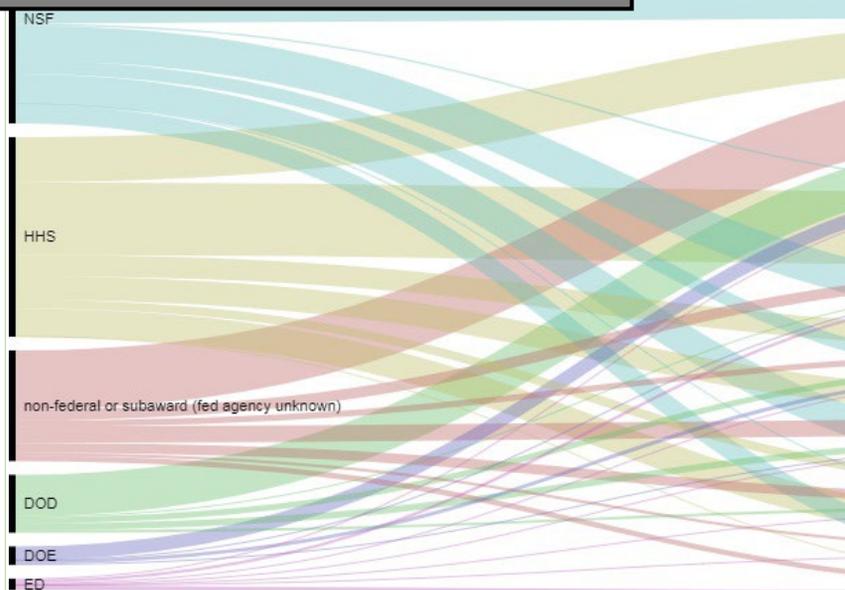


# Data Dictionary for IRIS UMETRICS 2019 Data Release



# Data Dictionary for 2019 IRIS UMETRICS Data Release

*Data fields are alphabetically ordered.*

<b>core_award</b>					
<b>Field Name</b>	<b>Column Name</b>	<b>Data Type</b>	<b>Set Length</b>	<b>Max Length</b>	<b>Field Definition</b>
Award Title	award_title	varchar	500	496	Title of award
Campus ID	campus_id	varchar	10	8	IRIS-generated identifier assigned to each campus for de-identification purposes. Each ID is a combination of institution_id and a serial number helpful to identify the campus to which a given award is made and by which the award is being administered
CFDA	cfda	varchar	10	6	A five-digit CFDA (Catalog of Federal Domestic Assistance) number retrieved from the unique_award_number that universities submit. This number is assigned to awards to represent the source of funding. If the first two digits are between 10 and 98, awards are federally funded. If the first two digits range between 00 and 09, or 99, sources of awards are most likely non-federal
Federal Award Number Standardized	fed_awd_num_std	varchar	100	50	Unique identifier specifying a federal award; this is the same as the unique_award_number but with the CFDA number and white spaces removed. Null if CFDA is missing or the CFDA number indicates a non-federal award, i.e., a CFDA starting with two digit numbers between 00 and 09, or 99
Funding Source Name Clean	funding_source_name_clean	varchar	200	98	Cleaned name of funding source assigned to each project; if university is a recipient of a prime award, the funding source name is the direct sponsor; if university is a subrecipient of an award, this is the name of the pass-through entity (exceptions noted in documentation)

Funding Source Name Raw	funding_source_name_raw	varchar	200	104	Raw name of funding source assigned to each project; if university is a recipient of a prime award, the direct sponsor is the name of funding source; if university is a subrecipient of an award, this is the name of the pass-through entity (exceptions are noted in documentation)
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Overhead Charged	overhead_charged	numeric	9	9	Actual overhead dollars charged to the award in the specified period
Period End Date	period_end_date	date	3	3	End of period in which a monthly expense transaction took place; each period end day is the last day of a month: e.g., 3/30/2008 or 12/31/2014
Period Start Date	period_start_date	date	3	3	Beginning of period in which a monthly expense transaction took place; each period start day is the first day of a month: e.g., 4/1/2009 or 10/1/2015
Recipient Account Number	recipient_account_number	varchar	50	18	A university's internal account number to uniquely identify each project; typically an accounting code used to allocate funds received from an award
Sub-organization Unit	sub_org_unit	varchar	20	12	IRIS-generated identifier assigned to sub-organization units to which each funded project is assigned, such as a particular college within a given IRIS member university. (This is not at the level of individual departments.) Each ID is a combination of campus_id described above and a serial number assigned to each sub-organization unit within each campus. This ID helps to identify the college or unit to which a given award is made and by which the award is being administered at a lower-level
Total Direct Expenditures	total_direct_expenditures	numeric	9	9	Total direct expenditures charged to the award in the specified period

Unique Award Number	unique_award_number	varchar	100	88	<p>University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)</p>
---------------------	---------------------	---------	-----	----	--

