

7-b: Digital Library Curriculum Development

Module: Reference Services

Draft: 7 October 2009

1. **Module name:** Reference Services

2. **Scope:** This module presents an overview of human-mediated assistance for information seeking and question answering, and addresses issues in integrating these types of services into DLs.

3. **Learning objectives:** Students will be able to:
 - a. Articulate the pros and cons of varieties of services for meeting different types of user needs.
 - b. Articulate the differences between human-mediated reference and automated information retrieval (IR) and question answering (QA) services.
 - c. Identify the appropriate forms of service that may be integrated into DLs in different contexts and different user communities.

4. **5S characteristics of the module:**
 - a. Spaces: Reference is provided in an information space, physical or online.
 - b. Scenarios: Reference is an info seeking technique employed by users in specific situations / contexts / anomalous states of knowledge.
 - c. Societies: Reference is provided by a community of answerers (usually librarians and/or subject experts) and used by communities of users.

5. **Level of effort required:**
 - a. Prior to class: Approximately an hour, for a homework assignment (see Exercises, below).
 - b. In class: 1½ hours. Each of the following sections may take approximately ½ hour.
 - i. Human intermediated reference
 - ii. Automated reference
 - iii. Use of DL resources

6. **Relationships with other modules:**
 - a. 6-a: Info needs, Relevance: Module 6-a is a prerequisite to 7-b.
 - b. 6-b: Search strategy, info seeking behavior, user modeling: Module 6-b is a prerequisite to 7-b.
 - c. 7-a: Search engines, IR, indexing methods: Modules 7-a and 7-b should be taught around the same time in the semester.

7. Prerequisite knowledge required:

- a. General understanding of library reference services.
 - i. In LS programs: Completion of the introductory, general library reference course.
 - ii. In CS programs: Familiarity with and prior use of a library reference service.

8. Introductory remedial instruction:

- a. Introduction to library reference (This section may be skipped in ILS programs where most students have already taken the introductory Reference course.)
 - i. Depth of service provided: citations or links to resources, and/or an answer to the question.
 - ii. Steps in the reference interview (from Bopp, 2001):
 1. Opening the interview:
 - a. Librarian should have a welcoming manner and seem approachable.
 - b. Librarian should have opening questions to prompt the user.
 - c. Librarian's response to the user's question should be encouraging.
 2. Negotiating a question:
 - a. Librarian should use open-ended questions to elicit information from the user about the information need.
 - b. Active listening is important to elicit information from the user.
 - c. Asking why the user needs the information may be useful, but is controversial.
 3. The search process:
 - a. In what types of sources, or specific sources will information be found?
 - b. Opportunity for the librarian to instruct the user in the use of library resources.
 4. Communicating information to the user:
 - a. How much work the user will have to do to make sense of the information depends on how complicated the question is.
 - b. At an appropriate level, given the user's age, language, previous knowledge of the topic, etc.
 - c. Differences in providing information in person (desk reference) vs. at a distance (digital reference)
 5. Closing the interview:

- a. Ask if the question has been answered completely and to the user's satisfaction.
 - b. Encourage the user to return if she has more questions.
 - iii. The role of the librarian in assisting the user to locate information and information sources.
 - 1. Instruction in the use of the library / navigating within the physical space and the cataloging system.
 - 2. Bibliographic instruction: instruction in the use of specific sources.
 - 3. Information literacy instruction: instruction in the evaluation and use of information.
 - 4. Reference expert systems & QA systems presently do not provide instruction.

9. Body of knowledge:

- a. Major themes / tensions in reference service for DLs:
 - i. Human intermediation vs. automation
 - ii. For human intermediation: synchronicity of the interaction, and media in which the interaction is conducted
 - iii. For automation: reference expert systems vs. question-answering (QA) systems
- b. Human-intermediated digital reference
 - i. Asynchronous reference
 - 1. Current forms: email, web
 - 2. Time-delayed
 - a. Duration of the delay may be set by the service's policy (e.g., 2 business days)
 - b. Delay allows the librarian to conduct research & hopefully craft a response.
 - ii. Synchronous reference
 - 1. Current forms: commercial chat application (similar to CRM software), instant messenger (IM) applications
 - 2. Conducted in real-time
 - a. Closer in speed to a face-to-face conversation
 - b. Immediacy may be preferable for users; more similar to search engine use.
 - iii. Expensive to maintain & not easily scalable:
 - 1. Human staff is expensive.
 - 2. Answering questions takes time.
- c. Automation
 - i. Reference expert systems
 - 1. Attempt to automate a librarian's decision-making process in identifying resources to fulfill an information need.

