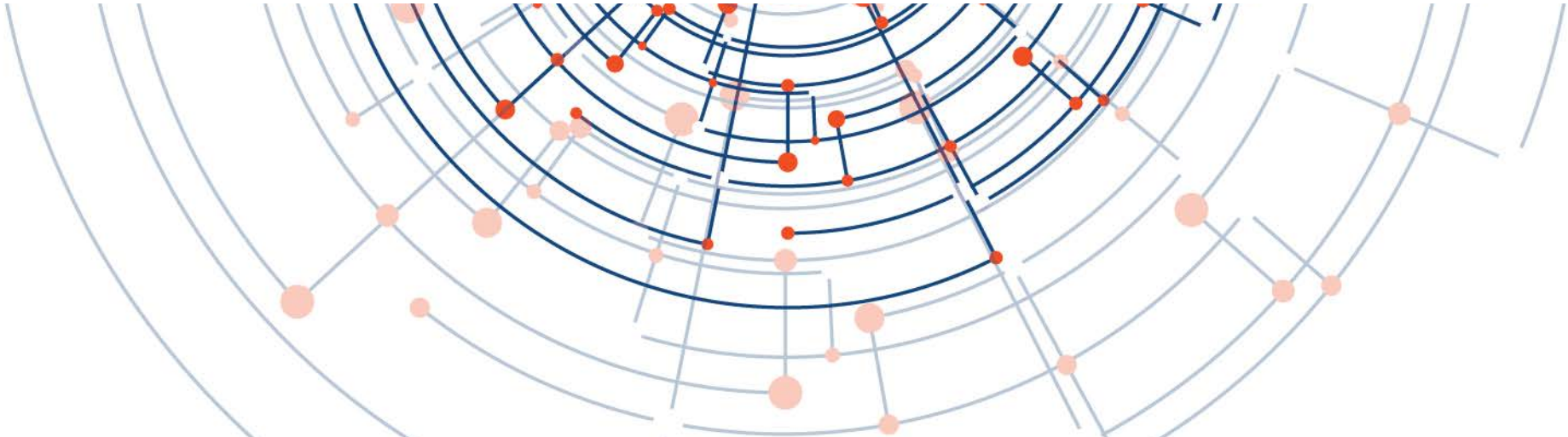
INSTITUTE FOR RESEARCH ON INNOVATION & SCIENCE

Data Submission

File and Field Descriptions

April 2020



Contents

Introduction 9

Data Submission to IRIS 10

IRIS UMETRICS Data Diagram 13

File Descriptions 14

AWARD 14

File Summary 14

Why is this important? 15

AWARD ID 16

File Summary 16

Why is this important? 18

AWARD CO-INVESTIGATORS 20

File Summary 20

Why is this important? 21

EMPLOYEE 22

File Summary 22

Why is this important? 23

EMPLOYEE NAME 24

File Summary 24

Why is this important? 25

VENDOR 26

File Summary 26
Why is this important? 26
SUBAWARD 28
File Summary 28
Why is this important? 28
OBJECT CODE 30
File Summary 30
Why is this important? 30
ORGANIZATION UNIT 31
File Summary 31
Why is this important? 31
Field Details 32
Field Lengths 32
Linking Fields Across Files 32
Definitions for Linked Fields 33
'PeriodStartDate' 33
'PeriodEndDate' 34
'UniqueAwardNumber' 35
'RecipientAccountNumber' 38
'CampusID' 39
'SubOrgUnit' 39
'DeIdentifiedEmployeeIDNumber' 40
'ObjectCode' 41

'InternalAwardId' 42
Field Definitions by File 43
Award File 43
'FundingSource' 43
'AwardTitle' 44
'OverheadCharged' 44
'TotalDirectExpenditures' 45
Award ID File 47
'PrimeAwardId' 47
'PrimeAwardTitle' 47
'PrimeSponsor' 48
'FedFlag' 49
'FAIN' 49
'CFDAcode' 50
'PrimeFlag' 51
'DirectSponsor' 51
'InternalAwardTitle' 52
'Dept' 53
'TotalAwardAmt' 53
'AwardBeginDate' 54
'AwardEndDate' 54
'PrimePIFirstName' 55
'PrimePILastName' 55

'SubPIFirstName' 56
'SubPILastName' 56
'OrcidID' 57
'CoIFlag' 58
Award Co-Investigators File 59
'InternalAwardTitle' 59
'CoFirstName' 59
'CoLastName' 60
'CoNumber' 60
'CoRole' 61
'CoAffiliation' 61
Employee File 63
'OccupationalClassification' 63
'JobTitle' 64
'SOCCode' 64
'FTEStatus' 65
'ProportionofEarningsAllocated' 65
Employee Name File 67
'LastName' 67
'MiddleName' 67
'FirstName' 68
'BirthYear' 69
'BirthMonth' 69

Vendor File 71
'OrganizationID' 71
'EIN' 71
'DUNS' 72
'VendorPaymentAmount' 73
'OrgName' 74
'POBox' 74
'BldgNum' 75
'StrNumber' 75
'StrName' 76
'Address' 77
'City' 77
'State' 78
'DomesticZipCode' 79
'ForeignZipCode' 79
'Country' 80
Subaward File 81
'OrganizationID' 81
'EIN' 81
'DUNS' 82
'SubAwardPaymentAmount' 83
'OrgName' 83
'POBox' 84

'BldgNum' 85
'StrNumber' 85
'StrName' 86
'Address' 87
'City' 87
'State' 88
'DomesticZipCode' 88
'ForeignZipCode' 89
'Country' 90
Object Code File 91
'ObjectCodeText' 91
Organization Unit File 92
'CampusName' 92
'SubOrgUnitName' 92
Appendix A: OFS (Other Funding Sources) Codes 94
Appendix B: Pre-Submission Checklist 98

Tables

Table 1. Sample Award Data 15
Table 2: Sample Award ID Data 17
Table 3: Sample Award Co-Investigators Data 20
Table 4. Sample employee file entries 23
Table 5. Sample employee name file entries 24
Table 6. Sample vendor file entry 27
Table 7. Sample subaward file entries 29
Table 8. Sample object code file entries 30
Table 9. Sample organization unit file entries 31
Table 10. Elements of the 'UniqueAwardNumber' 36
Table 11. Award number format for selected agencies 37
Table 12. Common data issues and how to avoid them 98

Introduction

This document is intended to supplement the IRIS UMETRICS data dictionary, providing additional context, description and examples. The UMETRICS dataset is comprised of data from dozens of research institutions, each with their own human resources, grants management, and procurement systems. In order for IRIS to provide accurate reports and consistent datasets for researchers, data submitted from members must adhere to the definitions in this document and the data dictionary. IRIS has collected some code snippets used by member universities for pulling their data that may be useful, depending on the software used at your institution. Contact IRIS if you believe this code would be helpful. In rare cases, it may not be possible to adhere completely to the structure laid out in this documentation due to how your institution stores research-related data. If this is the case, please reach out to IRIS.