# core\_employee

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
CFDA	cfda	varchar	10	6	A five-digit CFDA (Catalog of Federal Domestic Assistance) number retrieved from the unique_award_number that universities submit. This number is assigned to awards to represent the source of funding. If the first two digits are between 10 and 98, awards are federally funded. If the first two digits range between 00 and 09, or 99, sources of awards are most likely non-federal
Employee Number	emp_number	varchar	100	32	IRIS-generated unique identifier assigned to all personnel being paid by awards
Federal Award Number Standardized	fed_awd_num_std	varchar	100	86	Unique identifier specifying a federal award; this is the same as the unique_award_number but with the CFDA number and white spaces removed. Null if CFDA is missing or the CFDA number indicates a non-federal award, i.e., a CFDA starting with two digit numbers between 00 and 09, or 99
FTE Status	fte_status	numeric	9	5	Designation of the status of the funded personnel (e.g., full time = 1.0, half time = .5); FTE is a university specific, not an award specific field; the value ranges between 0 and 1
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Job Title	job_title	varchar	200	123	Job or occupation title assigned to the funded personnel by IRIS member universities
Object Code	object_code	varchar	50	39	Internal object code or other expense type category assigned to a transaction to identify payment purposes or resources
Occupational Class	occupational_class	varchar	50	50	Job classification provided by IRIS member universities
Period End Date	period_end_date	date	3	3	End of period in which a monthly expense transaction took place; each period end day is the last day of a month: e.g., 3/30/2008 or 12/31/2014

Period Start Date	period_start_date	date	3	3	Beginning of period in which a monthly expense transaction took place; each period start day is the first day of a month: e.g., 4/1/2009 or 10/1/2015
Proportion of Earnings	proportion_earnings_allocated	numeric	9	9	Calculated portion of earnings charged by funded personnel to the award in the specified period. This is not actual salary or dollar amounts, the value ranges between 0 and 1 depending on how much of the salary is derived from an award
Recipient Account Number	recipient_account_number	varchar	50	15	A university's internal account number to uniquely identify each project; typically an accounting code used to allocate funds received from an award
SOC Code	soc_code	varchar	50	30	Standard Occupational Classification codes that are required for federal agency reporting ( <a href="http://www.bls.gov/soc/">http://www.bls.gov/soc/</a> ); each occupation in the SOC is placed within one of 23 major groups
UMETRICS Occupational Class	umetrics_occupational_class	varchar	50	22	Job classification generated by IRIS; jobs are categorized into 6 major aggregate groups (Faculty, Staff, Post Graduate Research, Graduate Student, Undergraduate, and Other). The Staff group is further classified into 6 categories (Clinical, Research, Research Facilitation, Technical Support, Instructional, Other Staff)
Unique Award Number	unique_award_number	varchar	100	50	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)

# core\_vendor

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
CFDA	cfda	varchar	10	6	A five-digit CFDA (Catalog of Federal Domestic Assistance) number retrieved from the unique_award_number that universities submit. This number is assigned to awards to represent the source of funding. If the first two digits are between 10 and 98, awards are federally funded. If the first two digits range between 00 and 09, or 99, sources of awards are most likely non-federal
Federal Award Number Standardized	fed_awd_num_std	varchar	100	86	Unique identifier specifying a federal award; this is the same as the unique_award_number but with the CFDA number and white spaces removed. Null if CFDA is missing or the CFDA number indicates a non-federal award, i.e., a CFDA starting with two digit numbers between 00 and 09, or 99
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Object Code	object_code	varchar	50	39	Internal object code or other expense type category assigned to a transaction to identify payment purposes or resources
Period End Date	period_end_date	date	3	3	End of period in which a monthly expense transaction took place; each period end day is the last day of a month: e.g., 3/30/2008 or 12/31/2014
Period Start Date	period_start_date	date	3	3	Beginning of period in which a monthly expense transaction took place; each period start day is the first day of a month: e.g., 4/1/2009 or 10/1/2015
Person Organization Flag	person_org_flag	varchar	1	1	A binary code ('P' for person or 'O' for organization) to differentiate type of vendors. This dichotomous category was utilized to mask vendors' personally identifiable information
Recipient Account Number	recipient_account_number	varchar	50	15	A university's internal account number to uniquely identify each project; typically an accounting code used to allocate funds received from an award

Unique Award Number	unique_award_number	varchar	100	50	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)
Vendor Address	vendor_address	varchar	200	129	Address of the vendor. IRIS has replaced vendor address with the string of 'masked' if vendor addresses are provided by universities (thus not null) and vendors are individuals
Vendor City	vendor_city	varchar	50	37	City of the vendor associated with the vendor address
Vendor Country	vendor_country	varchar	50	16	Country of the vendor associated with the vendor address
Vendor Domestic Zipcode	vendor_domestic_zipcode	varchar	50	15	US ZIP code of vendor associated with the vendor address
Vendor DUNS	vendor_duns	varchar	50	12	A vendor's nine-digit (DUNS) number to identify business entities on a location-specific basis—the Data Universal Numbering System or D-U-N-S Number is copyrighted and provided by Dun & Bradstreet (D&B). If a DUNS was provided (i.e., not null) by universities for vendors that are individuals, IRIS has replaced DUNS with the string 'masked'
Vendor EIN	vendor_ein	varchar	50	30	A vendor's nine-digit Employer Identification Number (EIN). If an EIN was provided (i.e., not null) by universities for vendors that are individuals, IRIS has replaced EIN with the string 'masked'
Vendor Foreign Zip Code	vendor_foreign_zipcode	varchar	50	17	Foreign ZIP/postal code of vendor associated with the vendor address
Vendor ID	vendor_id	varchar	200	32	IRIS-generated unique identifier assigned to the vendor (an organization or individual) that provides goods or services paid by an IRIS member university's award. IRIS cleans vendor name records from the data submitted by universities and generates this identifier based on the cleaned names

