Guide for Public Use Data Files

National Science Foundation's Federal Facilities Research and Development Survey:

Fiscal Year 2024

August 2026

National Science Foundation
National Center for Science and Engineering Statistics

Table of Contents

This data users' guide provides a detailed description of the public use data files for the FY 2024 Federal Facilities Research and Development (FFRD) survey. The FFRD survey is sponsored by the National Center for Science and Engineering Statistics (NCSES) within the U.S. National Science Foundation. For additional information about the survey population and procedures, refer to the FY 2024 Survey Methodology Report (available upon request from NCSES). NCSES has reviewed this product for unauthorized disclosure of confidential information and approved its release (NCSES-DRN25-004).

1. ABOUT THE FFRD SURVEY

The FFRD Survey collects information on research and development (R&D) expenditures, funding, and personnel for all federally owned and operated facilities in the United States. This survey is the unique source of federal performer-reported data on R&D expenditures and R&D personnel. In conjunction with performer-reported data for the remaining sectors of the economy, FFRD survey data can be used to estimate total national R&D performance and other critical statistics that are increasingly important in the analysis of global R&D competitiveness.

The survey was piloted for FY 2021, with full implementation beginning for the FY 2022 cycle. The FY 2023 cycle was skipped to improve the timing of data collection. The FFRD survey is an annual survey; therefore, the next survey cycle will be FY 2025. Microdata are only available for public use beginning with the FY 2024 data collection cycle.

1.1. SURVEY POPULATION, INSTRUMENT, AND DATA AVAILABILITY

1.1.1. Target Population and Frame

The population for the FY 2024 FFRD survey consisted of federally owned and operated facilities in the United States that performed R&D in FY 2024, excluding facilities of the Central Intelligence Agency (due to the classified nature of their work), federally funded research and development centers (surveyed separately in the Federally Funded Research and Development Centers R&D Survey), and University Affiliated Research Centers or facilities associated with universities, both of which may be surveyed as part of the Higher Education Research and Development Survey.

The facilities for the FY 2024 FFRD survey were identified from the list of federal agencies with R&D obligations based on the Survey of Federal Funds for R&D and the Federal Lab Consortium list of laboratories. A facility is defined as a unit within the agency that is responsible for performing R&D, generally with its own distinct budget and leadership.

1.1.2. Key Variables

There are two versions of the survey: one for Department of Defense (DOD) facilities and one for non-DOD facilities (Standard). The FY 2024 FFRD survey consisted of the following:

- An **initial section** requested that facilities provide contact information for the individual responsible for the survey, an alternate contact, and up to three additional contacts. This section

also requested facility information and acknowledgement of the expenditure reporting requirement for this survey.

- **Question 1** on the Standard version of the survey requested the facility's total R&D expenditures (also called 'outlays'). **Question 1** on the DOD version of the survey requested the facility's total R&D expenditures (also called 'outlays'; budget activities 1-6), operational system development and software and digital technology pilot programs expenditures (budget activities 7-8), and total research, development, testing, and evaluation (RDT&E) expenditures.

- **Question 2** requested the facility's total R&D expenditures by type of cost.

- **Question 3** asked what types of expenditures were included in their reporting of internal R&D expenditures in Question 2c.

- **Question 4** asked facilities to report the amount of their federal and nonfederal R&D expenditures for basic research, applied research, and experimental development.

- **Question 5** requested that facilities report their total expenditures by source of funds (e.g., federal, business, nonprofit organizations).

- **Question 6** requested that facilities report the amount of their R&D expenditures by federal agency sources (i.e., the U.S. Departments of Agriculture, Commerce, Defense, Energy, Health and Human Services, Homeland Security, Interior, Transportation, and Veterans Affairs; the Environmental Protection Agency; the National Aeronautics and Space Administration [NASA]; and all other federal agencies).

- **Question 7** asked facilities to report whether any of their R&D projects were funded through public-private partnerships.

- **Question 8** requested that facilities report the amount of their R&D expenditures by field (i.e., agricultural sciences and natural resources and conservation; biological, biomedical, and health sciences; computer and information sciences; geosciences, atmospheric sciences, and ocean sciences; mathematics and statistics; physical sciences; psychology; social sciences; engineering; and other fields).

