left for basic facts about the patient, information sources, history of the present illness, observation of the patient, and family information, as well as other types of contextually and historically sensitive information. In more Giddensian terminology, she shows relatively little discursive penetration into the conditions of the perspective that organizes her clinical reality. And in terms of dualism-duality typology, on the paradigmatic level, she relates to the DSM dualistically in a taken-for-granted performative way. The instantiation of the rules and resources of the DSM are the medium and outcome of its use. On the syntagmatic plane, she relates in terms of dualism as something external to her over which she has little or no effect or control.

Most importantly, those aspects of the clinical evaluation suggest an acceptance of the belief that DSM, as a classification of mental disorders, is atheoretical; it presupposes both pure perception uninfluenced by thought, raw facts free of interpretation, and an atheoretical observational language. Unfortunately, "there is no perceptual experience that does not involve cognitive processing directed by assumptions, no fact that is not constituted by theory-guided interpretation of sensory stimuli, and no observational language that can describe experience without involving some theoretical background, whether explicit or implicit" (Goodman, 1994, p. 295; for a highly influential treatment of the theory laden-ness of observation, see Kuhn, 1970). The problem with an implicit commitment to atheoretical description (in addition to its falsity) is that, by accepting only those theories (paradoxically, atheorism is of course also a theory) or particular worldviews based on descriptive realism, a clinician will only consider an unnecessarily restricted range of options when contemplating alternative, and perhaps equally valid, conceptions of clinical reality. To the extent that such factors operate tacitly or covertly without being subject to examination, clinicians will unintentionally further and reproduce the presuppositions that subserve the sectional interests of such hegemonic groups as the neo-Kraepelinians.

The second, and more important, way in which Gerring relies on the DSM is not only to analyze the information that she gathered, but also to authorize her specific diagnoses, referred to by McCarthey as "DSM-III-backed analysis." In her analysis, Gerring refers to diagnostic criteria and
DSM categories in a taken-for-granted manner without explanation. The audiences for which this evaluation is intended require DSM-based diagnostic analyses. Only in this legitimized and legitimating nomenclature can Gerring authoritatively communicate with the other rehabilitation unit medical personnel, other mental health researchers, and insurers and legal personnel. Thus, her conclusions are validated by the same document that generated the type and form of her data. The same document, the DSM, determined not only the data and information that were collected but also their communication, interpretation, and authority.

Giddens (1984) would rightly call such a system a reproduction circuit: "By circuits of reproduction, I mean fairly clearly defined "tracks" of processes which feed back to their source, whether or not such feedback is reflexively monitored by agents in specific social positions" (p. 192). According to Giddens, these circuits of reproduction are implicated in the "stretching" of institutions across time and space. In the case of a psychiatrist who gathers data on the basis of the same system of diagnostic categories in terms of which she analyzes them, we have a relatively closed impermeable circuit in which the structures instantiated in the DSM are both the medium and outcome of her practice.

Subsequent to her work with McCarthey, the psychiatrist Gerring co-authored a paper on the revision process leading to DSM-IV (McCarthey & Gerring, 1994), a paper that evinces a radically different view of the DSM from the orientation of McCarthey (1991). While not a clinically-oriented study like the latter paper, the former offers a rhetorical analysis of the DSMs, along the more critical lines of Kirk and Kutchins (1992). Specifically, in the 1994 paper, the authors analyze the revision of DSM-III-R by observing work groups, by textually analyzing documents, and by interviewing principals in the revision process. The paper concludes with a detailed analysis of work group deliberations about the conception and inclusion of a new diagnostic category, BED (Binge Eating Disorder).

After critically discussing the theoretical and sociopolitical background of the DSMs, McCarthey and Gerring (1994) move to a detailed rhetorical analysis of the "selling of DSM-IV." Without repeating their argument, they bisect the persuasion strategy used to "sell" DSM-IV into strategic use of two rhetorical repertoires. First, the contingent repertoire is used when task force leaders attempt to distance themselves from DSM-IV's predecessors. Invoking the rhetoric of contingency allows them to account for mistakes made in past revisions in terms of the personalities and biases of the individuals involved.