Vendor Name Clean	vendor_name_clean	varchar	200	170	Cleaned name of the vendor. IRIS has replaced vendor names with the string of 'masked' if names are provided by universities (thus not null) and vendors are individuals
Vendor Name Raw	vendor_name_raw	varchar	200	142	Raw name of the vendor. IRIS has replaced vendor names with the string of 'masked' if names are provided by universities (thus not null) and vendors are individuals
Vendor Payment Amount	vendor_payment_amt	numeric	9	9	Funds charged to the award by the vendor in the specified period
Vendor State	vendor_state	varchar	50	23	State of the vendor associated with the vendor address

# core\_subaward

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
CFDA	cfda	varchar	10	6	A five-digit CFDA (Catalog of Federal Domestic Assistance) number retrieved from the unique_award_number that universities submit. This number is assigned to awards to represent the source of funding. If the first two digits are between 10 and 98, awards are federally funded. If the first two digits range between 00 and 09, or 99, sources of awards are most likely non-federal
Federal Award Number Standardized	fed_awd_num_std	varchar	100	98	Unique identifier specifying a federal award; this is the same as the unique_award_number but with the CFDA number and white spaces removed. Null if CFDA is missing or the CFDA number indicates a non-federal award, i.e., a CFDA starting with two digit numbers between 00 and 09, or 99
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Object Code	object_code	varchar	50	7	Internal object code or other expense type category assigned to a transaction to identify payment purposes or resources
Period End Date	period_end_date	date	3	3	End of period in which a monthly expense transaction took place; each period end day is the last day of a month: e.g., 3/30/2008 or 12/31/2014
Period Start Date	period_start_date	date	3	3	Beginning of period in which a monthly expense transaction took place; each period start day is the first day of a month: e.g., 4/1/2009 or 10/1/2015
Person Organization Flag	person_org_flag	varchar	1	1	A binary code ('P' for person or 'O' for organization) to differentiate type of subawardees. This dichotomous category was utilized to mask subawardees' personally identifiable information
Recipient Account Number	recipient_account_number	varchar	50	15	A university's internal account number to uniquely identify each project; typically an accounting code used to allocate funds received from an award

Subaward Address	subaward_address	varchar	200	123	Address of the subawardee. IRIS has replaced subawardee address with the string of 'masked' if subawardee address are provided by universities (thus not null) and subawardees are individuals
Subaward City	subaward_city	varchar	50	50	City of the subawardee associated with the subawardee address
Subaward Country	subaward_country	varchar	50	16	Country of the subawardee associated with the subawardee address
Subaward Domestic Zipcode	subaward_domestic_zipcode	varchar	50	12	US ZIP code of subawardee associated with the subawardee address
Subaward DUNS	subaward_duns	varchar	50	14	A subawardee's nine-digit (DUNS) number to identify business entities on a location-specific basis—the Data Universal Numbering System or D-U-N-S Number is copyrighted and provided by Dun & Bradstreet (D&B). If a DUNS was provided (i.e., not null) by universities for subawardees that are individuals, IRIS has replaced DUNS with the string 'masked'
Subaward EIN	subaward_ein	varchar	50	30	A subawardee's nine-digit Employer Identification Number (EIN). If an EIN was provided (i.e., not null) by universities for subawardees that are individuals, IRIS has replaced EIN with the string 'masked'
Subaward Foreign Zip Code	subaward_foreign_zipcode	varchar	50	19	Foreign ZIP/postal code of subawardee associated with the subawardee address
Subaward ID	subaward_id	varchar	200	32	IRIS-generated unique identifier assigned to the subaward recipient organization to which an IRIS member university provides program awards / subgrants / subcontracts. IRIS cleans subawardee name records from the data submitted by universities and generates this identifier based on the cleaned names
Subaward Name Clean	subaward_name_clean	varchar	200	199	Cleaned name of the subawardee. IRIS has replaced subawardee names with the string of 'masked' if names are provided by universities (thus not null) and subawardees are individuals

Subaward Name Raw	subaward_name_raw	varchar	200	200	Raw name of the subawardee. IRIS has replaced subawardee names with the string of 'masked' if names are provided by universities (thus not null) and subawardees are individuals
Subaward Payment Amount	subaward_payment_amt	numeric	9	9	Funds charged to the award by the subawardee in the specified period
Subaward State	subaward_state	varchar	50	11	State of the subawardee associated with the subawardee address
Unique Award Number	unique_award_number	varchar	100	66	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)

# aux\_suborganization

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Campus ID	campus_id	varchar	10	8	IRIS-generated identifier assigned to each campus for de-identification purposes. Each ID is a combination of institution_id and a serial number helpful to identify the campus to which a given award is made and by which the award is being administered
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Main Campus Flag	main_campus	int	4	4	A binary value to indicate the main campus of each member university; valued 1 if a given campus is the main campus (flagship campus); valued 0 otherwise; if a given university has one campus in UMETRICS data, then that is the main campus
Sub-organization Unit	sub_org_unit	varchar	20	12	IRIS-generated identifier assigned to sub-organization units to which each funded project is assigned, such as a particular college within a given IRIS member university. (This is not at the level of individual departments.) Each ID is a combination of campus_id described above and a serial number assigned to each sub-organization unit within each campus. This ID helps to identify the college or unit to which a given award is made and by which the award is being administered at a lower-level
Sub-organization Unit Name	sub_org_unit_name	varchar	100	100	Sub-organizational unit name that maps to sub-organizational unit code, e.g., the college of natural sciences, the medical school, or the college of engineering. If the sub-org unit name provided by IRIS member universities contains identifiable information, IRIS replaced all information in this field with the string 'masked'

