2-2003

Question Negotiation and the Technological Environment

Joseph Janes
University of Washington

Joanne Silverstein
Syracuse University

Follow this and additional works at: https://surface.syr.edu/istpub

Part of the Library and Information Science Commons

Recommended Citation

This Article is brought to you for free and open access by the School of Information Studies (iSchool) at SURFACE. It has been accepted for inclusion in School of Information Studies - Faculty Scholarship by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
The information world is a very different place in 2003 than it was a few years ago. It is possible, for example, to find information more quickly and easily than ever before, using new tools, and drawing on sources of information unavailable or even nonexistent not so very long ago. People seeking answers or providing them now have many more options for "question negotiation", in the broadest Tayloresque sense [1]. Consulting an information professional continues to be one of those options. In fact, such professionals can now be much more accessible via these new tools and technologies.

There has been a great deal of discussion in the professional literature about the reference interview and question negotiation. Among the most important are Taylor's notions of question negotiation (encompassing self-help and a variety of other modes of satisfying information needs). As yet, however, no universally accepted and satisfying definition exists for the term "question negotiation".

We therefore propose the following as a conceptual definition for "question negotiation":

an interaction between a person with an information need and an information service; its purpose is to refine the information need so that it can be usefully responded to by the information service.

Given that definition, we now investigate modes or channels in which question negotiation occurs, and how the technology of those modes affects the tasks and opportunities of the digital reference professional and, perhaps, vice versa.

1. Characteristics of Some Modes of Digital Reference

In this section, we'll examine synchronous and asynchronous modes of conducting a digital reference transaction, with an eye for their affects on digital reference practice.

One of the difficulties in studying the reference interview in the past has been the ephemeral nature of the interview. Since traditional reference interviewing has taken place in person or on the phone, no objective record existed unless the interview was recorded or transcribed. Even then, one is left to wonder whether the recording process itself materially affected the nature of the interview.

In the digital environment, records—such as transcripts and email archives—exist and can be of great use in studying the reference interview. To date, however, little has appeared in the research literature to describe the nature of these interviews. Abels' oft-cited article examined purely email interview
exchanges [Abels, 1996], and a number of "how we do it" articles have appeared in the practitioner literature.

A more recent study examined the range and scope of digital reference services in public libraries in 2002, and some findings are reported below [Janes, in press].

Email
Findings from the Janes study show that the proportion of public libraries offering reference services via email has risen sharply, almost tripling in a two-year time frame, but it is still lower than the proportion of academic libraries offering such services in 1999. The increase in digital reference volume over all types of libraries is dramatic and may perhaps be related to new methods of question submission. While the proportion of libraries providing email addresses is roughly comparable to what has been observed in the past, the percentage of libraries offering detailed forms (anything more than the simple form questions) has doubled, largely at the expense of the very bare-bones, simple forms. In addition, the number of libraries offering real-time services (chat, instant messaging, call center software based) now exceeds 20%, and a similar number provide a page summarizing their services and offering guidance as to which type of reference service might be more appropriate.

Web Forms
Further examination of the web forms used for digital reference yields this list of the most frequently observed questions asked on web forms. (Nearly all forms asked for name and email address.)

<table>
<thead>
<tr>
<th>Question</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>phone number</td>
<td>53</td>
</tr>
<tr>
<td>affiliation</td>
<td>22</td>
</tr>
<tr>
<td>street/mailing address</td>
<td>19</td>
</tr>
<tr>
<td>fax number</td>
<td>17</td>
</tr>
<tr>
<td>deadline</td>
<td>14</td>
</tr>
<tr>
<td>sources already consulted</td>
<td>10</td>
</tr>
<tr>
<td>subject area</td>
<td>8</td>
</tr>
<tr>
<td>preferred info delivery method</td>
<td>6</td>
</tr>
<tr>
<td>school assignment?</td>
<td>6</td>
</tr>
</tbody>
</table>

A number of other questions were asked, but none more than 5 times. It is interesting to note omissions here: the rationale or motivation behind the query—one of the main reasons for doing a reference interview—was asked on only 3 forms, as was the preferred answer type (factual, sources, etc.). The number of items or amount of information desired was asked only once. Most of the frequently asked questions, then, are functional and personal (name, email address, phone number and so on), rather than based on the nature of the query itself. These questions serve more as a profiling activity than an interview aimed specifically on refining the nature of the query.

