Preprint of: Lavender, K., Nicholson, S., & Pomerantz, J. (2005, March - forthcoming). Building bridges for collaborative digital reference between libraries and museums through an examination of reference in special collections. *Journal of Academic Librarianship.*

Building Bridges for Collaborative Digital Reference between Libraries and Museums through an Examination of Reference in Special Collections

Kenneth Lavender, Research Scientist, Information Institute of Syracuse University
Scott Nicholson, Assistant Professor, Syracuse University School of Information Studies
Jeffrey Pomerantz, Assistant Professor, University of North Carolina at Chapel Hill School of Information and Library Science

ABSTRACT
While a growing number of the digital reference services in libraries have become part of collaborative reference networks, other entities that serve similar information-seeking needs such as special collections and museums have not joined these networks, even though they are answering an increasing number of questions from off-site patrons via the Internet. This article examines the differences between questions asked electronically of traditional reference services and those asked of special collections services; it further explores how a better understanding of digital reference in special collections will facilitate the development of the tools and models needed to create a bridge between digital human intermediation at general academic libraries, special collections, and museums.

INTRODUCTION
For many years, general reference departments in academic libraries accommodated distant users through acceptance of questions by letter or by phone. If by letter, the request was routed to the staff member best able or most available to answer it, and the reply was mailed back generally within few days time. If by phone, a library staff member answered the question while the user was "on hold" or arranged to have the answer delivered to the user at a later time, depending on the urgency of the question, the availability of staff time, and the number and type of resources involved. Today, although inquiries by letter or phone still exist, the great majority of off-site questions to general reference departments in academic libraries arrive via the Internet, which offers some of the functions of both these traditional modes of communication. Much like writing a letter, a patron may choose to compose a lengthy e-mail describing a research inquiry, which is then received by the reference service and sent to the best available staff member or referred to an Ask-An-Expert (commonly known as AskA) service or other collaborative reference service for asynchronous reference service . Or, just as with a telephone call, the user may receive synchronous reference service by using a chat-based or other live Web reference service. All of these methods of inquiry have in common

three fundamental factors: they accommodate the asker's need for information; they involve human intermediation on the part of the answerer (Lam, 2003)[1]; and they are concerned primarily with finding the answer, whatever the source.

From this context of general reference services a few generalizations may be derived:

1) the typical user is not concerned with a specific collection within a specific library, but rather with his or her information need and consequently with getting an answer to their question from any collection;
2) a collaborative general reference service is effective because the answer may be found in many sources, or because the relevant sources can be found in many subcollections, or because appropriate subject experts are available within the consortium;
3) human intermediation in search strategy development or in subject expertise is valuable to a patron.

The recent work by Pomerantz, Nicholson, and Lankes (2003)[2] codifies these generalizations from the perspective of establishing a partially automated collaborative reference service. In a Delphi study using a panel of reference experts, 15 factors were identified that were considered important by the majority of the respondents in the decision-making process for triaging questions to the appropriate subject specialists. Overall, those factors relating to sources and collections were rated lowest, and those relating to reference and subject expertise were rated highest. In addition, many factors relating to the user (as distinct from the question) were voted out early in the selection process. The study by Pomerantz, Nicholson, and Lankes thus emphasizes that for digital general reference services question content is not dependent upon the specific questioner and answer content is not dependent upon specific collections and sources. These considerations allow for the development of an effective triage system within a general reference service in an academic library or within a library consortium.

Benefits of Expanding Collaborative Systems for General Reference Services

Many digital reference services currently participate in question-swapping consortia that allow them to seamlessly refer patrons to other services. It would be beneficial to general reference services to have the ability to include subject matter experts from museums and special collections in these referral networks. One advantage of a pre-established network is that there is an agreed-upon protocol for passing questions and answers between services, and the developing NETREF standard will make this even easier.[3] Another is that the relationships between the library and the system have already been negotiated. When contacting a museum or special collection with a question, however, it can be difficult for another service to discover the appropriate entry point for submitting that question. Having a collaborative reference system in place that connects general academic libraries to museums gives libraries a pre-established technical and personal networking infrastructure to aid referral of patrons.

These bridges, however, have not yet been constructed, and initial attempts to include museums in the cooperative networks have not been very successful. One of the

problems is that the structure and purpose of libraries differ from those of museums; specifically, the well-supported reference departments in libraries do not always have an equivalent department in museums. This situation makes it challenging to understand how the question-answering function works in museums in order to create interoperable reference systems. For example, the recent ASIST Bulletin Special Section on Museum Informatics (2004) contains articles on the roles of information specialists in museums and how they must change to meet users' needs; however, there is little mention of the question-answering function of museum information professionals. Further, even though Coburn and Baca state that "it won't be long before the library, archive and museum communities can create successful models of interoperability and integrated access," their emphasis is on "standards and best practices of data-driving publishing," not on reference services.[4]

Special Collections as a Bridge between General Reference and Museums

What is needed to continue is a bridge between general reference desks and museums. Within many academic libraries, there is a "museum with the library" – the special collections department. This department usually has a well-established reference function while also having collections and exhibitions much like those in museums. Examining the questions asked of the reference department of special collections allows us to understand the type of questions typically asked of museums. How do these services compare with the generalizations outlined above for a general reference service, and, consequently, how may these services be accommodated in a collaborative reference system? The following analysis of digital reference questions received by special collections shows that these generalizations are not accurate in describing those users asking questions of special collections or in taking specialized materials or expertise into consideration.

There are significant intellectual and practical reasons for including special collections in a collaborative reference system, but the distinct differences between general and special collections must be accommodated for both to receive benefits from such a system. Within a library structure "special collections" most often signifies departments of rare books and archives, but as academic libraries also look to the outside for financial support, museums must often be included as partners in establishing interoperable systems that provide the broadest access to cultural resources. To build towards this outcome, the present paper discusses established methods of analysis for general reference transactions, applies these to the specific situation of special collections, and presents the challenges and rewards of creating a system that involves the broad spectrum of general and special collections electronic reference needs.

GENERAL REFERENCE TAXONOMIES

Classification is one of the fundamental tasks of library work. The entities that are classified, however, are usually physical objects: books, bound periodicals, maps, and a

variety of other materials. Once the leap was made in libraries to thinking about other types of artifacts (such as art and architectural objects, and archival objects) as entities within the purview of a library's collection, and therefore as entities to be classified, then it was a smaller step to thinking about questions (non-print and, indeed, immaterial pieces) as entities that could be classified.

Since that time, a number of classification schemes have been re-purposed or developed specifically for the task of classifying questions received by library reference services. Pomerantz reviewed the reference literature to identify classification schemes that have been used to classify reference questions[5]; the three most germane to this present study are:

1. Subjects of questions
2. Functions of expected answers to questions
3. Forms of expected answers to questions

Each of these three question classification schemes will now be explored in more detail.

Subjects of questions

Organization by subject has been a common means for classifying documents since Melvil Dewey first conceived of his subject scheme in 1873. Perhaps the earliest example of a classification scheme for questions dates back three-quarters of a century: Conner uses the ten main classes of the Dewey Decimal Classification (DDC) to classify questions recorded by the reference department of the Carnegie Library of Pittsburgh.[6] What is perhaps most interesting about Conner's classification is that she applies the same scheme used to classify materials in the library's collection to also classify reference questions. The assumption made by Conner in this approach to question-answering is that questions received by the reference desk are best thought of in terms of the arrangement of the library's collection.

R. S. Taylor[7]describes five "filters through which a question passes" in the mind of the reference librarian that enable the librarian to interpret the question, understand the patron's information need, and proceed to formulate an answer. The first of these steps is the determination of the subject of the question. Although Taylor does not discuss classification of questions by subject, such a classification scheme is implied by this first filter. As innovative as Taylor was in the reference community, still he was heir to the tradition of thinking about reference questions first and foremost in terms of their subjects.

To be fair, reference services to this day tend to think about reference questions primarily in terms of their subjects. Many instruments for collecting statistics about reference transactions require that the subjects of questions be recorded.[8, 9] Part of the function of these instruments is to collect data about subjects on which the reference service answers questions, in order to identify subjects on which the library may need to expand its collection, as well as to identify subjects on which it is difficult for reference librarians to

answer questions, or to which reference librarians frequently cannot give accurate answers.[10, 11] The organization of libraries according to subject classification schemes (LCSH, DDC, and others) tends to promote this mode of thinking about all materials in terms of their subjects.

