67th IFLA Council and General Conference
August 16-25, 2001

Collaborative E-Reference: A Research Agenda

Zorana Ercegovac
InfoEN, UCLA
E-mail: zercegov@ucla.edu

Introduction

Throughout the history of our profession, information-related activities have been often seen as diametrically opposing forces. Such have been pulling relationships between technical processing and public services, between library catalogs, periodical indexes, and finding aids, books and non-books, printed and born-digital carriers, between "fixed" objects and others that are dynamic, between librarianship and education, between libraries, archives, museums, and other information agencies. Recently, the notion of digital libraries has introduced a new competitor, computer scientists and engineers (see, for instance, *Communications of the ACM* May issue 2001), and threatened to displace or alter some of the professional library skills. Examples include skills in question negotiation, analytical bibliography, knowledge of gaps, crossovers, and resources, expertise in collection development, discipline-specific seeking behaviors, scholarly communication, and many other types of skills and expertise that I will be discussing at the IFLA Conference. While it is difficult to gain and maintain expertise in multiple tracks in parallel, joint workshops would help convene people together to discuss how we can as a group of professionals achieve the common goal: to organize collective knowledge and human memory for exploration and discovery.

Today, we need again to take a fresh and holistic look at what we are trying to do, to examine the nature of all types of resources that we acquire, describe, organize, arrange, preserve, access, select, and use. We need to understand characteristics of users, their abilities, inquiries, and

information tasks, retrieval models, and operational standards. Finally, we need to educate current and future professionals about the evolving nature of virtual public service components.

We all read papers on important issues of digital libraries, on declining uses of *cumbersome* traditional reference services, on sharply increased uses of *convenient* search engines, on early experiences of collaborative digital reference services, and on uses of various chat technologies (e.g., Apple VideoPhone Kit, CU-SeeMee, LivePerson, E-gain, MOO). (e.g., Morgan)

In contrast, I will outline several research areas that need to be investigated if we are to close the gap between seemingly inconvenient "reference desk" access to answers and those that are obtained relatively easy over the Internet. The selected areas are based on studies from diverse body of literature as well as from my own research, experience and affinity.

Keywords:

Knowledge maps, Models of E-reference, Neutral question answering, Representation of realities, Research in E-reference, Search vocabularies, Skills of reference librarians,

Task allocation, User studies

Research agenda: Unanswered questions

In the time that I have here, three inter-related areas will be introduced, each about 5 minutes in length, and each will be viewed at the intersection between traditionally different sides on the information plane. Each intersection is a moment of truth between users' behaviors and the following areas:

➢ **How finely do we need to represent the reality** (the IFLA *FRBR Study* and our call for a "Bibliographic Genome Project")?

➢ **How well do we communicate the reality** (interoperability between and among (controlled->search) vocabularies)?

➢ **How well do we divide labor between human reference experts and mass collaborative Internet-based programs**? Identifying those reference tasks that humans can consistently outperform machine intelligence can help us design optimal interfaces between people and machines in collaborative digital real-time reference services.

1. The FRBR (IFLA) entity relationship model for works, expressions, manifestations, and items ought to be examined with an eye of a reference provider and from the point of view of the consumer. Consumers will be increasingly remote rather than in-house, diverse in their capabilities and needs, and will have high expectations from 24/7 real-time virtual reference (similar to the ones in banking, automated gas stations, food services, and other self-help industries). For different populations of customers, can we start thinking of what would be optimal (good enough) display elements and relationships between the different entity groups? (Bibliographic Genome proposed by Ercegovac, 2001a; also see this author's keynote paper at http://www.cs.ucla.edu/Leap/zer/maribor.htm). This area has been mainly studied by the cataloging community and at a generic level only. In addition, we need better linkages between different types of resources. For example, at one level, there is the 856 field, electronic location and access, between a MARC catalogue record and networked documents. It links a bibliographic record that is displayed on WebOpacs with full text documents. Other issues will be related to attaching tags for quality assurance, especially to fluid digital-born documents. Examples include authenticity, provenance, permanency,

methodological integrity--reliability and validity. Other variables relate to the concept of genre in digital libraries (Beghtol; Kwasnik et al.; Toms), to education (Sutton), access (Lawrence and Giles), and the host of legal issues such as intellectual property, security, and use.

