Content Analysis: A Flexible Methodology

Marilyn Domas White and Emily E. Marsh

Abstract

Content analysis is a highly flexible research method that has been widely used in library and information science (LIS) studies with varying research goals and objectives. The research method is applied in qualitative, quantitative, and sometimes mixed modes of research frameworks and employs a wide range of analytical techniques to generate findings and put them into context. This article characterizes content analysis as a systematic, rigorous approach to analyzing documents obtained or generated in the course of research. It briefly describes the steps involved in content analysis, differentiates between quantitative and qualitative content analysis, and shows that content analysis serves the purposes of both quantitative research and qualitative research. The authors draw on selected LIS studies that have used content analysis to illustrate the concepts addressed in the article. The article also serves as a gateway to methodological books and articles that provide more detail about aspects of content analysis discussed only briefly in the article.

Introduction

As a research methodology, content analysis has its roots in the study of mass communications in the 1950s.[1] Based on a basic communications model of sender / message / receiver, initially researchers emphasized mak-

LIBRARY TRENDS, Vol. 55, No. 1, Summer 2006 ("Research Methods," edited by Lynda M. Baker), pp. 22–45

ing inferences based on quantified analysis of recurring, easily identifiable aspects of text content, sometimes referred to as *manifest content*. Since then, researchers in many fields, including anthropology, library and information studies (LIS), management, political science, psychology, and sociology, have used content analysis. In the process, they have adapted content analysis to suit the unique needs of their research questions and strategies and have developed a cluster of techniques and approaches for analyzing text grouped under the broad term of *textual analysis*. A significant change has been a broadening of text aspects to include syntactic, syntagmatic, and pragmatic aspects of text, although not always within the same study. Merten (as cited by Titscher, Meyer, Wodak, & Vetter, 2000) notes that "the range of procedures in content analysis is enormous, in terms of both analytical goals and the means or processes developed to pursue them" (p. 55). The variants include, for example, besides content analysis, conversational analysis, discourse analysis, ethnographic analysis, functional pragmatics, rhetorical analysis, and narrative semiotics.[2] Although these approaches are alike in their reliance on communicative material as the raw material for analysis, they vary in the kinds of questions they address and in their methods.

This article focuses only on content analysis, not on all forms of textual analysis. It distinguishes, however, between quantitative and qualitative approaches to content analysis since both are used in information studies. Content analysis is a flexible research method that can be applied to many problems in information studies, either as a method by itself or in conjunction with other methods. Table 1 provides a selective list of research studies in LIS using content analysis published within the past fifteen years (1991–2005).

After defining content analysis, the article goes through the basic steps in a content analysis study. It does this first for quantitative content analysis, then notes the variations that exist for qualitative content analysis. Throughout the article draws on the LIS studies in Table 1 for examples. Although only certain aspects of the LIS studies are mentioned in the article, they constitute a rich trove showing the broad applicability of content analysis to many topics. The article closes with a brief bibliographical note leading to sources providing more detail about the content analysis aspects treated only briefly here.

Definition

Not surprisingly, multiple, nuanced definitions of content analysis exist that reflect its historical development. This article accepts a broad-based definition in a recent content analysis textbook by Krippendorff (2004).[3] For the purpose of this article, content analysis is "a research technique for making replicable and valid inferences from texts (or other meaningful

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
Agosto & Hughes-Hassell, 2005	To describe the everyday life information-seeking patterns of urban young adults	Written activity logs and transcribed semi-structured group interviews[c]	Qualitative	
Buchwald, 2000	To explore the role of a public interest group in federal information policy making	Recorded and transcribed informant interviews, observation notes, group documents	Qualitative	
Croneis & Henderson, 2002	To identify the changing nature of work in academic libraries, focusing on positions emphasizing electronic and digital resources	Job advertisements in *College & Research Libraries* for a ten-year period	Qualitative	
Dewdney, 1992	To determine the impact of training in communications on reference interviews	Recorded and transcribed reference interviews	Quantitative	E. Z. Jennerich's Interviewing Skills Rating Scale (1974); modified to 4-point scale for each skill
Dilevko & Gottlieb, 2004	To determine the portrayal of librarians in popular culture	Obituaries in the *New York Times*	Qualitative	
Green, 1991	To identify conceptual models of information	Abstracts in the *LISA* database	Qualitative	
Haas & Grams, 1998a	To analyze Web pages and the links they contain to develop a classification system for both	1,500 links contained within 75 Web pages	Qualitative	
Haas & Grams, 1998b	To analyze links in Web pages and develop a taxonomy of functions 75 Web pages and 1,500 links contained therein	75 Web pages and 1,500 links contained therein.	Qualitative	
Haas & Grams, 2000	To discuss four overarching questions arising from past two studies of Web pages and links	75 Web pages and 1,500 links contained therein	Qualitative	
Hahn, 1999	To investigate how authors, editors and readers viewed development of electronic journals; to identify how authors decided to become involved in electronic publishing and how social structures influenced the process	Recorded and transcribed interviews	Qualitative	

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
Kracker & Wang, 2002	To investigate students' perceptions of research and research paper anxiety	Student writings on research paper experience	Qualitative/ Quantitive	
Lynch & Smith, 2001	To identify how computer technology is changing academic library positions	Job advertisements in *College & Research Libraries News*	Quantitative/ Qualitative	Yuan's (1996) Checklist of Computer-related Codes
Maloney-Krichmar & Preece, 2005	To develop an in-depth understanding of the dynamics of online group interaction and relationship between individual participants' online and offline lives	Messages on an electronic healthsupport list; interviews with participants.[c]	Qualitative/ Quantitative	Bales Transactional Analysis Schema (1951) Benne and Sheats's (1948) Group Membership Role Classification (1948)
Marsh & White, 2003	To develop a thesaurus of image-text relationships	Published research articles and books	Qualitative/ Quantitative	
Nitecki, 1993	To identify conceptual models of libraries among three groups	Letters and opinions in *Chronicle of Higher Education*	Qualitative	
Schoch & White, 1997	To compare communication patterns of participants in consumer health electronic lists for a chronic and an acute disease	Messages on two consumer health electronic lists	Quantitative/ Qualitative	Bales Transactional Analysis Schema (1951)
Stansbury, 2002	To determine the nature of problem statements in LIS articles	Research-based articles in eight core LIS journals	Qualitative/ Quantitative	Hernon & Metoyer-Duran (1993) attributes of problem statements
Wang & Gao, 2004	To analyze the extent and nature of technical services–oriented Web pages in academic libraries; to determine variations by type of institution	Technical services Web pages at U.S. academic libraries	Qualitative	
Wang & Soergel, 1998	To explore how real users select documents for their research projects from bibliographic searches	Interviews with participants as they made judgments about items in an online search	Qualitative	
Wang & White, 1999	To determine how real users make relevance judgments during use and citation phases of research projects	Interviews with participants about decision factors for using and citing documents in a research project	Qualitative	Wang & Soergel's (1998) Criteria for Document Selection (preliminary; modified substantially during coding)