For a pre-submission checklist to maximize the quality and consistency of your data relative to other member universities' submissions, please see Appendix B: Pre-Submission Checklist.

Data Submission to IRIS

Submission Deadlines

April 15
October 15

This document contains a detailed description of each of the nine files that make up the IRIS UMETRICS data submission and their corresponding fields. A standard data submission sent to IRIS will consist of up to nine comma-separated text files:

1. **Award***
2. **Award ID***
3. **Award Co-Investigators***
4. **Employee***
5. **Employee Name***
6. **Subaward***
7. **Vendor***
8. **Object Codes**
9. **Organization Unit**

In a standard submission of data, the first seven UMETRICS files listed above are required (*). These files contain all the data necessary for IRIS to provide the analysis that goes into the reports and feedback used by your university. Additionally, these data support the research projects being conducted by IRIS once the submitted data have been scrubbed by IRIS internal processes. The **Object Code** and **Organization Unit** files should be submitted during the very first transmission of data to IRIS and anytime thereafter if updates need to be made. This will allow for proper matching and understanding of the submitted codes and units.

Where do I submit files?

Files are submitted through the IRIS member portal. Contact IRIS or your university's IRIS portal administrator if you need an account.

When do I submit files?

There are two deadlines for data submission per year—April 15 and October 15. This is done to coincide with IRIS data transmissions to the U.S. Census Bureau. It is recommended that data are submitted prior to the deadline to allow more time for data analysis by IRIS, and for possible resubmission if data anomalies are found.

What time period should each file cover?

The IRIS member portal accepts files of any size that cover any span of time. Some members submit files for each month, others quarterly, and new members often submit multi-year files for their first transmission. However, the majority of our members submit one file of each type per fiscal year.

What should I name files?

IRIS does not require submitted files to follow a particular naming convention—the portal will scan the file and determine what type it is based on its structure. Still, certain simple guidelines are recommended. Information such as the time period covered in the file, campuses covered in the file, or revision numbers (in the case of a resubmission) can be helpful when diagnosing and communicating about data anomalies.

Example: IRISuniversity_vendor_FY19_medcampus_version2.csv

What file type should I use?

Currently, comma-delimited files (*.csv) are required for the portal. If exporting your CSV from Excel, pay extra attention to fields that may contain leading zeros as Excel automatically eliminates these by default. Importing data as text or custom formatting will help avoid this issue. Microsoft provides documentation on how to do so here.

How should I handle values with commas?

If you submit a comma-delimited file, commas that appear in these underlying data need to be escaped so that the portal can properly recognize where fields start and end. We recommend wrapping values containing commas with double quotes (“”) in order to avoid this issue. Values not containing commas can still be wrapped in double quotes without causing issues, so it may be easier to wrap all values in double quotes. Wrapping in double quotes is the default behavior of Excel when exporting to CSV, but other software may need to be adjusted to export in this form.

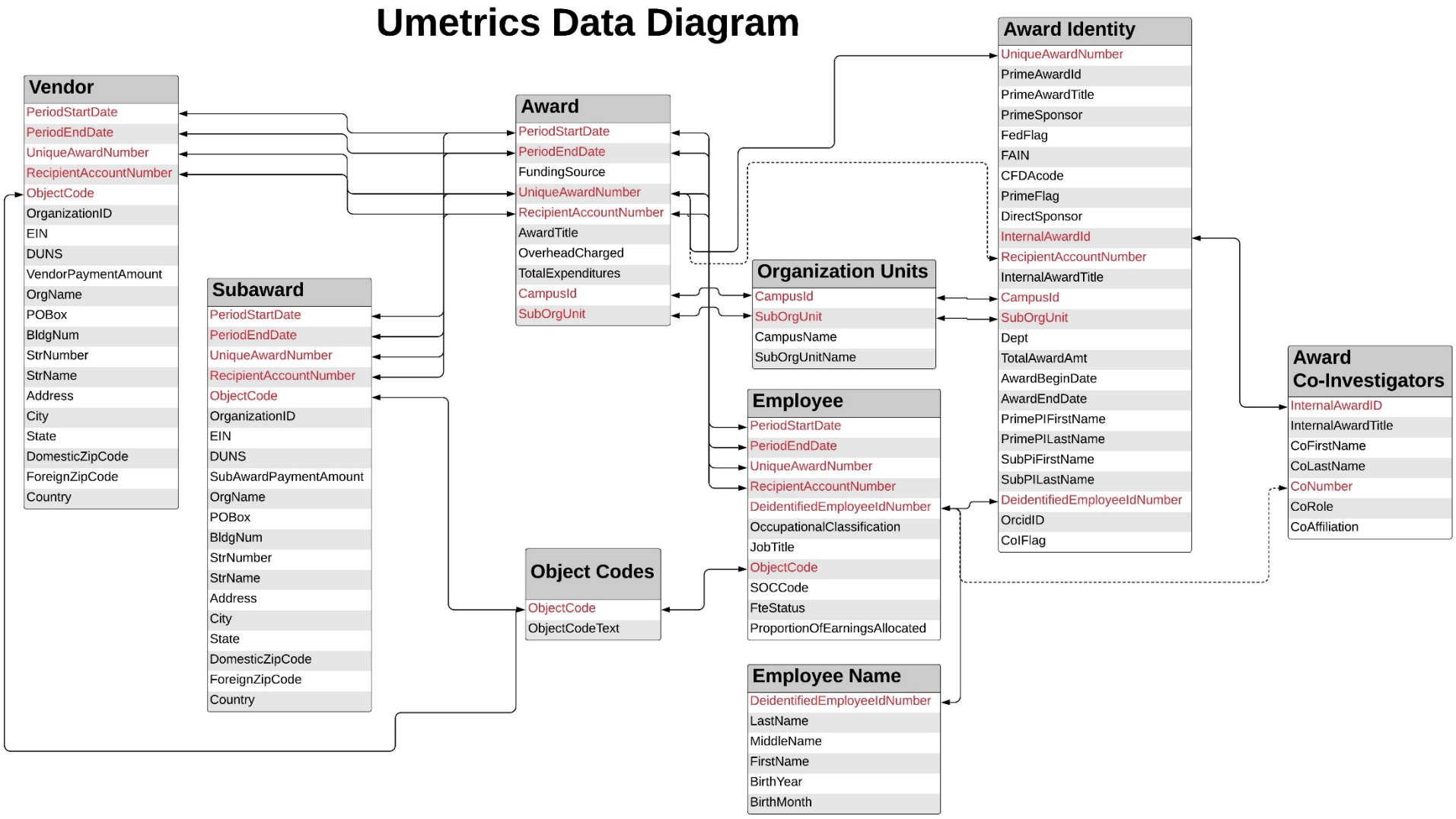
How do I know if my upload is successful?

If a file upload fails, the portal will return the message “Your upload file format does not match the expected file type. Please see sample <FileType> file.” If this happens, first check that your submission file is in a CSV format. If your file is a CSV, check that the file header matches the file headers in the provided sample files. If everything looks correct and the file still won’t upload, please contact IRIS at IRIStechrequests@umich.edu.