2. Another area that will be equally important to consider is the capability to use seamless languages both by the reference provider and the consumer in various question negotiations and answering settings. The word seamless here means that clients need a series of gentle transformation of search languages on the Web, ranging from a pictorial representation of a "thing" for the elementary school child, to conceptual maps and DDC for the high school student. These would be linked to LCC for general population, and to domain specific thesauri and codes (e.g., UMLS, IEEE, AAT, ULAN, TGN, SIC). (Ercegovac, 2001b). More than ever before, we need a search language that would ensure consistency, accuracy, precision, and negotiation power between the remote parties. Controlled languages typically include subject headings, classifications, names (personal, corporate, genres, geographic), and unified titles. This area has been traditionally studied by catalogers and classification researchers; it has recently been studied by others and renamed into knowledge organization tools, ontologies, and meta languages. Research is needed to investigate level of compatibility between and among languages for different user groups. Here, we have not touched upon communication languages that are needed for disadvantaged users if we are to provide equitable CDRS for all potential clients.

3. An estimated only 6 percent of the total number of Web documents contains scientific or educational content; some 83 percent contains commercial pages. Search engines are likely to index more popular (number of pages that link to a given page) and US sites (Lawrence and Giles). With billions of digital documents out there, even with embedded tiny networked sensors and actuators ('smart dust'), it will be mandatory to delineate reference tasks to be performed by trained and experienced people and others to be computer mediated at different levels of intervention. We need to address questions such as: (a) How to capture subtleties of the face-to-face neutral question negotiation (Abels; Dervin and Dewdney)? (b) How to design "dissolving" customized interfaces (van Dam)? (c) Which instructional / explanation modalities are optimal for which users at the point-of-need? (d) How to provide "high-touch" in evolving "high-tech" worlds? A model for different levels of human/computer interventions will be discussed at the Conference.

References

Abels, Eileen G. (1996) "The email reference interview." *RQ*, 35(3): 345-358.

Beghtol, Clare. (Dec 2000/Jan 2001) "The concept of genre and its characteristics." *Bulletin of the American Society for Information Science*, 27(2):17-19.

Dervin, Brenda and Patricia Dewdney. (Summer 1986) "Neutral questioning: A new approach to the reference interview." *RQ*, 25: 506-513.

Ercegovac, Zorana (2000) "Toward a global access to bibliographic information: Converging patterns, new paradigms." Invitational keynote presentation given at the *COBISS/SICRIS Annual*

Meeting. November 29th, Maribor, Slovenia. *Obvestila*, 5(4): 4-28. Also available in English at: http://www.cs.ucla.edu/Leap/zer/maribor.htm

Ercegovac, Zorana (2001a) *Accessing engineering global information for engineers: Phase 2*. Technical Report supported in part by the Engineering Information Foundation grant # EiF-99.8. Department of Computer Science, Henry Samueli School of Engineering and Applied Science (HSSEAS). The developmental part, Engineering Information Sources and Access (EISA) is at: http://www.cs.ucla.edu/Leap/Eisa/

Ercegovac, Zorana (2001b) "Accessing engineering global information for engineers: Phase 2." *Proceedings of the 64th Annual Meeting of the American Society for Information Science and Technology*. Vol 38. Washington DC., Oct 31-Nov 4, 2001. Medford, NJ: Information Today, published for the American Society for Information Science.

Functional Requirements for Bibliographic Records: Final Report. IFLA Study Group on the Functional Requirements for Bibliographic Records. (1998). International Federation of Library Associations and Institutions. IFLA Universal Bibliographic Control and International MARC Programme, Deutsche Bibliothek, Frankfurt am Main. Approved by the Standing Committee of the IFLA Section on Cataloguing. UBCIM Publications, New Series Vol. 19. Munich: K. G. Saur.

Kwasnik, Barbara H., Kevin Crowston, Michael Nilan, and Dmitri Toussinov. (Dec 2000/Jan 2001) "Identifying document genre to improve Web search effectiveness." *Bulletin of the American Society for Information Science*, 27(2): 23-26.

Lawrence, Steve and Lee Giles. (1998) "Accessibility and distribution of information on the Web." *Nature*, 400: 107-109.

Morgan, E. L. (August 1999) *See You See A Librarian Final Report*. Available at: http://sunsite.berkeley.edu/~emorgan/see-a-librarian/

Sutton, Stuart A. (1999) "Conceptual design and deployment of a metadata framework for educational resources on the Internet." *Journal of the American Society for Information Science*, 50(13) :1182 -1192. In: Ercegovac, Zorana, Guest ed., Special Topic Issue: *Integrating Multiple Overlapping Metadata Standards*. 50th Anniversary of JASIS.

Toms, Elaine G. (Dec 2000/Jan 2001) "Recognizing digital genre." *Bulletin of the American Society for Information Science*, 27(2): 20-22.

Van Dam, Andries. (March 2001) "User interfaces: Disappearing, dissolving, and evolving." *Communications of the ACM*, 44(3): 50-52.