## aux\_object\_code

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Object Code	object_code	varchar	50	7	Internal object code or other expense type category assigned to a transaction to identify payment purposes or resources
Object Code Description	object_code_desc	varchar	500	469	Description of internal object code or other expense type category assigned to a transaction; maps to object code

## aux\_emp\_age

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Emp Number	emp_number	varchar	100	32	IRIS-generated unique identifier assigned to all personnel being paid by awards
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Year of Birth Category	yob_category	int	4	4	To simplify age bands in the field of "yob_range" above, we converted age bands into integers between 1 and 11. If yob_range value is "na" it is coded as 99, and if "masked" it is coded as 98
Year of Birth Range	yob_range	varchar	50	21	The range in which an employee's year of birth falls, e.g., if one's birth year falls between 1988 and 1992, it is coded as "between_1988_and_1992". Employee age is mapped to unique employee ID found in the Employee File, and due to disclosure risk, 11 age bands are used to aggregate years. Additional values include "na" and "masked." If no information on an employee's birth year was provided by universities, it is coded as "na". Some age bands are rolled up into one and coded as "masked" due to disclosure risk

## aux\_institution\_fastfacts

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Carnegie Classification	carnegie_classification	varchar	50	50	Derived from the 2018 Classification Update of the traditional Carnegie Classification Framework for each academic institution; data retrieved Carnegie Classification of Institutions of Higher Education website: <a href="http://carnegieclassifications.iu.edu/">http://carnegieclassifications.iu.edu/</a>
Carnegie Code	carnegie_code	int	4	4	A numerical code for the Carnegie Classification. Based on the 2018 Carnegie Classification Update of the traditional Carnegie Classification Framework for each academic institution; Description of codes are available on the Carnegie Classification of Institutions of Higher Education website: <a href="http://carnegieclassifications.iu.edu/">http://carnegieclassifications.iu.edu/</a>
Distance between medical school and main campus	main_med_distance	int	4	4	The geographical distance (in miles) of the medical school from the main campus if an IRIS member university has a medical school
Fall Enrollment	fall_enrollment	int	4	4	The number of students enrolled in courses that are creditable toward a degree, diploma, certificate, or other formal award, or are part of a vocational or occupational program including any students enrolled in off-campus centers; data retrieved from the Integrated Postsecondary Education Data System (IPEDS) Fall Enrollment Survey
Federally Financed R&D Expenditures in All Fields	fed_rd_expenditures	int	4	4	R&D expenditures in all fields, including direct and recovered indirect costs, funded by all agencies of the Federal government; data retrieved from NSF HERD
Institution Control	institution_control	int	4	4	Defined for academic institutions as private or public (not applicable to biomedical institutions); values include 1 (Private) and 0 (Public); data retrieved from NSF HERD
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Land Grant	land_grant	decimal	9	5	Indicator for whether an institution is a Land Grant institution; values include 1 (institution is a Land Grant institution) and 0 (not a Land Grant institution); data retrieved from IPEDS and verified

					on <a href="https://nifa.usda.gov/land-grant-colleges-and-universities-partner-website-directory">https://nifa.usda.gov/land-grant-colleges-and-universities-partner-website-directory</a>
Medical School	med_school	int	4	4	Indicator for each institution having a medical school included as part of its reporting unit; Values include 1 (has medical school) and 0 (does not include medical school); data retrieved from NSF HERD
Number of Doctorate Recipients	number_doc_recipients	int	4	4	All earned doctorates granted by universities; data retrieved from the NSF Survey of Earned Doctorates (SED)
Number of Graduate Students	number_grad_students	int	4	4	The number of graduate students enrolled in GSS-eligible science, engineering, and health (SEH) units in the fall of the data collection year; data retrieved from the NSF-NIH Survey of Graduate Students & Postdoctorates in Science and Engineering (GSS)
Number of Other Personnel	number_other_personnel	int	4	4	All other personnel paid from the R&D salaries, wages and fringe benefits reported on the NSF Higher Education Research and Development Survey (HERD) who are not categorized as principal investigators; data retrieved from NSF HERD
Number of Postdocs	number_post_docs	int	4	4	Personnel defined as postdocs, namely recent doctorate recipients with limited-term appointments primarily intended for training under the supervision of a senior scholar. Data retrieved from the NSF-NIH Survey of Graduate Students & Postdoctorates in Science and Engineering (GSS)
Number of Principal Investigators	number_pis	int	4	4	Personnel paid from the R&D salaries, wages and fringe benefits reported on the survey (NSF Research and Development Expenditures at Universities and Colleges/Higher Education Research and Development Survey), and designated by the institution to direct the R&D project or program and be responsible for the scientific and technical direction of the project; Co-investigators (co-PIs) may be designated for this role and are also included. Missing data for this question were not imputed, therefore aggregate totals represent an undercount; data retrieved from NSF HERD
Total R&D Expenditures in All Fields	total_rd_expenditures	int	4	4	R&D expenditures from the institution's current operating funds that were separately accounted for, including expenditures for organized research as defined by 2 CFR 220 Part 200 Appendix III and expenditures from funds designated for research.

