

Mr. RUGGLES. I would agree completely.

Mr. GALLAGHER. Has anybody given this any consideration, Mr. Dunn?

Mr. DUNN. Let me add a few comments on this matter of the uses. First, of all, I don't think there is really any way, as a matter of practical possibilities, to determine something that we might call the request incidence or frequency of requests upon statistical files in the Federal Government, and certainly no way of identifying specifically what they would be upon an improved service and capability of this kind.

Now, one of the principal reasons why this is so is because most of the distribution of the information which is generated by this process is undertaken through published documents like the Census Monograph, and so forth. These documents go out to all sorts of places, including standard library, reference, and referral services of various kinds. People come to these records and documents for all sorts of reasons without any way of ever tracing who tried to get into the record for what, you see, for what purpose. Here is the important thing.

Well, let me make another comment on the same point and that is this: To try to anticipate what the frequency of requests would be upon a system which would improve the service and capabilities of matching records in various ways and so forth is likewise impossible at the present time because it would be roughly identical with the kind of problems DuPont would face when trying to decide whether to produce nylon or not.

Mr. GALLAGHER. Not exactly. We already have data centers set up. I think you should determine what use is made of this information at this point.

Mr. DUNN. No, there is a distinction here and it is a very difficult distinction to convey and I am not sure I can be successful in conveying this point. May I take a few minutes to try to convey this and do it in terms of a simple kind of example?

Mr. GALLAGHER. Please do.

Mr. DUNN. I would like to start, first of all, from the point of view that is very fundamental here in the whole operation of using numbers for analysis in statistics, and so forth, and that is that no number will convey any information to anyone, without being associated with some other number, except as it can be related to other numbers and except as it can be identified with an attribute or characteristic of that number.

Let me illustrate: Suppose I were to write on the wall over here the symbol "2" and ask you what it means to you. You would probably say nothing. You might try to give it significance by saying that it is a number in an array of numbers that falls between 1 and 3, but if you would say that, you are already trying to associate it with some other number, you see. However, if you are really quick, you will immediately say "two what?"

Mr. HORTON. \$2 million for this system. That is what I thought the minute you said it.

Mr. DUNN. It might be that, but if I said "two apples," the information content jumps immediately. If I go on and associate the apples with the behavior of people and say two apples are eaten by the average person each week, the information takes another jump. If I