Chat
Only very cursory examination has been made of the web pages describing and presenting chat-based services. The following are subjective observations (from the same study referenced previously), and more substantial analysis of these pages may well yield more sophisticated and trustworthy conclusions. However, these glimpses into what those pages say and do are intriguing:
• A handful of the live services provide web forms for posing the initial set of questions; these are usually very few in number and the boxes in which users can type their questions are almost always very small, which tends to encourage people to type relatively little.

• A number of these pages describe technological requirements to use the service but also include restrictions, warnings and barriers (e.g., Macintosh computers can't be used; what to do if the service freezes; etc.).

• Several of the services described in some detail who would be responding (e.g., a librarian) to the user's question—something not often seen in other digital reference environments.

• Somewhat more mention is made here about privacy and confidentiality policies than in other digital reference environments.

• Almost all specified the hours their services are available, and some also described the process used by the service to respond to questions.

General Characteristics for All Modes

Another source of potentially revealing evidence about the reference interview in the digital environment is that of comments made by librarians. The DIG_REF listserv, maintained by the Virtual Reference Desk project, is probably the best-known and most widely trafficked listserv in this area. While a comprehensive survey of the DIG_REF discussions on the reference interview is beyond the scope of this article (and would be a fascinating and worthwhile venue for further research endeavors), a discussion thread from October of 2001 seemed to encapsulate many of the issues of recent concern to reference librarians about the nature of the reference interview and how it's changing. A posting by Bernie Sloan on October 22, 2001, began the discussion by asking the following: "I'm looking for practitioners' perspectives on how the reference interview is affected by the medium it's conducted in…From your experience, how does the digital reference interview differ from the traditional face-to-face reference interview in the library?"

A large number of responses to Sloan's question were posted over the next several days. Among the discussion points of interest here are:

Regarding asynchronous (email, web form based) interviewing:

• loss of nonverbal clues (tone of voice, eye contact, etc.)

• the disappearing questioner, someone who never responds when asked for clarification

• receipt of more responses when clarifying questions were presented with some information, or a guess as to what the original question was about, as opposed to simply receiving questions probing the information need alone

• the longer time needed for an interview using repeated email exchanges, and the additional strain on staff time that this implies

• the creation of a permanent record and ability to track reference performance

• lack of feedback on whether responses were correct, appropriate, and/or satisfactory

• the opportunity for both user and librarian to think through the question and response, to let the question "percolate" as the librarian works on it

• the importance and role of a form as a simulation of the reference interview, as well as commentary on the perception of reluctance of some users to filling out forms

• the ability to clarify by phone questions received via email or form

Regarding synchronous (chat, call center software) interviewing:

• Users send a probe query, asking an initial simple, vague query, and upon receiving a response then proceed to ask much more detailed questions.

• The person responding to the question needs the ability to make educated guesses (based on grammar, syntax, email address, time query was sent) on what an initially vague query might really be.

• People using chat are in a hurry and don't want to take time to fill out forms.

• Chat is better for the back-and-forth of an interview, and thus using it provides more interaction. However, because of perceived time pressures, little of what typically makes up a
traditional reference interview might take place.

• Chat places more of an onus on patrons to make their needs clear quickly and to make it clear also when responses aren't what they want.
• Dead time is a problem; there is a perceived need to stay active with users to keep them online and interested.
• Chat can be difficult for users if they're not familiar with the chat medium.
• It can be helpful to estimate the time a response will take in order to let users know what to expect.
• Chat sessions aren't necessarily quick — estimates of the average time they take range from 10 to 20 minutes, and some are as long as 45 minutes or more. There is also the threat of a "virtual line" that can form if too many chat requests come in simultaneously.
• "You can't chat the way you talk."
• Students may have very different views of the written word than librarians have (on punctuation, spelling, etc.), which affects chat behavior.
• Both parties may be somewhat disoriented by chat because of the novelty of the communication mechanism.

From the above, we can begin to draw a few conclusions. In the asynchronous world of digital reference, there is more concern about the structure of the inquiry, in the use of forms and guidelines for email messages but that is a phenomenon seen less frequently in synchronous services. Asynchronous service transactions also take more time to complete, and this increase in time needed permeates the process: more time (and, perhaps more thought) is required to formulate and submit a query; time may elapse before the query is read and responded to; and more time is required to formulate and send a response. Considerably more data is collected regarding an inquirer's personal characteristics than is collected about the nature of the inquirer's query itself, and the use to which this personal information is put is not known.

It appears that people designing synchronous digital reference services are attempting to make them more like traditional services in the use of a real-time interview and response process. In general, the services provide hours of service and give fuller descriptions of what the service entails as well as its parameters. Time also seems to be an important concern, although here it is the pace and flow of the interaction that is discussed.