					Expenditures came from internal or external funding and included recovered and unrecovered indirect costs. Funds passed through to subrecipient organizations were also included. R&D was excluded if it was conducted by university faculty or staff at outside institutions and was not accounted for in the reporting institution's financial records. Data retrieved from NSF HERD
Year	year	int	4	4	Year (ranging between 2010 and 2017) is defined in two ways: 1) Academic year: doctorate recipients, fall enrollment, and other personnel-related data; 2) Fiscal year: R&D expenditures as defined in the data source, the NSF Higher Education R&D Survey (NSF HERD)

# aux\_comprehensive\_award\_list

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Award File	award	int	4	4	A binary code to differentiate the file from which a given award originates; coded 1 if the award is present in Award file; coded 0 otherwise
CFDA	cfda	varchar	10	6	A five-digit CFDA (Catalog of Federal Domestic Assistance) number retrieved from the unique_award_number that universities submit. This number is assigned to awards to represent the source of funding. If the first two digits are between 10 and 98, awards are federally funded. If the first two digits range between 00 and 09, or 99, sources of awards are most likely non-federal
Employee File	employee	int	4	4	A binary code to differentiate the file from which a given award originates; coded 1 if the award is present in Employee file; coded 0 otherwise
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Subaward File	subaward	int	4	4	A binary code to differentiate the file from which a given award originates; coded 1 if the award is present in Subaward file; coded 0 otherwise
Unique Award Number	unique_award_number	varchar	100	88	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)
Vendor File	vendor	int	4	4	A binary code to differentiate the file from which a given award originates; coded 1 if the award is present in Vendor file; coded 0 otherwise

# aux\_cfda

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Active CFDA as of FY2017	active_cfda_FY2017	int	4	4	A categorical value to indicate the status of the five digit CFDA numbers associated with domestic assistance programs; valued 1 if a given five digit CFDA number is mapped to a currently active domestic assistance program as of June 2018; valued 0 if inactive; valued 99 if a given CFDA number is not found in the historical file. See below for the definition of historical file
Active File	active_file	int	4	4	This field is a binary value to differentiate the file from which a given CFDA number originates; valued 1 if the CFDA number is present in active file; valued 0 otherwise. The active file was downloaded from cfda.gov (now sam.gov) and contains all currently active (as of June 2018) domestic assistance programs
Active First Two Digits of CFDA as of FY2017	active_two_FY2017	int	4	4	A binary value to indicate the current status of the first two digit numbers of CFDA numbers which is mapped to assisting federal agencies; valued 1 if a given two digit number is actively used as of June 2018; valued 0 otherwise
CFDA	cfda	varchar	10	7	A CFDA (Catalog of Federal Domestic Assistance) number used to identify and sort federal financial assistance programs. Each CFDA number contains five digits and appears in the following format: ##.### (e.g., 10.001 or 93.301)
Current Status	current_status	varchar	200	157	Details about current status of each domestic assistance program if a given CFDA number was changed, transferred to or incorporated into other CFDA numbers, or archived. This field is null if no detail is available from the historical file or CFDA numbers are present only in the UMETRICS file
First Two Digits of CFDA	first_two	int	4	4	The first two digits of CFDA numbers that appear in the CFDA field described above. The first two digit number can be used to map each federal financial assistance program to funding agencies. The first two digit numbers (in integer format) range between 10 and 98
Government Unit Code	gov_unit_code	varchar	50	13	The official acronym of the assisting federal agency at the highest level of a given domestic assistance program. For example, if a domestic assistance program is administered and managed by the National

					Institutes of Health, this field indicates its parent organization acronym, HHS
Government Unit Name	gov_unit_name	varchar	200	62	The name of the assisting federal agency at the highest level of a given domestic assistance program. For example, if a domestic assistance program is administered and managed by the National Institutes of Health, this field indicates its parent organization, the Department of Health and Health Sciences
Historical File	historical_file	int	4	4	This field is a binary value to differentiate the file from which a given CFDA number originates; valued 1 if the CFDA number is present in historical file; valued 0 otherwise. The historical file was downloaded by IRIS staff in December 2017 and includes basic information about all domestic assistance programs that have existed from the 1960s to date. This file is no longer publicly available from the new website (sam.gov) that replaced cfda.gov. The raw file is archived by IRIS
Program Title	program_title	varchar	500	238	The program title that is available and downloaded from the website, cfda.gov (sam.gov). This field is null for CFDA's that are not found in the historical file. See above for the definition of historical file
Sub-unit Code	sub_unit_code	varchar	50	21	The official acronym of the federal agency that directly provides and administers a given domestic assistance program. For example, if a domestic assistance program is provided and administered by the National Institutes of Health, this field indicates NIH
Sub-unit Name	sub_unit_name	varchar	200	86	The name of the federal agency that directly provides and administers a given domestic assistance program. For example, if a domestic assistance program is provided and administered by the National Institutes of Health, this field indicates National Institutes of Health. NOTE: In the historical file, the equivalent field is named "office"
Type of Assistance	type_of_assistance	varchar	500	500	Each program is identified in terms of one or more of the 15 types of assistance provided. This field has values only for active CFDA's
UMETRICS File	umetrics_file	int	4	4	This field is a binary value to differentiate the file from which a given CFDA number originates; valued 1 if the CFDA number is present in umetrics file; valued 0 otherwise. The UMETRICS file was created by IRIS staff and includes 1879 unique five-digit CFDA numbers that were retrieved from core award, employee, vendor, and subaward files in December 2018
Program Established Year	year_established	int	4	4	The year when a given domestic assistance program (associated with a unique CFDA number) was established. The field is null if CFDA

					numbers are not found in the historical file. See above for the definition of historical file
Program Modified Year	year_modified	int	4	4	The year when a given domestic assistance program (associated with a unique CFDA number) was modified, if modified. The field is null if CFDA numbers are not found in the historical file. See above for the definition of historical file