However, since there is direct and virtually unbroken continuity between DSM-IV and its predecessors (in fact, the former can only be discussed and understood in terms of the latter), too severe criticism of the past would inevitably undermine their current efforts. Consequently, DSM-IV task leaders can securely position themselves as being "in a direct line
with DSM-III and DSM-III-R, by using the empiricist repertoire to describe the development of all three manuals" (McCarthey & Gerring, 1994, p. 166). In what they call the progress of the science repertoire, the false steps of past revisions can be "presented as vital stepping stones in the increasing adoption of the empirical method by the mental health field, as the best and only thing that could have been done under the circumstances" (p. 167). This rhetorical move allows the framers of DSM-IV to represent their work as another logical and essential step towards "a time when mental disorders will be understood well enough to be classified according to their pathogenesis, that is, their causal mechanisms [i.e., etiology], rather than just according to their symptomatology, as at present" (p. 167).

Finally, on the basis of observations of the deliberations of the BED work group, as well as of discussions with the participants, the authors draw four conclusions. First, the work group followed a strategy designed to present psychiatry as a mature biomedical discipline. Second, work group deliberations were shaped by scientific and clinical data, conceived as professionalizing concerns. Third, just as the texts themselves are influenced and shaped by unacknowledged personal and sociopolitical agendas, so were the work group discussions of BED. For example, "work group conversations were shaped by members' differing assumptions about the maturity of the field and the role DSM should play in either stimulating new research or slowing change and stabilizing current knowledge in psychiatry" (p. 171). Finally, they found evidence in the work group deliberations both of the contingent repertoire to describe their predecessors' erroneous work and of the empiricist progress of science repertoire to account for their more scientific work on BED.

Comparing this rhetorical analysis with the description of the conventional reproductive rule-following of Gerring in McCarthey (1991) suggests that a transformation in her orientation to the manual has occurred. In terms of the four-fold typology, we see, on the paradigmatic level, a movement from duality to dualism—i.e., from a natural-performative to a strategic-theoretical orientation to the rules and resources that structure the manual. On the syntagmatic level, there is movement in the opposite direction from dualism to duality—from a situation in which the actor is inseparable from, and whose actions constitute, the system to one in which the system is perceived as external to the agent. In other words, by distancing herself from the manual, rather than merely enacting the presuppositions of it, the psychiatrist is able to critically analyze and perhaps transform her interactions with it.

Thus, by examining the perception and use of the manual by a psychiatrist over time, the utility (necessity) of maintaining the duality/dualism distinction to truly empower the knowledgeable and capable agent has, it is hoped, been demonstrated. As Mouzelis (1995) correctly states,
if one opts exclusively for a subject/object duality approach, the only way of conceiving the relationships between subject and structure is to see the latter as medium/outcome—which means conflating agency and structure, and eliminating the possibility of actors distancing themselves from rules and resources in order to view them strategically. (p. 123)

While Giddens claims that his construal of reflexivity encompasses both that of the agent (as social theorist) in the flow of action and that of “the institutionalization of an investigative and calculative attitude towards generalised conditions of social reproduction” (Giddens, 1993, p. 6), a theory of knowledgeable and capable agency must allow the actor the necessary distance to strategically “stand back” from institutionalized rules to be able to attack or defend them or their variously perceived contradictions and incompatibilities. Consequently, agent-structure dualism, while problematic if not rigorously conceptualized, cannot be eliminated from structuration theory without paying too high a price, that is without sacrificing the agent to the constraints and affordances of structure.

Discussion

Structuration theory has several theoretical (as well as meta-theoretical) implications for classification research in general and for construction of diagnostic classifications in particular. As mentioned earlier, structuration theory allows not only for theorizing processes leading to change and continuity within theoretical systems but also facilitates theory-guided specification of generative mechanisms, processes underlying system dynamics that account for their surface manifestations. Such a distinction exists in psychiatric classification as the ongoing interlevel debate between etiological explanation and symptom description.

Second, not only does structuration theory focus our attention on situated practices as constitutive frames for understanding structures, but it also maintains that classification is an inherently social practice and, as such, cannot be understood without reference to the larger forces in which it, as a social practice, is embedded. The manual must be understood not only as the official nomenclature and classification of the American Psychiatric Association but also as a field of competing and intersecting forces, including psychiatrists, psychoanalysts, clinical psychologists, clinical social workers, psychiatric nurses, and insurers, each striving to control its ultimate form and content.