The taxonomy of subjects of questions is the only one identified in the reference literature that is purely a classification scheme for questions. The following taxonomies actually classify answers, or, rather, they classify certain aspects of answers as they relate to the question. These classification schemes have been used to classify questions because questions generally do not exist in isolation in reference services; the purpose of a reference service is to provide answers to questions. An information-seeking question is an attempt by the questioner to elicit a response from the person questioned, and reference librarians are trained to think ahead to the answer when speaking to the patron about his or her question. It is therefore only natural that reference questions would come to be classified according to aspects of the answer, as the librarian expects it to take shape. The taxonomy of subjects of questions is an a priori classification scheme, in that a question can be accurately classified by subject before it is answered. The following two taxonomies are a posteriori classification schemes: a question can be classified by aspects of the expected answer before it is answered, but it cannot be accurately classified by aspects of the actual answer until after it is answered.

Functions of expected answers to questions

The taxonomy of the functions of expected answers classifies questions according to the possible functions of an answer in fulfilling a questioner's information need: verification of a fact, comparing or contrasting two objects, determining causality, etc. This classification scheme was originally developed not for library reference work, but rather for a story-understanding system named QUALM that attempted to automate the process by which humans understand and answer questions.[12] Subsequently, this taxonomy was adopted by Graesser and colleagues for several studies analyzing questions asked by individuals in a variety of real-world settings: while reading texts, while learning a new computer system, and while watching television news.[13]

Over time, Graesser and colleagues developed a theoretical model of question-asking behavior,[14] This taxonomy reached its most fully developed form in Graesser, McMahen, and Johnson.[15] In this developed form, this taxonomy is divided into classes that require short versus long answers, and then further divided by the type of question asked. White utilizes this taxonomy to analyze questions asked at reference desks. White's study, therefore, is a landmark both for the development of this taxonomy and for the literature on library reference.[16]

Forms of expected answers to questions

The need for standards for measurement and evaluation of reference services has been recognized in the library profession for some time. In the mid-1970s the American

Library Association (ALA)'s Library Administration and Management Association (LAMA) created standard definitions for two types of reference transactions for inclusion in their Library General Information Survey (LIBGIS).[17] These two types of reference transactions are as follows:

- "A reference transaction . . . involves the knowledge, use, recommendation, interpretation, or instruction in the use of one or more information sources by a member of the reference/information staff."
- "A directional transaction . . . provides assistance in finding and using library services, collections and facilities."[18]

The LIBGIS definitions were the first standardization of types of reference transactions, and for the first time provided a classification (simple as it is) of the types of services provided at a reference desk. Also for the first time, LIBGIS enabled reference services at different libraries, holding different collections and serving different communities of patrons, to share reference statistics.[19]

The classes "reference" and "directional" are, however, extremely broad. As a result, some researchers and libraries divided these classes into a variety of subclasses. [20, 21] Rothstein, presaging the classification to come, discusses grouping questions into the following types: directional, ready reference, search (or research), and readers' advisory.[22] Seng discusses three question types: direction, and two subclasses of the LIBGIS reference class, but which Seng defines in a unique way: information (a question that "is concerned with information resources and/or their use" – what might today be called bibliographic instruction), and general (a reference question "answered through the use of information resources" – what might today be called ready reference).[23] Brown drops the directional class entirely, and divides questions into informational (any question that can be answered using ready reference sources such as the card catalog or telephone directory) and reference (any question that requires non-ready reference sources to answer it).[24] Fogarty discusses the following four types: directional, instructional, ready reference, and extended reference.[25]

These variations on the LIBGIS theme demonstrate that even given a standard, different services will modify and extend that standard to accommodate their specific situation and requirements. Even more interesting is the amount of "convergent evolution" that has occurred surrounding this taxonomy of question types. Several researchers and libraries explicitly modified the LIBGIS classes. Equally many, however, independently developed question classification schemes that resembled the LIBGIS scheme; either these authors did not know of the existence of the LIBGIS scheme, or they simply did not mention it.[26] Looking across all of these variations on a theme, the following taxonomy of the forms of the expected answer to a question was developed:

Table 1: The Taxonomy of Forms of Expected Answers

Each of these taxonomies has added to our understanding of the complex relationship of question and answer in the environment of general reference services. It is this Forms of Expected Answers taxonomy, however, that is the most useful for the purpose of analyzing queries received by special collections and thus for creating a bridge between digital reference services in general academic libraries, special collections, archives, and museums.

SPECIAL COLLECTIONS DIGITAL REFERENCE

One impetus for the present paper comes from the strong effect that e-mail reference has already had on rare book departments, archives, and museums.[27] Typical of the increase in digital reference queries is that found at the Special Collections Research Center (SCRC), Bird Library, Syracuse University. In 1996 the number of remotely received research queries was 840, with e-mail accounting for 240, or 29% of the total. In 2003 the number rose to 1070, with e-mail accounting for 1032, or 87% of the total.[28] A similar increase has been noted in many archival collections. At the University of North Texas e-mail queries to the University Archives now account for over 70% of the total remote queries,[29] and the Southern Historical Collection at the University of North Carolina has seen an increase of over 40%.[30] Many museums have also taken a proactive role in making their collections more accessible to a wider audience over the Internet, and these virtual patrons now form a considerable part of the clientele.[31] At the John D. Rockefeller Library of the Colonial Williamsburg Foundation, for example, e-mail requests have risen 166% since 1999 and now account for 83% of remote research inquiries.[32]

The following discussion of the digital reference function in a special collections department is based on an analysis of inquiries received via e-mail at the SCRC. These transactions occurred between June 16th and September 15th, 2003, and represent a typical number for a three-month period in 2003. Since the SCRC does not possess a formal mechanism for allowing patrons to submit questions electronically (i.e., a question submission Web form or an e-mail address specifically for reference service), these reflect e-mails sent from patrons directly to the staff of the public services unit or, in a very few cases, forwarded from the library's main Reference Department.[33] During this period there were 251 transactions with 308 separate queries. Because the emphasis of this paper is on the types of questions received relating to special collections, those of a purely personal nature have been eliminated.[34]

There are four basic stages in this analysis of digital reference in special collections:

- Overview of digital reference questions from SCRC;
- Adjustment and application of "Forms of Expected Answers" taxonomy;

- Creation and application of new taxonomy for special collections;
- Comparison of questions using both taxonomies.

Overview of digital reference questions from Special Collections

A preliminary examination was made of the 308 digital reference queries from SCRC. The SCRC does not classify the questions that it receives by subject, and this also is not a major factor in determining which librarian should answer a question. Thus, given the questions that were received by the SCRC, and the approach to answering those question by SCRC librarians, the Forms taxonomy was the most appropriate. The forms taxonomy is also the only taxonomy discussed above that explicitly takes the physical presentation of an answer into consideration, which is important for a service that deals specifically with unique artifacts such as works of art.

The questions are received at the main departmental e-mail address.[35] Most of the inquiries are handled by the public services staff, with two notable exceptions: those pertaining to rare book bibliographical matters are routed to the curator, and those concerning permissions and loans are brought to the attention of the department head. What became clear by the end of this overview were these factors:

1) most of the queries concerned specific collections (including individual authors or artists);
2) many queries concerned individual titles or objects;
3) many queries asked about reproductions or permissions;
4) a few queries asked about visiting policies; and
5) a very few queries were unrelated to the collections or policies of the SCRC.

These last two factors comprised 1.6% of the total number of queries and have been eliminated from the rest of the calculations, but they do reveal aspects of searching techniques that are important for understanding how patrons access special collections. For example, in one case the same question was asked by three different patrons inquiring what real estate the SCRC had for sale. From the information asked for, it was obvious that the patrons had accessed a particular document deep within a collection, one that contained notes of land ownership. This situation reveals one of the possible complications with searchable finding aids, particularly when patrons are unfamiliar with the structure of archival collections.

From the above five factors arose two broad but fundamentally important observations: 1) almost all patrons already knew what they were looking for, whether through knowledge of the collections themselves or through perusal of the finding aids; 2) no other collections or sources were likely to hold their answers.

Figure 1 shows a high-level breakdown of the queries, where questions were labeled based upon their content. About one-third of the questions were related to the holdings of the collection and another one-third involved the acquisition of reproductions or gaining permissions to use a portion of the collection. The percentage of reproduction questions

clearly shows the necessity of having librarians knowledgeable in handling this important aspect of special collections. Since these questions relate to individual authors and artists and their works, however, they have been subsumed into these categories in further analyses. The most obvious fact to be learned from this breakdown is that holdings and reproductions, which both refer to specific collections or items, together account for 62% of the total; both of these categories are focused on specific portions of a specific collection. (The category of "other" is broken down into seven forms in Figure 2.)

