Second, the question of costs is confused by the possible existence of joint costs in this program area. The general problem set of which the reference problem forms a part has several other dimensions that extend beyond the restricted problem this committee has taken for discussion. The costs of establishing a reference service of the kind discussed here would be quite different if it were established as a discrete function of, if it were developed as one component of, a more generalized user service capability. If the provision of reference services were combined with other statistical services addressed to establishing and maintaining archives, servicing administrative records, or providing tape translation, disclosure bypasses and other file management services, the professional staff and program facilities including computer facilities could perform many joint functions. The increment costs of providing a reference capability as a part of a more extensive user service capability could possibly be a great deal smaller than would be required for a separate and independent function.

Despite these obstacles to cost estimation, the committee noted that the operating budgets for two agencies performing related (though in many ways basically different) functions ranged from $400,000 a year for the National Science Referral Center in the Library of Congress to $2 million a year currently for the Science Information Exchange of the Smithsonian Institute. The committee speculated that it probably would take as much as $2 million a year on the average over a 5-year period to develop a meaningful reference service capability including enough resources to involve agency programs in the way necessary.

The committee also discussed the question of the demand for reference services and whether anticipated needs were appropriate to justify such levels of expenditure. The need for the service was judged to be sufficiently great to justify a serious effort.

There is no way, at this preliminary stage, that anything more than a judgment can be offered. In the first place, nothing in present agency experience can serve as a guide to demand levels for a service that has never existed in anything like the form indicated in this prospectus. The judgment rested on several considerations. First, there is considerable evidence of frustration and inefficiency because of the absence of such a service. Second, there already exist other programs, such as the ones referred to, that have been judged essential and for which expenditures substantially exceeding the sums of money mentioned here have been undertaken. Most of these have been undertaken in the interest of improving the efficiency of documentation in the physical sciences. In the circles where public and business policy are made, and social science, management and marketing research undertaken, the most compelling needs is not for a way to handle better the traditional documentary materials generated, but for a way to acquire efficient access to relevant numerical files that constitute the main bodies of evidence and of research inputs. The need for expanding the services in this area seems equally as compelling as those in the field of physical science documentation that are already receiving extensive attention.

APPENDIX B

OCTOBER 1, 1965.

From: Rudolph C. Mendelssohn.
To: Mr. Edgar Dunn.
Subject: Report on data inventory.

DATA BANK REQUIREMENTS

An estimated $3 to $3.5 million and 3 to 5 years are needed to stock the proposed data center with Federal statistics now in existence. These data would probably comprise about 20,000 reels of magnetic tape. However, a Federal center could be stocked with a respectable volume and variety of data relatively quickly for about $260,000. At the rate of about $27 per reel, a bank of 9,000 tape files could be established in about a year. Such a course would provide a fairly representative selection of significant data including, for example, 750 reels of the census housing data; census current population data on 375 reels; the BLS Consumer Expenditure Survey on 43 reels; the BLS industry hours, earnings, and labor turnover data on 36 reels; the OBE national income accounts on 2 reels; as well as IRS tax data on 5,300 reels and BOASI social security data on 1,900 reels. An insight into the volume of data readily available is gained by the rough calculation that the equivalent of nearly 1 billion punched cards would be included (table 1).