COLLECTION DEVELOPMENT AND USAGE:
An Analytical Approach

M A Gopinath and I K Ravichandra Rao, *Documentation Research and Training Centre, Indian Statistical Institute, 31 Church Street, Bangalore 560 001.*

Discusses the importance of collection development process in the library service. Highlights the variables having influence on the collection development. The analytic techniques helpful in the measurement of document usage are mentioned. The helpfulness of such studies for collection development is indicated.

1 Collection Development

The development of a good collection is an essential prerequisite of any library and information centre. The library's efficiency and effectiveness is reflected in its collection development programmes. Methods, for formulating such programmes in turn depend on the techniques for analysis of a collection. In a sense, it is a performance analysis. We may assess it on the basis of the cost of purchase and replacement, circulation data, hours of accessibility, number of users, citation of documents available, collection in the texts of theses, technical papers, reports and books, dependence of the academic research and production programmes of the collection. In recent years, studies on the collections have been made to set norms on the (a) absolute size of the collection; (b) size in relation to the variables such as subject, date of publication, language, type of documents, number of volumes *per capita*, number of volumes per documents circulated; (c) growth rates of publications; (d) variety of media in which it is published; (e) transmission speed of the information contained; (f) usage of documents on the basis of circulation statistics; and (g) expenditure on the collections.

2 Size of the collection

The size of a collection creates a sense of proportion in its utility.

The minima and maxima of the size of collection have to be determined for effective library service. It is obvious that a collection in a library depends on its host organisation and the care the latter bestows on the library. However, we can derive the norm to determine the optimum size of the collection. Regression analysis would help us to do it. For example, we may have the data on the number of users who have borrowed at least once and the size of the collection which was used at least once in an academic year in different types of libraries. We can then identify whether or not there exists a linear relation, such as $y = a + bx$. Such a correlation would help us express the size of the collection as a variable of usage. The actual computation of the size also depends on several other variables, composition of faculties, curriculum, teaching method, geographic location of the campus and physical facilities. For instance, an equation may be formulated as follows:

$$Y = a0x0 + a1x1 + a2x2 + \ldots + a_n x_n + e$$

where a's and e are constants; and Y gives the required size of the collection for the given a_i s $(i = 0, 1, 2, \ldots n)$. $x0 \ldots x_n$ may refer to a variety of variables mentioned earlier.

3   Growth Rate

We have found that the phenomenon of growth is visible in a large measure in the publication of documents. It has projected a tendency of exponential growth. It has also reflected in the growth of collections. They have been growing in size and varieties. The computation of growth rate requires data on the collection size for a period of at least ten to twelve years — immediate preceding years. Let us assume that we have data on Y_t (the collection size in time t; $t = 0, 1, 2, 3, \ldots$). Let us also assume that t (collection) is growing exponentially at a constant rate of g. We can derive an equation as follows:

$$g_t = \frac{Y_{t+1} - Y_t \cdot}{Y_t} = \frac{b}{Y_t}$$

g_t, the growth rate, is a function of time. From the values of the constants b and Y_t, we can obtain the value of g_t where $t = 0, 1, 2, 3, \ldots n$. In such cases, g can be either based on the arithmetic mean or geometric mean.

4 Collection of periodicals

The periodical publications are an important part of the collection development. Their nature of periodicity and currency in relation to knowledge transfer dynamics have thrust this importance to a high level. But their increasing cost and contraint on the library budget have imposed the necessity for an objective analysis of the collection. Here comes the value of the Bradford's law. Bradford first described his Law in an article published in 1934 (1) and then in 1948 (2). It has been applied to several management problems of libraries. The Law states, "If scientific journals are arranged in the order of decreasing productivity of articles on a given subject, they may be divided into a nucleus of periodicals more particularly of articles on a given subject, they may be divided into a nucleus periodicals devoted to a subject and several groups or zones containing the number of articles as the nucleus, where the zones will be $1:n:n^2 \ldots$". In other words, if the top n periodicals contribute (or publish) 33.33% of the total number of articles on a given subject, the next top n^2 periodicals contribute only 33.33% of the literature, and the rest of the literature (that is, 33.33%) is scattered in n^3 periodicals. To be more specific, let us assume that $n = 10$. To obtain 33.33% of the articles, it is enough if we subscribe to the top 10 periodicals to obtain the next 33.33% of the articles. We have to acquire 100 periodicals. The rest of the articles (33.33%) is scattered over 1000 periodicals. It has been found that the Bradford phenomenon holds good in almost every field. Hence it is suggested that a proper quantitative analysis to determine the core collection of periodicals on a given subject should be carried out to assist a decision on the selection of documents.