- **Question 9** requested a headcount of research personnel by three R&D job categories.

- **Question 10** requested the full-time equivalents (FTEs) by three R&D functions.

1.1.3. Changes to the Survey for FY 2024

Changes to the FFRD survey questionnaire included the following:

- The questionnaire was reorganized to ask for total expenditures for R&D activities both within a facility by federal personnel or contractors and funding provided to others to conduct R&D outside of a facility at the beginning of the questionnaire versus separating

it into two sections beginning with expenditures within the facility only. For DOD facilities, a breakout of RDT&E expenditures by budget activities 1-6 and budget activities 7-8 is requested.

- R&D funding to other organizations (FY 2022 questions 6-8) was removed from the survey. Total funding provided to other organizations is now requested as part of the new question.

- A question was added to request information on what types of expenditures are included in survey reporting: salaries and fringe for federal personnel, onsite contractors, and military personnel; indirect costs; equipment, materials, and supplies; other.

- Additional funding agencies were added to federal source of funding, reducing the number of write-in agencies.

- When reporting personnel headcounts by job category, facilities have the option to provide only total counts if a breakout by job category and/or function is not possible. Similarly, facilities can provide only total federal personnel FTE if a breakout by R&D function is not possible.

1.1.4. Excluded Data

Some data collected as part of the FFRD survey are not included in the public use data files. Information categorized as part of the initial section (e.g., respondent contact information, facility information, expenditure acknowledgement). The exception is the facility state information which was recoded into the facility state variable.

Responses that are categorized as information helpful for data provided (i.e., open-ended comments, unable to separate checkboxes on questions 9 and 10) are excluded from publicly available data. Information categorized as which expenses are included in the total R&D expenditures (from Question 3) are not published or released.

Responses for Question 6 as part of additional or other agencies not specifically listed are excluded from public use data files. A combined other variable (fed_all_oth) contains the sum of responses provided on these rows.

1.2. DATA PROCESSING AND QUALITY MEASURES

Completed questionnaires were carefully examined by survey staff upon receipt. Reviews focused on unexplained missing data, expenditures that significantly differed from the FY 2022 FFRD Survey, expenditures that significantly differed from the Survey of Federal Funds for R&D intramural obligations, and other data anomalies. If additional explanations or data revisions were needed, respondents were sent personalized e-mail messages asking them to provide any necessary revisions before the final processing and tabulation of data.

1.2.2. Estimation Techniques

The FY 2024 FFRD Survey imputation process used a combination of multiple imputation models and deterministic proportional allocation for item totals and item details, respectively. Auxiliary information from The FY 2022 FFRD Survey was used in modeling and calculation of proportions where appropriate.

1.2.3. Survey Quality Measures

Sampling error. Because the FY 2024 survey was distributed to all eligible agencies performing R&D, there was no sampling error.

Coverage error. Under the total survey error framework, coverage error describes the difference between reporting units and units in the target population that the frame was developed to reach. The subset of agencies is based on those reporting non-zero intramural R&D obligations on the Survey of Federal Funds for R&D. Risk of coverage bias is small if an agency had federal R&D expenditures but did not have reported obligations.

Nonresponse error. Of the 300 eligible reporting units, 287 responded for a response rate of 95.78%. For unit nonresponse, multiple follow-ups were conducted with nonresponding facilities, and multiple contact and data collection modes were used (i.e., phone and e-mail) to mitigate nonresponse error. The imputations include predictions of unit nonresponse to reduce the risk of nonresponse bias in the final estimates.

Measurement error. The largest risk of measurement error is likely respondents' interpretation of the definition of R&D activities and variations in record-keeping procedures used by respondents to answer the survey questions. In order to reduce measurement error, the FFRD Survey contained various ways for respondents to explain their survey responses.

1.3. DATA IMPUTATION

In order to provide totals of all R&D expenditures, it was necessary to develop estimates for the facilities that did not respond to the survey as well as for items not answered by survey respondents. Data imputation is an automated procedure used to estimate missing data for both nonresponding and partially responding facilities.

The imputation process for the FY 2024 FFRD Survey used model-based multiple imputation for item totals and deterministic proportional allocation for item details. Auxiliary information from the FY 2022 FFRD Survey was used in modeling and calculation of proportions where appropriate. The imputation includes five main steps:

First, the total intramural R&D expenditures were estimated using a linear regression model with the expenditures from the FY 2022 FFRD Survey. The model included separate intercepts for each agency and log-transformed R&D totals. Multiple imputation was used to create multiple sets (m = 5) of plausible values for the missing values. The average of the multiple imputed values are used for reporting and provided in the public use data files. After imputing total onsite R&D expenditures, the imputed value was assigned to other fields representing the same value

including total type of R&D-type expenditures (Question 4), total field of R&D expenditures (Question 5), and within facility costs (subset of Question 2).

Second, the intramural total R&D expenditures were then allocated to the subcategories of type of R&D, source of funding and federal agency source of funding, and field of R&D. Type of R&D and field of R&D were imputed for nonresponding facilities based on the percent distributions using the following information in order:

1. FY 2022 FFRD Survey observed distribution for the facility,

2. responding FY 2024 facilities within the same agency aggregated to the agency level, and

3. responding FY 2024 facilities within the same agency aggregated to the department level.

All imputed R&D amounts were assigned as a federal source of R&D. This follows the same procedure in FY 2022 based on the majority of R&D in FY 2024 coming from federal sources for each type (98% basic research, 97% applied research, and 99% experimental development.)

Third, total R&D expenditures (Question 1) and personnel item totals (Questions 9 and 10) were imputed together using Multivariate Imputation by Chained Equations (MICE). For every imputation for the first group of variables (e.g., R&D expenditures), four sets of imputations are created for the second group (e.g., personnel), holding the first group imputed values fixed. This two-stage nested multiple imputation produced a total M = *5 x 4* imputed data sets (M = 20).

Fourth, after the second imputation model, the total R&D expenditures value was assigned to the cost type total (Question 2), since these two variables represent the same amount. R&D expenditures performed outside facilities (subset of Question 2) were calculated as the difference between total R&D expenditures and total R&D within facilities.

For DOD facilities, when some facilities within the agency responded (partial agency response), total research, development, testing, and evaluation (RDT&E) (Question 1) for nonresponding DOD facilities were imputed based on a ratio adjustment using the ratio of total expenditures at the agency level and applying it to the imputed R&D total for the facility. When no facilities responded within the agency surveyed by the FY 2024 FFRD Survey, the ratio adjustment was based on the department level from the FY 2024 FFRD Survey. Finally, operational system development expenditures were calculated as the difference between total RDT&E and total R&D**.**

Lastly, personnel (Questions 9 and 10) totals were allocated to function (i.e., researcher, technician, support) based on the percentage distribution using the following information in order:

1. FY 2022 FFRD Survey observed distribution for the facility,

2. Responding facilities within the same agency aggregated to the agency level,

3. Responding facilities within the same department aggregated to the department level, and

4. Responding facilities within the same DOD or non-DOD category aggregated to the DOD or non-DOD level.

The imputed job counts and FTE for federal employees, contractors, and other employment status for each role were summed to calculate the imputed total job count and FTE.

1.4. PUBLIC USE DATA FILES

Survey data is collected from facilities and checked and cleaned as necessary. Missing values are imputed as described above. After completion of the annual data processing cycle, the collected data are stored in files for public use and archival purposes. These files include any corrections to prior-year data that survey respondents have submitted. For each survey year, the data file contains two types of information: facility information and questionnaire response values. The inst_id is a variable unique to the FFRD survey that can be used to link a facility's records across years. Next are the variables in the order they appear on the questionnaire. Questionnaire response values contain the numerical responses to the survey questionnaire items. All expenditure data items are expressed in dollars. Variables for personnel and full-time equivalents are counts. The imputation flags are next in the same order with variable names corresponding to the imputed variable with an "_s" suffix. The level variables show the hierarchy of the agency-department-subagency, and state as reported in the tables. For each data line on the questionnaire to which a nonzero response has been received, a data record is present. For total rows in each survey question, a data record is present regardless of whether a zero or nonzero response has been received. A status code associated with each data column indicates the source of and/or special treatment for the data.

Variables that are skipped due to survey logic are coded with a value of -1. Imputed flags have a value of "i" if the value of the corresponding variable was imputed, and they are left blank when values not imputed. The public private partner flag (public_priv_partner_flag) has values of 1=Selected, 0=Not selected, or Missing=Not answered.

National Institute of Standards and Technology and National Aeronautics and Space Administration facilities were unable to separate costs for onsite contractors. These costs are included with all other costs performed by the facility (cost_type_other_within). The variable for onsite contractors (cost_type_onsite_contractors) for these facilities has been coded as -999999999.

1.4.1. Survey Format

Table 1 lists the variables included in the FY 2024 FFRD survey public use data files, with a description of each variable, question number, and question text for key variables.

Table 1. Variables descriptions for FY 2024 FFRD survey public use data files

Variable name	Description	Question	Question text
year	Fiscal year	-	
inst_id	Institution id, used for respondent ID	-	
facilityname	Facility name	-	
toi_code	Survey version (1 Standard, 3 DOD)	-	
total_expend_rd	R&D expenditures	Q1.a	What were your facility's total FY 2024 expenditures for R&D? (Standard) / What were your facility's total FY 2024 expenditures for RDT&E, by budget activity? (DOD)
total_expend_op_sys_dev	Operational system development expenditures	Q1.b (DOD)	
total_expend_tot	RDT&E expenditures	Q1.c (DOD)	
cost_type_onsite_contractors	Onsite contractors	Q2.a	What were your facility's total FY 2024 expenditures for R&D by the following types of costs?
cost_type_other_within	All other costs within facility	Q2.b	
cost_type_within_tot	Total within facility	Q2.c	
cost_type_outside	Expenditures for work outside facility	Q2.d	
cost_type_tot	Total	Q2.e	
rd_type_basic_fed	Basic research - Federal	Q4.a.1	What were your facility's total FY 2024 expenditures for R&D (question 2, row c) by type of work?
rd_type_basic_nonfed	Basic research - Nonfederal	Q4.a.2	
rd_type_basic_tot	Basic research - Total	Q4.a.3	
rd_type_applied_fed	Applied research - Federal	Q4.b.1	
rd_type_applied_nonfed	Applied research - Nonfederal	Q4.b.2	
rd_type_applied_tot	Applied research - Total	Q4.b.3	
rd_type_dev_fed	Experimental development - Federal	Q4.c.1	
rd_type_dev_nonfed	Experimental development - Nonfederal	Q4.c.2	
rd_type_dev_tot	Experimental development - Total	Q4.c.3	
rd_type_all_fed	Total - Federal	Q4.d.1	
rd_type_all_nonfed	Total - Nonfederal	Q4.d.2	
rd_type_all_tot	Total - Total	Q4.d.3	
source_fed	U.S. federal government	Q5.a	Of the FY 2024 R&D expenditures you reported in Question 2, row c, how much came from the following sources?
source_state	State and local government	Q5.b	
source_bus	Businesses	Q5.c	
source_np	Nonprofit organizations	Q5.d	

Variable name	Description	Question	Question text
source_oth	All other organizations	Q5.e	
source_tot	Total	Q5.f	
fed_usda	Department of Agriculture	Q6.a	Of the federally funded FY 2024 R&D expenditures you reported in Question 5, which agencies funded this R&D and how much of the reported amount was from each agency?
fed_commerce	Department of Commerce	Q6.b	
fed_dod	Department of Defense	Q6.c	
fed_doe	Department of Energy	Q6.d	
fed_hhs	Department of Health and Human Services	Q6.e	
fed_dhs	Department of Homeland Security	Q6.f	
fed_doi	Department of the Interior	Q6.g	
fed_dot	Department of Transportaton	Q6.h	
fed_va	Department of Veterans Affairs	Q6.i	
fed_epa	Environmental Protection Agency	Q6.j	
fed_nasa	National Aeronautics and Space Administration	Q6.k	
fed_all_oth	All Other	Q6.l - Q6.o	
fed_tot	Total	Q6.p	
public_priv_partner_flag	Public private partnership flag	Q7	Were any of your facility's FY 2024 R&D projects funded through public-private partnerships?
rde_fields_agr	Agricultural sciences and natural resources and conservation	Q8.a	Of the total FY 2024 R&D expenditures within your facility you reported in Question 2, row c, what were your expenditures in each field below?
rde_fields_bio	Biological, biomedical, and health sciences	Q8.b	
rde_fields_comp	Computer and information sciences	Q8.c	
rde_fields_geo	Geosciences, atmospheric sciences, and ocean sciences	Q8.d	
rde_fields_math	Mathematics and statistics	Q8.e	
rde_fields_phys	Physical sciences	Q8.f	
rde_fields_psych	Psychology	Q8.g	
rde_fields_soc	Social sciences	Q8.h	
rde_fields_eng	Engineering	Q8.i	
rde_fields_oth_field	Other fields	Q8.j	
rde_fields_all_tot	Total	Q8.k	
job_count_fed_rsr	Federal employees and military personnel - Researchers	Q9.a.1	

Variable name	Description	Question	Question text
job_count_fed_tech	Federal employees and military personnel - R&D technicians	Q9.a.2	How many of your facility's total R&D personnel (headcount) worked in the job categories listed below in FY 2024?
job_count_fed_suppt	Federal employees and military personnel - R&D support staff	Q9.a.3	
job_count_fed_tot	Federal employees and military personnel - Total	Q9.a.4	
job_count_cont_rsr	Contractors - Researchers	Q9.b.1	
job_count_cont_tech	Contractors - R&D technicians	Q9.b.2	
job_count_cont_suppt	Contractors - R&D support staff	Q9.b.3	
job_count_cont_tot	Contractors - Total	Q9.b.4	
job_count_other_rsr	All other R&D personnel - Researchers	Q9.c.1	
job_count_other_tech	All other R&D personnel - R&D technicians	Q9.c.2	
job_count_other_suppt	All other R&D personnel - R&D support staff	Q9.c.3	
job_count_other_tot	All other R&D personnel - Total	Q9.c.4	
job_count_all_rsr	Total - Researchers	Q9.d.1	
job_count_all_tech	Total - R&D technicians	Q9.d.2	
job_count_all_suppt	Total - R&D support staff	Q9.d.3	
job_count_all_tot	Total - Total	Q9.d.4	
fte_rsr	Researchers	Q10.a	How many federal full-time equivalents (FTEs) worked in the functions listed below in FY 2024?
fte_tech	R&D technicians	Q10.b	
fte_suppt	R&D support staff	Q10.c	
fte_tot	Total	Q10.d	
total_expend_rd_s	total_expend_rd imputation flag		
total_expend_op_sys_dev_s	total_expend_op_sys_dev imputation flag		
total_expend_tot_s	total_expend_tot imputation flag		
cost_type_onsite_contractors_s	cost_type_onsite_contractors imputation flag		
cost_type_other_within_s	cost_type_other_within imputation flag		
cost_type_within_tot_s	cost_type_within_tot imputation flag		
cost_type_outside_s	cost_type_outside imputation flag		
cost_type_tot_s	cost_type_tot imputation flag		
rd_type_basic_fed_s	rd_type_basic_fed imputation flag		
rd_type_basic_nonfed_s	rd_type_basic_nonfed imputation flag		

Variable name	Description	Question	Question text
rd_type_basic_tot_s	rd_type_basic_tot imputation flag		
rd_type_applied_fed_s	rd_type_applied_fed imputation flag		
rd_type_applied_nonfed_s	rd_type_applied_nonfed imputation flag		
rd_type_applied_tot_s	rd_type_applied_tot imputation flag		
rd_type_dev_fed_s	rd_type_dev_fed imputation flag		
rd_type_dev_nonfed_s	rd_type_dev_nonfed imputation flag		
rd_type_dev_tot_s	rd_type_dev_tot imputation flag		
rd_type_all_fed_s	rd_type_all_fed imputation flag		
rd_type_all_nonfed_s	rd_type_all_nonfed imputation flag		
rd_type_all_tot_s	rd_type_all_tot imputation flag		
source_fed_s	source_fed imputation flag		
source_state_s	source_state imputation flag		
source_bus_s	source_bus imputation flag		
source_np_s	source_np imputation flag		
source_oth_s	source_oth imputation flag		
source_tot_s	source_tot imputation flag		
fed_usda_s	fed_usda imputation flag		
fed_commerce_s	fed_commerce imputation flag		
fed_dod_s	fed_dod imputation flag		
fed_doe_s	fed_doe imputation flag		
fed_hhs_s	fed_hhs imputation flag		
fed_dhs_s	fed_dhs imputation flag		
fed_doi_s	fed_doi imputation flag		
fed_dot_s	fed_dot imputation flag		
fed_va_s	fed_va imputation flag		
fed_epa_s	fed_epa imputation flag		
fed_nasa_s	fed_nasa imputation flag		
fed_all_oth_s	fed_all_oth imputation flag		
fed_tot_s	fed_tot imputation flag		
rde_fields_agr_s	rde_fields_agr imputation flag		

Variable name	Description	Question	Question text
rde_fields_bio_s	rde_fields_bio imputation flag		
rde_fields_comp_s	rde_fields_comp imputation flag		
rde_fields_geo_s	rde_fields_geo imputation flag		
rde_fields_math_s	rde_fields_math imputation flag		
rde_fields_phys_s	rde_fields_phys imputation flag		
rde_fields_psych_s	rde_fields_psych imputation flag		
rde_fields_soc_s	rde_fields_soc imputation flag		
rde_fields_eng_s	rde_fields_eng imputation flag		
rde_fields_oth_field_s	rde_fields_oth_field imputation flag		
rde_fields_all_tot_s	rde_fields_all_tot imputation flag		
job_count_fed_rsr_s	job_count_fed_rsr imputation flag		
job_count_fed_tech_s	job_count_fed_tech imputation flag		
job_count_fed_suppt_s	job_count_fed_suppt imputation flag		
job_count_fed_tot_s	job_count_fed_tot imputation flag		
job_count_cont_rsr_s	job_count_cont_rsr imputation flag		
job_count_cont_tech_s	job_count_cont_tech imputation flag		
job_count_cont_suppt_s	job_count_cont_suppt imputation flag		
job_count_cont_tot_s	job_count_cont_tot imputation flag		
job_count_other_rsr_s	job_count_other_rsr imputation flag		
job_count_other_tech_s	job_count_other_tech imputation flag		
job_count_other_suppt_s	job_count_other_suppt imputation flag		
job_count_other_tot_s	job_count_other_tot imputation flag		
job_count_all_rsr_s	job_count_all_rsr imputation flag		
job_count_all_tech_s	job_count_all_tech imputation flag		
job_count_all_suppt_s	job_count_all_suppt imputation flag		
job_count_all_tot_s	job_count_all_tot imputation flag		
fte_rsr_s	fte_rsr imputation flag		
fte_tech_s	fte_tech imputation flag		
fte_suppt_s	fte_suppt imputation flag		
fte_tot_s	fte_tot imputation flag		

Variable name	Description	Question	Question text
level1	Reporting level 1 (Agency)		
level2	Reporting level 2 (Department)		
level3	Reporting level 3		
level4	Reporting level 4		
level5	Reporting level 5		
level6	Reporting level 6		
facility_state	Facility state		

2. ADDITIONAL RESOURCES

2.1. Survey Questionnaires

Copies of survey questionnaires are available on the NSF website at https://ncses.nsf.gov/surveys/federal-facilities-research-development/, within the Methodology section.

2.2. Publications

Data from the FFRD survey is published in analytic reports and data tables available by fiscal year on the NCSES website: https://ncses.nsf.gov/surveys/federal-facilities-research-development/. Information from this survey will also be included in future versions of a congressionally mandated report produced by NCSES: Science and Engineering Indicators.[1]

2.3. Electronic Access

All data from this survey are available on the NCSES website. Selected aggregate data are provided in public use data files upon request.

Public use files beginning with the FY 2024 survey are available at https://ncses.nsf.gov/explore-data/microdata/federal-facilities-research-development.

2.4. Contact for More Information

Additional information about this survey can be obtained by contacting:

National Center for Science and Engineering Statistics
U.S. National Science Foundation
Randolph Building, 401 Dulany Street
Alexandria, VA 22314
Phone: (703) 292-8780
E-mail: ncsesweb@nsf.gov

[1] Publications from NCSES can be found at http://www.nsf.gov/statistics.