## link\_nsf\_xwalk

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Award ID	award_id	varchar	10	7	The NSF unique award identifier assigned to each NSF award that is matched to UMETRICS award number
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Match Step	match_step	varchar	10	3	Indicator of which matching step generated each matched pair between UMETRICS and federal agency award record; values include: 1, 2, 3, 4, and 5 as the code is written to match award numbers through five different methods
Unique Award Number	unique_award_number	varchar	100	49	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)

# link\_nsf

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Abstract Narration	abstract_narration	varchar	8000	8000	Abstract of the award
ARRA Amount	arra_amount	varchar	50	12	Amount of funding obligated designated as ARRA funding
Award Amount	award_amount	numeric	9	5	The amount obligated to date for the project
Award Effective Date	award_effective_date	date	3	3	Effective date of the award
Award Expiration Date	award_expiration_date	date	3	3	The date on which the award expires
Award ID	award_id	varchar	10	7	The agency assigned award number (a seven digit number)
Award Instrument	award_instrument	varchar	100	33	Type of Award
Award Instrument Code	award_instrument_code	varchar	50	4	Code associated with type of award
Award Title	award_title	varchar	500	181	Descriptive title of the project
Directorate	directorate	varchar	100	60	Department of NSF funding the award
Division	division	varchar	200	74	Division of NSF funding the award

# link\_nih\_xwalk

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Core Project Number	core_project_num	varchar	50	11	The NIH core project number assigned to each NIH-funded project that is matched to the core project part of UMETRICS award number
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Unique Award Number	unique_award_number	varchar	100	61	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)

# link\_nih

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Abstract	abstract	varchar	8000	8000	Abstract of the funded project
Activity	activity	varchar	8000	3	A 3-character code identifying the grant, contract, or intramural activity through which a project is supported. Within each funding mechanism, NIH uses 3-character activity codes (e.g., F32, K08, P01, R01, T32, etc.) to differentiate the wide variety of research-related programs NIH supports
Administrator IC	administering_ic	varchar	10	2	Administering Institute or Center - A two-character code to designate the agency, NIH Institute, or Center administering the grant
Application ID	application_id	int	4	4	A unique identifier of the project record
Application Type	application_type	varchar	10	3	A one-digit code to identify the type of application funded: 1) New Application; 2) Competing continuation; 3) Application for additional support; 4) Competing extension for an R37 award or first non-competing year of a Fast Track SBIR/STTR award; 5) Non-competing continuation; 7) Change of grantee institution; 9) Change of NIH awarding Institute or Division
ARRA Funded	arra_funded	varchar	10	1	"Y" indicates a project supported by funds appropriated through the American Recovery and Reinvestment Act of 2009
Award Notice Date	award_notice_date	date	3	3	Award notice date or Notice of Grant Award (NGA) is a legally binding document stating the government has obligated funds and which defines the period of support and the terms and conditions of award
Budget End	budget_end	date	3	3	The date when a project's funding for a particular fiscal year ends
Budget Start	budget_start	date	3	3	The date when a project's funding for a particular fiscal year begins
CFDA Code	cfda_code	varchar	50	17	A CFDA (Catalog of Federal Domestic Assistance) number used to identify and sort federal financial assistance programs
Core Project Number	core_project_num	varchar	100	27	Core project number
Direct Cost	direct_cost_amt	numeric	9	5	Total direct cost funding for a project for a given fiscal year. Available only for NIH awards funded in FY 2012 onward and not available for SBIR/STTR awards
ED Inst Type	ed_inst_type	varbinary	200	124	Institution type

FOA Number	foa_number	varchar	50	14	The number of the funding opportunity announcement, if any, under which the project application was solicited. Funding opportunity announcements may be categorized as program announcements, requests for applications, notices of funding availability, solicitations, or other names depending on the agency and type of program. Funding opportunity announcements can be found at <a href="http://Grants.gov/FIND">Grants.gov/FIND</a> and in the NIH Guide for Grants and Contracts
Full Project Number	full_project_num	varchar	100	35	Commonly referred to as a grant number, intramural project, or contract number. For grants, this unique identification number is composed of the type code, activity code, Institute/Center code, serial number, support year, and (optional) a suffix code to designate amended applications and supplements
Funding ICs	funding_ics	varchar	1000	291	The NIH Institute or Center(s) providing funding for a project are designated by their acronyms (see Institute/Center acronyms). Each funding IC is followed by a colon (:) and the amount of funding provided for the fiscal year by that IC. Multiple ICs are separated by semicolons (;). Project funding information is available only for NIH projects awarded in FY 2008 and later fiscal years
FY	fy	int	4	4	The fiscal year appropriation from which project funds were obligated
IC Name	ic_name	varchar	200	79	Full name of the administering agency, Institute, or Center
Indirect Cost	indirect_cost_amt	numeric	9	5	Total indirect cost funding for a project for a given fiscal year. Available only for NIH awards funded in FY 2012 and onward and not available for SBIR/STTR awards
NIH Spending CATS	nih_spending_cats	varchar	8000	2232	Congressionally-mandated reporting categories into which NIH projects are categorized. Available for fiscal years 2008 and later. Each project's spending category designations for each fiscal year are made available the following year as part of the next President's Budget request. See the Research, Condition, and Disease Categorization System for more information on the categorization process
Org Dept	org_dept	varchar	100	30	The departmental affiliation of the contact principal investigator for a project, using a standardized categorization of departments. Names are available only for medical school departments
Org District	org_district	varchar	50	4	The congressional district in which the business office of the grantee organization or contractor is located. Note that this may be different from the research performance site