Figure 1: High-level Analysis of Question Types

Adjustment of "Forms of Expected Answers" taxonomy

In order to place these queries within the context of general reference schema, the queries were analyzed and compared to the Forms of Expected Answers taxonomy. There were some adjustments that needed to be made to this schema to better fit the questions asked of special collections. Table 2 lists the original Forms of Expected Answers taxonomy and the taxonomy as adjusted for special collections.

Table 2: Original and Adjusted Taxonomies

The major adjustments and expansions are as follows:
Holdings: To help define the difference between "holdings" as a category in the original taxonomy and "holdings" in the analyses of special collections, the category "known-item search" was created to identify "questions about whether a specific information source or document is owned by the library (i.e., SCRC)." Several inquiries concerned whether SCRC might want to receive as a gift or purchase a specific title or work of art from the patron. This aspect places the special collections reference librarian in the role of collector and agent for the department.
Directional: Most of the directional questions concerned visiting the collections and thus also inquired about policies and costs.
Reproduction: Reproduction questions in special collections almost always entail permission for publication, sometimes also associated with loans for exhibits. Each of these may involve complex issues, depending upon who holds the legal rights. It is instructive for reference as well as collection maintenance to track these requests to determine those items that are consistently requested as opposed to those that have arisen because of an event such as national or international centennial celebrations.
Critique/Appraisal: Both of these aspects depend upon the librarian's judgment. Because of the legal ramifications and the possibility of suit, however, most special collections librarians are prohibited from giving appraisals. In fact, many departmental Web sites have statements that inform the public upfront not to ask for appraisals. Links are then often provided to the Antiquarian Bookseller's Association of America and other reputable sites that offer discussions of values and addresses of rare book dealers.

Research: This was eliminated as a separate category because, with the exceptions of visiting, reproduction/permission, and some factual questions, all special collections inquiries are "research" oriented.

Application of Adjusted "Forms of Expected Use" Taxonomy

After adjusting the taxonomy, the questions were sorted into the appropriate categories. Figure 2 shows the percentage of transactions in each category of the taxonomy, as revised for special collections. While holdings questions and reproduction questions remained stable, the delineation of "other" into the additional categories allowed greater insight into the types of transactions. In this expanded version, the greatest difference is the category of "factual," which received the third highest number of queries. Factual questions ask specific details about a person or work (such as date of birth or publication), and are those queries that are most likely to be able to be answered by more than one person or source. Since one of the research questions posed in the paper is how to include special collections in any kind of collaborative reference system, this category becomes crucial to such an analysis.

Figure 2: Forms of Expected Answers

Creation of new taxonomy for special collections

The "Forms of Expected Answers" taxonomy is very useful in classifying the content of questions, but a further delineation will help explain how special collections are defined and thus known and accessed by the public. The most appropriate scheme for identifying these factors is granularity. "Granularity is the relative size, scale, level of detail, or depth of penetration that characterizes an object or activity."[36] It is used in a number of disciplines, such as astronomy, photography, and information technology. In the present context it describes the level of specificity in the queries to special collections. A closer analysis of the SCRC's e-mail reference queries reveals the following four levels of granularity:

General: These are queries about the overall holdings of the library without specific collection designations. Example: How many incunabula do you have?

Collection: These are queries about a collection not devoted to a particular individual. Example: What depot plans do you have in the Erie Railroad collection?

Creator: These are queries about the holdings by a specific author or artist, or factual queries about that individual. Example: What photographs do you have of Margaret Bourke-White's trip to Russia?

Work: These are queries about a named or identifiable work. Example: What is the collation formula for volume 4 of your copy of Corpus juris canonici (Rome 1582)? In your photograph of Stephen Crane in front of his home, is his father facing to the right or to the left?

Figure 3: Granularity of Subject

Figure 3 quite clearly substantiates the general observations made from the preliminary analysis and further defines the "specialness" of special collections and their patrons. Indeed, separating out the "collection," "work," and "creator" factors makes it clear that the finer level of granularity best defines the needs of the patrons because it reveals that almost all patrons have already formed ideas of what they are looking for, whether for a generally named collection, an individual-specific collection, or an individual work. This level of granularity is familiar to archivists, whose collections are most usually arranged by provenance.[37]

This level of granularity also underscores another important distinction between general library collections and special collections. As Pomerantz, Nicholson, and Lankes have shown, general reference triage is based on the fact that the same answer may be found by different librarians in a number of sources. The emphasis is on the information itself as distinct from the source.[38] In special collections, however, the specificity of a query most often points to a specific collection, and this means that the answer can be found in only one source. (Factual questions are discussed below.) The emphasis here is thus on the container of the information, that is, the "unique" source. This finding is not only evidenced by the data from the SCRC study but is also indicative of collection usages in archives [39] and museums.[40]

Understanding the Users of Special Collections

What has also become obvious from an analysis of the content of the queries to SCRC is that many patrons have already searched the finding aids available on the department's Web site before they formulate and submit a query. The fact that the most popular search engine, Google, allows phrase searching means that a great many more finding aids are now findable and, consequently, the collections they describe.[41] As an example, one academic archivist reports that 70% of the queries about his collections are collection specific, and that "most of the users through e-mail have stated it is our websites and online finding aids that have prompted their queries."[42]

Users of special collections, as evidenced by the inquiries to SCRC, are seeking information about specific collections, and in this they play the role of researcher. The librarians thus play the role of subject expert. There is one category, however, that reveals the use of different roles and thus is best separated out in our analysis. This category encompasses business transactions, in which the users play the role of buyers and the librarians the roles of purveyors and assessors. Included are permission requests, photocopy costs, licensing costs, and the like. This category amounted to 12.4% of the total e-mail queries. Obviously, these queries may be handled only by the relevant institution and not by another member of a collaborative reference system.

An analysis of the research patrons submitting e-mail queries reveals the following general breakdown:

Table 3: Patron Categories

These classifications may be described and elaborated as follows:
Non-Academic Researcher: Primarily includes authors on contract to a publisher for a book or an independent researcher engaged in a project that will most likely lead to publication.
General Public: Includes patrons who happened to "hit" on the SCRC collections through an Internet search, patrons with specific inquiries about family history, and patrons who wish an appraisal or an opportunity to sell some item.
Academic: Restricted to patrons who identified themselves as members of an academic institution or who have used an ".edu" address. "Academic" includes institutions of higher education, museums, and archives.
Librarian: Restricted to librarians with research inquiries on behalf of their libraries.
Student: Restricted to patrons who clearly identified themselves as students of institutions.

Perhaps the two most surprising elements of this analysis are the high percentage of non-academic researchers and the low percentage of students using the digital reference services. The former obviously reflects the importance of general search tools such as Google and patrons' familiarity with using them. The latter probably reflects the type of assignments and the professors' stressing that students must come to the special collections department in order to use its materials. Likewise, very few of these students were from the associated university or surrounding institutions, who would be more apt to visit the collections in person. Similar to the non-academic researcher is the general public patron in the use of major online search tools. These factors become clearer when off-site reference use is compared to on-site reference use, as seen in Figure 4.

Figure 4: Off-site Reference Use Compared to On-site Reference Use

The great difference in percentages for the general public category (23% vs. 4%) points to the importance of online search tools to these patrons, as well as to the fact that they did not return to special collections once they had received an answer. Contrariwise, both the academic and non-academic researchers returned multiple times to use the collections. Many of them had already contacted the department via e-mail to determine that the collections held relevant materials and that their research projects could be helped only by the materials themselves, not by the finding aids or online reference help. Almost half of the academic researchers who visited the department were from the associated university or neighboring institutions and thus underscored their need of the materials themselves rather than finding aids or electronic surrogates. Of the students, over half were from the associated university and indicated that they had come to work

on assignments. There were also several doctoral students from other institutions doing dissertation research, all of whom had previously contacted the SCRC via e-mail to determine the relevance of the collections. The classifications of "Business Transactions" and "Librarian" had no uses because all of the relevant non-electronic inquiries of these two categories were handled via phone or letter, not in a face-to-face situation.

With these descriptions and comparisons in mind, the analyses by granularity of subject and forms of expected answers take on added significance.