Table 1. Selected Examples of Content Analysis in LIS Research, 1991–2005

Article[a]	Purpose	Data	Type of Analysis[b]	External Coding Scheme
White, 1998	To characterize questioning behavior in reference interviews preceding delegated online searches and to relate it to questioning behavior in other interviews/settings	Reference interviews	Qualitative/ Quantitative	Graesser's Typology of Questions (Graesser, McMahen & Johnson, 1994); White's (1985) Typology of Reference Interview Content (modified to add categories)
White, 2000	To characterize questioning behavior on a consumer health electronic list	Messages on an electronic list	Qualitative/ Quantitative	Graesser's Typology of Questions (Graesser, McMahen & Johnson, 1994); Roter's (1984) Typology of Medical Content of Patient Questions
White, Abels & Gordon-Murnane, 1998	To characterize adopters and non-adopters of Web for e-commerce	Web sites of publishers of business information	Qualitative	
White, Abels, & Agresta (in process)	To assess the relationship between chat interview quality and response quality	Transcripts of chat interviews	Quantitative	White, Abels & Kaske Typology of Turn Functions in Reference Interviews (2003)
White & Iivonen, 2001	To identify the reasons for selecting initial strategy in Web searches	Brief responses to questionnaire question about reasoning for decision	Qualitative	

[a]Complete references for items referred to in the first and last columns are in the bibliography.
[b]Studies are sometimes hybrids, with characteristics predominant to one type of content analysis but with some from the other type. For these, both types are sometimes noted with the predominant form first.
[c]Authors used other data in broader project; only data covered in this article or analyzed by content analysis are mentioned here.

matter) to the contexts of their use" (Krippendorff, 2004, p. 18). The notion of inference is especially important in content analysis. The researcher uses *analytical constructs,* or rules of inference, to move from the text to the answers to the research questions. The two domains, the texts and the context, are logically independent, and the researcher draws conclusions from one independent domain (the texts) to the other (the context). In LIS studies the analytical constructs are not always explicit.

The analytical constructs may be derived from (1) existing theories or practices; (2) the experience or knowledge of experts; and (3) previous research (Krippendorff, 2004, p. 173). Mayring (2000), the author of a standard German-language text on qualitative content analysis, suggests using a model of communication to determine the focal point for the inferences. Conclusions can be drawn about the communicator, the message or text, the situation surrounding its creation—including the sociocultural background of the communication—and/or the effect of the message. For example, Nitecki (1993) focuses on characterizing the communicator. She draws inferences about academicians' conceptual models of libraries based on analyzing the metaphors they used when they referred to libraries in published letters to the editor and opinion articles.

Content analysis involves specialized procedures that, at least in quantitative content analysis, allow for replication. The findings of a good study using quantitative content analysis, therefore, do not rely solely on the authority of the researchers doing the content analysis for their acceptability. They can be subjected to independent tests and techniques for judging their validity and reliability. Indeed, the extent to which validity and reliability can be judged are significant issues in evaluating a research methodology, and they are considered in subsequent sections in relation to both quantitative and qualitative content analysis.

Data

What constitutes data that can be used for content analysis studies? Most important is that the data provide useful evidence for testing hypotheses or answering research questions. Another key factor is that the data communicate; they convey a message from a sender to a receiver. Krippendorff's definition expands text to include "other meaningful matter" (2004, p. 18). Pictures on a Web site, for example, are used to convey one or more meanings, often in combination with text (Marsh & White, 2003) and, as such, can be subjected to content analysis either by themselves or by looking at the relationships between images and text, as Marsh and White have done. Both Bell (2001) and Collier (2001) discuss the content analysis of visual images.

Beaugrande and Dressler (1981) suggest seven criteria for defining a text, which is the more common form of data for content analysis: cohesion, coherence, intentionality, acceptability, informativity, situationality,

and intertextuality. In other words, text appropriate for content analysis is composed of linguistic elements arranged in a linear sequence that follows rules of grammar and dependencies and uses devices such as recurrence, anaphora and cataphora, ellipsis, and conjunctions to cause the elements to "hang together" to create a message (cohesion). The text has meaning, often established through relationships or implicature that may not be linguistically evident, and draws on frameworks within the recipient for understanding (coherence). The writer or speaker of the text intends for it to convey meaning related to his attitude and purpose (intentionality). Conversely, recipients of the message understand the text as a message; they expect it to be useful or relevant (acceptability). The text may contain new or expected information, allowing for judgments about its quality of informing (informativity). The situation surrounding the text affects its production and determines what is appropriate for the situation and the culture (situationality). The text is often related to what precedes and follows it, as in a conversation (one interpretation of intertextuality), or is related to other similar texts, for example, others within a genre, such as transcripts of chat sessions (another meaning of intertextuality).

The texts used in the LIS studies in Table 1 vary significantly. Some are generated in connection with the immediate research project; other texts occur naturally in the conduct of normal activities and independent of the research project. The former include responses to open questions on questionnaires (Kracker & Wang, 2002; White & Iivonen, 2001, 2002) and interviews with participants (Buchwald, 2000; Hahn, 1999). The latter include reference interviews (Dewdney, 1992), segments of published articles and books (Green, 1991; Marsh & White, 2003; Nitecki, 1993), obituaries (Dilevko & Gottlieb, 2004), problem statements in published articles (Stansbury, 2002), job advertisements (Croneis & Henderson, 2002; Lynch & Smith, 2001), messages on electronic lists (Maloney-Krichmar & Preece, 2005; White, 2000), and Web pages (Haas & Grams, 1998a, 1998b, 2000; Wang & Gao, 2004). Some studies use a combination of the two. For example, Buchwald (2000) analyzed recorded and transcribed informant interviews, observation notes generated during the research, and existing group documents in studying Canada's Coalition for Public Information's role in the federal information policy-making process.

Neuendorf (2002) proposes a useful typology of texts that takes into consideration the number of participants and/or setting for the message: individual messaging, interpersonal and group messaging, organizational messaging, and mass messaging. Individual responses to an open question on a questionnaire or in an interview are examples of individual messaging; the objective of content analysis is usually to identify that person's perspective on the topic. Reference interviews are a form of dyadic, interpersonal communication (Dewdney, 1992). Messages on electronic lists (Schoch & White, 1997) offer an example of group messaging; the person sends the

message to the group, any member of which can reply. The objective, in this case, is to characterize the communications of the group. Technical services Web sites (Wang & Gao, 2004), often existing only on Intranets, are examples of organizational communication. Job advertisements in LIS journals (Croneis & Henderson, 2002) are examples of mass messaging.

All of these types of text can occur within various applied contexts. For example, within the context of consumer health communication, studying messages on consumer-oriented health electronic lists (informal, group messaging) can provide insights into information needs that are not satisfied through doctor-patient interviews (more formal, interpersonal, dyadic communication) (White, 2000). Analyzing job advertisements (Croneis & Henderson, 2002) is similar to studying personal ads in other fields (Cicerello & Sheehan, 1995).