- a. Does not provide answers, just resources.
 - 2. Work dropped off at the end of the 1980s, early 1990s, coincident with trends in AI research away from modeling human cognition as a whole and focusing instead on human thought processes in specific contexts.
 - ii. QA systems
 - 1. TREC conference, QA track
 - a. “The goal in the QA task is to retrieve small snippets of text that contain the actual answer to a question rather than the document lists traditionally returned by text retrieval systems. The assumption is that users would usually prefer to be given the answer rather than find the answer themselves in a document” (Voorhees, 1999, p. 77).
 - b. Five-year roadmap for QA research (Burger, et al., 2001)
 - i. Six key directions for QA research:
 - 1. Scope: open vs. narrow domain
 - 2. Context: disambiguation of
 - 3. Judgment: answers must be justified
 - 4. Multiple sources: information for an answer may be spread across multiple documents
 - 5. Fusion: an answer must be formulated from information from multiple documents
 - 6. Interpretation: disambiguating unclear questions
 - 2. Distinct from IR systems:
 - a. IR systems provide links to documents, not answers.
 - b. QA systems provide “snippets” from documents that hopefully contain answers.
 - iii. Expensive to build, but more easily scalable:
 - 1. Automation requires large up-front costs, but ongoing costs are less.
 - 2. Economy of scale: answering many questions may be no more expensive or problematic than answering few.
- d. Use of DL resources in responses / Collection development
 - i. Both human intermediaries and automated QA systems may mine the DL’s collection for materials.
 - 1. May also mine other DLs, or any site online.
 - 2. Legal & copyright issues involved in using others’ materials.
 - ii. Questions & responses are annotations to the materials in the DL.
 - 1. E.g., Object X in the DL is useful for answering questions on topic Y.

2. Over time, materials in the DL will develop a “profile”: usage data about what types of questions or what information needs a resource may be used to answer.
 3. Sub-collections may be built of the most popular materials, by “harvesting” the materials pointed to in responses.
 - a. For a similar idea in automated collection building: see Bergmark (2002).
- iii. Responses are resources in their own right & may be added to the DL collection.
1. Store answers “raw” or cleaned?
 - a. Privacy issues involved in making questions and responses searchable.
 - i. Note: There are very serious privacy issues involved here. See, e.g., news about AOL’s 2006 release of 650,000 users’ search histories
 - b. Methods for deidentifying questions and responses: see Nicholson & Arnott-Smith (2005).
 2. Collection development-related issues:
 - a. Collection is built based on gaps in users’ knowledge, rather than based on the mission and goals of the DL.
 - b. Requirements for collection development policies.

10. Resources

- a. Required readings:
 - i. Pomerantz, J. (2003). Integrating Digital Reference Service into the Digital Library Environment. In R. D. Lankes, S. Nicholson & A. Goodrum (Eds.), *The Digital Reference Research Agenda* (pp. 23-47). Chicago: Association of College and Research Libraries.
 - ii. Choi, Y. (2006). Reference services in digital collections and projects. *Reference Services Review*, 34(1), 129-147.
- b. Reference in general
 - i. Reference and User Services Association (RUSA)’s Definitions of a Reference Transaction:
<http://www.ala.org/ala/mgrps/divs/rusa/resources/guidelines/definitionreference.cfm>
 - ii. RUSA’s Guidelines for Information Services:
<http://www.ala.org/ala/mgrps/divs/rusa/resources/guidelines/guidelinesinformation.cfm>
 - iii. Bopp, R. E. (2001). The Reference Interview. In R. E. Bopp & L. C. Smith (Eds.), *Reference and Information Services: An Introduction* (pp. 47-68). Englewood, CO: Libraries Unlimited, Inc.
- c. Digital reference in general