Equally applicable to library classifications, Kwasnik (1993), echoing the introduction to DSM-III, states that the DSM, as a conceptual structure intended to coordinate and articulate interaction (Schmidt & Bannon, 1992), is meant to facilitate and further the intradisciplinarily necessary functions of providing a common language, an accurate diagnostic tool, and a standardized vocabulary. However, she asserts unequivocally that:

The mandate for [the DSMs were] . . . politically and economically
motivated: government agencies, insurance companies, benefits programs, and others wanted to be able to differentiate and "tag" patients with mental disorders unambiguously for the purpose of reimbursement, legal action, confinements and so on. (Kwasnik, 1993, p. 64)

Consequently, to nontrivially understand the development, amplifications, and uses of a discipline's powerful conceptual structures—its official classifications—nomenclatures—it is imperative to account for both intradisciplinary and professional as well as societal, cultural, and historically situated forces and contingencies. As Bowker and Star (1991) said of the *International Classification of Diseases (ICD)*, an even more widely used and thus consequential conceptual scheme, "the list cannot be made homogeneous, neutral and appeal to all parties" (p. 77) because different categories of developers and users have (often incommensurable) different needs and impose conflicting demands on its design.

De Grolier (1982), employing the seminal observation of cultural anthropologists Durkheim and Mauss (1903/1963) that conceptual classification systems depend upon and reflect social conditions, bibliometrically investigated classification structures from medieval times to the present as cultural artifacts, and suggests that library classifications both are correlated with conceptual structures prevailing at their respective times and are thus inextricably historically situated. Additionally, Britain (1975) and Batty (1969) looked at the cultural context and embeddedness of classification and subject indexing, especially differences between British and American classificatory practices.

Britain (1975) states that there continue to be strongly held, contentious opposing points of view about, and dissident reactions to, classification as a tool and as the basis for subject analysis precisely because there are neither transcultural nor panhistorical acceptance of any consistent set of underlying principles. He quotes A. C. Foskett approvingly that "practically any classification scheme one would care to examine, far from being objective as it should be according to the emphasis of classification theorists is likely 'to reflect both the prejudices of its time and those of its author' [i.e., the classificationist]" (p. 34).

He rightly concludes that librarianship, being a historically situated professional subculture, has "its own ideas, its norms, and its tools . . . [which] will always tend to reflect the larger culture of which it is a part—its ideas, its laws and mores and even its aberrations" (Britain, 1975, p. 35).

Batty (1969) asserts that, in addition to such extrinsic (or external) factors as the sociopolitical system (e.g., democratic versus oligarchic class structure) of a culture, the intrinsic meaning of indexing and classification systems vary and must be understood if such differences are to be appreciated rather than judged or blindly repeated. He concludes that:

To the Western European, classification is an almost inevitable method of expression: it seems so natural to order subjects or ideas into
groups, each with a group name that therefore allows the further collection of groups-as-units into higher classes still. To the Americans, classification has meant only one thing: shelf—"marking and parking" . . . . It is not that there is any inability to understand how complex numbers are put together, or even how facet theory can be used to make a classification scheme: it is rather an inability to understand why they should be. [emphasis in original] (Batty, 1969, p.6)

Clearly, Batty is alluding to deeply held beliefs about the what, why, and how of library classification, and not merely to the more superficial (and probably more cross-culturally stable) technical abilities of classifiers to master any given scheme.

In sum, what Grob (1991) said of psychiatric nosology can be said of all attempts to classify and order, including LIS classificatory activities:

Classification systems are neither inherently self-evident nor given. On the contrary, they emerge from the crucible of human experience: change and variability, not immutability, are characteristic. Indeed, the ways in which data are organized at various times [and in various places] reflect specific historical circumstances. (p.421)

Nosologies and classificatory schemes are rarely, if ever, etched permanently in stone. They ineluctably grow out of specific historical contexts and reflect the various Zeitgeist of the times and places in which they were, and are, developed.