Figure 5: Off-site Usage by User Type and Granularity

As can be seen in figure 5, for both the academic and non-academic researchers, and the general public, the level of specificity was very high, that is, they already knew the subjects of the collections or, indeed, the individual creator and his or her works. In fact, users in all of the above categories revealed a knowledge of the collections and asked very few "give me everything you've got" types of questions The variations in the "Student" category can be explained by the popularity of several distinct collections (not associated with an individual), such as the Erie Railroad and the Oneida Community.

Figure 6 shows the breakdown of question by the form of expected answer. For the three top-ranked users of online reference services in special collections, the most important category is "holdings," defined earlier as "whether a specific information source or document is owned by the library." The size of this category probably reflects the level of cataloguing of manuscript collections, particularly the many large ones that are catalogued only down to the box level. This situation is quite typical of academic research collections with substantial manuscript holdings. An interesting difference, however, is the number of "factual" questions asked by Non-Academic Researchers, as compared to the "Academic" category. A possible explanation for this is that many of these users are at the beginning of their research and thus have more questions of a factual nature.

Figure 6: Off-site Usage by User Type and Form of Expected Answer

Since a primary concern of this study is to show how electronic reference services in a special collections environment may serve as a bridge between those of library general reference departments and museums, comparative analyses of the SCRC data were made between the "Forms of Expected Answers (Adjusted)" and the new "Granularity of Subject" taxonomy. The resultant observations thus compare those factors indicative of e-mail queries of general reference and those of special collections and archives. Table 4 shows these raw data, removing the 5 Out of Scope questions.

Table 4: Taxonomy by Granularity of Subject and Forms of Expected Answers

The earlier analysis of granularity revealed that "Collection," "Creator," and "Work" had by far the greatest number of queries. Most important for the present study, however, are the specific forms of expected answers that create the greatest number of queries for each of these categories. The top two "Forms of Expected Answers" categories for each of the three "Granularity" categories are highlighted in gray in Table 4. These common areas of overlap between the two taxonomies have been extracted to Table 5, in order to aid a conceptual understanding of how special collections and archives might fit in with a collaborative reference system.

Table 5: Common Areas of Overlap between Granularity and Forms of Expected Answers

Table 5 thus represents the largest areas of overlap between the taxonomy of general reference queries and that of special collections. What does this table show about the feasibility of combining special collections, archives, and museums in a collaborative digital reference system? By removing the common areas of transactions not related to a particular collection or item, the only remaining area represents factual questions about a creator (the one cell that is not shaded grey).

Each of the grayed areas is one that can be handled only by the special collection to which the query was addressed, since the answer was to be found only in this collection. This is true even though certain aspects of "reproduction" and "permission" may eventually be handled by another entity, such as a business office or institutional attorney. The query itself still refers to a specific item within a collection. This leaves only the "factual" questions as possibly being able to be answered by more than one person or source from these top categories.

In order to better understand this group of questions, we performed a qualitative examination of these factual questions. About one-third of these questions requested information that might be found in general reference sources. The other two-thirds of the questions were so specialized that it was highly likely that they, too, were "unique source" answers. In the context of the entire set of queries received by SCRC, this finding means that actually only 8% of all transactions analyzed could be referred to other members of a digital reference consortium. This is an important finding of this study, as it has serious implications regarding the interest of museums and special collections in joining cooperative reference service

Figure 7: General Reference Compared to Special Collections

BUILDING THE BRIDGES: ISSUES IN INTEROPERABILITY

Comparing General Reference and Special Collections

The above analyses and discussions in emphasizing the "specialness" of special collections, archives, and museums hint at the challenges inherent in combining them within a collaborative general digital reference system. An important comparison to consider is the difference between this taxonomy as applied to questions from a general reference service and as applied to special collections. Pomerantz,[43] when developing his taxonomy based on expected answers, applied it to a set of digital reference questions from a general reference service. Figure 7 shows the breakdown of questions using this taxonomy for both general reference and special collections. While the majority of questions in special collections related to holdings or reproduction topics, most of the questions posed to a general reference service were either factual in nature or requested a list of citations.

This difference, therefore, further supports the claim that the underlying nature of questions posed to a general reference service is fundamentally different than that of questions posed to special collections. The general reference questions focus on information that can usually be found in multiple sources or, at the very least, a single source with wide distribution. The questions posed to special collections most often focus on a unique aspect of that particular service, and thus are not answerable by another service. To use terminology from information science, many general reference questions are focused on information, regardless of source, while the special collections questions require a specific information container to answer correctly. These differences are summarized in Table 6.

Table 6: Comparison of Requirements

Special Collections and Collaborative Reference Systems

All of these findings point to the same conclusion: most digital reference questions posed to special collections cannot be handled by a general reference service. Since one of the major selling points for collaborative reference services is that another group of librarians can answer questions for your library, it does not benefit special collections departments (and, by extension, archives and museums) to be part of these networks. Many of the question types cannot be handled by a general reference staff, and of the remaining questions, many are too specific to be answerable by non-specialized information resources.

A driving factor in understanding this situation is that a special collections librarian in the role of question-answerer is playing a different role than that of a general reference librarian. Since the materials held by special collections are valuable and unique, the patron often assumes that the librarian may act as curator, owner, and distributor of the materials. In addition, the special collections librarian is expected to be a subject expert as well as an expert about the physical information containers. These are both roles that general reference staff are not trained or authorized to perform.

However, it is probable that some of the questions posed to special collections could be handled by other special collections or museums that gather materials on the same specialized topic. While reproduction and permission questions are based on the policy of an individual institution, questions about a particular item or creator might be able to be answered elsewhere. Union catalogs, such as the National Union Catalog of Manuscript Collections, or censuses of specific artists and their works enable different institutions to help researchers in their quest for information. As collaborative reference services are extended to include more special collections and museums, the chances that questions could be referred easily and archived improve. As special collections and museums are added to the collaborative systems, a directory of experts would emerge that would aid all participating services in referring questions to the most appropriate service.

One additional benefit to special collections and museums of such a service is that it would serve as a gateway to their own original digitized materials. As these services justify their existence partially through use of their digital collections, joining collaborative reference services would direct more patrons to their online materials. Another benefit is that it would make it much easier for special collections and museums to access the finding aids for other services. These finding aids are the key to understanding what artifacts exist in a collection, and can be difficult to get access to. By creating standard methods of connecting finding aids to the reference services, those interested in a finding aid for a similar collection can easily discover these useful tools.

Strategies for involving Special Collections (and Museums) in Collaborative Reference Services

The first strategy to draw special collections and museums into participating in collaborative reference services is that their out-of-scope questions can be quickly forwarded to more appropriate services. While this does not make up a large portion of questions, it is one fewer patron that needs to be accommodated. The second, and perhaps more important strategy, is that these services can improve cooperation between special collections and museums that focus on the same topic areas. As no two collections have the exact same focus, being able to have easy referral tools for questions on the fringe of the holdings of one special collection will allow for better information services for users of the services.

One issue to consider is the interface used in these services. Most interfaces used in collaborative reference services are fairly sparse; Janes has shown that the more fields the user has to fill out, the lower the chance the user will finish the form.[44] However, special collections and museums often require more information about the query. Since many of the questions are about specific items, then a more specific form is needed to capture all of the required information from the patron.

Based on the taxonomy shown in Table 2, such forms need to allow for entry of at least the following information:
- Collection;
- Creator; and
- Work.

In these categories, libraries and museums can register as experts in specific categories under this taxonomy. The patron's question can then be easily referred to the appropriate question-answering group.

As an example of this type of specialized interface, figure 9 contains the question submission form from the Ask Joan of Art service, which is a service of the Smithsonian American Art Museum:

Figure 8: Form from Ask Joan of Art service

As the Smithsonian has developed this service, it has realized the need to create a more specialized interface to deal with the questions asked by these specific patrons. Similar interfaces would need to be developed for collaborative reference services that include libraries, museums, and special collections. In order to develop these interfaces specialized metadata standards need to be implemented.

SUMMARY AND CONCLUSIONS

The goal of this paper was to gain a better understanding of the types of users of and questions submitted to a special collections department. As special collections are similar in structure to museums, this understanding will allow developers of systems for digital reference services to create more robust systems that can encompass the needs of libraries and museums.

The overall findings of this study are that users of special collections have much more specific information needs than users of general reference services; these needs are so specific, in fact, that general reference desks would not be able to successfully answer many of the questions. This creates a problem for the growing trend of shared and collaborative reference services: if most of the questions cannot be handled by another service, then why does it benefit that museum/special collection to participate in a collaborative environment?

