Technical notes

The first section below contains summary statistics and the second section contains a description of our coding procedures.

<u>SUMMARY STATISTICS</u>

We coded 3,050 quotations in 86 profiles. The richness of the profiles varied greatly. Excluding two outliers (Cal-Berkeley with only 8 quotations, and British Columbia with 93) the other 84 profiles ranged from 11-58 quotations each, with an average of 35.

We ended up with 136 codes, with a wide variety in the amount of data coded to each. Excluding outlier codes (one code, "special collections", was coded to 218 quotations, and 17 codes attracted five or fewer quotations) the range was wide, from 6-132 quotations per code, with an average of 37.

The codes are organized in four topics: "Management & assessment" and "Services", each accounting for approximately 30% of the coded quotations; "Collaboration & support" accounting for approximately 25%; and "Collections" accounting for the remaining 15%. Each topic has several sub-topics. Topic and sub-topic codes begin with a "#". Below each sub-topic are the codes that comprise that sub-topic.

In Appendix A the "Topical framework" is presented through a series of eight charts that lays out all the codes according to their frequency of use. The first chart, called "Overview", displays the four topics with their main sub-topics. The following chart displaying the detail of each topic ("Collaboration & support" needs two charts, and "Services" needs three, so there are seven in all).

The boxes in the charts contain up to three pieces of information: (a) a percentage indicating the proportion of quotations a sub-topic represents of its larger topic, (b) the number (in parentheses) of quotations coded to that code, and (c) the percentage [in square brackets] of profiles that had quotations coded to that code.

The boxes are arranged from left to right and from top to bottom in descending order of frequencies. So for the row of boxes at the top of each chart, the ones on the left are the most frequent, and as you go vertically down each column, codes with fewer and fewer quotations are displayed.

Also a spreadsheet called "Frequency explorations" includes analysis of the variability of coding among the four different types of libraries, different regions, by index grouping, as well as a detailed presentation of the distribution of codes within each profile. The spreadsheet is

only available on the web at: . The numbers in the spreadsheet refer to the percentage of profiles that have quotations coded to a code. Therefore the purpose of this analysis is to see which codes had the most variation, by type of library or region, in the proportion of profiles coded to the code. Briefly, the analysis by type of library and by region group the codes by topic for example, following the same organization as the eight charts. Within each topic, the codes are divided into two sections, first the sub-topic codes, which accumulate groups of codes as a single number, and then each individual code. The percentage of profiles that had codings are then indicated for each type of library, or each region. There is also a column of standard deviations, and each group of codes is sorted by standard deviation.

CODING PROCEDURES

We have analyzed 86 library profiles. Our analysis plan followed these steps:

1. Codebook development. An initial set of 6 profiles were coded by Colleen Cook, producing 102 initial codes and code descriptions. Bill Potter and Martha Kyrillidou reviewed the codes, and added their comments to Colleen's code descriptions. Nick Woolf then synthesized these codes into a draft hierarchical codebook. The draft was reviewed by Bill, Colleen and Martha, and revised to a final version consisting of four main topics (*collaboration & support*, *collections*, *management & assessment*, and *services*) which subsumed 28 sub-topics in two or three levels of hierarchy. This hierarchy in turn subsumed 139 topical codes, the original 102 codes plus 37 new added over the course of data analysis.

2. First round of coding. Nick prepared a training document and held two conference calls to train the coders in use of ATLAS.ti data analysis software, and the intended use of the hierarchical codebook. An initial set of 22 profiles were divided among four coders, David Green, Michael Maciel, Jennifer Rutner, and Martha, and coded based on the training and the descriptions of the 139 topical codes.

3. Initial reliability assessment. After the first round of coding one coder reviewed the coding of one other profile, and wrote notes on any coding disagreements. This was followed by a conference call to discuss discrepancies in the use of codes and the amount of text allocated to each meaning unit, serving as a second training in use of the codebook.

4. Post-first round coding. The remainder of the profiles were allocated to the coders for several additional rounds of coding, until all 86 profiles were coded.

5. Subsequent reliability assurance. Recognizing that with available resources it was not feasible to adopt formal reliability assurance procedures, we adopted a plan to have a single individual (Jennifer) review all profiles, thereby assuring consistency in the coding. It is

possible that in some areas the scope and usage of some codes in some branches of the code hierarchy will differ from the original intent for those codes. Yet we feel assured of coding consistency overall. Additionally, based on the initial reliability assessment after the first round of coding, and the concreteness and general comfort of the coders with the great majority of the codes, we are satisfied that our procedures have produced a high quality of coding.

Following coding we took one additional step to code earlier profiles for later identified codes. After coding began, a total of 37 new codes were added at various stages as new topics were identified. The great majority were identified after the first two rounds of coding, that is, after approximately one third of the profiles had been coded. Resources did not permit recoding all earlier coded profiles each time later codes were identified, although we assume that in most cases if these topics had been found in earlier profiles then the new codes would have been proposed at that time, and in many instances the new topics were picked up in the earlier coded profiles during the review process. However, to assure ourselves that this was not a significant issue, on completion of coding Nick analyzed the frequencies of coding to identify the codes that had a significantly greater than 2/3 proportion of codings in the 2/3 of profiles coded after the first two rounds. We identified 14 such codes. Nick reviewed the quotations already coded to these codes, and in approximately one third of them found one or more keywords used when describing that topic. Nick then used the autocode feature of ATLAS.ti to find instances of these keywords in the text of the earlier 1/3 of profiles, and coded these examples to the 14 codes. This process generated 30 new meaning units. While not 100% exact, we feel this procedure has sufficiently corrected the coding for later identified codes.