

First, consider the technical form of the record. You see in my hand a piece of computer tape. This is the form of the record. You and I can't read it. Although there may be several advanced computer technicians in this room, they can't read it. To get any information out of this tape requires (1) a machine, (2) a codebook, (3) an appropriate set of instructions, (4) a technician. You see, when we put information into such a system we start with an inherent technical advantage over traditional records when it comes to protecting personal privacy. Information on paper in agency files can be very easily prostituted by any human being who can read and write and can be bought. The number of people who could get a specific piece of information out of a file of this type are very limited and very easy to identify. This facilitates the establishment of security procedures.

Second, I just said an appropriate set of instructions was necessary to get any information off the tape. This prompts me to make two points.

A statistical system will have the data formatted and stored and will have developed a set of instructions—known in the lingo nowadays as computer software—that facilitate the generation of statistical services. It may, in consequence, not be a very efficient system for intelligence purposes because neither the format of the record nor the standard instructions are developed for intelligence use.

However, there is a more important point. In addition to designing instructions that tell the machine what to do, you can give it instructions that can tell it what not to do. You can teach the machine to distinguish appropriate inquiries—statistical questions—from inappropriate inquiries—intelligence questions or individual data. Furthermore, you can go further than that. You can teach the machine to identify “trick” inquiries—either accidental or purposeful. That is, you can teach the machine to say, “This is a statistical inquiry but it is framed in such a way that the population or group you have defined contains only one individual or less than some specified number of individuals.”

Thus you can design a system in ways that prohibit any output other than a legitimate statistical analysis. The machine itself and the design system that operates it can be used to monitor the use procedures in ways that greatly increase the efficiency and assurance of security.

In addition, through system design and the kinds of instructions you give the machine you can teach it ways to disguise records. For example, you can replace all individual identification codes with a special security code that only the machine can use and which it uses only for the purpose of associating records. Under such circumstances, no one engaged in any part of the productive operation of the system would be able to identify any individual record without access to the set of translation codes that could be protected by special provisions of law and intensive security devices. The machine can also generate data useful for statistical purposes that are randomized or modified into “prototype” records that retain certain useful statistical properties while losing all identity as individuals.

We have had experience with all of these devices, legal and technical; they can be further perfected and extended as we attempt to improve the efficiency of the system for legitimate statistical use.