Reviewing three comprehensive, and currently used, library classifications, the Dewey Decimal Classification, Library of Congress Classification, and Bliss Classification, second edition, Langridge (1995) makes the apposite and salutary observation that:

The number, scope, and order of main classes represents a conscious or unconsciously held view of the world. . . . Yet all three systems, samples of a liberal humanist attitude, look alike when compared with Marxist schemes devised for Russia or China or with mediaeval schemes. [It seems likely that] . . . changes over long periods of time make different classifications appropriate to different epochs. The knowledge of the ancient world, the middle ages, and modern times are best accommodated by different schemes. (pp. 12-13)

Programmatically, structuration theory affords the study of library classification what it offers the study of the development and use of the conceptual structures and schemes of any other discipline: discursive penetration into the sociocultural conditions of the multiple perspectives that organize the context within which historically situated practitioners act.

**Conclusion**

In closing, the inability of Giddens's structuration theory, at least the version of it presented above, to theorize intentional transformative action has an unfortunate and particularly paralyzing relevance to the
ongoing revision of a living, yet institutionalized, text such as the manual. In general, change is problematic and only some sorts of it are always everywhere realistically possible. Unless we maintain the distinction between duality and dualism, allowing for agent-structure distanciation, implications of Giddens's structuration theory are that only an arbitrarily limited range of options will be possible for particular agents, that of possible changes only some will be known and desired, and that only an unrealistically limited range of those may be realized as the unintended consequences of agents' otherwise (but contentless and sterile) knowledgeably directed action. While Giddens emphasizes the importance of a critical reading and application of theory (see, for example, Giddens, 1984), he pays much more attention to the unintended consequences of social reproduction than to intentional creative transformation, which emphasis itself seems an ineluctable (unintended) consequence of his misguided and unsuccessful attempt to transcend agent-structure dualism.

According to New (1994), we intentionally change social structures by identifying them, the activities in which they are used, and their role in the reproduction of the social system to determine their liability or susceptibility to change. Moreover we, as knowledgeable agents, need to understand how these social structures simultaneously enable and constrain various position-practices, and how, by offering channels for agents' purposes, those generative rule-resource sets themselves consciously motivate. Consequently, New (1994) rightly concludes:

> [effective “reflexive appropriation” requires agents to recognise their own structural capacity and to use it to the full, or act to increase it. ... increasing our understanding of all the “unacknowledged conditions of action,” which would include unconscious sources of motivation, is likely to reduce the proportion of unintended consequences. ... The better these are theorised, the more likely that the chosen policy will fulfill its [intended] purposes. (p. 203)

In the final analysis, knowledgeable transformative action presupposes intentionality. Otherwise, we will be left in the ironic and unenviable position that “society is transformed by knowledgeable agents,” that this represents an “achievement,” and that nevertheless these knowledgeable agents know not what they do, since they both change and reproduce society by mistake, unintentionally, as a side effect of everyday social life” (New, 1994, p. 200). Unless we successfully theorize intentional processes of social change, for example, by acknowledging the situated reality of agent-structure dualism (or subject-object distance), agents, such as the psychiatrists of this paper’s title, are unhappily and unnecessarily reduced to Garfinkel’s “judgmental dopes,” despite Giddens’s protestations of knowledgeability, producing invalid diagnoses for seemingly valid organizational and professional reasons.
NOTES

1 This, of course, is a rather broad transposition of Marx's celebrated aphorism, "Human beings make their own history, but not in circumstances of their own choosing" (cited in Cohen 1987, p. 273). However, Giddens takes Marx's point very seriously; in fact, one could cogently argue that a large part of the Giddensian project is directed at explicating the full import of that aphorism. Additionally, throughout the paper the acronym DSM will be used to refer to various editions of the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders.

2 According to structurational analysis it is essential to distinguish between the study of structure-agency interdependencies at both the social and systemic levels. Consequently, there are two principal ways to study social system properties, each of which is separated out by a methodological epoche:

To examine the constitution of social systems as strategic conduct is to study the mode in which actors draw upon structural elements—rules and resources—in their social relations. "Structure" here appears as actors' mobilisation of discursive and practical consciousness in social encounters. Institutional analysis, on the other hand, places the epoche upon strategic conduct, treating rules and resources as chronically reproduced features of social systems (Giddens, 1979, p. 80).