Data: Unitizing

At an early point in a content analysis study, the data need to be "chunked," that is, broken into units for sampling, collecting, and analysis and reporting. *Sampling units* serve to identify the population and establish the basis for sampling. *Data collection units* are the units for measuring variables. *Units of analysis* are the basis for reporting analyses. These units may be, but are not necessarily, the same. In many cases, the sampling unit is the documentary container for the data collection unit and/or units of analysis. It is the naturally occurring vehicle that can be identified and retrieved. In Dewdney (1992), for example, the entire interview serves as all three units. In White (2000) the message is the sampling unit; she has several different units of analysis in her study of questions in electronic lists: the message as a whole and individual questions within the messages. She also breaks the questions down into the form and content of the question, focusing on different segments of the question as phrased for categorizing.

In separate studies, Green (1991) and Nitecki (1993) focus on two words (*information* and the stem *librar,* respectively) and analyze the phrase immediately surrounding each occurrence of the word (data collection units) in two types of documents (sampling units) (for Green, abstracts in the *LISA* database; for Nitecki, letters and opinion articles in the *Chronicle of Higher Education*) to identify the metaphors surrounding use of these terms. They subsequently analyze the phrases to generate atomized phrases and then collapse them into metaphors (the unit of analysis). Each then interprets the metaphors as evidence of conceptual models held by the writers of the documents. In comparison to Dewdney (1992), who also studied reference interviews, White, Abels, and Agresta (2004) analyze turns (the unit of analysis) within chat reference interviews (the sampling unit). In Marsh and White (2003) the emphasis is on relationships between images and text, so the unit of analysis is the image-text pair, defined as the image and its related text segment (p. 652).

Pragmatism determines the sampling and data collection unit; the research question or hypothesis determines the unit of analysis. In all of the studies mentioned above, the unit of analysis is naturally related to the research question or hypothesis being addressed.

Procedures: Quantitative Content Analysis

Before discussing distinctions between qualitative and quantitative content analysis, it is useful to identify and explain the steps involved in content analysis. The focus initially is on the steps for a study using quantitative content analysis. The steps are as follows:

1. Establish hypothesis or hypotheses
2. Identify appropriate data (text or other communicative material)
3. Determine sampling method and sampling unit
4. Draw sample
5. Establish data collection unit and unit of analysis
6. Establish coding scheme that allows for testing hypothesis
7. Code data
8. Check for reliability of coding and adjust coding process if necessary
9. Analyze coded data, applying appropriate statistical test(s)
10. Write up results

Generating Hypotheses

Quantitative content analysis flows from a positivist research tradition and is deductive in its approach. Its objective is to test hypotheses, not to develop them. Drawing on related research and existing, relevant theory, a researcher first establishes one or more hypotheses that can be tested using content analysis. These hypotheses flow from what is already known about the problem and the extant research questions. In Dewdney, for example, "the hypothesis predicted, essentially, that interviews conducted by librarians who had received training in either neutral questioning or in the use of microskills would contain more examples of effective use of the skills taught, respectively, than interviews conducted by these same librarians before training, or than interviews conducted by librarians who had received no direct training" (1992, p. 131).

Determining Data for Analysis

The hypotheses, in turn, serve to guide subsequent decisions in the methodology. For example, they determine the nature of the data that would be required to test the hypotheses. In Dewdney (1992) it is clear that, to test her hypothesis, she needs to collect reference interviews under different situations: from librarians with training (1) before and (2) after the training, and (3) from librarians with no direct training.

Sampling

A major objective of social science research is generalizability, that is, the ability to generalize from the specific to the general—for example, to study the sample but infer from the sample's findings something about the population from which the sample is drawn. With a relatively nonstratified population, the ideal is random sampling, that is, sampling in which the probability of any unit within the population being selected is the same. To do this effectively, it is essential to know all units that exist in the population, such as all research articles published during a particular time period within a set of journals (Stansbury, 2002). Sometimes it is not possible to know all units beforehand, but a list can be generated as the sample is drawn. For example, to obtain a representative sample, randomly selected, from messages on two electronic lists and to ensure that the sampling period was sufficiently long to allow for getting a range of topics, messages, and participants, Schoch and White (1997) first did a preliminary study, based on archives of the lists, to determine the rate of messaging per list, or the average number of messages per month. At the start of the data-gathering period, all messages were downloaded and numbered separately for each list, and a sample of 1,000 messages was randomly chosen from the first 3,000 messages on each list written from the onset of the data-gathering period. Based on the messaging rate, the data-gathering period should have lasted approximately two months, but, because the rate of messaging actually varied across the two lists, data-collecting continued slightly longer in one list than in the other to achieve the same number of messages per list.

Coding

In quantitative content analysis the coding scheme is determined a priori, that is, before coding begins. A coding scheme operationalizes concepts that may in themselves be amorphous. It establishes categories that are relevant and valid. Relevant means that they allow for testing the hypotheses. Validity refers to "the extent to which a measuring procedure represents the intended, and only the intended, concept" (Neuendorf, 2002, p. 112). Validity can be assessed in several ways. Face validity, which is common in content analysis, refers to the extent to which a measure "gets at" the essential aspects of the concept being measured. Face validity is inherently subjective. To determine face validity, researchers assess as objectively as possible the correspondence between what they measure and how they measure it. One way of corroborating face validity is to have judges work backwards from the measure to determine the concept being measured (Neuendorf, 2002, p. 115). Other means of assessment are criterion validity, which relies on assessing the correspondence between the code and criteria, such as concurrent or predictive behavior or norms of behavior; content validity, which looks at the completeness of representation of the concept; and construct validity, which refers to "the extent to which

a measure is related to other measures (constructs) in a way consistent with hypotheses derived from theory" (Neuendorf, 2002, p. 117). Construct validity is more difficult to assess than criterion or content validity but is a worthy objective.

In addition, a good coding scheme has categories or levels that are exhaustive, that is, all relevant aspects of the construct are represented, are mutually exclusive, and are measured at the highest possible scale of measurement based on the four scales of measurement (nominal, ordinal, interval, and ratio).[4] The coding scheme should have clear definitions, easy-to-follow instructions, and unambiguous examples. All of these features promote the reliability of the coding, that is, the likelihood that all coders will code the same item the same way or that a coder will code the same item the same way at different points in time.[5] (For examples of coding schemes, see Haas & Grams, 2000, pp. 191–192; Hahn, 1999, Appendix B, pp. 229–237; Kracker & Wang, 2002, Appendices A-C, pp. 304–305; and Marsh & White, 2003, pp. 666–672.) If the coding scheme is modified during the coding, it must be re-applied to the data already coded so that all data are coded according to the same coding scheme.

The complexity of the coding scheme varies, and individual codes may be combined after the coding to develop a composite measurement, such as an index, or otherwise grouped to show relationships among the measures. Kracker and Wang (2002), for example, initially identified affective words that expressed emotions and subsequently clustered the categories into an affective classification scheme indicating negative and positive clusters for three major areas. Marsh and White (2003) grouped the image-text relationships into three categories: functions expressing little relation to the text; functions expressing close relation to the text; and functions going beyond the text.

Many content analysis studies do not develop their own coding scheme but rely instead on coding schemes devised by other researchers. Stansbury (2002) used the problem statement attributes identified by Hernon and Metoyer-Duran (1993) as a code for analyzing problem statements in LIS journals. Maloney-Krichmar and Preece (2005) and Schoch and White (1997) used Bales's Transactional Analysis Schema (Bales, 1951) to analyze messages on consumer health electronic lists. Using the same coding scheme across studies allows for easy comparisons among the studies. For example, after applying Graesser's Typology of Questions (Graesser, McMahen, & Johnson, 1994) to questions in reference interviews, White (1998) compared the incidence of questions and types of questions in reference interviews with similar question incidence data in tutoring sessions and decision support systems. In another study (White, 2000) coding the content of questions on consumer-health electronic lists with Roter's (1984) typology of questions in physician-patient interactions allowed for comparisons across the two settings. The last column in Table 1 shows

the content analytic schemes from other researchers used in quantitative content analysis studies.

Several coding schemes developed by LIS researchers have potentially broad use in LIS: (1) Haas and Grams' (1998a, 1998b, 2000) taxonomies for Web pages and links; (2) the two sets of categories developed by Kracker and Wang (2002) reflecting affective and cognitive aspects of Kuhlthau's (1993) Information Search Process (ISP) model; and (3) Marsh and White's (2003) taxonomy for analyzing text-image relationships in a variety of settings.

Just because coding schemes are developed a priori does not mean that the instances of the categories become immediately obvious and, as a result, easy to code. As in qualitative content analysis, the analysis often requires careful, iterative reading of the text. Marsh and White (2003) include several examples of image-text pairs, their codes, and the thinking surrounding coding each pair with their taxonomy of image-text relationships. These examples illustrate the complexity and depth of thinking that may be necessary in coding, even with an a priori coding scheme.

Analyzing the Coded Data

After the coding, which in itself is analytical, the researcher undertakes several additional steps. These steps, too, are done within the framework of the hypotheses or research questions. First, he[6] summarizes the findings identified during the coding, formulating and restating them so that they can be understood easily and are applicable to his hypotheses or research questions. Second, he identifies and articulates the patterns and relationships among his findings so that he can test his hypotheses or answer his research questions. Finally, he relates these more involved findings to those in other situations or other studies. The last step allows him to put his findings into perspective.

In the analysis, the content analyst chooses from among a variety of statistical approaches or techniques for presenting and testing his findings. They range in complexity and demands for different scales of measurement for the variables. The approach he selects takes into consideration not only the questions he is addressing but also the nature of the data and may include tabulations; cross-tabulations, associations, and correlations; multivariate techniques, such as multiple regression analysis; factor analysis and multidimensional scaling; images, portrayals, semantic nodes, and profiles; contingencies and contingency analysis; and clustering. Often, decisions about using these techniques are made in the planning phase of the project since they influence and build on decisions that, of necessity, must occur earlier in the project, such as establishing the level of measurement for a particular variable. The output of these techniques can be presented, in most cases, both in tabular and graphic form. Not all of these techniques are used in the LIS content analysis studies in Table 1. Tabulations, cross-tabulations, associations, and correlations are common (see, for example,

Schoch & White, 1997; Stansbury, 2002; White, 1998). White, Abels, and Gordon-Murnane (1998) use clustering techniques to develop a typology of innovators in a study of the content of publishers' Web sites and use it to profile publishers along a spectrum from traditionalist to innovator.

Procedures: Qualitative Content Analysis

Proponents of qualitative and quantitative content analysis often emphasize their differences, yet many similarities exist as well. Noting four common elements, Krippendorff, who covers both variants in his text, points out "the proponents of both approaches: [1] sample text, in the sense of selecting what is relevant; [2] unitize text, in the sense of distinguishing words or propositions and using quotes or examples; [3] contextualize what they are reading in light of what they know about the circumstances surrounding the text; and [4] have specific research questions in mind" (2004, p. 87). Table 2 characterizes the two types of content analysis along several dimensions. The most significant differences are the foci of this section.

Formulating Research Questions

In contrast with quantitative content analysis, qualitative content analysis flows from a humanistic, not a positivistic, tradition. It is inductive. Qualitative content analysis may yield testable hypotheses but that is not its immediate purpose. Replacing the hypotheses are foreshadowing questions, that is, open questions that guide the research and influence the data that are gathered. In qualitative content analysis, however, the text plays a slightly different role in that, as the researcher reads through the data and scrutinizes them closely to identify concepts and patterns, some patterns and concepts may emerge that were not foreshadowed but that are, nevertheless, important aspects to consider. In that case, the researcher may legitimately alter his interests and research questions to pursue these new patterns. For example, in Hahn's study of the author and editor as early adopters of electronic journals, she initially had three open, foreshadowing research questions, based, to some extent, on diffusion theory (Rogers, 1995): "1) How do authors and editors working closely with an electronic journal perceive electronic journals?; 2) What is the decision process that authors are using to decide to publish in an electronic journal?; 3) How do social factors influence the adoption decision?" (Hahn, 1999, p. 6). As her coding and analysis evolved, she added: "4) What key relations between the scientific community and the publishing system are affected by electronic publishing?" (p. 122). Krippendorff refers to this iterative process of "recontextualizing, reinterpreting, and redefining the research until some kind of satisfactory interpretation is reached" (2004, pp. 87–88) as a *hermeneutic loop*. This procedure may actually occur in quantitative content analysis studies but only at the development phase of the research design; the development phase is followed by adherence to the practices specified earlier.

Table 2. Characteristics of Quantitative and Qualitative Content Analysis

Category	Quantitative	Qualitative
Research approach	Deductive; based on previous research, which allows for formulating hypotheses about relationships among variables	Inductive; research questions guide data gathering and analysis but potential themes and other questions may arise from careful reading of data
Research tradition or orientation	Positivist	Naturalist or humanist; hermeneutics
Objective	To make "replicable and valid inferences from texts . . . to the contexts of their use" (Krippendorff, 2004, p. 19)	"To capture the meanings, emphasis, and themes of messages and to understand the organization and process of how they are presented" (Altheide, 1996, p. 33); "Search for multiple interpretations by considering diverse voices (readers), alternative perspectives (from different ideological positions), oppositional readings (critiques), or varied uses of the texts examined (by different groups)" (Krippendorff, 2004, p. 88)
Data: Nature	Syntactic, semantic, or pragmatic categories; naturally occurring texts or text generated for project	Syntactic, semantic, or pragmatic categories; naturally occurring texts or text generated for project
Data: Selection	Systematic, preferably random, sampling to allow for generalization to broader population; data selection usually complete prior to coding	Purposive sampling to allow for identifying complete, accurate answers to research questions and presenting the big picture; selection of data may continue throughout the project
Categorization schema	Coding scheme developed a priori in accord with testing hypotheses; if adjustments are made during coding, items already coded must be recoded with the revised scheme; may use coding scheme(s) from other studies	Coding scheme usually developed in the process of close, iterative reading to identify significant concepts and patterns
Coding	Objective; tests for reliability and validity	Subjective; in some cases, use of memos to document perceptions and formulations; techniques for increasing credibility, transferability, dependability, and confirmability of findings

Table 2. Characteristics of Quantitative and Qualitative Content Analysis

Category	Quantitative	Qualitative
Argument basis for proof	Frequency, indicating existence, intensity, and relative importance; data allow for statistical testing of hypotheses; objectives are usually to generalize to broader population and to predict; interpretations may be supported by quotations from text	Deep grounding in the data; if numbers are presented, they are usually presented as counts and percentages; description of specific situation or case accurately and thoroughly; may involve triangulation based on multiple data sources for same concept; may use techniques to develop grounded theory to relate concepts and to suggest hypotheses that can be tested deductively; presentation "Support[s] interpretations by weaving quotes from the analyzed texts and literature about the contexts of those texts into their conclusions, by constructing parallelisms, by engaging in triangulations, and by elaborating on any metaphors they can identify" (Krippendorff, 2004, p. 88)
Use of computers	For dictionary-based content analysis or for developing environments prior to dictionary-based content analysis; also for statistical tests; representative software for content analysis: VBPro, WordStat	As annotation and searching aids; representative software: Atlas. TI or NVivo

Sampling

Both qualitative and quantitative content analysis researchers sample text and choose text that is relevant for their purpose, but qualitative researchers focus on the uniqueness of the text and are consciously aware of the multiple interpretations than can arise from a close perusal of it. The need for close, reiterative analysis itself usually limits the size of the sample.

In addition, since the object of qualitative research is not generalizability but transferability, sampling does not need to insure that all objects being analyzed have an equal or predictable probability of being included in the sample. Transferability refers to a judgment about whether findings from one context are applicable to another. Instead, the sampling should be theoretical and purposive. It may have as its objective providing the basis for identifying all relevant patterns in the data or characterizing a phenomenon. It may even present the findings quantitatively through numbers

and percentages but not through inferential statistics. Some cases may be selected prior to initiating coding, but the selection and coding may also occur in tandem, with subsequent case selection influenced by discoveries during the coding process. Analyzing new cases may continue until no new patterns or findings related to the concept under analysis become apparent in the coding process. If no new patterns are being found, generally the presumption is that all relevant patterns have been discovered and additional work would only confirm that finding. If at this point there is interest in noting the prevalence of a particular pattern, the researcher may move to using the pattern or range of patterns as a coding scheme and analyzing a body of documents. But, because the sampling is purposive, the researcher cannot extrapolate from the sample to the population.

Coding

For qualitative coding, the researcher's initial foci are not a priori codes but the initial foreshadowing questions he aims to answer through his research. The questions guide his initial approach to the data, but the process is inductive, not deductive. The evidence plays almost as significant a role in shaping the analysis as the initial questions. It is not unusual to have a person doing qualitative content analysis read through the data initially with the intent of trying to see the big picture. As he reads through the documents, he begins to tag key phrases and text segments that correspond to those questions, notes others that seem important but are unexpected, sees similarities in expressing the same concept, and continues iteratively to compare the categories and constructs that emerge through this process with other data and re-reading of the same documents. In the process, he may be looking for diversity of ideas, alternative perspectives, oppositional writings, and/or different uses of the texts, perhaps by different groups.

Data collection units and units of analysis vary. The researcher continually checks his growing interpretation of answers to his research questions against the documents and notes, especially situations that do not fit the interpretation or suggest new connections. In this way, he looks not only at confirming evidence of his emerging construct(s) but also at disconfirming evidence that needs to be considered as he presents his case for his interpretation. The overall process may suggest new questions that were not anticipated at the start of the analysis. Glaser and Strauss (1967) refer to the *constant comparison* approach to data analysis, in which the emerging relationships and categories are continually refined and emerging theory or patterns tested as new data are compared with old (see also Boeije, 2002).

To keep track of the developing concepts and the models that are emerging about how the concepts relate to each other, the researcher records his decisions and comments in memos. Two types of memos are common: concept memos, which logically focus on emerging concepts, the distinctive ways in which these are phrased, and his own interpretation of the con-

cepts; and theory memos, in which he focuses on relationships among the concepts and gradually integrates these concepts into a workable model. Memos reveal the subtleties of the researcher's interpretation and understanding of the constructs over time. In a conceptual memo, for example, Hahn (1998) comments:

> Thinking over some of the features of discussions that I feel are recurring but not previously captured by existing coding structures, I initially considered the concept of advantages and disadvantages. However, it seems like a more useful organizing conceptual structure is one of optimizing characteristics. The idea is that these are characteristics of the journal perceived by the community. The editors and publishers try to optimize these to encourage both submissions and readership. Authors also try to make an optimal match with these characteristics given the nature of the paper they have in hand ready for submission. (n.p.)

Qualitative content analysis has developed approaches similar to validity and reliability for assessing the rigor of the coding and analysis process. Qualitative content analysis focuses on creating a picture of a given phenomenon that is always embedded within a particular context, not on describing reality objectively. Lincoln and Guba (1985) describe four criteria used to assess the degree to which a qualitative study will have "truth value," that is, "confidence in the 'truth' of the findings of a particular inquiry" (Guba & Lincoln, 1981, p. 246): credibility, transferability, dependability, and confirmability.[7] Credibility, the equivalent of internal validity, calls for identifying all important factors in the research question and accurately and completely describing the ways in which these factors are reflected in the data gathered. Transferability, or external validity, is essentially a judgment about the applicability of findings from one context to another. Generally a qualitative researcher tries to situate his findings within a relevant theoretical paradigm, understanding that findings sensible within it can be applied to other, comparable contexts with greater confidence. Similarly, he usually tries to collect data on a single factor or question aspects from multiple sources with the understanding that findings based on multiple data sources can be transferred with greater confidence. Collecting, analyzing, and cross-checking a variety of data on a single factor or aspect of a question from multiple sources, and perhaps perspectives, as Buchwald (2000) did, is termed *triangulation* and is a way to heighten a qualitative study's credibility and confirmability.

Dependability addresses the notion of replicability and defines it as "stability after discounting . . . conscious and unpredictable (but rational and logical) changes" (Guba & Lincoln, 1981, p. 247) in findings during repetitions of the study. Confirmability relates to objectivity and is measured in quantitative content analysis by assessing inter-rater reliability. In qualitative research findings are confirmed by looking at the data, not the researcher(s), to determine if the data support the conclusions. The

important criterion is not numeric correspondence between coders but conceptual consistency between observation and conclusion.

Method of Analysis

Analysis is integrated into coding much more in qualitative content analysis than in quantitative content analysis. The emphasis is always on answering the research questions but considering as well any transformations that the initial foreshadowing questions may have undergone during the coding or any new questions or themes that emerge during the coding. Often the result of qualitative analysis is a composite picture of the phenomenon being studied. The picture carefully incorporates the context, including the population, the situation(s), and the theoretical construct. The goal is to depict the "big picture" of a given subject, displaying conceptual depth through thoughtful arrangement of a wealth of detailed observations.