Org FIPS	org_fips	varchar	50	2	The country code of the grantee organization or contractor as defined in the Federal Information Processing Standard
PHR	phr	varchar	8000	8000	Submitted as part of a grant application, this statement articulates a project's potential to improve public health
PI IDs	pi_ids	varchar	1000	261	A unique identifier for each of the project Principal Investigators. Each PI in the RePORTER database has a unique identifier that is constant from project to project and year to year, but changes may be observed for investigators that have had multiple accounts in the past, particularly for those associated with contracts or sub-projects
Project End	project_end	date	3	3	The current end date of the project, including any future years for which commitments have been made. For subprojects of a multi-project grant, this is the end date of the parent award. Upon competitive renewal of a grant, the project end date is extended by the length of the renewal award
Project Start	project_start	date	3	3	The start date of a project. For subprojects of a multi-project grant, this is the start date of the parent award
Project Terms	project_terms	varchar	8000	8000	Thesaurus terms assigned by NIH CRISP indexers, only applicable to projects funded prior to the fiscal year 2008
Project Title	project_title	varchar	1000	200	Title of the funded grant, contract, or intramural (sub)project
Serial Number	serial_number	varchar	20	8	A six-digit number assigned in serial number order within each administering organization
Study Section	study_section	varchar	20	4	A designator of the legislatively-mandated panel of subject matter experts that reviewed the research grant application for scientific and technical merit
Study Section Name	study_section_name	varchar	200	94	The full name of a regular standing Study Section that reviewed the research grant application for scientific and technical merit. Applications reviewed by panels other than regular standing study sections are designated by "Special Emphasis Panel"
Sub Project ID	subproject_id	int	4	4	A unique numeric designation assigned to subprojects of a "parent" multi-project research grant
Suffix	suffix	varchar	20	6	A suffix to the grant application number that includes the letter "A" and a serial number to identify an amended version of an original application and/or the letter "S" and serial number indicating a supplement to the project

Support Year	support_year	int	4	4	The year of support for a project, as shown in the full project number. For example, a project with number 5R01GM0123456-04 is in its fourth year of support
Total Cost	total_cost	numeric	9	5	Total project funding from all NIH Institute and Centers for a given fiscal year. Costs are available only for: 1) NIH and CDC grant awards (only the parent record of multi-project grants) funded in FY 2000 and later fiscal years; 2) NIH intramural projects (activity codes beginning with "Z") in FY 2007 and later fiscal years; 3) NIH contracts (activity codes beginning with "N") in FY 2007 and later fiscal years. For multi-project grants, Total Cost includes funding for all of the constituent subprojects. This field will be blank on subproject records; the total cost of each subproject is found in Total_Cost_Sub_Project (FY 2000 and later fiscal years only)
Total Cost Sub Project	total_cost_sub_project	numeric	9	5	Applies to subproject records only. Total funding for a subproject from all NIH Institute and Centers for a given fiscal year. Costs are available only for NIH awards funded in FY 2000 and later

# link\_nih\_pub\_xwalk

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Core Project Number	core_project_num	varchar	50	11	The NIH core project number assigned to each NIH-funded project that is matched to the core project part of UMETRICS award number
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
PMID	pmid	int	4	4	PubMed unique identifier assigned by the NIH National Library of Medicine to papers indexed in PubMed (index of abstracts). The number is 1- to 8- digits with no leading zeros
Unique Award Number	unique_award_number	varchar	100	61	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)
Year	year	int	4	4	Year of publication, from pub_date field in the link_nih_pub file