However, the introduction of the duality/dualism typology enables a more nuanced interpretation than the binary (and, perforce, reductionistically false) methodological bracketing advocated by Giddens.

3 For full treatments of structuration theory, it is essential to read Giddens's evolving and variously nuanced accounts, which can be found in Giddens (1979, 1984, 1993) among others.

4 According to Cohen (1989), "structure 'exists' in manifest form only when it is instantiated in social practices. It otherwise persists between instances of social reproduction only as 'memory traces' sustained by knowledgeable social agents" (p.46).

Mouzelis (1995, p. 138) correctly observes that:

"a proper study of the linkages between a micro and a macro approach should not take the form

Institutional Structures (macro level)

Participants (micro level)

but the more complex configuration

Macro Institutional Actors

Macro Actors

Micro Institutional Actors

Micro Actors

Simply put, Mouzelis's more complex configuration posits that the consequentiality of an actor's actions for others can be large (macro) or small (micro), whether the actor is a single individual or a collective.

5 One of Giddens's important contributions to social science research is the realization that such analysis always involves a double hermeneutic:
The intersection of two frames of meaning as a logically necessary part of social science, the meaningful social world as constituted by lay actors and the metalanguages invented by social scientists; there is a constant "slippage" from one to the other involved in the practice of the social sciences (Giddens, 1984, p. 374).

One of Giddens's important contributions to social science research is the realization that such analysis always involves a double hermeneutic:

The intersection of two frames of meaning as a logically necessary part of social science, the meaningful social world as constituted by lay actors and the metalanguages invented by social scientists; there is a constant "slippage" from one to the other involved in the practice of the social sciences (Giddens, 1984, p. 374).

The so-called neo-Kraepelinians comprised a relatively small group of research oriented psychiatrist-nosologists, who, in the 1960s and 1970s, promulgated tenets of diagnostic classification first advanced by nineteenth century German nosologist, Emil Kraepelin. Their credo can be summarized as follows: psychiatry is a branch of medicine and should seek to establish scientific knowledge; psychiatry treats people who are sick, and who can be reliably distinguished from those who are well; psychiatry should conceive of mental illnesses biologically, as discrete isolatable entities; and classification and diagnosis are important and legitimate areas of research within the biomedical science of psychiatry. For detailed treatments of the neo-Kraepelinians see, for example, Blashfield (1984) and Klerman (1978).

According to Abbott (1988), the successful advancement of a profession's knowledge base is central to its jurisdictional strength. He states that:

the academic knowledge system of a profession generally accomplishes three tasks—legitimation, research, and instruction—and in each it shapes the vulnerability of professional jurisdiction to outside interference. Legitimacy provides a central foundation for jurisdiction, and its absence provides a central line for attack. . . . The academic knowledge system also provides new treatments, diagnoses, and inferences for working professionals; if it fails in this function, professional jurisdictions gradually weaken. (pp. 65-67)

Giddens (1979, 1984, 1993) distinguishes two types of resources: allocative, arising from command over objects and material phenomena, and authoritative, arising from capabilities to organize and coordinate the activities of social actors. These clearly have implications for use of the DSM. For example, allocative resources pertain to the eligibility of DSM diagnoses for third party reimbursement, while allocative resources refer to the power of the DSM to construct mentally ill identities.

Sewell (1992) makes the useful point that if structures are virtual, then they cannot include both rules and resources, and if they include both, they cannot be virtual. Resources, as media of power, and particularly allocative resources, must exist materially, and thus cannot be considered virtual. Thus, Sewell (1992) suggests that structure should refer only to rules or schemas, not to resources, which are better seen as effects of structures, as "media animated and shaped by structures, that is, by cultural schemas" (p. 11).

According to Giddens (1984), the dialectic of control is characterized by "the two-way character of the distributive aspect of power (power as control); how the less powerful manage resources in such a way as to exert control over the more powerful in established power relationships" (p. 374).