- i. Penka, J. T. (2003). The Technological Challenges of Digital Reference: An Overview. *D-Lib Magazine*, 9(2). <http://www.dlib.org/dlib/february03/penka/02penka.html>
 - ii. RUSA's Guidelines for Implementing and Maintaining Virtual Reference Services: <http://www.ala.org/ala/mgrps/divs/rusa/resources/guidelines/virtrefguidelines.cfm>
- d. Digital reference in digital libraries
 - i. Pomerantz, J. (2003). Integrating Digital Reference Service into the Digital Library Environment. In R. D. Lankes, S. Nicholson & A. Goodrum (Eds.), *The Digital Reference Research Agenda* (pp. 23-47). Chicago: Association of College and Research Libraries.
 - ii. Choi, Y. (2006). Reference services in digital collections and projects. *Reference Services Review*, 34(1), 129-147.
- e. Asynchronous human-intermediated digital reference
 - i. Janes, J., Hill, C., & Rolfe, A. (2001). Ask-an-Expert Services Analysis. *Journal of the American Society for Information Science and Technology*, 52(13), 1106-1121. <http://dx.doi.org/10.1002/asi.1177> .
 - ii. Lankes, R. D. (1998). AskA's (K-12 digital reference services). *Reference & User Services Quarterly*, 38(1), 63-71.
 - iii. Lankes, R. D. (2004). The Digital Reference Research Agenda. *Journal of the American Society for Information Science and Technology*, 55(4), 301-311. <http://dx.doi.org/10.1002/asi.10374> .
 - iv. Pomerantz, J., Nicholson, S., Belanger, Y., & Lankes, R. D. (2004). The Current State of Digital Reference: Validation of a General Digital Reference Model through a Survey of Digital Reference Services. *Information Processing & Management*, 40(2), 347-363. [http://dx.doi.org/10.1016/S0306-4573\(02\)00085-7](http://dx.doi.org/10.1016/S0306-4573(02)00085-7)
- f. Synchronous human-intermediated digital reference
 - i. Francoeur, S. (2001). An Analytical Survey of Chat Reference Services. *Reference Services Review*, 29(3), 189-203. <http://dx.doi.org/10.1108/00907320110399547> .
 - ii. Pomerantz, J. (2005). A Conceptual Framework and Open Research Questions for Chat-based Reference Service. *Journal of the American Society for Information Science and Technology*, 56(12), 1288-1302. <http://dx.doi.org/10.1002/asi.20215>
- g. Reference expert systems
 - i. Richardson Jr., J. (1995). *Knowledge-Based Systems for General Reference Work: Applications, Problems, and Progress*. San Diego: Academic Press. In particular, ch. ?.

- ii. Richardson Jr., J. (1989). Toward an Expert System for Reference Service: A Research Agenda for the 1990s. *College & Research Libraries*, 50(2), 231-248.
- h. Question-answering systems
 - i. Burger, J., et al. (2001). Issues, Tasks and Program Structures to Roadmap Research in Question & Answering (Q&A). Gaithersburg, MD: National Institute of Standards and Technology. http://www-nlpir.nist.gov/projects/duc/papers/qa.Roadmap-paper_v2.doc
 - ii. Voorhees, E. M. (1999, November 16-19). *The TREC-8 Question Answering Track Report*. The Eighth Text REtrieval Conference (TREC 8), Gaithersburg, MD. http://trec.nist.gov/pubs/trec8/papers/qa_report.pdf
 - iii. Voorhees, E. M., & Tice, D. M. (1999, November 16-19). *The TREC-8 Question Answering Track Evaluation*. The Eighth Text REtrieval Conference (TREC 8), Gaithersburg, MD. <http://trec.nist.gov/pubs/trec8/papers/qa8.pdf>
 - iv. Also: all Overviews of the TREC Question Answering Tracks from 2000-present. See the TREC publications at: <http://trec.nist.gov/pubs.html>
- i. Other related topics
 - i. Bergmark, D. (2002). Collection synthesis. Proceedings of JCDL 2002. pp. 253-262. <http://doi.acm.org/10.1145/544220.544275>
 - ii. Neuhaus, P. (2003). Privacy and Confidentiality in Digital Reference. *Reference & User Services Quarterly*, 43(1), 26-36. <http://search.epnet.com/direct.asp?an=11298320&db=afh>
 - iii. Nicholson, S. (2005). A framework for Internet archeology: Discovering use patterns in digital library and Web-based information resources. *First Monday*, 10(2). http://firstmonday.org/issues/issue10_2/nicholson/index.html
 - iv. Nicholson, S., & Arnott-Smith, C. (2005, October 28 - November 2). *Using Lessons from Health Care to Protect the Privacy of Library Users: Guidelines for the De-Identification of Library Data based on HIPAA*. Paper presented at the The American Society for Information Science & Technology 2005 Annual Meeting: Sparking Synergies: Bringing Research and Practice Together, Charlotte, NC.

11. Concept map

12. Exercises / Learning activities

- a. Homework to be completed prior to this lesson:
 - i. Students should submit the same question to one desk reference service and two digital reference services that utilize different media:

email, web, chat, IM, etc. The three services may be provided by the same library.