In presenting the results the researcher may use numbers and/or percentages, either in simple tabulations or in cross-tabulations to show relationships, but he may also rely simply on the gradual accretion of details within his textual presentation without resort to numbers. Often the analysis results in both graphic and tabular presentation of models elicited during the analysis. Wang and White (1999), for example, present a graphic model of document use at three different stages in a research project, showing the criteria and decision rules the researchers applied at each stage (see Figure 6, p. 109). This table incorporates data from a previous study, which covered the first stage (Wang & Soergel, 1998), and is supported in the second study by data in Tables 2 and 4 (Wang & White 1999, pp. 104, 107, respectively) for criteria and decision rules in the second and third stages, respectively. The tables present, for each criterion and decision rule, the number of users mentioning each and the number of documents about which they were mentioned.

The text may be a narrative of findings about the phenomenon being studied with quotations to illustrate the conclusions. In the same study, for example, the authors refer to the participants' use of reputation as a criterion in determining relevance:

> Participants comment on whether or not the document is written by a reputable author or organization or published in a reputable journal. An example related to the document's author is "It is by a very minor person, X; Y [co-author] is a little better known than X. I know them by reputation. I don't know them personally." Another example comments on the authority of the publisher or the author's affiliation: "I was looking for something which wouldn't have a bias. The World Bank is accepted by all countries. We already know that the World Bank is very involved in sending technical support or funding for such projects" (Wang & White, 1999, p. 105).

Ahuvia (2001) suggests that reviewers can better judge the confirmability or public credibility of a qualitative content analysis if the researcher submits his original data set, codings, and justification for particular codes if necessary along with a manuscript. In a published study, the data, or at least a random subset, can be included as an appendix.

Using Computer Software

Depending on the number of documents, content analysis can be tedious and benefits enormously from the use of computers for a variety of tasks. Collectively, the software programs serve in several capacities:

- As a research assistant, making it easy to markup the data, divide them into chunks for analysis, write notes, group together multiple instances of the same classification, and allow for global editing and coding.
- As a manipulator and extractor of data, matching the text against specialized dictionaries for coding purposes.
- As data collections, maintaining the electronic and coded versions, keeping track of all steps in the analysis, and, in the latter case, allowing for replicating the analysis.
- As a means for doing or facilitating quantitative analyses, such as frequency counts and percentages, either within the program itself or by exporting data to statistical packages, thereby eliminating errors that would occur in multiple inputs of the data. The statistical packages would usually allow for inferential statistics. (Mayring, 2000)

The programs arrange themselves on a spectrum from simply facilitating a human's coding of the electronic data to direct involvement in analyzing the document; matching terms to an electronic dictionary, which is a coding scheme; and coding the data. In the latter human input occurs primarily in developing the dictionary and in interpreting the results of the coding. In the middle is a set of programs that facilitates developing the dictionaries used in the latter. Lowe (2002) refers to these respectively as annotation aids, developing environments, and programs for dictionary-based content analysis. Examples of the first are NVivo (2003–2005), QSR N6 (2005) and Atlas-TI (Muhr, 2004). These programs now allow for storing not only textual documents but also images and audio in electronic form. Qualitative content analysis relies more on annotation aids. Dictionary-based content analysis programs rely on several basic functions: word and category counts and frequency analyses, visualization (including clustering), and sometimes concordance generation. DIMAP-4 (Litkowski & McTavish, 2001) and KEDS (Schrodt, 1996), and TABARI (Schrodt, 2000) are examples of developing environments. WordStat 5.0 (Peladeau, 2005), VBPro (Miller, 2003), and the General Inquirer (Stone, 2002; The General Inquirer, n.d.) are examples of dictionary-based content analysis programs. LIS researchers do not always identify the software used in analyses. Agosto

and Hughes-Hassell (2005) mention NVivo; Marsh (2002) uses Atlas-TI; White (1998) and Kracker and Wang (2002) use QSR NUD*IST, renamed, in its latest version, as QSR N6.

Conclusion

Content analysis is a highly flexible research method that has been widely used in LIS studies with varying research goals and objectives. The research method is applied in qualitative, quantitative, and sometimes mixed modes of research frameworks and employs a wide range of analytical techniques to generate findings and put them into context. The LIS studies referred to in this article are not always purist but occasionally use a hybrid approach, incorporating elements of qualitative and quantitative content analysis for good reason. This article characterizes content analysis as a systematic, rigorous approach to analyzing documents obtained or generated in the course of research. It briefly describes the steps involved in content analysis, differentiates between quantitative and qualitative content analysis, and shows that content analysis serves the purposes of both quantitative research and qualitative research. In addition, the article serves as a gateway to selected LIS studies that have used content analysis and to methodological books and articles that provide more detail about aspects of content analysis discussed only briefly here.

Bibliographic Note

Two recent texts on content analysis are Krippendorff (2004) and Neuendorf (2002). Krippendorff covers both quantitative and qualitative content analysis; Neuendorf focuses on quantitative content analysis. Neuendorf (2005) maintains a text-related Web site with many useful resources: the Content Analysis Guidebook Online (http://academic.csuohio.edu/kneuendorf/content). Titscher, Meyer, Wodak, and Vetter (2000) provide chapters for specific types of textual analysis not covered in this article; Schiffrin (1994) discusses various types of discourse analysis. Additional useful methodological chapters are Bell (2001) and Collier (2001) for content analysis of visual images and Evans (2002) for dictionary-based content analysis.

Articles reviewing software are useful but become dated quickly; Skalski's (2002) review in Neuendorf's (2002) text has a tabular presentation of software features in addition to paragraphs describing about twenty programs; his table establishes a useful framework for evaluating software. Several Web sites maintain more current reviews and/or links to content analysis software publisher pages. See, for example, the "Classification of Text Analysis Software" section of Klein's (2002–2003) Text Analysis Info Page (http://www.textanalysis.info) and the content analysis resources listed under the software section of Evans's (n.d.) Content Analysis Resources (http://www.car.ua.edu). Krippendorff's (2004) chapter 12 on computer aids is also useful for showing how computers can aid content analysis.

The Web sites mentioned above (Neuendorf, Klein, and Evans) are the most useful for content analysis researchers. Contents analysis researchers in all fields communicate informally via the Content Analysis News and Discussion List (2006) (content@bama.ua.edu). Its archives are available at http://bama.ua.edu/archives/content.html.

Notes

The authors are grateful to Karla Hahn for permitting a quotation from a concept memo, Susan Davis for comments, and the authors whose works are mentioned in this article for their careful and clear presentation of their methodology.

1. Berelson's (1952) *Content Analysis in Communications Research* is considered the "first systematic presentation" of the conceptual and methodological elements of content analysis and "codified the field for years to come" (Krippendorff, 2004, p. 8).
2. For a useful discussion and explanation of each type, see Krippendorff (2004), Schiffrin (1994), and Titscher, Meyer, Wodak, and Vetter (2000). Titscher et al. includes a map of theories and methods that is notable for illustrating relationships among them (Figure 4.1, p. 51).
3. Krippendorff's (2004) text considers both quantitative and qualitative content analysis. Another recent text by Neuendorf (2002) focuses on quantitative content analysis.
4. Any statistics text should discuss scales of measurement. See, for example, StatSoft, Inc.'s (2004) *Electronic Statistics Textbook.*
5. See Lombard, Snyder-Duch, and Bracken's (2005) *Practical Resources for Assessing and Reporting Intercoder Reliability in Content Analysis Research Projects.* This paper is invaluable in discussing the reasons for assessing and reporting intercoder reliability, the proper steps involved in doing so, the preferred statistical tests, and the information to be reported, among other topics. Krippendorff (2004) also includes useful sections on reliability (chap. 11, pp. 211–256) and validity (chap. 13, pp. 313–338).
6. Throughout this article, when *he, his,* and *him* are used without the context of a specific researcher, they refer to researchers of both genders.
7. Lincoln and Guba (1985) apply these to qualitative research studies generally, not just to coding, but they are also applicable in the narrower context.

References

Agosto, D. E. & Hughes-Hassell, S. (2005). People, places, and questions: An investigation of the everyday life information-seeking behaviors of urban young adults. *Library & Information Science Research, 27,* 141–163.

Ahuvia, A. (2001). Traditional, interpretive, and reception based content analyses: Improving the ability of content analysis to address issues of pragmatic and theoretical concern. *Social Indicators Research, 54,* 139–172.

Altheide, D. L. (1996). *Qualitative media analysis.* Thousand Oaks, CA: Sage.

Bales, R. F. (1951). *Interaction process analysis: A method for the study of small groups.* Cambridge, MA: Addison-Wesley Press.

Beaugrande, R. D., & Dressler, W. U. (1981). *Einführung in die textlinguistik.* Tübingen: Niemeyer.

Bell, P. (2001). Content analysis of visual images. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 10–34). Thousand Oaks, CA: Sage.

Benne, K. D., & Sheats, P. (1948). Functional roles of group members. *Journal of Social Issues, 41,* 41–49.

Berelson, B. (1952). *Content analysis in communications research.* New York: Free Press.

Boeije, H. (2002). A purposeful approach to the constant comparison method in the analysis of qualitative interviews. *Quality & Quantity, 36,* 391–409.

Buchwald, C. C. (2000). A case study of Canada's Coalition for Public Information in the information highway policy-making process. *Library & Information Science Research, 22,* 123–144.

Cicerello, A., & Sheehan, E. P. (1995). Personal advertisements: A content analysis. *Journal of Social Behavior and Personality, 10,* 751–756.

Collier, M. (2001). Approaches to analysis in visual anthropology. In T. Van Leeuwen & C. Jewitt (Eds.), *Handbook of visual analysis* (pp. 35–60). Thousand Oaks, CA: Sage.

Content Analysis News and Discussion List. (2006). Archives of CONTENT@BAMA.UA.EDU. Retrieved March 1, 2006, from http://bama.ua.edu/archives/content.html.

Croneis, K., & Henderson, P. (2002). Electronic and digital librarian positions: A content analysis of announcements from 1990 through 2000. *Journal of Academic Librarianship, 28,* 232–237.

Dewdney, P. (1992). Recording the reference interview: A field experiment. In J. D. Glazier & R. R. Powell (Eds.), *Qualitative research in information management* (pp. 122–150). Englewood, CO: Libraries Unlimited.

Dilevko, J., & Gottlieb, L. (2004). The portrayal of librarians in obituaries at the end of the twentieth century. *Library Quarterly, 74*(2), 152–180.

Evans, W. (2002). Computer environments for content analysis: Reconceptualizing the role of humans and computers. In O. V. Burton (Ed.), *Computing in the social sciences and humanities* (pp. 67–86). Urbana, IL: University of Illinois Press.

Evans, W. (n.d.). *Content analysis resources.* Available at the University of Alabama. Retrieved March 1, 2006, from http://www.car.ua.edu.

The General Inquirer [computer software]. (n.d.). Princeton, NJ: The Gallup Organization. Retrieved November 2, 2005, from the WebUse Project at the University of Maryland Web site, http://www.webuse.umd.edu:9090.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine Press.

Graesser, A. C., McMahen, C. L., & Johnson, B. K. (1994). Question asking and answering. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 517–523). San Diego, CA: Academic Press.

Green, R. (1991). The profession's models of information: A cognitive linguistic analysis. *Journal of Documentation, 47,* 130–148.

Guba, E. G., & Lincoln, Y. S. (1981). Epistemological and methodological bases of natural inquiry. *Educational Communication & Technology Journal, 30*(4), 233–252.

Haas, S. W. & Grams, E. S. (1998a). Page and link classifications: Connecting diverse resources. In I. H. Witten, R. M. Akscyn, & F. M. Shipman (Eds.), *Proceedings of Digital Libraries '98—Third ACM Conference on Digital Libraries, June 23–26, 1998, Pittsburgh, PA* (pp. 99–107). New York: Association for Computing Machinery.

Haas, S. W. & Grams, E. S. (1998b). A link taxonomy for Web pages. In C. M. Preston (Ed.), *Information access in the global information economy, Proceedings of the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA, October 26–29, 1998* (pp. 485–495). Medford, NJ: Information Today.

Haas, S. W., & Grams, E. S. (2000). Readers, authors, and page structure: A discussion of four questions arising from a content analysis of Web pages. *Journal of the American Society for Information Science, 51*(2), 181–192.

Hahn, K. (1998). *Memo: Optimizing characteristics of journal titles.* Unpublished memo provided to author, October 2005.

Hahn, K. (1999). *Electronic journals as innovations: A study of author and editor early adopters.* Unpublished doctoral dissertation, University of Maryland at College Park, Maryland.

Hernon, P., & Metoyer-Duran, C. (1993). Problem statements: An exploratory study of their function, significance, and form. *Library & Information Science Research, 15*(1), 71–92.

Jennerich, E. Z. (1974). *Microcounseling in library education.* Unpublished doctoral dissertation, University of Pittsburgh, 1974.

Klein, H. (2002–2003). *Welcome to text analysis info.* Rudolstat: Social Science Consulting. Retrieved November 2, 2005, from http://www.textanalysis.info.

Kracker, J., & Wang, P. (2002). Research anxiety and students' perceptions of research: An experiment. Part II. Content analysis of their writings on two experiences. *Journal of the American Society for Information Science & Technology, 53*(4), 294–307.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Thousand Oaks, CA: Sage.

Kuhlthau, C. (1993). *Seeking meaning: A process approach to library and information services.* Norwood, NJ: Ablex.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry.* Beverly Hills, CA: Sage.

Litkowski, K., & McTavish, D. (2001). DIMAP-4 (DIctionary MAintenance Programs) [computer software]. Damascus, MD: CL Research.

Lombard, M., Snyder-Duch, J., & Bracken, C.C. (2005). *Practical resources for assessing and reporting intercoder reliability in content analysis research projects.* Retrieved May 27, 2005, from Temple University, School of Communications and Theater, Doctoral Program in Mass Media & Communication, http://www.temple.edu/mmc/reliability.

Lowe, W. (2002). *Software for computer analysis—A review.* Retrieved October 14, 2005, from Harvard University, The Institute for Quantitative Social Science, http://people.iq.harvard.edu/~wlowe/Publications/rev.pdf.

Lynch, B. P., & Smith, K. R. (2001). The changing nature of work in academic libraries. *College & Research Libraries, 62*(5), 407–420.

Maloney-Krichmar, D. & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction, 12*(2), 201–232.