On the syntagmatic level [actual relationships], subject-object dualism refers to situations where a subject's participation in a game does not seriously affect its outcome, whereas duality refers to situations where the opposite is true.
On the paradigmatic level [virtual linkages], actors can, for strategic/monitoring reasons, distance themselves from rules (paradigmatic dualism); or they can use rules in a taken-for-granted manner (paradigmatic duality). (p. 156)

Thus, on the one hand, in terms of practice, dualism (separation) connotes little consequentiality, while duality describes situations wherein the consequences of an actor’s practice for others are large and compelling. On the other hand, in terms of the structural properties of social practices, duality (closeness) refers to a performative relationship of actor to object, while dualism describes situations in which actors distance themselves from formal structures for strategic purposes. Only by considering the consequences of the full range of relationships that actors have with rules and resources at both strategic conduct and institutional analytic levels can we fully account for the irreducible logics of the dispositional, interactive-situational, and positional dimensions of social action.

An admittedly arbitrary overview of social scientific realism posits that knowledge is a social product and lacks any sort of secure foundations; that there is a knowable external world; that while the social world is a construction, it is profoundly constrained by a specific history that provides agents with the materials for continued reproduction and, less frequently, transformation; and that valid social science aims to explain rather than predict. As to whether Giddens is a realist, there seems to be little doubt, but what kind of realist he is has been the subject of some debate. Some complain that he emphasizes structure over agency, others that he privileges agency over structure, and finally some accuse him of merely conflating agency and structure, explaining neither. In addition to Mouzelis’s (1995) critique, as presented in this paper, for differing but suggestive viewpoints, see, for example, Archer (1982), Layder (1987, 1990), Pawson (1989), the collected essays in Bryant & Jary (1991) and Held & Thompson (1989), and finally the special issues of Theory, Culture, and Society (1982), 2(2) and Journal for the Theory of Social Behavior (1983), 13.

REFERENCES


About the Contributors

HANNE ALBRECHTSEN is a Consultant in the Royal School of Library and Information Science in Denmark. She is President of the International Society for Knowledge Organization (ISKO). Ms. Albrechtsen’s interests are support and education within information technology, society, and libraries; cultural aspects of information systems and networks; system design and knowledge organization; information science, and science and technology studies; new challenges for the information profession including international collaboration. Her recent publications include “Database 2001: Design af en multimedie Web katalog i Ballerup Bibliotekerne” (Danmarks Biblioteksskole, 1997, 67 p.) [In Danish: Database 2001—the Design of a Multimedia Web catalogue in Ballerup Public Library]; Public Libraries and the Information Society edited with Jens Thorhauge, Gitte Larsen og Hans-Peter Thun—Luxembourg: Office for Official Publications of the European Communities (1997).

GEOFFREY C. BOWKER is Associate Professor in the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign. His book, Science on the Run: Information Management and Industrial Science at Schlumberger, 1920-1940, is available from MIT Press. He has just completed, with Leigh Star, a book on the history and sociology of medical classifications (Sorting Things Out: Classification and Practice—to be published by MIT Press in September 1999) and has recently co-edited a volume on Computer Support Cooperative Work (Social Science, Technical Systems and Cooperative Work: Beyond the Great Divide (LEA Press, 1997).

MARY L. GILLASPY is the manager of The Learning Center, a consumer health information resource at the University of Texas, M.D. Anderson Cancer Center, Houston, Texas. She was formerly the librarian at the AIDS Resource Center in Dallas, Texas, and has a particular interest in HIV disease in women.
JEFFREY T. HUBER holds dual appointments in the Informatics Center at Vanderbilt University Medical Center, Nashville, Tennessee: Research Information Scientist in the Active Digital Library, Eskind Biomedical Library, and Research Assistant Professor in the Division of Biomedical Informatics, School of Medicine. He specializes in HIV/AIDS and contributes regularly to the body of knowledge concerning the disease.

ELIN K. JACOB is an Associate Professor in the School of Library and Information Science at Indiana University in Bloomington, Indiana, where she teaches in the area of knowledge organization and representation. Her research, which adopts a neopragmatic, poststructuralist approach to formal systems of organization, focuses on the sociocognitive nature of categorization; on the linguistic and cultural origins of classification structures; and on the process of knowledge potentiation and the development of organizational tools that will extend current capabilities through the implementation of flexible representational frameworks.