## link\_nih\_pub

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
ISSN	issn	varchar	20	9	The International Standard Serial Number, an eight-character value that uniquely identifies the journal.
Journal Issue	journal_issue	varchar	200	75	An issue, part, or supplement of the journal in which the article was published.
Journal Title	journal_title	varchar	1000	282	Full journal title, taken from the NIH National Library of Medicine's cataloging data.
Journal Title Abbreviation	journal_title_abbr	varchar	500	108	Standard abbreviation for the title of the journal in which the article appeared.
Journal Volume	journal_volume	varchar	200	99	Volume number of the journal in which the article was published
Language	lang	varchar	20	3	Three-letter abbreviation representing the language(s) in which an article was published. List of abbreviations is available at: <a href="https://www.nlm.nih.gov/bsd/language_table.html">https://www.nlm.nih.gov/bsd/language_table.html</a> .
Page Number	page_number	varchar	500	138	Pages for the article, including document numbers for electronic articles.
PMCID	pmc_id	varchar	20	9	A unique identifier for the article in PubMed Central (index of full-text papers). The PMCID or PMC Identifier, is assigned to each full-text paper in PubMed Central by the National Library of Medicine.
PMID	pmid	int	4	4	PubMed unique identifier assigned by the NIH National Library of Medicine to papers indexed in PubMed (index of abstracts). The number is 1- to 8- digits with no leading zeros; this is the field one should use to retrieve publication details when using the award-level NIH - publication - UMETRICS crosswalk.
Publication Date	pub_date	varchar	50	23	Date on which the issue of the journal was published. The standardized format includes a 4-digit year, a 3-character abbreviated month, and a 1 or 2-digit day, but the data are taken as published in the journal issue so not every record contains all elements.
Publication Title	pub_title	varchar	8000	2000	Title of the article; if originally published in a non-English language this is a translation enclosed in brackets.
Publication Year	pub_year	int	4	4	Year of publication, from pub_date

# link\_usda\_xwalk

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Award ID	award_id	varchar	50	16	USDA award identifier assigned to each USDA award that is matched to UMETRICS award number; this identifier could be grant number or project number if grant number is not available in UMETRICS or project number is matched
Institution ID	institution_id	int	4	4	IRIS-generated unique identifier assigned to each IRIS member university for de-identification purposes. Values are four or five digit numbers
Match Step	match_step	varchar	10	3	Indicator of which matching step generated each matched pair between UMETRICS and federal agency award record; values include: 1, 2, 3, 4, and 5 as the code is written to match award numbers through five different methods
Unique Award Number	unique_award_number	varchar	100	32	University-generated unique identifier specifying an award and its funding source, made up of the 5-digit funding source code (e.g., CFDA number) and an award identifier. Award identifier may include the awarding agency's federal award ID (e.g., federal grant number, contract number, or loan number) or an internal award ID for non-federal awards. Values may include a space or dash in between them: e.g., "10.310 2010-12345-54321" (USDA example), "47.050 1234567" (NSF example), "93.865 2-R01-DK-012345-15-S1" (NIH example), "00.000 1234567" and "00.200 State Award 1" (Non-federal grant examples)

# link\_usda

Field Name	Column Name	Data Type	Set Length	Max Length	Field Definition
Abstract	abstract	varchar	8000	8000	Abstract of the funded project
Accession Number	accession_number	int	4	4	A unique 7-digit identifier assigned by USDA to all Grants, Cooperative Agreements, and other types of funding instrument.
Funding Source	funding_source	varchar	100	26	This field indicates the funding mechanism for a project from a USDA and/or NIFA perspective. These include: 3D Grant; Animal Health; Cooperative Agreement; Evans-Allen; Hatch; McIntire-Stennis; NRI Competitive Grant; Other Extension Grant; Other Grant; RREA; SERD Grant; Small Business Grant; Special Grant; State; USDA Contract; USDA Grant, and; USDA In-house
Grant Number	grant_number	varchar	100	18	A unique award identifier assigned by USDA; formatting differs by sub-agency. For example, USDA grants funded through NIFA will be assigned a Grant Number consisting of a 4-digit year followed by a hyphen, five-digit financial data code, another hyphen, and five-digit sequential number, e.g., 2004-45066-03027. If not a grant, NIFA assigns 'N/A ('Not Applicable') as the Award ID/Grant Number. If USDA awards are funded through ARS, the Grant Number consists of a two-digit number followed by a hyphen, four-digit number, hyphen, one-digit number, hyphen, and three-digit number, e.g., 58-3620-0-125. If not a grant, NIFA assigns 'N/A' ('Not Applicable') as Award ID.
Grant Year	grant_year	int	4	4	Year in which a project is funded
Multistate Number	multistate_number	varchar	50	15	An identifier associating multistate (regional) research projects that is assigned by the sponsoring regional association or committee. Codes include a regional association abbreviation, e.g., NC (North Central), NE (North East), S (South), W (West) or NSRP (National Research Support Program) or associated committee abbreviation (e.g., NCA, NCERA) with a serial number
Program Code	program_code	varchar	50	5	The associated program (area priority) code or program code for an award
Project End Date	project_end_date	date	3	3	The end date of a project
Project Number	project_number	varchar	100	21	An alpha-numeric identifier assigned and maintained by the reporting institution. Project numbers may be reused within an institution, but the

					project number is prefixed with a three-or four-character institution code (controlled by REEport Site Administrators) to insure a unique identity across institutions.
Project Start Date	project_start_date	date	3	3	The start date of a project
Project Status	project_status	varchar	50	11	The identification of the status of a project, e.g., active, extended, new, pending, revised, or terminated
Project Title	project_title	varchar	500	252	The title of award
Proposal Number	proposal_number	varchar	100	12	A Proposal No. consists of a four-digit year followed by a hyphen and a five-digit sequence number (e.g., 2004-01478).
Sponsoring Institution	sponsoring_institution	varchar	100	42	The sponsoring institution name, e.g., State Agricultural Experiment Station; Forest Service/USDA; National Institute of Food and Agriculture; Other Cooperating Institutions; Economic Research Service/USDA; Cooperating Schools of Veterinary Medicine; Rural Business-Cooperative Service; Agricultural Research Service/USDA