Marsh, E. E. (2002). *Rhetorical relationships between images and text in Web pages.* Unpublished doctoral dissertation, University of Maryland at College Park, Maryland.

Marsh, E. E., & White, M. D. (2003). A taxonomy of relationships between images and text. *Journal of Documentation, 59*(6), 647–672.

Mayring, P. (2000). Qualitative content analysis. *Forum Qualitative Social Research/Forum Qualitative Sozialforschung, 1*(2). Retrieved September 24, 2005, from http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-e.htm.

Miller, M. M. (2003). VBPro (Verbatim Protocol) [computer software]. Freeware.

Muhr, T. (2004). Atlas.ti5 [computer software]. Berlin: ATLAS.ti Scientific Software Development GmbH.

Neuendorf, K. A. (2002). *The content analysis guidebook.* Thousand Oaks, CA: Sage.

Neuendorf, K. A. (2005). *The content analysis guidebook online: An accompaniment to the content analysis guidebook by Kimberly A. Neuendorf.* Cleveland, OH: Cleveland State University. Retrieved March 1, 2006, from http://academic.csuohio.edu/kneuendorf/content.

Nitecki, D. A. (1993). Conceptual models of libraries held by faculty, administrators, and librarians: An exploration of communications in the *Chronicle of Higher Education*. *Journal of Documentation, 49*(3), 255–277.

NVivo 2.0 [computer software]. (2003–2005). Doncaster, Victoria, Australia: QSR International.

Peladeau, N. (2005). WordStat 5.0 [computer software]. Montreal: Provalis Research.

QSR N6 [computer software]. (2005). Durham, UK: QSR Software.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York: Free Press.

Roter, D. L. (1984). Patient question asking in physician-patient interaction. *Health Psychology, 3*(5), 395–409.

Schiffrin, D. (1994). *Approaches to discourse.* Cambridge, MA: Blackwell.

Schoch, N. A., & White, M. D. (1997). A study of the communications patterns of participants in consumer health electronic discussion groups. In C. Schwartz, & M. Rorvig (Eds.), *Digital collections, implications for users, funders, developers, and maintainers, Proceedings of the 60th Annual Meeting of the American Society for Information Science, Washington, DC, November 1–6, 1997* (pp. 280–292). Medford, NJ: Information Today.

Schrodt, P. A. (1996). KEDS (Kansas Event Data System) [computer software]. Lawrence: Department of Political Science, University of Kansas.

Schrodt, P. A. (2000). TABARI 0.5.1 (Textual Analysis By Augmented Replacement Instructions) [computer software]. Lawrence: Department of Political Science, University of Kansas.

Skalski, P. (2002). Computer content analysis software. In K. A. Neuendorf, *The content analysis guidebook* (pp. 225–235). Thousand Oaks, CA: Sage.

Stansbury, M. C. (2002). Problem statements in seven LIS journals: An application of the Hernon/Metoyer-Duran attributes. *Library & Information Science Research, 24*(2), 157–168.

StatSoft, Inc. (2004). *Electronic Statistics Textbook.* Tulsa, OK: StatSoft. Retrieved October 31, 2005, from StatSoft, Inc., http://www.statsoft.com/textbook/stathome.html.

Stone, P. J. (2002). *Welcome to the General Inquirer Home Page.* Cambridge, MA: Harvard College. Retrieved March 1, 2006, from http://www.wjh.harvard.edu/~inquirer.

Titscher, S., Meyer, M., Wodak, R., & Vetter, E. (2000). *Methods of text and discourse analysis* (B. Jenner, Trans.). Thousand Oaks, CA: Sage.

Wang, J., & Gao, V. (2004). Technical services on the net: Where are we now? A comparative

study of sixty Web sites of academic libraries. *Journal of Academic Librarianship, 30*(3), 218–221.

Wang, P., & Soergel, D. (1998). A cognitive model of document use during a research project. Study I. Document selection. *Journal of the American Society for Information Science, 49*(2), 115–133.

Wang, P., & White, M. D. (1999). A cognitive model of document use during a research project. Study II. Decisions at the reading and citing stages. *Journal of the American Society for Information Science, 50*(2), 98–114.

White, M. D. (1985). Evaluation of the reference interview. *RQ, 25,* 76-84.

White, M. D. (1998). Questions in reference interviews. *Journal of Documentation, 54*(4), 443–465.

White, M. D. (2000). Questioning behavior on a consumer health electronic list. *Library Quarterly, 70*(3), 302–334.

White, M. D., Abels, E. G., & Agresta, J. (2004). The relationship between interaction characteristics and answer quality in chat reference service. In *Online proceedings of the Virtual Reference Desk Conference, Cincinnati, OH, November 8–9, 2004.* Retrieved November 2, 2005, from http://www.vrd.org/conferences/VRD2004/proceedings/presentation.cfm?PID=376.

White, M. D., Abels, E. G., & Gordon-Murnane, L. (1998). What constitutes adoption of the Web: A methodological problem in assessing adoption of the World Wide Web for electronic commerce. In C. M. Preston (Ed.), *Information access in the global information economy, Proceedings of the 61st Annual Meeting of the American Society for Information Science, Pittsburgh, PA, October 26–29, 1998* (pp. 217–226). Medford, NJ: Information Today.

White, M. D., Abels, E. G., & Kaske, N. (2003). Evaluation of chat reference service quality: Pilot study. *D-LIB Magazine, 9.* Retrieved November 2, 2005, from http://www.dlib.org/dlib/february03/white/02white.html.

White, M. D., & Iivonen, M. (2001). Questions as a factor in Web search strategy. *Information Processing & Management, 37,* 721–740.

White, M. D., & Iivonen, M. (2002). Assessing the level of difficulty of Web search questions. *Library Quarterly, 72*(2), 207–233.

Yuan, Z. (1996). Analysis of trends in demand for computer-related skills for academic librarians from 1974 to 1994. *College & Research Libraries, 57,* 259–272.

Marilyn Domas White is an associate professor in the College of Information Studies at the University of Maryland, College Park, where she teaches in the area of information access. Her Ph.D. is in library science from the University of Illinois. Her current areas of research are information behavior, especially questioning behavior; information access, especially to electronic images; and scholarly communication. Recent publications include "A taxonomy of relationships between images and text," *Journal of Documentation, 59*(2003), 647–672, with Emily E. Marsh; "Evaluation of chat reference service quality: Pilot study," *D-Lib Magazine, 9*(2003), with Eileen G. Abels and Neal Kaske; and "Assessing the level of difficulty of Web search questions," *Library Quarterly, 72*(2002), 207–233, with Mirja Iivonen. Until recently, she was co-PI for the Computational Linguistics for Metadata Building (CLiMB-2) Project at the University of Maryland.

Emily E. Marsh is a consultant in information science and has been an adjunct faculty member at the College of Information Studies at the University of Maryland, College Park, where she teaches in the area of user behavior. Her Ph.D. is in library and information science from the University of Maryland (2002). Her areas of research interest are information design, illustration, and research methods. Recent publications include "A taxonomy of relationships between images and text," *Journal of Documentation, 59*(2003), 647–672, with Marilyn Domas White.