HOPE A. OLSON is Associate Professor and Graduate Coordinator at the School of Library and Information Studies, University of Alberta, where she teaches organization of information including cataloging and classification, feminism, and library and information studies and globalization and diversity issues. Ms. Olson's research is on subject access in library catalogs and uses post-structural and other critical theoretical stances. She has published articles on classification and subject headings, on deconstruction as a methodology, and on teaching cataloging.

MARK A. SPASSER is Informatics Coordinator at the Center for Botanical Informatics, L.L.C. at the Missouri Botanical Garden in St. Louis, Missouri. Mr. Spasser's recent publications include: "The Enacted Fate of Undiscovered Public Knowledge" (Journal of the American Society for Information Science, 48[8], 707-717, 1997), "Mapping the Terrain of Pharmacy" (Scientometrics, 39[1], 77-97, 1997), and "Managing Cognitive Overload in the Flora of North America Project" (with Kay L. Tomlinson, J. Alfredo Sanchez, and John L. Schnase) (Proceedings of the Thirty-First Annual Hawaii International Conference on System Sciences, 1998).

SUSAN LEIGH STAR is Professor at the Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign. Her research is in the area of social aspects of computing, information science, and technology. She is the author of Regions of the Mind: Brain Research and the Quest for Scientific Certainty (Stanford, 1989) and, with Geoffrey Bowker, of Sorting Things Out: Classification and Practice, to be published by MIT Press in September 1999. She is the editor of Ecologies of Knowledge: Work and Politics in Science and Technology (SUNY 1995) and The Cultures of
Computing (Blackwell, 1995). Professor Star teaches classes in social informatics, writing, qualitative methods, and race and gender issues in information science. She was a student of both Anselm Strauss and Barney Glaser during her doctoral studies at the University of California in San Francisco.

JENNIFER TOBIAS is an Associate Librarian in the Reference Department at the Museum of Modern Art Library in New York.
INDEXING
AND ABSTRACTING
IN THEORY AND PRACTICE
2nd edition
By F. W. Lancaster

SECOND EDITION FEATURES
MULTIMEDIA SOURCES AND THE INTERNET
Award-winning author F.W. Lancaster has revised his widely used text to address growing complexities in the field. Featured in the second edition of *Indexing and Abstracting in Theory and Practice*:
- New multimedia sources chapter
- New indexing within the Internet chapter
- Updated chapters on text searching, automatic processing methods, and the future of indexing and abstracting
- Nine updated chapters on basic principles and theories
- Modified practical exercises

In addition to use as a text, *Indexing and Abstracting in Theory and Practice* holds value for managers of information services and others concerned with indexing, abstracting, and all related issues of content analysis.

The Publications Office
Graduate School of
Library and Information Science
University of Illinois
501 East Daniel
Champaign, IL 61820
(217) 333-1359
(217) 244-7329 fax
puboff@alexia.lis.uiuc.edu

Orders must be prepaid to
The University of Illinois
Major credit cards and checks accepted
ISBN 0-87845-102-1
426 pages
cloth
$47.50 plus shipping
"Library Trends has become the premier thematic quarterly journal in the field of American Librarianship."

*Library Science Annual*

Both practicing librarians and educators use *Library Trends* as an essential tool in professional development and continuing education. They know *Library Trends* is the place to discover practical applications, thorough analyses, and literature reviews for a wide range of trends. See for yourself the breadth of topics covered in the 47th volume.

**Service to Remote Users**
(Summer 1998) Edited by Lorraine Haricombe

**How Classifications Work: Problems and Challenges in an Electronic Age**
(Fall 1998) Edited by Geoff Bowker and Susan Leigh Star

**Folkloristic Approaches to Library and Information Science**
(Winter 1999) Edited by Betsy Hearne.

**Human Response to Library Automation**
(Spring 1999) Edited by Janice Kirkland

Institutional subscription price $75 (plus $7 for international subscribers). Individual subscription price $50 (plus $7 for international subscribers). Student subscription price is $25 (plus $7 for international subscribers). Single copies are available for $18.50, including postage. Order from the University of Illinois Press, Journals Department, 1325 S. Oak St., Champaign, IL 61820-6903. Telephone 217-333-8935, Mastercard, Visa, American Express, and Discover accepted.