The needs of users of special collections tend to focus on specific items or specific collections. Many queries are answered by a specific item, by providing a reproduction of that item, providing information about the specific item, or setting up a time for the patron to visit the item. The focus of the question, therefore, is on a specific information container. Digital reference systems must allow for the more detailed metadata useful in answering questions that focus on specific items instead of the information contained within the items.

There are obvious reasons for including special collections and museums in a collaborative digital reference system, among which are the ease of referring queries to appropriate experts in specialized fields, greatly increased access to digitized materials, including both text and images, and access to well-researched finding aids. These benefits are not always apparent to those working in these specialized information services, and modified publicity materials, question forms, and infrastructure will be needed to make collaborative reference systems attractive to these other services.

Interoperable systems have been created that take advantage of vast holdings of digitized images, such as the Colorado Digital Initiative, ArtsConnectEd [45], and that of Western Reserve Historical Society, the Ohio Historical Society, and the Cincinnati Historical Society.[46] Through integrated searching interfaces these offer user access to print and non-print collections, and the broad offerings of general libraries and the specialties of archives and museums. These exist to make available the best combination of sources to respond to their patrons' information needs, and they are excellent examples of what a large-scale collaborative system could provide to librarians and patrons alike.

FIGURES

Figure 1: High-level Analysis of Question Types

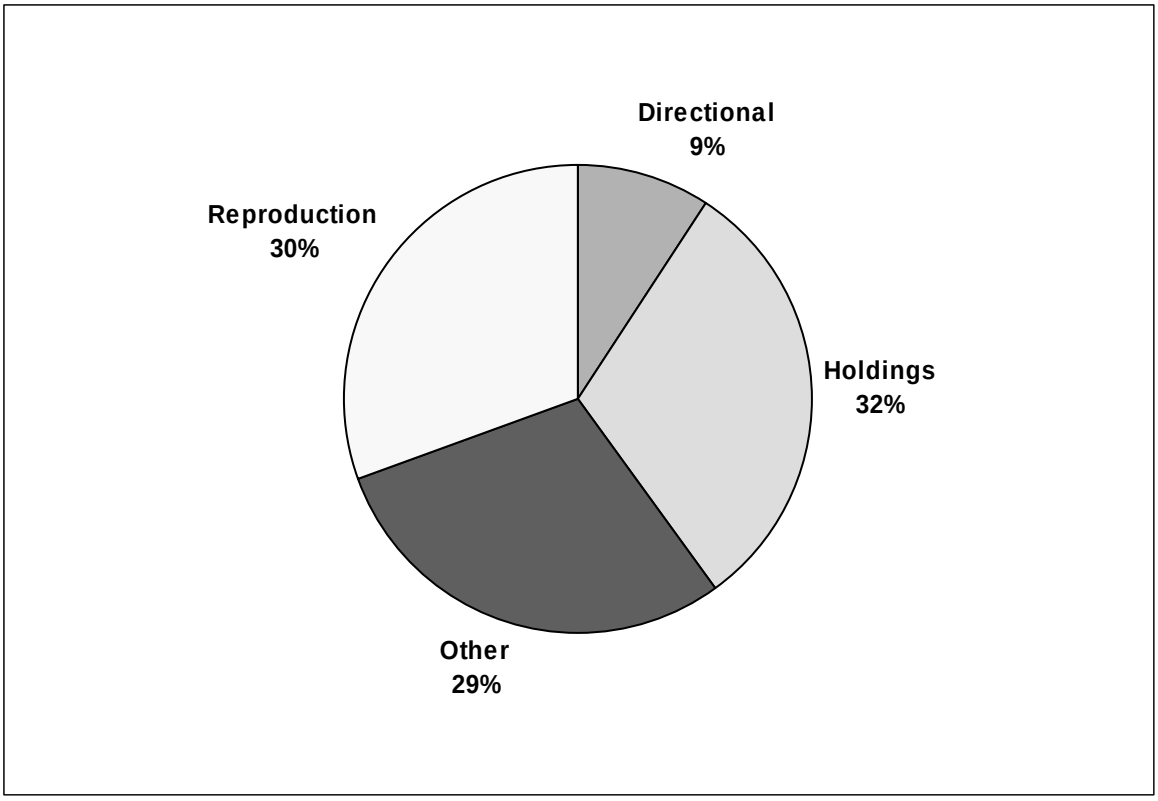


