Z39.50 and Navigation

Ray Denenberg, Library of Congress

Libraries and other institutions are creating collections, organized thematically (e.g. by subject, creator, historical period) with numerous, diverse objects, both digital and physical. These collections are often organized hierarchically and distributed across institutions and servers.

Significant resources may be invested in digitization and in the intellectual efforts of aggregation, organization, and description of the information in a collection. Yet to a user or client, a collection often appears to be simply an accumulation of undifferentiated data, because there is no agreed-upon semantics for navigating the collection, to locate and retrieve objects of interest. Coherent organizational structures must be imposed on the data, to provide a view that supports navigation.

A key obstacle to effective navigation is the inability to distinguish content from description. A primary goal of navigation is to locate and retrieve objects of interest; a vital step in that process is to locate relevant descriptive information. Thus it is useful to navigate among descriptive information as well as content, and consequently, to be able to distinguish content from description.

Various ad hoc descriptive formats have been developed, to describe collections as well as objects (e.g. finding aids, encoded archival descriptions, exhibition catalogs). Often these so-called *descriptive aids* do not have a well-defined structure and cannot be used alone for reliable navigation. At many institutions though, descriptive aids of various types have been created, at significant expense. It is therefore imperative that an application relying on descriptive information for collections and objects be able to utilize these available aids, rather than mandate the creation of new, redundant structures. In order to exploit these descriptive aids, a higher level enveloping structure is necessary, to allows users and clients to navigate among and select descriptive aids of interest.

Another navigational problem is posed by the various and often complex relationships among and between objects and collections. For a given object, there may be other objects that are duplicates or variations (e.g. different resolutions) or which bear other relationships (e.g. thumbnails). For any given collection there may be superior, subordinate, related, and context collections. These relationships among objects and collections must be modelled coherently.

Different digital objects, even within a single collection, may be retrievable by different protocols (e.g. Z39.50, HTTP, FTP). It is therefore important that structures describing these objects, and how they may be accessed, be defined independent of any specific protocol.

All of these problems amplify when a collection is distributed across institutions and servers. In particular some normalization of query semantics is necessary for coherent navigation of collections. Clients must be able to formulate queries that will be interpreted consistently across servers.

Z39.50 provides rich tools for effective navigation of digital collections. A Z39.50 *profile*: Z39.50 Profile for Access to Digital Collections has been developed over the past several months and is nearly complete. (Approval is expected this summer. The latest draft is available at ftp://ftp.loc.gov/pub/z3950/profiles/draft6.ps). The profile addresses the problems cited above; it models collections, object, descriptive aids, and defines an enveloping structure for describing an object or collection.

The profile, though limited in scope, explicitly anticipates the development, of several *companion profiles*, addressing specific disciplines, for example, museum collections and objects. The Digital Collections profile, together with the companion profiles, have several objectives. Among these are: to allow a server to clearly designate what is content and what is description; to allow a client to navigate among descriptive information; to model relationships among and between collections and objects; to provide semantics for queries to be interpreted consistently across servers, and data structures that provide semantics for navigation, ultimately to allow a user to locate and retrieve objects of interest.