1. Discussion points for this assignment:
 - a. How effective was the librarian / answerer in responding to your question? (Effectiveness may be operationalized as some or all of: completeness of the answer, accuracy, speed, user satisfaction.)
 - b. What differences were there in the responses that you got to your question in the different media?
 - c. Is one medium superior or inferior to others for responding to this question / this type of question? Why?
 2. Some suggestions for locating services:
 - a. The Virtual Reference Desk's AskA+ Locator (asynchronous services):
http://www.eduref.org/Resources/Reference/AskA_Services.html
 - b. Three lists of libraries providing synchronous reference services:
 - i. http://liswiki.org/wiki/List_of_libraries_providing_virtual_reference_services
 - ii. <http://www.public.iastate.edu/~CYBERSTACKS/LiveRef.htm>
 - iii. http://www.libsuccess.org/index.php?title=Online_Reference#Libraries_Using_Virtual_Reference_Services
 - c. Additionally, most academic and public libraries provide digital reference services.
- b. Homework to be completed prior or subsequent to this lesson:
- i. Data analysis of previously-answered digital reference questions.
 1. These may be from synchronous or asynchronous services. The form of the previously-answered questions will differ for different types of services: email will produce email threads; chat will produce transcripts.
 2. These may be archives from the university library's services, or from the Internet Public Library (<http://www.ipl.org/>), if the instructor's school is an IPL member.
 3. Analysis should be performed on a small-ish corpus: maybe 50-100 questions, depending on the number of students.
 4. Data analysis can take a variety of forms:
 - a. Simple descriptive statistics: Number of questions categorized by topic, from particular user groups, submitted by day of the week or week of the semester, etc.

- i. The IPL for example has its own pre-defined topic and user categories, or the instructor may choose other schemes.
 - b. Use of DL materials: What materials from the IPL (or other DL) are used in answers? What percentage of materials provided in answers are from the IPL?
 5. IRB approval may be required for this assignment. The IPL has IR approval at Drexel, which will assist any other institution in getting local IRB approval.
- c. Discussion in class:
 - i. Look at DLs that integrate reference services (see Choi, 2006). Look at archived questions, if available. Some question archives that are available online include:
 1. The Internet Public Library, Ask A Question: <http://www.ipl.org/>
 2. The National Science Digital Library, AskNSDL: <http://ask.nsdlib.org/>
 3. The MAD Scientist Network: <http://www.madsci.org/>
 - ii. Look at QA systems with no associated DL. E.g.:
 1. START: <http://start.csail.mit.edu/>
 2. Yahoo Answers: <http://answers.yahoo.com/>
 3. AnswerBus: <http://www.answerbus.com/>
 - iii. Look at other DLs with no reference service or QA system in place.
 - iv. Discussion points for class:
 1. What sorts of questions do users submit to services of different types? What sorts of questions are answered best by services of different types?
 2. How are materials from DLs and other online resources used in answers?
 3. What would be the appropriate form(s) of service to integrate into several different DLs? Why?
- d. Lab in class:
 - i. Demo to students or have students use different applications for delivering digital reference.
 1. Students can use these applications as users, or as question answerers.
 2. Look to local libraries for specifics on available services.
 - ii. Alternative to a lab: It may not be possible to get applications for students to use for trials. So guest speakers from local libraries or online reference services would be appropriate. A speaker could demo one or more of the applications in use.
 1. Many states have statewide chat reference consortia, and most of these have an individual who is the service administrator. This individual would make an excellent speaker.

- a. List of libraries with chat reference services:
http://liswiki.org/wiki/List_of_libraries_providing_virtual_reference_services

13. Evaluation of learning outcomes

- a. Homework to be completed subsequent to this lesson:
 - i. Students should formulate a question to submit to a digital reference service, and reformulate the question as a query to submit to an automated IR or QA system.
 - 1. Discussion points for this assignment:
 - a. How effective was the service / system in responding to your question?
 - b. What differences were there in the responses that you got?

14. Glossary

- a. Asynchronous: Lag time exists between sending a message and receipt of a response.
- b. Expert system / Reference expert system: Automated system that identifies resources to fulfill an information need, by attempting to emulate a librarian's decision-making process for selecting resources.
- c. Human-intermediated: Questions are responded to by a human being, often a librarian.
- d. Question-answering system: An IR-like system that retrieves document snippets in response to a user query.
- e. Scalability: The ability and ease with which a service may increase in size; where size may be defined as number of users, throughput of questions and responses, etc.
- f. Synchronous: No lag time exists between sending a message and receipt of a response; more conversational and "real-time" pace.

15. Useful links

16. Contributors

- a. Initial author: Jeffrey P. Pomerantz
- b. Evaluators: