

detailed review of the condition of the more important data files that might form an archival or file management center, concluded that the loss of data because of the failure to support good file management is distressing. (Appendix B.) The report of the study group formed at the National Bureau of Standards also emphasized the problem of file maintenance (appendix C) as did the Ruggles Committee report.

The reference problem

One of the serious obstacles to the fullest utilization of the information resources of the Federal Statistical System and to the effective association of its records in use, is the absence of any clearly defined reference function. The inadequate nature of this kind of service is directly traceable to the production orientation of the agencies stemming from their primary missions as data publishers. The reference function has generally been thought of as the responsibility of the documentary centers. To the extent that the agencies attempt to provide occasional referencing assistance, the task falls to an individual whose primary mission is defined in terms of the publication mission. The inadequate nature of this service is also traceable to the fragmented nature of the records of the Federal statistical system growing out of the divided responsibility for their generation and maintenance. The reference problem is made especially complex because of the decentralized character of the Federal statistical program. No agency has been in a position to perform a reference service in relation to the total file. The problem is becoming more acute as records are frequently demanded in disaggregate or special form not met by traditional documentary formats and are frequently used in ways that require extensive knowledge and understanding about the compatibility of records in several dimensions and the circumstances that condition their accessibility.

The problem of administrative records

Some of the most useful general purpose numerical records are generated as a by-product of administrative and regulatory procedures of Federal agencies. These agencies rarely interpret their missions to include the capacity to provide general purpose statistical services. Even when they might like to do so they experience serious difficulty because of the traditions, program priorities, budgetary procedures, and legislative authorities peculiar to their agency. Users who need to acquire from these agencies, and especially those who need to match these records with other files, find their task difficult. There is an important need to provide for the management and servicing of these records for general purpose statistical use.

The disclosure problem

The legal and administrative regulations on the disclosure of information supplied by individual respondents are becoming increasingly restrictive to the user. Only rarely is this because policy or research requires specific information about individual respondents per se. It is usually because of the need to associate sets of data in the interest of determining the interrelationship between two or more variables. The strain upon disclosure arises because matching several sets of data for consistency at levels of aggregation appropriate to the problem requires a retreat to elemental units in the process of constructing the necessary aggregates. It is not widely understood that the interest in micro-data and the existing pressures and constraints do not grow out of an interest in information about the specific respondent.

The fact that the strains upon the disclosure rules usually are of this form is fortunate because there are possible a number of servicing procedures based upon computer technology that can satisfy the needs of the user in most cases without violating disclosure regulations. Currently, however, the agencies of the Federal statistical system have only a very limited capability for performing the kinds of services that would lead to disclosure by-passes. The usage of the data is, thus, severely constrained, and valuable information is lost by aggregation at too early a stage in the analysis.

Problems stemming from the procedures used to generate data

It would be a mistake to conclude that the serious obstacles to effective use of the Federal statistical system under modern conditions is solely a product of its present inability to perform a series of user oriented services. Some of the most serious anomalies arise out of current practice in the production of data.