Figure 2: Forms of Expected Answers

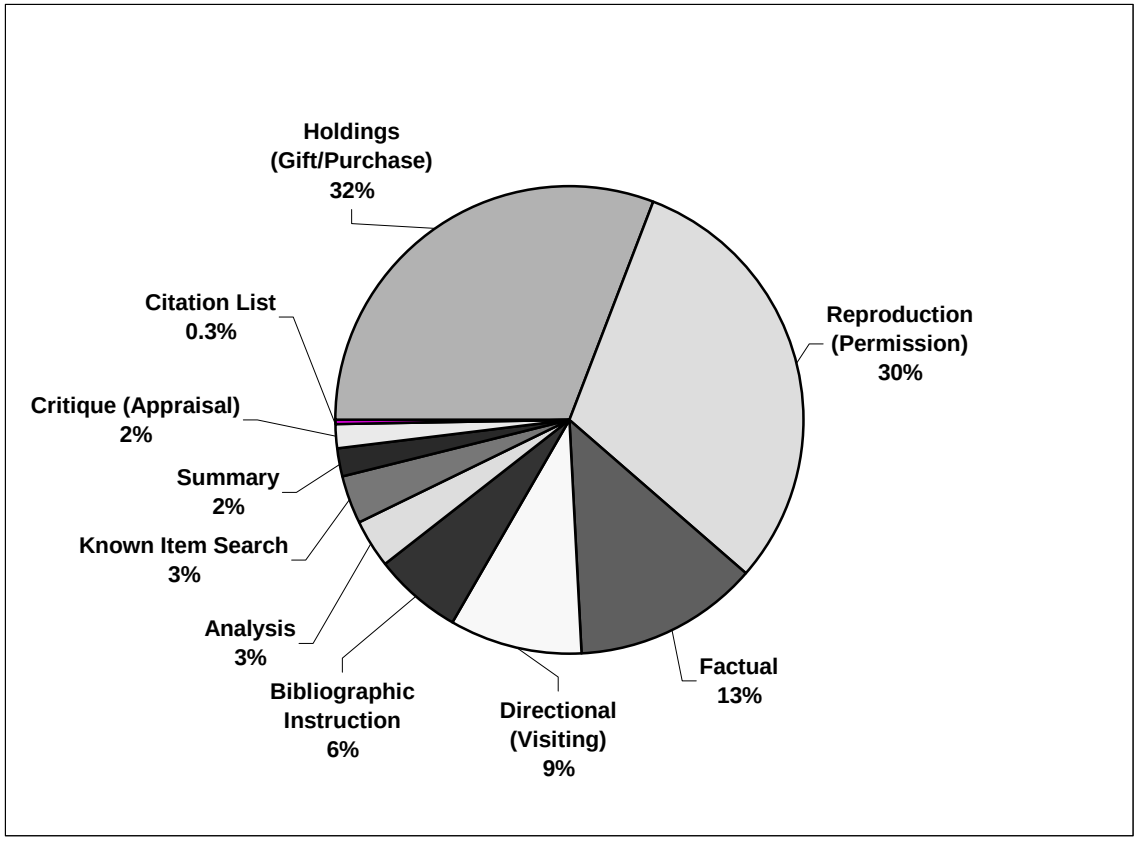


Figure 3: Granularity of Subject

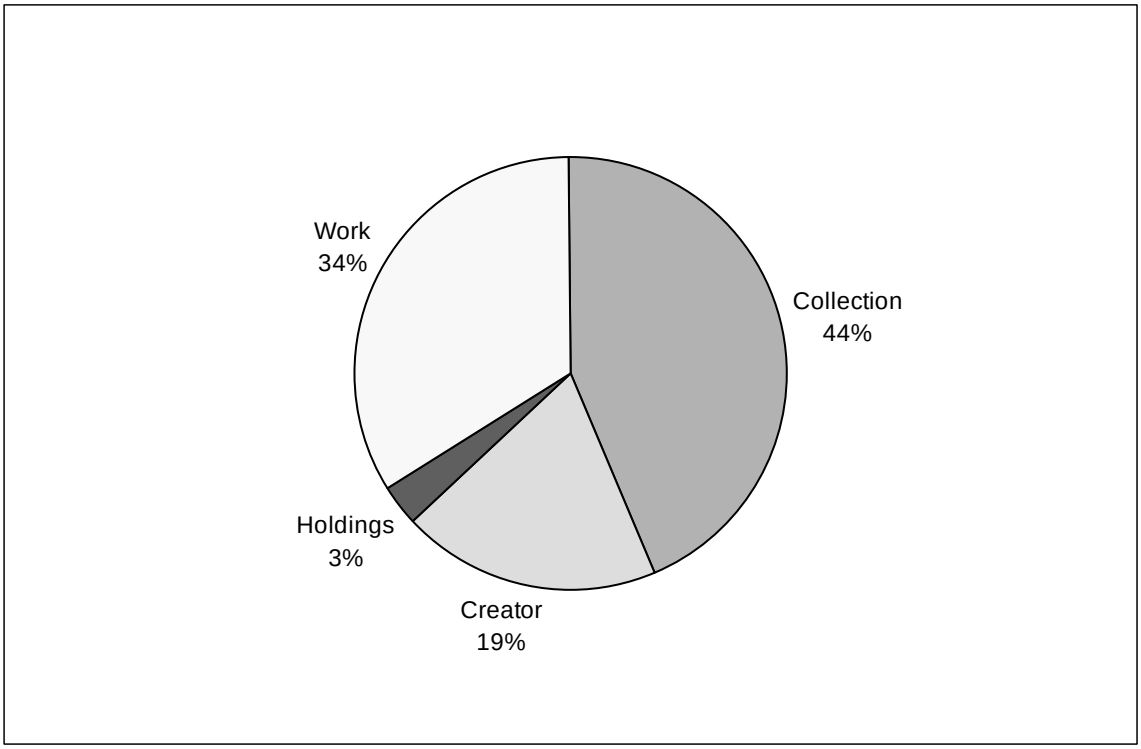


Figure 4: Off-site Reference Use Compared to On-site Reference Use

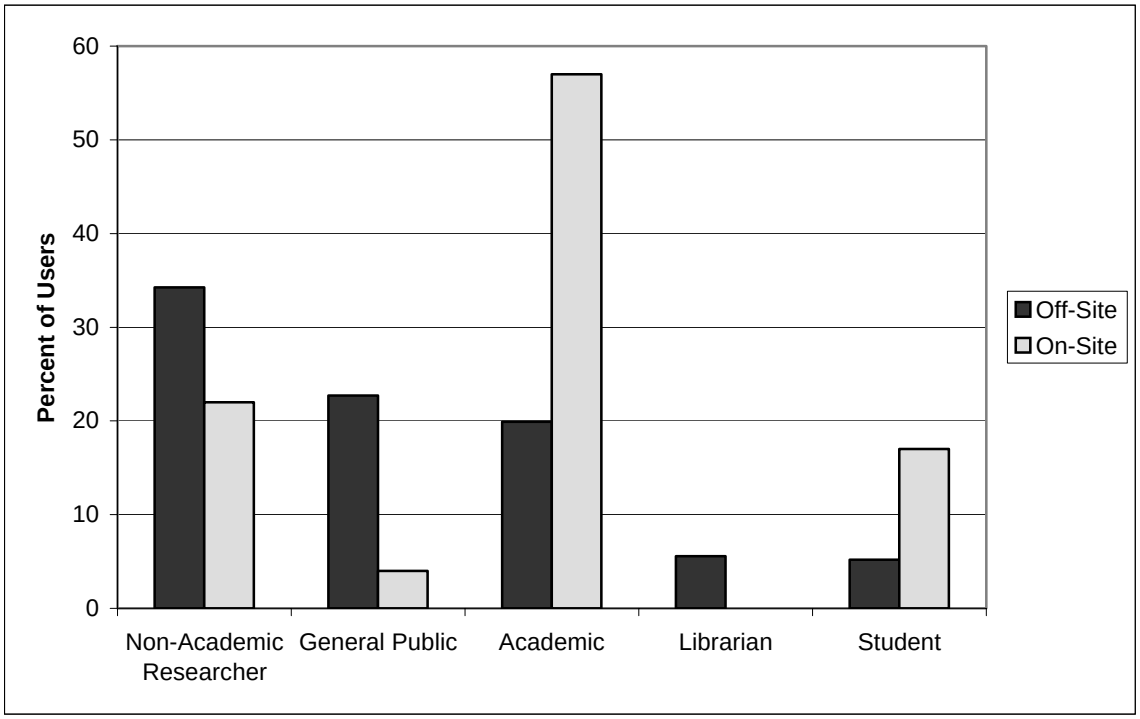


Figure 5: Off-site Usage by User Type and Granularity

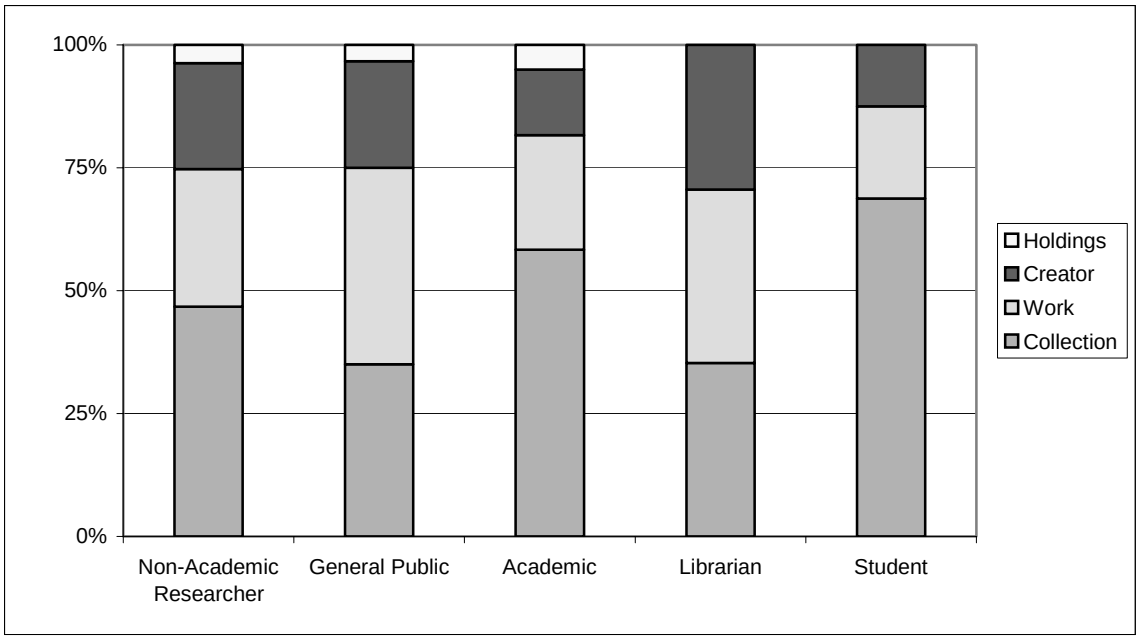


Figure 6: Off-site Usage by User Type and Form of Expected Answer

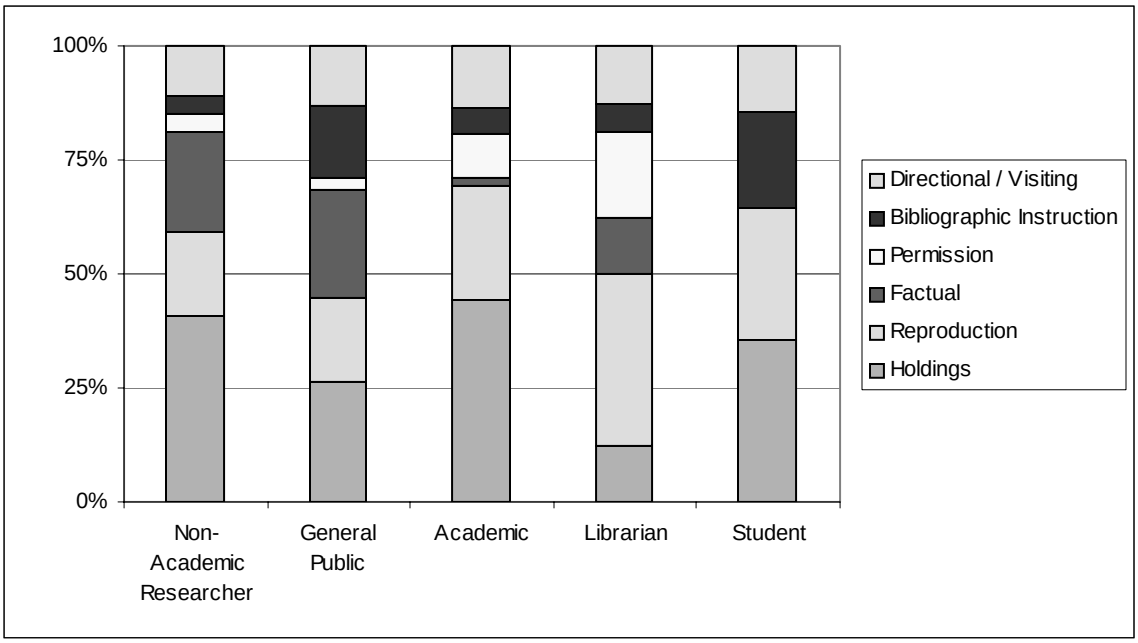


Figure 7: General Reference Compared to Special Collections

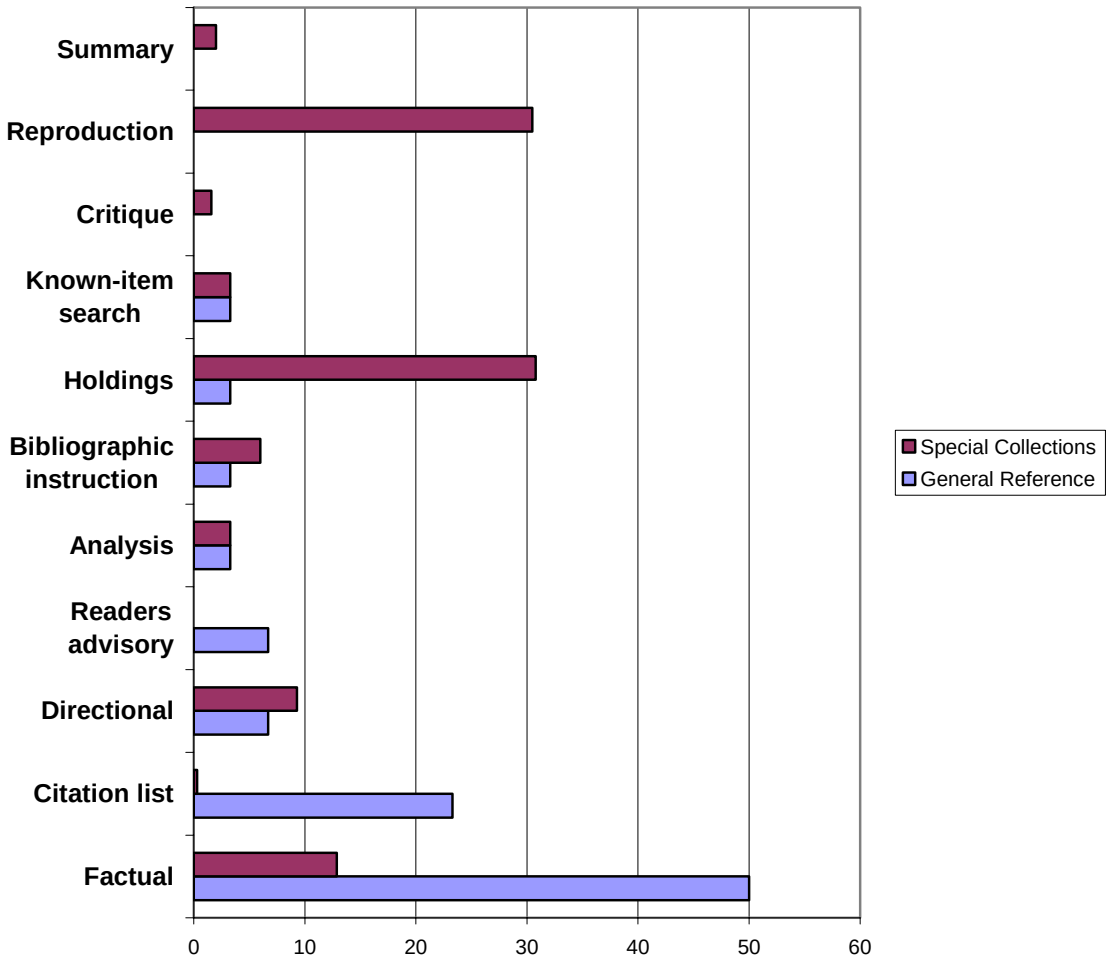
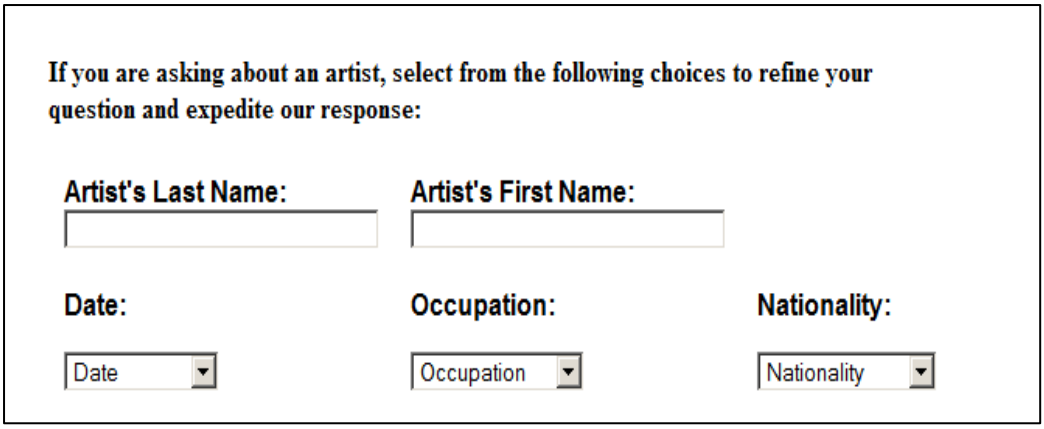


Figure 8: Form from Ask Joan of Art service



TABLES

Table 1: The Taxonomy of Forms of Expected Answers

Class	Scope note
Directional	Questions asking about the location of a specific information source.
Holdings	Questions about whether a specific information source or document is owned by the library.
Ready reference	Questions asking for simple, factual answers; the answer should be readily ascertainable from available information sources.
Exact reproduction	Questions asking for pictorial and textual materials, taken directly from an information source and unchanged.
Description	Questions asking for a description of something, briefer in length than the original thing (basically, an abstract).
Readers advisory	Questions asking for assistance in the choice of books or the gathering of data.
Bibliographic instruction	Questions asking for assistance in use of information source(s).
Research	Questions asking for involved answers; the answer should require some effort and wide use of information sources to formulate.
Citation list	Questions asking for a list of information sources on a particular subject.
Analysis	Questions asking for some form of data analysis, whatever that data might be – scientific, social, financial, etc. Questions of this type might ask for trends, pro or con arguments, cause and effect, compare and contrast, etc.
