

purpose. The producers and users of data find their requirements and their missions intersecting in ways unknown a short time ago. Neither can continue to live the independent life formerly customary. The computer and other dimensions of social change have performed a shotgun wedding and both parties are in the process of discovering their incompatibility.

The procedural and program difficulties that led to this report are the product of this incompatibility. It is not uncommon in cases of this type for the assessment of difficulty to be one sided. The community of users has been sensitive to and vocal about many of the limitations of the producers of general purpose data—their mission concepts and institutional forms. The producers, in turn, can perceive many inadequacies on the part of the users. This report continues to be one sided in emphasis because it addresses itself primarily to the problems of use. The charge that framed it was couched in terms of the missions of the producing agencies of the Federal statistical system. It would be very worthwhile for some group to produce a companion evaluation of the anomalies in the production and usage of statistical information that arise from the practices and concepts of the user groups.

It is sometimes true that people who have concerned themselves with this problem are content with a superficial level of diagnosis and prescription. There is considerable attachment to the notion that most of our problems can be solved by computerizing all of the data we have in the backroom. This may be characterized as the "naive data bank" notion and its widespread acceptance is a source of some concern. I should emphasize here that this is not a characteristic of the Ruggles Committee report which was produced by knowledgeable and sophisticated people. However, the tendency to see the solution to the problems in relatively cheap technical programs has led to some misinterpretation of the Ruggles Committee report on the part of both those looking for additional support for data bank schemes and those reacting to the naive data bank concept. These evaluations and solutions are not based upon an adequate understanding and appreciation of the realities of the production processes essential to data generation or the institutional forms appropriate to their purpose.

The Ruggles Committee report gave us a healthy beginning toward an evaluation of this problem in realistic terms. However, this group did not have the time or staff resources to spell out the total problem set in a way that seems essential to support a more detailed consideration of program options. There is also a tendency in this report to see the problem primarily in terms of the accessibility of existing records and the solution in terms of the extension of user services. There is much that is valid in this representation but it gives insufficient attention to the important fact that accessibility is bound up with all of the production procedures and is inseparable in a number of fundamental respects from the issues related to the quality and scope of the existing records. It seems useful, therefore, to attempt a more precise formulation of the problem set to the solution of which the data center concept is addressed.

THE PROBLEM

The central problem of data use is one of associating numerical records. No number conveys any information by itself. It acquires meaning and significance only when compared with other numbers. The greatest deficiency of the existing Federal statistical system is its failure to provide access to data in a way that permits the association of the elements of data sets in order to identify and measure the interrelationship among interdependent activities. This deficiency has been partially overcome in a few vital areas where we need to trace and analyse the performance of the economy, by the establishment of special programs to bring together data sets in the form of national accounts, special index series, etc., but remains a debilitating constraint for most uses of data for analysis and planning. This is true for virtually all levels of use and for all purposes. It is a problem that plagues the research analysts inside and outside of the National Government who, for example, are engaged in building models of the economy in the interest of analyzing and projecting the major dimensions of economic growth and stability. It has been the principal obstacle to the administration's attempt to build a postattack revaluation and recuperation model. It is just as serious a problem for the uses that do not take the form of integrating data sets into a complex and formal model structure. The organization like EDA that wishes to establish a measurable test of eligibility for its program benefits faces the same problem. The business analyst who simply