28 April 1978, Volume 200, Number 4340 **SCIENCE**

Needed: Better Data About Academic Science

The information system which provides data for policy decisions about U.S. academic science has fundamental flaws. Yet the quality of such decisions rest, to some degree, on the quality of these data.

Paradoxically, fewer information requests could be made by federal agencies (and related national organizations) but the data would be more useful if presently diverse retrieval activities were better coordinated. Spriestersbach and Farrell* recently documented the burden of federal demands for information on one university. Redundant requests for data about graduate science—enrollment, Ph.D.'s produced, expenditures, and so on—constitute a modest portion of this burden.

More to the point, it is difficult, if not impossible, to relate data from different agencies. In two recent studies—an evaluation of the National Science Foundation's Science Development Program and research commissioned by the President's Biomedical Research Panel—my colleagues and I have attempted to create such a merged data file. This was necessary because different agencies or groups retrieve the best (or the only) data about different characteristics of science departments. For example, NSF has the best data on federal expenditures by discipline, while the National Research Council has virtually complete information about Ph.D.'s and their job placement. The following are some of the problems we encountered.

- Academic fields are defined and classified differently. For example, in the NRC Doctorate Record File Ph.D.'s indicate their own fields (which may not match their departments), while the key file about National Institutes of Health funding references actual departments, and NSF's funding file aggregates departments into "disciplines."
- During the past decade many state universities developed strong branch campuses with graduate programs. Agencies vary as to whether they report activities at the main campus only or at the main campus and branch campuses identified separately, or lumped together. As a further complexity, the point at which a branch campus is considered sufficiently active to be recognized by an agency (or, in fact, by the university itself) varies.
- Some files fail to separate data about the medical or agriculture schools from data about the main unit. Thus, all federal expenditures for biochemistry are reported as one datum for a particular university.

The astute reader will realize that successive sections of *Science Indicators 1976* had to be based on data derived from different definitions.

Policy analysts frequently discover that indicators of the same phenomenon provided by two agencies do not agree. One reason is that different organizations use different sources. For example, the NRC polls doctorate recipients. The NSF gets enrollments from department chairmen. The National Center for Educational Statistics retrieves both enrollment and Ph.D. statistics from university-level administrators such as registrars.

Each agency has developed its own definitions and retrieval techniques to be consistent with its organizational objectives. And agency officials may fear that the loss of control over these activities that standardization might require will undermine the realization of those objectives. Although it is problematic, this need not present an insurmountable obstacle.

Thoughtful policy-makers and analysts have reviewed these issues and suggested a variety of solutions. For example, last year both the Paperwork Commission and a federal interagency committee chaired by NSF's Robert Wright issued detailed recommendations for agency cooperation about this problem. But the implementation of these recommendations has been glacial at best.

In my opinion the university science community, which has much to gain from more consistent data and less burdensome requirements, should press for more government action on this problem.—DAVID E. DREW, *Rand Corporation, Santa Monica, California 90406*

*D. C. Spriestersbach and W. J. Farrell, *Science* **198**, 27 (1977).