Overall, the thinking and discussion about synchronous services reflects an attempt to adapt reference practice to a new set of technological environments. This is also, of course, the case in asynchronous environments. Perhaps the most intriguing question raised is whether the most fruitful and appropriate approach is to adapt reference practice to the technological domain, or vice versa.

Discussion
Several issues have emerged as central to a greater understanding of the development of techniques that efficiently and effectively identify and refine users' information needs. In general, this is a period of feeling through how best to develop and deploy reference services in an emerging technological and information environment, as well as how to offer those services in a professional and high-quality manner. In the process, new phenomena, such as the disappearing questioner and the "probe query" [2] described above, are being observed.

Among the most important matters arising recently in the literature and in the previous discussion are:

• **The growth of synchronous services:** There has been an increase in the number of synchronous services using a variety of technologies (chat, call center software, instant messaging, and discussions of future use of video and voice-over-IP), and that raises questions about what situations and settings are most appropriate and useful for them.

• **The multiple effects of time:** An important factor to consider is the perceived trade-off between response time and depth of response: the luxury of time in an asynchronous digital reference service to compose a thoughtful response to a possibly more thoughtfully composed question as opposed to the quicker but possibly more shallow response to queries in a synchronous digital
reference service.

- **New phenomena**: Digital reference phenomena such as the disappearing questioner and the probe query were not experienced by those providing traditional reference, and these phenomena must be considered.

- **More consortial services**: While consortial digital reference services—for example, QuestionPoint, statewide and provincial services (in Ohio, Washington, Alberta), or multi-institutional services (college libraries in Illinois and Pennsylvania)—provide potentially large benefits, such services also raise questions of equitable sharing of questions and resources, specialization of staff and services, trust and the transportability of the initial reference interview.

- **Greater scrutiny of digital reference**: Those organizations or communities that provide funding for reference services will ask many questions about costs, including, for example: Is it really worth the time and effort to mount real-time or collaborative services [Coffman, 2002]? Should their libraries provide the financial and human resources used to answer questions from people who are not in their communities?

We see then that digital technologies have affected digital reference professionals in several ways. However, it appears that the partnership of digital reference and technology is a marriage of convenience at best, and there are clearly strains in adopting practice to the currently available technologies. As digital librarians attempt to fit their familiar practices to a new world, they may find the digital reference practices inadequate, frustrating, or difficult. Perhaps we should consider instead how the technology can be adapted to and support librarians' needs, rather than have the technology dictate how librarians are to meet those needs.

There is much work to be done in developing effective ways to satisfy user information needs and assist digital reference professionals to manage questions. Several organizations have begun this work, and examples are discussed below.

### 2 Technology to Support Digital Question Negotiation

Although technology is obviously an important aspect of the question negotiation discussion, to this point technology has been seen primarily as a communications medium, a networked means to solicit and respond to user queries.

It might be worthwhile, however, to think about how the process of question negotiation, again broadly understood, might otherwise be supported by technological means. This leads to a new question:

*In what ways could automation augment the digital reference professional's capabilities in question negotiation?*

Email was originally used for digital reference, more as an extension of the familiar use of email than as a new method of outreach by librarians or a new way of thinking about reference work. As the Web gained greater acceptance, email became a principal tool in the arsenal of the digital reference librarian. (It was at this stage that the notion of "digital reference", "virtual reference" or "electronic reference" began to be discussed as a new kind of service.) In the last several years, web forms have found widespread use, adding richness and complexity to what were previously simple exchanges of personal email.

Services have evolved, and there is growing recognition of the need for supporting the intake, management, response and archival functions of reference services. One of the first special purpose solutions was QRC (produced by the Internet Public Library), which was designed to carry out the functions necessary to manage digital reference transactions. These functions included:

- Taking queries via email or web form,
- Analyzing the queries,
- Performing query triage,
- Posting queries,
- Claiming responsibility for answering the queries,
- Formulating answers,
- Communicating and sharing information internally (between librarians),
- Posting responses to inquirers,
- Archiving responses, and
- Enabling search of archived, completed responses.

(See the article by Lagace and McClennen for further details [Lagace & McClennen, 1998].)

With varying degrees of success, other services have attempted to adapt commercial software to the needs of reference. In the last two years, 24/7 [McGlamery, 2001], LSSI [Coffman, 2001] and a variety of new vendors and players are entering the new digital reference software marketplace with standalone products and, more recently, with software that is interoperable with their integrated library systems. Rather than adapt existing software for new processes, however, other organizations have begun to create automation solutions that are driven by digital reference practice. What follows is a brief description of four such automation initiatives and an overview of how they support question management in digital reference services rather than impose technologically related constraints.

**QuestionPoint**

The most visible software development effort of late has been undertaken by OCLC as part of their contribution to the QuestionPoint service [Kresh, 2000; Quint, 2002]. An outgrowth of the Collaborative Digital Reference Service initiated by the Library of Congress several years ago, QuestionPoint is an attempt to build a global collaborative service drawing upon the skills, expertise and resources of many libraries and librarians. The QuestionPoint resources are intended to provide support to libraries in responding to questions beyond the scope of their local resources. This software is still in development, but it already encompasses many features familiar to users of QRC or 24/7, as well as many new ones.

One intriguing development, and a centerpiece of the QuestionPoint service, is its routing function. The designers of this service continue to develop and refine algorithms for sending a query to a particular service based on its subject matter and the collection and service strengths of members, as well as deadline, load, geography and a variety of other factors. Work on this function is ongoing, and while there have been some early difficulties, it seems to hold promise as an important and necessary feature of any large-scale cooperative service—streamlining the repetitive and time-consuming tasks of question routing.

**MULDER**

Full automation is an intriguing concept, very much in the long tradition of artificial intelligence and expert systems. (For more information, we refer the reader to Richardson; Alberico & Micco; and McCrank who have provided useful discussions about automating the reference process [Richardson, 1995; Alberico & Micco, 1990; and McCrank, 1993].)

In an article published in 2001, Kwok, Etzioni and Weld discuss MULDER, which they describe as "the first general-purpose, fully-automated question-answering system available on the web" [Kwok et al., 2001]. MULDER was designed to answer what these authors call "factual questions", and what most librarians would call "ready-reference questions", such as "Who was the first American in space?" or "What is the second-tallest mountain in the world?" While an extensive discussion of the system is beyond the scope of this article, a brief description follows.

The MULDER system has three components. The first is a retrieval engine that sits on top of the document collection and handles retrieval requests. In the context of the Web, this is a search engine that indexes web pages. The second component is a query formulation mechanism that translates natural-language questions into queries for the information retrieval (IR) engine in order to retrieve relevant documents from the collection, i.e., documents that can potentially answer a particular question. The third component, answer extraction, analyzes these documents and extracts answers from them.
In testing their system, Kwok, Etzioni and Weld used questions from the TREC-8 text retrieval competition and compared their system with the performance of Google and AskJeeves. MULDER outperformed both in terms of recall and "user effort" (actually a word distance metric) to achieve given levels of recall. Thus, although there has been no apparent attempt to extend this work or adopt it in any production-level service, MULDER shows promise for implementation and use in a practitioner's setting.

A major challenge to this system would be dealing with questions that require answers with high levels of synthesis and questions of broad scope, rather than factual questions within a narrow scope. And yet, broad scope, high—ambiguity questions are common to many digital reference service organizations that may find themselves in increasingly unfamiliar support scenarios.

**QABuilder Software**
Most digital reference services can be described within the VRD project's 5-step model [3], including question acquisition, triage, expert answer formulation, tracking and resource creation. Until recently, most of these steps, from the most repetitive tasks to the complex synthesis of answers, were performed by human intermediaries:

Several years ago, the Virtual Reference Desk project (VRD) at the Information Institute of Syracuse (IIS) created a software tool designed to relieve humans of some of the repetitive and low-level tasks involved in providing digital reference. Among the features of this software are:

- Questions are submitted within controlled categories, decreasing the time required to assign topics and increasing the accuracy of question routing.
- Included in the software is a built-in, Java-based Spell checker for the experts' convenience.
- Questions and answers in an QABuilder system are QUIP [4] enabled, thus the entire thread of a question (including claim, answer, addendum, feedback, follow-up, and edits) is preserved. System administrators have automated tools for viewing archived questions and changing their status, viewing questions in progress, changing expert profile information, and removing questions from the archive.
- QABuilder Software allows manipulation of the knowledge base, including sorting questions and answers by popularity, assigning one question to be stored in multiple categories, and searching and browsing.

More recent development of QABuilder Software includes further automation including:

**Question Acquisition**
Experts can now receive email notification of new or follow-up questions rather than having to log on to view available questions.

**Triage**
This system can match questions to experts based on audience and categories, and can route questions automatically to expert queues, or to the Administrator intermediary via email or QUIP.

**Experts Answer Formulation**
Experts can use this software to claim questions, to save them to finish at a later time and to send and receive private messages with administrators. They can also send problem questions to administrators along with comments and submit answers to the administrator for training assistance. Finally, this system allows administrators to customize automated outgoing emails and create pre-formulated answers to use when answering common questions.

**Tracking**
QABuilder Software enables real-time reporting on question activity, turnaround times, expert performance, user activity, and most viewed questions. It also checks for orphaned (i.e., questions with no matching experts) or overdue questions and allows viewing of question threads.
Resource Creation
This version of the software allows administrators to customize over 20 service-level options and to create unlimited categories and audience levels. Administrators can update profile information (change categories and audience levels), manage patron, expert, and administrator accounts, and customize all registration forms.

These automation initiatives described above represent some of the ways that automation is supporting the digital reference process, but many other organizations are working on these problems too, and even more sophisticated methods are being developed. The following description illustrates a research initiative that seeks to further relieve human intermediaries of repetitive tasks—in this case, sophisticated decision making regarding the routing of questions.

The National Science Digital Library
Like QuestionPoint, the National Science Digital Library (NSDL) [5] is interested in automating question routing. However, the NSDL researchers working on this initiative (researchers from the University of Massachusetts at Amherst, Syracuse University and the Wondir Foundation) have slightly different goals from those of QuestionPoint.

The NSDL Question Triage project seeks to enhance the NSDL by merging its information retrieval (IR) and digital reference components. By combining these functions, the team hopes to automatically determine whether incoming questions can be answered by online resources or need to be answered by human experts, and in the case of the latter, which expert. In this way, human intermediaries can handle questions beyond the scope of the current IR system. Therefore, high context queries, or those queries requiring expert judgment where even the best IR systems may fail, can be routed and handled by experts.

If a query is determined to need human intermediation, the system will use inductively developed profiles of available experts to route the user's query to the most closely matched online expert. The end result of the research described here will be an automatic process that takes a user's query and produces a multi-dimensional classification of the query that can be used as the basis for refining it, directing it to specific online resources, or directing it to human experts. This system would be able to address questions of higher ambiguity and context, further simplifying the experts' tasks and improving scalability.

3 Summary
Software is an important part of digital reference service development and maintenance and provides technological support for the processes of taking in and responding to inquiries.

We have pointed out that technology is sometimes the unintentional driver for software design in automated digital reference systems and has changed the role of digital reference professionals: they now must manage increasing numbers of questions from users with various levels of technical and searching experience, who have less time to search and who ask repetitive questions.

We note, too, that technology can be driven by digital reference practice, and we have briefly described automated components of some systems that complement the ways in which people are already managing question negotiation. Various systems have used automation to receive, route and answer questions. We believe that future initiatives—if based on the needs of the digital reference professional—will result in systems that are easy to use, high quality, and effective in reducing the number of repetitive or low-level tasks carried out by digital reference professionals.

Successful automation will allow more time for human intermediators to do the job that computers currently cannot—provide complex and highly synthesized answers to ambiguous or broad questions from their users.
4 Acknowledgements
Many thanks to Lorri Mon, who provided invaluable assistance and conducted the research study discussed in Section 2. The results of that research will be published under her name. Thanks also to the subscribers to DIG_REF, specifically the people who participated in the reference interview thread last October, in particular Sara Weissman at the Morris County Library, Camilla Baker at Canisius College Library, Pauline Lynch of AskERIC and Patricia Memmott of the Internet Public Library.

5 Notes
[1] Taylor's article (1968) is easily one of the most influential and cited works in the library literature, and describes four levels of question formulation (visceral, conscious, formalized and compromised needs) and the five filters that librarians use to help understand information needs (determination of subject, objective and motivation, personal characteristics of inquirer, relationship of inquiry to file organization, anticipated or acceptable answers).

[2] The "probe query" may not be all that new; it's described as a vague initial question that is followed by more specific ones. This is not unusual in face-to-face or telephone reference encounters, when people ask quite broad questions, either to feel out the librarian or based on a belief that all like information is located together and all they need to do is get to that place and they can find it themselves. Taylor (1968) among others describes this phenomenon [Taylor, 1996].


[4] QuIP is a threaded data format that relies on metadata to maintain, track, and store services' questions and answers in a consistent file format. See <http://vrd.org/Tech/QuIP.shtml>.

[5] The National Science Digital Library (NSD) is a National Science Foundation program to create a digital library for mathematics and science. The home page for the NSDL is at <http://www.nsdl.org>.

6 References


