Appendix A
The RaDiUS Database

RaDiUS, the database of Research and Development in the United States federal government, was developed to support the Science and Technology Policy Institute, a federally funded research and development center (FFRDC) operated by RAND. The institute (formerly known as the Critical Technologies Institute) was established by Congress in 1991 to improve public science and technology (S&T) policy by providing research and analysis to the White House Office of Science and Technology Policy (OSTP) and other government agencies (see 42 USC 6686). OSTP is the organizational and administrative force behind the National Science and Technology Council (NSTC).

Prior to the development of RaDiUS, the only information available on the contents of the federal R&D portfolio was either highly aggregated or very disaggregated and incomplete. For example, at the highly aggregated end of the spectrum is the report by the Office of Management and Budget (OMB) included in the annual Budget of the United States Government. This report comprises a chapter in the “Analytical Perspectives” volume of the federal budget and has been prepared annually beginning with the FY 1996 budget. The dollar amounts it contains are Outlays that do not match the official federal R&D baseline numbers, which are in Budget Authority. These numbers are complemented by a chapter in the main portion of each year’s budget, which are in Budget Authority, and which constitute the official federal R&D baseline. Also of a highly aggregated nature is the Federal Funds for Research and Development report series compiled and published annually since the 1950s by the National Science Foundation (NSF). The dollar amounts this report presents are Obligations, rather than Budget Authority or Outlays, so they do not match to anything presented in the federal budget.
In contrast, at the very disaggregated end of the spectrum is the Federal Research in Progress (FEDRIP) database that contains detailed information on a small portion of federal R&D activities. This database began in 1946 as the Smithsonian Science Information Exchange and has been maintained by the National Technical Information Service (NTIS) at the Department of Commerce since 1981. FEDRIP contains no information on defense R&D and covers only a portion of civilian R&D, all of which is of uneven currency. Because FEDRIP is a data warehouse, the contents of which are not systematically matched to the baseline of the federal budget, it is not easy to determine precisely how much of the federal R&D portfolio FEDRIP actually tracks. However, a careful comparison of the individual records in FEDRIP and the contents of the federal R&D portfolio indicates that, on an annual basis, FEDRIP contains data on only about one-fourth of all federal R&D activities.

RaDiUS is the first information system that connects the highly aggregated data on federal R&D with the very disaggregated data to provide a complete picture of all federal activities involving the conduct of R&D. Specifically, RaDiUS consists of five interconnected levels of increasingly detailed data on federal R&D. The least-detailed level of RaDiUS contains information on the 24 agencies that control and disseminate all R&D dollars spent by the federal government. The amounts carried in this level of RaDiUS are identical to those presented in the Federal Budget for Basic Research, Applied Research, and Development. These three “stages” of R&D collectively constitute the “Conduct of R&D” and constitute the major portion of the official “baseline” of the federal R&D portfolio. RaDiUS does not track funds spent on the construction and rehabilitation of federal R&D facilities or the purchase of major R&D equipment. When funds for these two activities are combined with those for the conduct of R&D, the result is the total budget for the federal R&D portfolio.

From their arrival at the 24 agencies, the dollars carried in the federal R&D budget that are spent on the “conduct of R&D” are tracked by RaDiUS as they are sequentially dispersed to three levels of successively smaller organizational units within the agencies (Levels 2, 3, and 4 in RaDiUS). The information carried in each of the top four levels of
RaDiUS tracks 100 percent of all federal dollars spent annually on the conduct of R&D. The fifth and most-detailed level of RaDiUS tracks these dollars to their final destination at the universities, laboratories, and centers located throughout the world, both inside and outside the federal government. These are the places where the hundreds of thousands of experiments and analyses actually “purchased” with federal R&D dollars take place. The records of these activities are by far the most difficult to obtain, for they are scattered throughout the federal government in a wide variety of forms and formats. In addition, these records represent the core of agencies’ missions, so some are reluctant to share such detailed information with other agencies for fear that they might use it to strategic advantage. RAND's neutrality has clearly been an advantage to getting agencies to make their detailed R&D information available for inclusion in RaDiUS. RAND's in-house technical capabilities have also enabled it to make sense of even the most arcane data records.

To date, over 400,000 such records have been assembled in RaDiUS covering all fiscal years since 1993 and providing vital details on close to 80 percent of the activities in the federal R&D portfolio. More such records are being added to RaDiUS on a regular basis, with the objective of bringing the coverage of the fifth level of RaDiUS as close to 100 percent as soon as possible. Some records in the level are restricted and, while not classified, cannot be accessed freely by everyone. Others are classified and, therefore, cannot be included in the current version of RaDiUS. Still others involve salaries of individual researchers and research managers, which to date have not been a top priority to obtain, because they do not describe specific R&D activities.

The organizing premise of RaDiUS has been to harvest data on federal R&D from information already gathered by the federal government, even if it has not traditionally been viewed as relevant to the tracking of federal R&D activities. For example, RaDiUS includes information from the Federal Assistance Awards Data System (FAADS), which has never before been used to track R&D. The harvested data have then been woven together using common data fields and codes to form a comprehensive picture of federal R&D. Taking this approach to building RaDiUS was essential to ensuring that the Science and Tech-
nology Policy Institute could quickly acquire the broad and flexible data capabilities needed to support OSTP and NSTC. Making RaDiUS easily available to a wide range of users was also a goal. To facilitate this and simultaneously permit valuable but restricted information to be included in the database, RaDiUS was specifically designed to be password-accessible from any personal computer through an encrypted link on the web (http://www.rand.org/radius). Each password enables a user to access only those records in RaDiUS that he or she is permitted to view. As a result, RaDiUS makes as much information as possible available to as many people as possible using a single system. To date, more than 30 reports prepared by the institute for OSTP and NSTC have been based in whole or in part on information obtained from RaDiUS. In addition, RaDiUS has been used by federal agencies and contractors to support R&D planning efforts, leverage R&D investments, and transfer technology from discovery sites to places of critical need.