

COL survey highlights

The Research and Assessment Unit prepared a survey of Council of Librarians (COL) members as a preliminary step in identifying perceptions about needs for a CUL data repository. The repository, similar in nature to the University of Pennsylvania Library's Data Farm, will support evidence-based decision making and enable the library to communicate more effectively with internal and external audiences. The survey ran for eight days (Feb. 12-19). 26 of 40 members surveyed completed the questionnaire for a response rate of 65%.

The following report is a brief encapsulation of the results. Further surveys and interviews with library decision makers and data customers will be conducted in the coming weeks to obtain a clearer picture of the responses.

Frequency of collected data use

52% of respondents indicated that they use CUL annual stats at least quarterly, with 20% indicating they use them at least once a month.

The same number (52%) used other internal metrics at least quarterly, but the number of users referring to these numbers at least once a month was higher at 32%. This suggests that while CUL annual stats are useful, other measures may be more useful to library decision-makers. The breakdown below provides some additional information on how the data are used.

Application of data

CUL annual stats data were considered most critical for internal communications, with 19% of respondents citing them as "essential". For decision-making and planning, only 4 % felt the annual stats were "essential"; two-thirds noted that the data were at best "somewhat useful". 38% of all respondents felt the annual stats were marginally useful, not at all useful, or not applicable for decision-making and planning.

Users had a different picture of the utility of other metrics they collect. For decision-making and planning, nearly 12% felt these data were "essential" (4% characterized the CUL annual stats this way for the same purpose.) 42% felt they were "very useful" for decision-making and planning. These results are consistent with the findings about the frequency of collected data use above and further suggest that the CUL annual statistics, while important, may have less impact for the library decision-making than other measures.

The CUL annual statistics were seen to be more valuable for external communications, with 35% rating them "essential" or "very useful" (only 12% of COL members rated other statistics that way for the same purpose.) Neither the annual stats nor other internal measures seemed to be exceptionally useful for external relations purposes. The question here, which needs further exploration, is where and how the data are falling short and what kinds of information could be used in their stead or as complements to the currently available data for more effective external communications. With the capital campaign and the establishment of a larger communications office, telling the library's story will be more critical. CUL may need to spend more time reflecting on what kind of

information will best allow the library to spread the message about its programs and services and their impact on the CU community.

Changes to data collection

When asked how data could be made more useful, respondents noted that archiving past data and allowing easy customization were most important. Consistent definitions of data categories and more convenient access to data closely followed in the rankings. Visual presentations, combining data in different ways, and clear narrative accompaniment did not rank as highly. 23% wanted other changes, which they specified in free-text comments. Some of the suggestions involved changes to the data collection process (e.g., employing more automation) or changes to data processing (e.g., the ability to sift out particular metrics at varying degrees of granularity). Interestingly, 3 respondents felt no change is needed to current data collection efforts.

Data mart “stock”

A working assumption of this investigation is that the data mart will be a long-term project. Information received from our colleagues at the University of Pennsylvania confirms that building a full-featured, scalable service with reasonable start-up and maintenance costs will not happen quickly. Thus respondents to the survey were asked to rank the kinds of data they thought should have the highest priority in “stocking” the data mart.

Nearly 40% of respondents ranked usage data (e.g., circulation transactions, e-resource usage statistics) as the highest priority. The second-highest ranking category was collection data (e.g., total acquisitions, breakdowns by formats, costs), with nearly 29% of respondents suggesting it as top priority. Other categories such as evaluative or impact data, cost data, and patron demographic data were considered far less important. Interestingly this kind of data is precisely the information CUL would need to communicate more effectively with external audiences. Yet COL members did not find it to be as critical.

Data mart features

Working on the assumption that not all desired features of a data repository would be available at the inauguration of the service, COL members were asked to prioritize which features were most important to them. The ability to generate reports was cited by 73% of respondents. It was by far the highest-ranked feature; the second-highest priority, the ability to add data and reports, was cited by 48% of respondents. Less than 25% of CUL members felt that other features –interface design, help functions, and customizable security levels—were high-priority items.

The clear preference for reporting functionality strongly suggests a desire to have more ready access to data and the ability to tailor that data to specific needs on an ad-hoc basis. However, there are significant technical and institutional obstacles to making this happen quickly or easily. Standardized category definitions and a mechanism for low-overhead ingest of data –not to mention the development of query structures and user-friendly reporting capabilities—are not now in place and will take considerable time and effort to identify, develop, and implement. The committee-driven decision-making model employed at CUL, which relies heavily on the time and energy of staff already committed

to numerous other projects, will almost certainly extend development time. Senior library decision makers should also consider that the strong desire for highly granular customization may complicate efforts to generate more high-level (but possibly higher impact) reports in the near term. Unit-specific needs and those of CUL as a whole will need to be carefully balanced to ensure the data mart's success.

New metrics to add

The most commonly mentioned new metric to add was usage data on e-resources, alone or in combination with cost data. Many also noted that statistics on use of traditional materials, especially in combination with some patron demographic data, would be useful for planning purposes.

Further steps

Some further steps that are currently planned in the investigation:

- a. Feedback from COL members on the results
- b. Follow-up with contacts (data collectors/users) on what file formats and other tools they use to collect data
- c. Ongoing discussions with IT staff (Adam Chandler et al.) about technical needs, capabilities, and costs
- d. Ongoing discussions with CUL stakeholders about options for the data mart
- e. Investigation into other data repositories either on campus or at other institutions. Continued dialogue with Penn colleagues will be especially important in this area.