Critique	Questions asking for an evaluative discussion of a particular subject. (E.g.: a movie review, Cliffs notes-like analyses of a book, etc.)

Table 2: Original and Adjusted Taxonomies

General Reference	Special Collections
Directional	Directional Visiting
Holdings	Holdings Gifts/Purchases
	Known-item search
Ready Reference	Factual
Reproduction	Reproduction Permission
Description	Summary
Readers Advisory	Readers Advisory
Bibliographic Instruction	Bibliographic Instruction
Research	
Citation List	Citation List
Analysis	Analysis
Critique	Critique/Appraisal

Table 3: Patron Categories

User type	Percent
Non-Academic Researcher	40.5%
General Public	24.2%
Academic	22.7%
Librarian	6.4%
Student	6.1%

Table 4: Taxonomy by Granularity of Subject and Forms of Expected Answers

	Granularity of Expected Answer				
Form of Expected Answer	**Collection**	**Creator**	**Work**	**Nonspecific**	**Total**
Holdings	47	15	14	8	84
Reproduction	25	6	26		57
Factual	13	14	12		39
Permission	7	5	23		35
Bibliographic Instruction	13	1	4		18
Visiting	13		2	2	17
Directional	3	5	4		12
Analysis	1	3	6		10
Known Item Search	2	3	5		10
Gift/Purchase	3	3	3		9
Summary	4	2			6
Appraisal		1	4		5
Citation List	1				1

Table 5: Common Areas of Overlap between Granularity and Forms of Expected Answers

	Collection	Creator	Work
Most Common	Holdings	Holdings	Reproduction
Second Most Common	Reproduction	Factual	Permission

Table 6: Comparison of Requirements

	General Reference	Special Collections
Type of Answer Required	Factual or Citation List	Holdings information or Reproduction permission
Type of Resource Required	Information	Information Container
Uniqueness of Resource	Multiple resources	Specific resource

References

[1] K. Lam, "Exploring Virtual Reference: What It Is and What It May Be," in *Implementing Digital Reference Services*, edited by R.D.Lankes, C. R. McClure, M. Cross, and J. Pomerantz (New York: Neal-Schuman, 2003), pp. 31-39.