5 User Studies

The quality of a library collection is tested on the basis of its extent and mode of use. There are two major methods for obtaining samples for use of library books (4).

  (a) The collection sample — selecting a sample need not necessarily a random one — of the total collection and gathering information on the past use of books in this sample;

  (b) The checkout sample — studying all books checked out from the library — during a specified period of time

Since most users remove books from the shelves without the knowledge of the stack maintenance staff and whatever has been removed are not necessarily read, it is rather difficult to quantify the use made of the docments. Many librarians, therefore, in the past suggested that circulation data as depicted by the transaction records of documents borrowing could be taken as an indicator of the use of library resources. A simple circulation statistics suggests that a document borrowed more frequently by the same or different borrower is more useful than the one borrowed less frequently. It demonstrates that circulation data can be used to measure the degree of the library resources. The borrowed documents data together with the borrower data provide information which can be used by the management for formulating and adopting policies for the acquisition and processing of documents.

For meaningful results, however, circulation data should be correlated with other variables such as data on teaching methods adopted, circulation change effected and university examination procedures adopted. In addition, it is possible to correlate the circulation data with the overall objectives of the library, physical facilities available in the library, hours of charging and discharging, etc.

While studies on the analysis of circulation data based on characteristics of users and documents, such as sex, status of users, subject of documents, etc. give a picture of library's uses, they can be further enriched by investigating.

1.   The pattern of relative frequency distribution of documents in the collection in a given period of time; and

2.   The proportion of the circulated documents to the total collection of the library.

This type of investigation would involve:

(a)   An analysis of the data in terms of its usefulness to the library management;

(b)   Developing a technique for collecting data; and

(c)   Identifying the types of deductions that can be made from the data.

The following two variables are usually considered in analysing the circulation data:

1. Number of times a book is borrowed; and
2. The last recorded circulation data.

The data collection in relation to the first variable involve the examination of due date slips attached to books. Such slips are most often filled up quickly owing to frequent usage of books and are then removed from the documents. In such cases, the new slips attached to them have only a few entries which may mislead the research workers. One can, however, preserve the old slips. But then, slips have to be organised properly to facilitate a proper commutation of frequency of use. Trueswell (8), therefore, suggested an alternative variable — the last recorded data. By plotting these data, we can show the relation of the percentage of circulation (having the last circulation data within the cumulative time period) to the specified time period, say, 90 per cent of the documents had a last circultion data within the preceding 12-month period.

So, by analysing the circulation data, we can also determine the size of the core collection and in turn, divide the collection into primary and secondary collection — one is the frequently used documents and the other is the infrequently used ones. We can then appropriately decide the storage pattern for these two collections.

6 Conclusion

The collection development is an efficiency audit aspect of a library. It is closely linked with the usage of the library. The usage is enhanced by appropriate organisation of collection, accessibility to such an organisation, circulation of documents and also the maintenance of documents collection. Thus, collection development is a synergistic programme calling in the best of talents of the library and information profession. With the development of quantitative techniques and the identification of the parameter of usage of documents, collection development process has been cruising towards objectivity. Ranganathan's First Law of Library Science "Books are for use" is a seminal guide for this purpose.

7 Bibliographical References

1. BRADFORD (S C). Sources of information on specific subjects. (Engineering); Jun. 26; 1934, 85–86.
2. ———. Documentation. 1948. Grosby, Lockwood, London. 156 p.

3. CLAPP (V W) and JORDAN (R T). Quantitative criteria for adequacy of academic library collection. (College and Research Libraries. 26; 1965; 371-80).

4. GOPINATH (M A). Forecasting techniques for library management. Paper presented to 12th IASLIC Conference (Roorkee) (1979).

5. LANCASTER (F W). Measurement and evaluation of library services. 1977. Washington. Information Resources Press.

6. McINNIS (R M). Formula approach to library size : An empirical study of its efficiency in evaluating research libraries. (College and Research Libraries. 33; 1972; 180-90).

7. RANGANATHAN (S R). Five laws of library science. Ed. 2. 1957. Bombay, Asia Publishing House.

8. RAVICHANDRA RAO (I K). Use of computers and sampling techniques for stock verification. (Paper presented at the 27th All India Conference held at Hyderabad during Jan 9-12, 1981 by the Indian Library Association).

9. _____. Document and user distribution in transaction records of Canadian university libraries. 1981. Ph.D. thesis, University of Western Ontario, London, Ontario, Canada.

10. TRUESWELL (R W). User circulation vs. size of holdings at three academic libraries. (College and Research Libraries. 30; 1969; 204-13).