[2] Jeffrey, Pomerantz, Scott R. Nicholson, and R. David Lankes, "Digital Reference Triage: Factors Influencing Question Routing and Assignment," *The Library Quarterly* (73) (2003): 103-120.

[3] Discussion may be accessed at www.loc.gov/standards/netref.

[4] Coburn, Erin and Murtha Baca, "Beyond the Gallery Walls: Tools and Methods for Leading End-Users to Collection Information," *ASIS&T Bulletin (Special Section)* (June/July 2004): 19.

[5] Jeffrey Pomerantz, "A Linguistic Analysis of Question Taxonomies," *Journal of the American Society for Information Science and Technology* (forthcoming).

[6] M. Conner, "What a Reference Librarian Should Know," *The Library Journal* 52(8) (1927): 415-418.

[7] R. S. Taylor, "Question-Negotiation and Information Seeking in Libraries," *College & Research Libraries* 29 (1968): 183.

[8] G. B. King, "Try It--You'll Like It: A Comprehensive Management Information System for Reference Service," *The Reference Librarian* 3 (1982): 71-78.

[9] M. E. Murfin and G. M. Gugelchuk, "Development and Testing of a Reference Transaction Assessment Instrument," *College & Research Libraries* 48(4) (1987): 314-338.

[10] S. Rothstein, "The Measurement and Evaluation of Reference Service," *Library Trends* 12(3) (1964): 456-472.

[11] K. D. Crews, "The accuracy of Reference Service: Variables for Research and Implementation," *Library & Information Science Research* 10(3) (1988): 331-355.

[12] W. G. Lehnert, *The Process of Question Answering: A Computer Simulation of Cognition* (Hillsdale, NJ: Lawrence Erlbaum Associates, 1978).

[13] A. C. Graesser, K. Lang, and D. Horgan, "A Taxonomy for Question Generation," *Questioning Exchange* 2(1) (1988) : 3-15.

[14] A. C. Graesser, N. Person, and J. Huber, "Mechanisms that Generate Questions," in *Questions and Information Systems*, edited by T. W. Lauer, E. Peacock and A. C. Graesser (Hillsdale, NJ: Lawrence Erlbaum Associates, 1992), pp. 167-187.

[15] A. C. Graesser, C. L. McMahen, and B. K. Johnson, "Question Asking and Answering," in *Handbook of Psycholinguistics*, edited by M. A. Gernsbacher (San Diego: Academic Press, 1994), pp. 517-538.

[16] M. D. White, "Questions in Reference Interviews," *Journal of Documentation* 54(4) (1998): 443-465.

[17] U.S. Department of Education Office of Educational Research and Improvement, *LIBGIS - Public Library Universe* (Washington D.C.: U.S. Department of Education, Office of Educational Research and Improvement, 1981).

[18] H. D. White, "Measurement at the Reference Desk," *Drexel Library Quarterly* 17(1) (1981): 33-34.

[19] N. K. Kaske, and R. Aluri, *Analysis of Reference Statistics Reported in 1977 Library General Information Survey*. Paper presented to the Library Research Round Table, annual conference of the American Library Association, June1980, New York. 31p. Also available from ERIC (ED 202 486).

[20] H. D. White, "Measurement," 33-34.

[21] Phenix, K. *Analysis of the 1981/82 HEGIS/LIBGIS Responses of Illinois Academic Libraries* (No. Illinois Library Statistical Report 9), 1983.

[22] Rothstein, "Measurement," p. 458.

[23] M. Seng, "Reference Service Upgraded Using Patrons' Reference Questions," *Special Libraries* 69(1) (1978): 21-28.

[24] D. M. Brown, "Telephone Reference Questions: A Characterization by Subject, Answer Format, and Level of Complexity," *RQ,* 24(3) (1985): 290-303.

[25] E. Fogarty, "Reference Questions: Who, What, Where, When, How, and Why?" *New Jersey Libraries* 28 (1995): 19-21.

[26] M. J. Lynch, "Reference Interviews in Public Libraries," *The Library Quarterly* 48(2) (1978): 119-142; C. A. Bunge, "Factors Related to Output Measures for Reference Services in Public Libraries: Data from Thirty-Six Libraries," *Public Libraries,* 29 (1990): 42-47; P. Dewdney and G. Michell, "Oranges and Peaches: Understanding Communication Accidents in the Reference Interview," *Reference & User Services Quarterly,* 35(4) (1996): 520-534; J. C. Stalker and M. E. Murfin, "Why Reference Librarians Won't Disappear: A Study of Success in Identifying Answering Sources for Reference Questions," *RQ,* 35(4) (1996): 489-503; D. S. Carter and J. Janes, "Unobtrusive Data Analysis of Digital Reference Questions and Service at the

Internet Public Library: An Exploratory Study," *Library Trends*, 49(2) (2000): 251-265.

[27] K. J. Underhill and B. Palmer, "Archival Content Anywhere@Anytime," *Internet Reference Services Quarterly*, 7 (2002): 19-30; L. Sowers and M.D.White, "National Museum of American Art Reference Desk: A Usage Analysis of a Digital Reference Service," in *Digital Reference Service for the New Millennium: Planning, Management, and Evaluation*, edited by R.D. Lankes, J. W. Collins, and A. Kasowitz (New York: Neal-Schuman, 2002), pp.153-178.

[28] Carolyn A. Davis, personal communication, February, 16, 2004.

[29] Richard Himmel, personal communication, October 24, 2003.

[30] K. E Martin, "Analysis of Remote Reference Correspondence at a Large Academic Manuscripts Collection," *American Archivist*, 64 (2001): 17-42.

[31] P. Marty, W.B. Rayward, and M.B. Twidale, "Museum Informatics," in *Annual Review of Information Science and Technology* (Medford, NJ: assis&t., 2003), pp.259-294.

[32] J. M. Clark, "Analyzing E-mail Reference Service in a Museum Library: The Experience of Colonial Williamsburg's John D. Rockefeller, Jr. Library," in *Implementing Digital Reference Services*, edited by R.D.Lankes, C. R. McClure, M. Gross, and J. Pomerantz (New York: Neal-Schuman, 2003), pp. 9-19.

[33] A series of question forms have since been mounted on the SCRC Website. It will prove instructive to compare the same three-month period in 2004 with the 2003 sample. They may be accessed at: http://libwww.syr.edu/information/spcollections/contact.htm#forms.

[34] Requests for classroom instruction are handled separately from reference queries and are thus not included in this analysis.

[35] Exceptions are those that are transferred to individual librarians from other library departments, or that result from personal correspondence between a librarian and a user.

[36] Definition is taken from "whatis.com" and can be accessed at: http://whatis.techtarget.com/definition/0,,sid9_gci212209,00.html (accessed 7/14/04).

[37] F. O'Donnell, "Reference Service in an Academic Archives," *Journal of Academic Librarianship*, 26 (2000): 110-118.

[38] J. Pomerantz, S. Nicholson, and R.D. Lankes, "Digital Reference Triage."

[39] W. Duff and C.A. Johnson, "A Virtual Expression of Need: An Analysis of E-mail Reference Questions," *American Archivist*, 64 (2001); 43-60; O'Donnell, op.cit.

[40] Marty, "Museum Informatics."

[41] Martin, "Analysis."

[42] Himmel, loc.cit.

[43] J. Pomerantz, S. Nicholson, and R.D. Lankes, "Digital Reference Triage."

[44] J. Janes, "Digital Reference Services in Public and Academic Libraries," in *Evaluating Networked Information Services: Techniques, Policy, and Issues*, edited by C. R. McClure and J. C. Bertot (Medford, NJ: Information Today, 2001), pp. 175-195.

[45] R. Dowden, S. Sayre, and S. Dietz, "ArtsConnectEd: Collaboration in the Integration and Access to Museum Resources," *First Monday* 5(6) (2000): unpaginated. Retrieved March 2004, from http://www.firstmonday.org/issues/issue5_6/dowden/index.html#author.

[46] M. McCormick, "An Amishman with a Computer," Presentation at American Library Association Annual Conference, Toronto, June 21, 2003.