If the file format is correct, you will be taken to the IRIS data ingest page. This page provides a variety of statistics on your upload, as well as a description of initial cleaning steps run on the upload. If something appears incorrect, hit the red “cancel” button to clear the upload out of our systems. You can then either resubmit an adjusted file or contact IRIS for more information. If everything is correct, hit the green “yes, proceed!” button to finalize the upload.

IRIS UMETRICS Data Diagram

This diagram shows the connections between files in the UMETRICS dataset. Fields marked in red indicate connecting columns. These must be consistent between files to maximize usability for reporting and research. Connections marked with a dotted line indicate fields that may or may not link two files, depending on the structure of your data.



File Descriptions

AWARD

File Summary

The **Award** file is the centerpiece of the UMETRICS dataset and directly connects to the **Vendor**, **Subaward**, **Employee**, **Organization Units**, and **Award ID** files. The **Award** file should contain transaction data on every awarded (funded) project that has direct or overhead (also commonly referred to as indirect) expenditures during the period of time covered in the file. These transactions are at the month (as indicated in the 'PeriodStartDate' and 'PeriodEndDate' fields) – 'UniqueAwardNumber' – 'RecipientAccountNumber' level. This means that for each month, there should be one record for each combination of 'UniqueAwardNumber' and 'RecipientAccountNumber'. The data in the 'OverheadCharged' and 'TotalDirectExpenditures' fields should be summed monthly for each awarded project that has received funding. This means that for a given funded project you could submit up to twelve separate entries per year for each unique award number-recipient account number pairing. See an example below.

Data Fields

PeriodStartDate
PeriodEndDate
FundingSource
UniqueAwardNumber
RecipientAccountNumber
AwardTitle
OverheadCharged
TotalDirectExpenditures
CampusId
SubOrgUnit

Note that, depending on how your dataset is structured, expenditures to the same award/account number may span multiple suborganizational units or even campuses. If this is the case for your university, this file should be at the 'PeriodStart/EndDate' -- 'UniqueAwardNumber' -- 'RecipientAccountNumber' -- 'CampusID' -- 'SubOrgUnit' level, meaning that for each combination of those fields there should be one value for 'OverheadCharged' and 'TotalDirectExpenditures'. This increased granularity helps with tracking collaboration and spending allocation.

Table 1. Sample Award Data

Period Start Date	Period End Date	Funding Source Name	Unique Award Number	Recipient Account Number	Award Title	Overhead Charged	Total Direct Expenditures	Campus ID	Sub Org Unit
2019-10-01	2019-10-31	National Science Foundation	47.049 1811889	F011324	Computations and Analysis of Efficient Snake Locomotion	1222.31	52478.23	1	1
2019-11-01	2019-11-30	National Science Foundation	47.049 1811889	F011324	Computations and Analysis of Efficient Snake Locomotion	1005.14	45187.26	1	1
2019-12-01	2019-12-31	National Science Foundation	47.049 1811889	F011324	Computations and Analysis of Efficient Snake Locomotion	832.79	23177.84	1	1

As seen in the above example of the content of the **Award** file, this particular award has three entries (rows), one for each of the three months it was active as defined by the period start and period end dates. In each entry (row) the information is only different in the 'PeriodStartDate' and 'PeriodEndDate' fields (indicating the transaction period), and the 'OverheadCharged' and 'TotalDirectExpenditures' fields (indicating expenditures during that month). The latter two fields are where the summing will take place. Everything else stays the same.

Why is this important?

The data in this file typically represent the financial transactional data that have been recorded by the member university. In order to accurately assess the submitted data and generate reports from IRIS that are provided to the contributing university, we need award data that can be broken into monthly amounts. Rolling transactions up monthly also provides consistency between universities for researchers using these data.

AWARD ID

File Summary

The **Award Identity** (**Award ID**) file is used to link your internal award identifiers to external identifiers used by sponsors, providing an expansion of the information available in the **Award** file. It connects directly with the **Award**, **Award Co-Investigator**, **Organization Units**, and **Employee Name** files. You might think of the **Award ID** file as the birth certificate for an award. Whereas the **Award** file contains information about transactions that happen through the lifetime of an award, often pulling data from financial systems at your institution, the **Award ID** file provides more specific details about these awards as they originally came to your institution, whether directly from prime sponsors or as subawards. **Award ID** file information may often be found in pre- or post-award proposal management systems. Note that while expenditures on awards may span multiple campuses or suborganizational units, in this file you should only include the 'CampusID'/'SubOrgUnit'/'Dept' that is directly administering the award. Similarly, while a given award may have multiple associated recipient account numbers, in this file only include a single primary account number. There should only be one record per 'UniqueAwardNumber' in this file.

The **Award ID** file should contain:

1) Values that conform to how funding agencies and other external data sources describe the awards (i.e., external award identifiers), typically found in an awards management system;
2) Values used within/across your university systems to refer to a unique award; and,
3) Additional data fields necessary to tie subawards back to their prime and direct sponsors.

By providing all three sets of information, your dataset creates a crosswalk from your university-specific identifiers to those used outside your university. If your university does not maintain prime award information anywhere in your systems, those fields can be left blank.

Data Fields

UniqueAwardNumber
PrimeAwardId
PrimeAwardTitle
PrimeSponsor
FedFlag
FAIN
CFDAcode
PrimeFlag
DirectSponsor
InternalAwardId
RecipientAccountNumber
InternalAwardTitle
CampusId
SubOrgUnit
Dept
TotalAwardAmt
AwardBeginDate
AwardEndDate
PrimePIFirstName
PrimePILastName
SubPIFirstName
SubPILastName
DeidentifiedEmployeeIdNumber
OrcidID
CoIFlag

The following two pages provide an example of four records in an **Award ID** file, with the file's 25 fields broken up to fit the page.

Table 2: Sample Award ID Data

Unique Award Number	Prime Award Id	Prime Award Title	Prime Sponsor	Fed Flag	FAIN
47.049 1811889	1811889	Computations and Analysis of Efficient Snake Locomotion	National Science Foundation	F	1811889
93.838 AWD005585	1 R01 HL136682-01		Health and Human Services, Department of-National Institutes of Health-Subcontracts	F	R01HL136682
00.630 Award Ltr dtd 1/24/08	Award Letter dated 1/24/08	The Role of Prostaglandin E2 in Mediating Intracellular Signaling Pathways in Neonatal Alveolar Macrophages	The Hartwell Foundation Corporation	N	
10.310 2012-68004-20028	2012-68004-20028	Examining Disparities in Food Access and Enhancing the Food Security of Underserved Populations in Michigan	Agriculture, Department of	F	20126800420028

CFDACode	Prime Flag	Direct Sponsor	Internal Award Id	Recipient Account Number	Internal Award Title
47.049	P	National Science Foundation	AWDF010039	F011324	Computations and Analysis of Efficient Snake Locomotion
93.838	S	Cornell University	AWDF005585	F011110	The CAPTURE Study: Validating a unique COPD screening tool in primary care
00.630	P	The Hartwell Foundation Corporation	AWDNF006464	NF090152	The Role of Prostaglandin E2 in Mediating Intracellular Signaling Pathways in Neonatal Alveolar Macrophages
10.310	P	Agriculture, Department of	AWDF000226	F022222	Examining Disparities in Food Access and Enhancing the Food Security of Underserved Populations in Michigan

Campus Id	Sub Org Unit	Dept	Total Award Amt	Award Begin Date	Award End Date	Prime PI First Name	Prime PI Last Name
1	1	LSA Mathematics	220063	8/15/2018	7/31/2021	Peter	Canisius
1	2	Hlth Behavior & Hlth Ed Dept	279670	9/1/2017	6/30/2021		
1	3	Int Med-Pulm./Critical Care	100000	1/1/2008	12/31/2019	Jenny	Williams
2	4	Sch for Environ and Sustain	3997207	9/1/2012	8/31/2018	Marian	Rosemont

Sub PI First Name	Sub PI Last Name	Deidentified Employee ID Number	Orcid ID	Co I Flag
		123456789	0001-0001-0001-0001	N
Charles	Norris	987654321	0000-0000-0000-0000	N
Thomas	Tutone	8675309	0000-0000-0867-5309	Y
		101010100	0000-0001-5000-0007	Y

Why is this important?

Providing these additional award identifiers allows IRIS to better clean and link award numbers, providing stronger links between UMETRICS files and external data. IRIS can better identify employees and build networks of people who are paid by same awards, identify subaward pass-through dollars on your university vendor and spending reports, and add funding source industry and employer industry to the employee report.

IRIS is also better able to identify federal and nonfederal awards in your data. For example, the new funding type indicator included in the Award ID pilot project allowed us to accurately identify significantly more federally funded awards than we were able to using a method that depended solely on university submitted Catalog of Federal Domestic Assistance (CFDA) numbers.

AWARD CO-INVESTIGATORS

File Summary

This file is used to provide information about the names and roles of co-investigators for each award included in the **Award ID** file. It connects directly with the **Award ID** file. Provide a distinct record for each co-investigator named on each award provided in the **Award ID** file, using the same internal award identifier and title as provided in that file.

Data Fields

InternalAwardID
InternalAwardTitle
CoFIrstName
CoLastName
CoNumber
CoRole
CoAffiliation

Table 3: Sample Award Co-Investigators Data

Internal Award ID	Internal Award Title	Co First Name	Co Last Name	Co Number	Co Affiliate	Co Role
AWDNF006464	The Role of Prostaglandin E2 in Mediating Intracellular Signaling Pathways in Neonatal Alveolar Macrophages	Micah	Silver-Paul	65432198	University of Michigan	Co-Principal Investigator
AWDNF006464	The Role of Prostaglandin E2 in Mediating Intracellular Signaling Pathways in Neonatal Alveolar Macrophages	Rachel	Derringer	OH12345	The Ohio State University	Co-Investigator
AWDF000226	Examining Disparities in Food Access and Enhancing the Food Security of Underserved Populations in Michigan	Jack	Ryan	123123123	University of Michigan	Co-Principal Investigator
AWDF000226	Examining Disparities in Food Access and Enhancing the Food Security of Underserved Populations in Michigan	Louise	Penny	321321321	University of Michigan	Co-Principal Investigator

Why is this important?

Any Principal Investigator (PI) and Co-Principal Investigator (Co-PI) names help to associate the inputs and outputs of each award, for example to link to external datasets that provide information on publications (authors), patents (inventors), etc.

EMPLOYEE

File Summary

The **Employee** file is used to provide information about the individuals working on awards at your university. It connects directly with the **Award**, **Employee Name**, and **Object Code** files. The **Employee** file should contain a record for every employee that received any type of compensation from an award or spent time working on an award. These transactions are at the month (as indicated in the 'PeriodStartDate' and 'PeriodEndDate' fields) – 'UniqueAwardNumber' – 'RecipientAccountNumber' – 'DeidentifiedEmployeeIDNumber' level. The 'ProportionOfEarningsAllocated' field should be calculated at this level, meaning that there should be one value in that field for each combination of month/award/account number/employee. There should also be only one 'FteStatus', 'OccupationalClassification', 'JobTitle', 'ObjectCode', and 'SOCCode' for each combination of month/award/account number/employee.

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
DeidentifiedEmployeeIdNumber
OccupationalClassification
JobTitle
ObjectCode
SOCCode
FteStatus
ProportionOfEarningsAllocated

Each employee should have only one entry per award/account number pairing per month, but an employee could work on several awards concurrently or sequentially. In these cases, the employee could have multiple values for 'OccupationalClassification', 'JobTitle', 'ObjectCode', and 'SOCCode' in a given month, or over time. 'FteStatus' refers to the employee's full time equivalent status at the time of the transaction, so it should be consistent for all awards for a given employee in a given month. It could change over time if, for example, an employee went from full time (1) to working 20 hours a week (.5).

The example below contains two records for one award, and one record for a different award. Notice how the same person (00482199) can appear twice in the same month on different awards. Also note that the same award/account number pairing (10.310 2012-68004-20028/ F022222) appears in multiple records for a given month, one for each employee.

Table 4. Sample employee file entries

Period Start Date	Period End Date	Unique Award Number	Recipient Account Number	Deidentified Employee ID Number	Occupational Classification	Job Title	Object Code	SOC Code	FTE Status	Proportion of Earnings Allocated
2013-10-01	2013-10-31	10.310 2012-68004-20028	F022222	00482199	Faculty	Associate Professor	3848	11-1000	1	.5
2013-10-01	2013-10-30	10.310 2012-68004-20028	F022222	00574321	Graduate Student	Research Analyst	3858	11-1000	.5	1
2013-10-01	2013-10-31	47.049 1811889	F011324	00482199	Faculty	Associate Professor	3848	12-1500	1	.5

Two important factors related to the **Employee** file are that every employee listed is associated with an award (through the 'UniqueAwardNumber' and 'RecipientAccountNumber' fields) in the **Award** file for the associated month, and that every employee listed has a corresponding entry in the **Employee Name** file (through the 'DeidentifiedEmployeeIdNumber' field).

Why is this important?

The employee information is used to track award personnel expenditures, but more importantly, it is used to link employees with other external datasets including those at the U.S. Census Bureau, allowing IRIS to report back to a university aggregated information about the industries in which their former research-funded employees have gone on to work and their average earnings in those industries. De-identified UMETRICS employee information is widely used by researchers from economics, sociology, and network science.

EMPLOYEE NAME

File Summary

The **Employee Name** file is used to match up the de-identified employee data that are associated with award transactions in the **Employee** file with the employee's actual name and date of birth. It connects directly with the **Employee**, **Award ID**, and **Award Co-Investigator** files. This file is integral to the matching process with data sources such as the U.S. Census Bureau and allows IRIS to provide detailed reports.

Data Fields

DeidentifiedEmployeeIdNumber
LastName
MiddleName
FirstName
BirthYear
BirthMonth

An example of a sample dataset for the included fields is shown below.

Table 5. Sample employee name file entries

Deidentified Employee ID Number	Last Name	Middle Name	First Name	Birth Year	Birth Month
123456789	Piper	P	Peter	1945	04
00482199	Xavier	Alan	Charles	1934	12
00574321	Banner	Blane	Bruce	1964	02

Every employee name and date of birth should be associated, via the de-identified employee ID number, with an entry in the **Employee** file. Any transactions in the **Employee** file that does not have corresponding name information in the **Employee Name** file cannot be used to create earnings reports. An employee should have only one entry in this file from a specific institution for a given de-identified employee ID number in order to avoid double counting people. If your university's employee numbers have changed over time, please contact IRIS.

Why is this important?

The employee information is used to link employees with external datasets such as those at the U.S. Census Bureau, allowing IRIS to report on the far-reaching impact of a university's funded research. Having all fields filled is critical to a successful linkage. The name information provides the primary linking criteria, while the month and year of birth are crucial for disambiguating common names.

VENDOR

File Summary

The **Vendor** file should include transaction information on all vendor purchases made on the awards contained in the **Award** file. It connects directly with the **Award** and **Object Code** files. The **Vendor** file should *not* include spending to collaborating institutions through subawards or subcontracts, as this type of expense is included in the **Subaward** file. No individual transaction should be included in both the **Vendor** and **Subaward** files. These transactions must be rolled up to the month (as indicated in the 'PeriodStartDate' and 'PeriodEndDate' fields) – 'UniqueAwardNumber' – 'RecipientAccountNumber' – vendor establishment level. This means the 'VendorPaymentAmount' field should contain the sum of all expenditures with the given vendor establishment in the given month off of the given unique award number/recipient account number. Vendor establishment in this instance refers to a single address of a given vendor. It is important to be as granular as possible when pulling vendor location information, as these data are used for mapping your university's expenditures.

Negative expenditures should be included when calculating 'VendorPaymentAmount'.

Why is this important?

The **Vendor** file allows for the tracking of vendors and their involvement with the funded awards that are reported in the **Award** file. The vendor information is crucial for analysis and reporting on the effectiveness of research investments and on the impact of research dollars spent on the local, regional, and state levels.

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
ObjectCode
OrganizationID
EIN
DUNS
VendorPaymentAmount
OrgName
POBox
BldgNum
StrNumber
StrName
Address
City
State
DomesticZipCode
ForeignZipCode
Country

The following page provides an example of three records in a **Vendor** file, with the file's 20 fields broken up to fit the page.

Table 6. Sample vendor file entry

Period Start Date	Period End Date	Unique Award Number	Recipient Account Number	Object Code	Organization ID
2009-01-01	2009-01-31	00.630 Award Ltr dtd 1/24/08	NF090152	711700	41234
2009-02-01	2009-02-28	00.630 Award Ltr dtd 1/24/08	NF090152	711700	12632
2009-01-01	2009-01-31	93.838 AWDF005585	F011110	712800	962575

EIN	DUNS	Vendor Payment Amount	Org Name	PO Box	Bldg Num
123456789	987654321	626.71	Fisher Scientific		
555555555	998877665	20.19	ITS Rental		
111222333	112233445	64.64	Truex, Robert		

Str Number	Str Name	Address	City	State	Domestic Zip Code	Foreign Zip Code	Country
		7071 N Tower Dr	Traverse	MI	23487		USA
		64 Peach Ln	Ann Arbor	MI	48123		USA
		123 Rue de Paris	Paris			75008	FR

SUBAWARD

File Summary

The **Subaward** file allows for the tracking and analysis of funding sent to subawardees (also commonly referred to as subrecipients) off of the awards reported in the **Award** file. It connects directly to the **Award** and **Object Code** files. The **Subaward** file should include transaction information on all subawards or subcontracts made off the awards contained in the **Award** file. It should *not* include payments to vendors, as those are contained in the **Vendor** file. No individual transaction should be included in both the **Vendor** and **Subaward** files. These transactions must be rolled up to the month (as indicated in the 'PeriodStartDate' and 'PeriodEndDate' fields) – 'UniqueAwardNumber' – 'RecipientAccountNumber' – subawardee level. This means the 'SubawardPaymentAmount' field should contain the sum of all expenditures with the given subawardee in the given month off of the given unique award number/recipient account number. It is important to be as granular as possible when pulling subawardee location information, as these data are used for mapping your university's expenditures.

Negative expenditures should be included when calculating 'SubawardPaymentAmount'.

Why is this important?

The subaward recipient information is crucial for analysis and reporting on the effectiveness of research investments and on the impact of research dollars being spent on the local, regional, and state levels.

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
ObjectCode
OrganizationID
EIN
DUNS
SubAwardPaymentAmount
OrgName
POBox
BldgNum
StrNumber
StrName
Address
City
State
DomesticZipCode
ForeignZipCode
Country

The following page provides an example of three records in a **Subaward** file, with the file's 20 fields broken up to fit the page.

Table 7. Sample subaward file entries

Period Start Date	Period End Date	Unique Award Number	Recipient Account Number	Object Code	Organization ID
2009-01-01	2009-01-31	47.049 1811889	F011324	711600	2462534
2009-02-01	2009-02-28	47.049 1811889	F011324	711600	5476765
2009-01-01	2009-01-31	10.310 2012-68004-20028	F022222	712200	4567565

EIN	DUNS	Subaward Payment Amount	Org Name	PO Box	Bldg Num
123456789	987654321	700.71	University of Michigan		
555555555	998877665	23.19	Ohio State University		
111222333	112233445	42.64	IRIS University		

Str Number	Str Name	Address	City	State	Domestic Zip Code	Foreign Zip Code	Country
		555 Park Place	Traverse	MI	24447		USA
		46 Pear Ln	Ann Arbor	MI	48133		USA
		321 Rue de Paris	Paris			75008	FR

OBJECT CODE

File Summary

This file is used to build an identification file of all the submitted objects so they can be matched up with submitted information in the **Subaward**, **Vendor**, and **Employee** files. This file must contain every object code that has values in any of these other files.

An example of sample data for the included fields is shown below.

Data Fields

ObjectCode

ObjectCodeText

Table 8. Sample object code file entries

Object Code	Object Code Text
711700	Equipment & Servicing
712800	Computer Parts
3848	Faculty Salary

Why is this important?

This information is critical for classifying purchases and employee occupations for use in reporting.

ORGANIZATION UNIT

File Summary

This file is used to build an identification file of all the submitted organizational codes ('CampusID' and 'SubOrgUnit') so they can be matched up with submitted information in the **Award** and **Award ID** files. This file should contain every campus and every suborganizational unit on those campuses that have values in either of those files. You could have many entries for each campus if there are many suborganizational units present on those campuses. An example of sample data for the included fields is shown below.

Data Fields

CampusId
SubOrgUnit
CampusName
SubOrgUnitName

Table 9. Sample organization unit file entries

Campus ID	Sub Org Unit	Campus Name	Sub Org Unit Name
1	1	University of Michigan – Ann Arbor	School of Arts and Sciences
1	2	University of Michigan – Ann Arbor	School of Public Health
2	4	University of Michigan – Flint	College of Physical Sciences

Why is this important?

This information will help IRIS perform analysis and reporting based on data organized around campuses and the schools or colleges within those campuses.

Field Details

Field Lengths

Why is my field length so different from the ones suggested?

The field lengths for the listed string fields are typically oversized in order to accommodate any potential variations in the length of the reported data. Just because a field length is considerably longer than the data you are supplying doesn't mean that you are missing anything, it just means that the field size is capable of holding longer data.

Linking Fields Across Files

In the UMETRICS dataset, there are many fields that share a common name and definition between files, and are used to link two or more files in the same way that a primary key and foreign key connect files in a database. In order to reduce redundancy, these fields will be defined once in this section rather than having the definition repeated for each file in which they appear. Linking fields have their names marked in red.

Fields that are not used to link files are provided in the sections that follow, organized by the file in which they appear.

Definitions for Linked Fields

'PeriodStartDate'

Full Name: Period Start Date

Matching Key For: Award File, Employee File, Vendor File, Subaward File

Description: The Period Start Date should be the first day of the month that a given transaction is reporting on. While many records may be stored in your system for a given month, in the IRIS submission files they must be rolled up monthly with the first day of the month indicated by the 'PeriodStartDate' field. For more information on how a transaction is defined for each file, see the file-level descriptions for the **Award**, **Employee**, **Vendor**, and **Subaward** files. A 'PeriodStartDate' value must be included with all transactions.

Type: Date

Code/Format: YEAR, MONTH, DAY (YYYY-MM-DD)
All dates must appear in the following manner: DD must be the number of a day between 01 and 31. YYYY represents the year. Each MM must be one of the following:

01 January	05 May	09 September
02 February	06 June	10 October
03 March	07 July	11 November
04 April	08 August	12 December

Purpose: The Period Start Date is used to join the **Award**, **Employee**, **Subaward**, and **Vendor** files so that award activity can be tracked on a monthly basis. For the **Award** file, this field defines the monthly start date of the awarded project that has received funding, which helps IRIS create calendar year analysis of funding activity. In the **Employee** file, this field defines the starting period when an employee works on a specified award, and the period start date is used in conjunction with the period end date to allow for the analysis of working time spent on a given award. In the **Vendor** and **Subaward** files, this field provides the means to track, compare, and provide analysis of transactional data for given awards.

Examples: The Period Start Date should always be the first day of a month (e.g. January 1st, February 1st, March 1st, etc.) since the data in the file should be summed (rolled up) on a monthly basis. For example, if an award has funding for three months from January to March then there should be three separate rows of data reported each with a period start date for each of the three months:

Row ID	PERIOD START DATE
Row1	2015-01-01
Row2	2015-02-01
Row3	2015-03-01

'PeriodEndDate'

Full Name: Period End Date

Matching Key For: Award File, Employee File, Vendor File, Subaward File

Description: The Period End Date should be the last day of the month that a given transaction is reporting on. While many records may be stored in your system for a given month, in the IRIS submission files they must be rolled up monthly with the last day of the month indicated by the 'PeriodEndDate' field. For more information on how a transaction is defined for each file, see the file level descriptions for the **Award**, **Employee**, **Vendor**, and **Subaward** files. A 'PeriodEndDate' value must be included with all transactions. It must be in the same month as the corresponding 'PeriodStartDate' value.

Type: Date

Code/Format: YEAR, MONTH, DAY (YYYY-MM-DD)
As with 'PeriodStartDate', all dates must appear in the following manner: DD must be the number of a day between 01 and 31. YYYY represents the year. MM corresponds to the month (01=January, 12=December).

Purpose: The Period End Date is used to join the **Award**, **Employee**, **Subaward**, and **Vendor** files so that award activity can be tracked on a monthly basis. In the **Award** file, along with the Period Start Date, this field helps IRIS provide analysis on a calendar year basis. For the **Employee** file, the period end date is used in conjunction with the period start date to allow for the analysis of working time spent on a given award. In the **Vendor** and **Subaward** files, this field provides the means to track, compare, and provide analysis of transactional data for given awards.

Examples: If an award has funding for three months from January to March then there should be three separate rows of data reported each with a period end date for each of the three months, for example:

Row ID	PERIOD END DATE
Row1	2015-01-31
Row2	2015-02-28
Row3	2015-03-31

'UniqueAwardNumber'

Full Name: Unique Award Number

Matching Key For: Award File, Award ID file, Employee File, Vendor File, Subaward File

Description: Identifier specifying an award and its funding source. The contents of the unique award number field will differ based on the origin of the award. The main groups for determining the structure of the unique award number are Federal, Federal Pass-Through, and Nonfederal. A Federal award is an award that your university received directly from a Federal Agency. A Federal Pass-Through award is an award that a different institution received from a Federal Agency and then subawarded a portion of that award to your university. A Nonfederal award is an award that originates from any source other than a Federal Agency. Depending on the award origin, the unique award number is defined by concatenating the 6-position funding source code—either the CFDA code or a STAR Other Funding Source (OFS) code—with an award identifier—either the federal award ID from the awarding Federal Agency (such as the federal grant number, federal contract number, or the federal loan number) or an internal award ID—with a space in between the two numbers. The file below describes how the unique award number should be built based on award origin.

Table 10. Elements of the 'UniqueAwardNumber'

	Funding source code	Award identifier
Federal award example	CFDA code	Federal award ID from the awarding federal agency (e.g., federal grant number, federal contract number, federal loan number)
Federal Pass-Through award example	CFDA code of originating federal agency	Federal award ID from the awarding federal agency (e.g., federal grant number, federal contract number, federal loan number). If this is not stored in your system, use the internal award ID.
Nonfederal award example	STAR Other Funding Source (OFS) code	Internal award ID

The CFDA code used should be associated with the primary sponsor even if the award is a subcontract and the direct sponsor name is included in the funding source field. If a 5-digit CFDA code is not available, use two digits (representing the funding agency) followed by '.000'.

For nonfederal awards, OFS (Other Funding Sources) codes must be used rather than leaving the field blank. (Please review Appendix A: OFS Codes at the end of this document.)

If an internal award ID is used, it is critical that the number is unique so that associated records in other files match correctly to the award.

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

[Funding source] [Award identifier]
XX.XXX Award identifier

Purpose: This field is used to link your submitted files. Consistency in format and content in this field across submitted files is critical.

Examples: 10.310 2010-12345-54321
47.050 1234567
93.865 2-R01-DK-012345-15-S1
00.200 State Award 1

What should a federal award ID look like?

Each federal agency has its own name and type of award / grant number. This number needs to be formatted as accurately as possible in the "unique award number" field. This field may include spaces and/or dashes. Below is the award format for selected major federal agencies.

Table 11. Award number format for selected agencies

Agency Name	Unique Identifier	Description	What it looks like
HHS / NIH	Full Project Number	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	1 R01 CA 654321 01 A1
NSF	Award ID	The agency assigned 7-digit award number	1234567
NASA	Project Number	This unique identification number is composed of: three alpha letters, two digit number, two alpha letters, two-digit number, and one alpha letter.	NNX10AB62G

'RecipientAccountNumber'

Full Name: Recipient Account Number

Matching Key For: Award File, Award ID File, Employee File, Vendor File, Subaward File

Description: Recipient account numbers are typically internal accounting codes used to allocate funds received from an award. Depending on the records management systems of your university, there may be one recipient account number per award (a "parent" account) or multiple account numbers per award ("child" accounts). Each account number should only be tied to a single award, however. This is a linking field for all four transaction files (**Award**, **Employee**, **Subaward**, and **Vendor**). As such, it is important to use codes that appear in all of the systems used to pull data for those files. If your university's accounting systems do not have a code that bridges these systems, provide a single "parent" account number per award across all of the files.

Note that this field is slightly different in the **Award ID** file. Whether or not you provide "child" accounts in the transaction files, the **Award ID** file should always contain the "parent" account number held in your grants management system. This is done to keep the **Award ID** file at the "Award" level, rather than the "Award" – "Account Number" level.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field is used as a linking element along with the Unique Award Number. For example, an institution's procurement system may have purchases linked to a recipient account number, and those recipient account numbers are linked to the award number found in an award management system. Providing both numbers from your internal systems helps to identify and link the different types of accounts associated with an award across university systems and across data files.

Examples: FS111222555
1234567890

'CampusID'

Full Name: Campus ID

Matching Key For: Award File, Award ID File, Organization Unit File

Description: The campus to which each award is assigned. Campus ID should be defined in the **Organization Unit** ancillary file. For example, if the Campus ID is 11 it is expected that a matching Campus ID of 11 will be found in the **Organization Unit** file complete with a matching 'CampusName'.

Type: String

Length: 30 Characters

Code/Format: No required format

Purpose: Provides refinement of reported data allowing IRIS to provide reports based on specific campuses for a given university. This is particularly important for university systems that have many different campuses all under a single university umbrella.

Examples: 11
UMF
UMA
UMD

'SubOrgUnit'

Full Name: Suborganizational Unit

Matching Key For: Award File, Award ID File, Organization Unit File

Description: The suborganizational unit of the university campus to which each project is assigned. This unit should be at the college or school level, not at the level of individual departments. (Department should be included in the 'Dept' field of the **Award ID** file.) 'SubOrgUnit' should be defined in the ancillary file, **Organization Unit**. For example, if the 'SubOrgUnit' is 15 it is expected that a matching 'SubOrgUnit' of 15

will be found in the **Organization Unit** file complete with a matching 'SubOrgUnitName'. 'SubOrgUnit' can also encompass research organizations that are not classified as colleges or schools, but are larger than departments and smaller than campuses. For example, the Institute for Social Research at the University of Michigan.

Type: String

Length: 30 Characters

Code/Format: No required format

Purpose: Provides refinement of reported data allowing IRIS to provide reports based on specific suborganizational units within a given university. This is particularly important for if contributing universities would like to examine research investments for specific subunits of a university (e.g., within the college of natural sciences, the medical school, or the college of engineering).

Examples: 15
CPS
CWL

'DeIdentifiedEmployeeIDNumber'

Full Name: De-Identified Employee ID Number

Matching Key For: Employee File, Employee Name File, Award ID File (Partial), Award Co-Investigator File (Partial)

Description:
Unique employee ID (*not* Social Security Number) of grant funded personnel. Typically, it is the internal institutional number that a university assigns to that employee — for example, the University of Michigan 'M-Number' that appears on university-issued identification cards. This field does not have a required structure other than being 50 characters or less in length.

It is important that the value used for this field has an identical and matching value in the **Employee Name** file in order to match the data between the two data files. It is also critical that the format of this field not change over time, as that would result in a single person being identified as two or more people.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field is needed in order to match employees up with their details found in the **Employee Name** file.

Examples: E998811
A00811542

'ObjectCode'

Full Name: Object Code

Matching Key For: Employee File, Vendor File, Subaward File, Object Code File

Description: Internal organization object code or other expense-type category assigned to transaction. This field is the contributing university's resident object code used for accounting purposes. This is a purely internal categorical field and will vary from university to university. Any data submitted in this field should have a corresponding entry in the **Object Code** file.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: It allows us to link **Vendor**, **Subaward**, and **Employee** transactions to the 'ObjectCodeText' field. That field allows us to aggregate types of expenses and categorize types of employees by the object code used for their earnings.

Examples: B3848.001
A5584

'InternalAwardId'

Full Name: Internal Award ID

Matching Key For: Award ID File, Award Co-Investigators File

Description: Award identifier used across systems in your university to identify a particular award and its activities.

Type: String

Length: 100

Code/Format: Text field using any characters.

Purpose: To help universities link awards between award systems and financial systems.

Examples: 111240587

Field Definitions by File

This section contains information on each non-linking field, organized by file. Most of these fields are unique to their given file. Some, most notably in the **Vendor** and **Subaward** files, share a common name and definition across files despite not serving as linking fields. The complete list of data fields for each file, including linking fields, are provided to serve as reference.

Award File

'FundingSource'

Full Name: Funding Source

Description: The name of the funding source assigned to each project – maps directly to funding (CFDA) code when the name is a federal sponsor. If the award is a subaward from another institution, this field will not map to the associated CFDA code (which remains the CFDA code of the federal sponsor) and should instead be populated with the name of the direct sponsor (the contractor). If it is a nonfederal source, include the name of the foundation or organization.

Type: String

Length: 200 Characters

Code/Format: Text field using any characters.

Purpose: This field indicates where the money for this award is originating, allowing IRIS to better categorize funders, in particular pass-through entities for subawarded funds as well as nonfederal funders. The name of the funding organization should go in this field for reporting purposes when grouping or searching for specific types of funders (e.g. foundations, companies, specific federal agencies, etc.).

Data Fields

PeriodStartDate
PeriodEndDate
FundingSource
UniqueAwardNumber
RecipientAccountNumber
AwardTitle
OverheadCharged
TotalDirectExpenditures
CampusId
SubOrgUnit

Examples: Packard Foundation
Institute for Research on Innovation and Science
Advanced Idea Mechanics

'AwardTitle'

Full Name: Award Title

Description: The complete title of the award. This should be pulled from your system based on the time of the transaction being reported (e.g., Study X - year 5). By contrast, the 'PrimeAwardTitle' found in the **Award ID** file will be the title as used with the original sponsoring agency, which may or may not differ from this Award Title field.

Type: String

Length: 500 Characters

Code/Format: Text field using any characters.

Purpose: Providing a complete award title gives IRIS the ability to verify matches to abstract data and allows for searching the dataset by keywords.

Examples: Collaborative Research: Empirical Analyses of Committee Voting

'OverheadCharged'

Full Name: Overhead Charged

Description: The 'OverheadCharged' is a monthly sum of overhead for each unique award number/account number combination. Depending on your institution, these expenditures may also be referred to as "indirect cost" or "F&A (facilities and administrative) cost". This includes any expense not specified by a grant but necessary for research operations to take place, such as heating, electricity, custodial services, and accounting. If you have any questions about whether a particular expenditure belongs in 'TotalDirectExpenditures' or 'OverheadCharged', please

contact IRIS. The value of the field should be a sum of dollar values from the 1st of a given month until the end of a given month (e.g. June 1st to June 30th, July 1st to July 31st). Do not include commas in the dollar value, but always include any of the decimal values (cents). Negative values can be indicated either with a preceding hyphen ("-30.21") or parentheses ("(30.21)").

A value should be generated if there were any transactional overhead charges for a given month.

Type: Number

Code/Format: Dollar value - numeric without $ sign or commas

Purpose: This field provides a monthly sum of charges to a specific award.

Examples: 128544.13
52711.00
699800.54

'TotalDirectExpenditures'

Full Name: Total Direct Expenditures

Description: The 'TotalDirectExpenditures' is a monthly sum of all non-overhead spending for each unique award number – account number combination. If you have any questions about whether a particular expenditure belongs in 'TotalDirectExpenditures' or 'OverheadCharged', please contact IRIS. Note that while we do not ask for salary expenditures directly in any file, salary expenditures are included in this sum. The value of the field should be a sum of dollar values from the 1st of a given month until the end of a given month (e.g. June 1st to June 30th, July 1st to July 31st). Do not include commas in the dollar value, but always include any of the decimal values (cents). Negative values can be indicated either with a preceding hyphen ("-30.21") or parentheses ("(30.21)").

A value should be generated if there were any transactional expenditures for a given month.

Type: Number

Code/Format: Dollar value - numeric without $ sign or commas

Purpose: All vendor payments, salaries, and other direct expenditures should go into this monthly sum of total direct expenditures to the specified award as defined by the unique award number. Providing the total direct expenditures charged to the award in the specified period allows us to know the overall amount spent by the research team for that month, which gives context and can help identify or correct data anomalies found in individual transaction records.

Examples: 128544.13
52711.00
699800.54

Award ID File

'PrimeAwardId'

Full Name: Prime Award ID

Description: The number or text representing the unique identifier assigned by the Prime Sponsor to a commitment of funds awarded for a specified time period for an awarded proposal. Please note that this field does NOT include CFDA. This award ID should be exactly as specified by the sponsor. So, for example, if the National Science Foundation funded an award for the University of Michigan, and the University of Michigan subcontracted your institution on that award, this field would contain the original NSF award number.

For HHS awards, this should be core project number, not full project number.

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: This field provides an identifier exactly as specified by the sponsor so that awards can be linked to sponsors and to outside datasets.

Examples: HHS example: R01 CA 654321
NSF example: 1234567
USDA example: 2004-45066-03027

'PrimeAwardTitle'

Full Name: Prime Award Title

Description: Formal name of the award as used with the sponsoring agency.

Data Fields

UniqueAwardNumber
PrimeAwardId
PrimeAwardTitle
PrimeSponsor
FedFlag
FAIN
CFDAcode
PrimeFlag
DirectSponsor
InternalAwardId
RecipientAccountNumber
InternalAwardTitle
CampusId
SubOrgUnit
Dept
TotalAwardAmt
AwardBeginDate
AwardEndDate
PrimePIFirstName
PrimePILastName
SubPIFirstName
SubPILastName
DeidentifiedEmployeeIdNumber
OrcidID
CoIFlag

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: This field provides a title exactly as specified by the sponsor so that awards can be linked to sponsors and to outside datasets. It helps to link the university-provided award data to federally funded awards through an external award title (full title) especially if the award is a multi-year award with different titles.

Examples: Effect of Gamma Rays on Man-in-the-Moon Marigolds

'PrimeSponsor'

Full Name: Prime Sponsor

Description: Name of the prime sponsoring federal agency, foundation, or private company originating this award. This should always be the original or "prime" funding agency. In the case of subawards it is the ultimate source of the funding, not the pass-through organization(s) that funded the subaward. So, for example, if the National Science Foundation funded an award for the University of Michigan, and the University of Michigan subcontracted funding to your institution, this field would contain "National Science Foundation" rather than "University of Michigan".

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: This field identifies the originating funders of subawards in addition to pass-through entities from which subawards are most directly transferred. The additional information on prime sponsors enables IRIS to better trace how much money is transferred through pass-through entities.

Examples: National Science Foundation
National Institutes of Health

'FedFlag'

Full Name: Federal/Nonfederal Indicator Flag

Description: An indicator flag of whether the award was funded by a federal agency (F) or a nonfederal agency (N). In the case of federal awards that the university receives as subawards from pass through organizations, the flag should reflect the nonfederal pass-through organization.

Type: String

Length: 1 character

Code/Format: F= Federal
N=Nonfederal

Purpose: This categorization helps distinguish federal and nonfederal awards.

Examples: F
N

'FAIN'

Full Name: Federal Award Identification Number

Description: The Federal Award Identification Number (FAIN) assigned by sponsoring federal agency. For federal awards only; nonfederal awards should be null. The FAIN is now mandated by the Federal Government for federal awards active in 2013 and later. For prime federal awards active in 2013 or later, this will be the same as External Award Identifier.

A FAIN is typically letters and digits. For example, NIH derives the FAIN from the core elements of the grant number. That is, the FAIN for 1R0lGM654321-01 would be R01GM654321 (core project number).

Type: String

Length: 100 characters

Code/Format: Differs by sponsoring federal agency.

Purpose: This field helps to maintain consistency in award numbers across IRIS member universities in terms of formatting, and to find accurate records in databases such as USASpending.gov and FSRS.

Examples: P375A081922

'CFDAcode'

Full Name: Catalog of Federal Domestic Assistance (CFDA) Code

Description: For federal awards, this Catalog of Federal Domestic Assistance (CFDA) code is assigned to the program by the sponsoring agency. Awards from nonfederal sources or for which your university is the subawardee should follow the Other Funding Source (OFS) coding convention as described in Appendix A. An official list of CFDA codes (now called Assistance Listings) can be found online at https://beta.sam.gov/help/assistance-listing

Type: String

Length: 50 Characters

Code/Format: XX.XXX

Purpose: The CFDA code helps to identify federal funding sources as well as agencies that administer Federal domestic assistance programs. With this information, IRIS can report on the federal programs, projects, services and activities that provide financial assistance to a given IRIS member university.

Examples: 93.001
47.123

'PrimeFlag'

Full Name: Prime Award Flag

Description: An indicator flag for whether the award is a prime award (P=prime) funded directly by a federal agency, foundation, or other source that is not a pass-through organization. In the case of a pass-through subaward, the award should be flagged as a subaward (S=subaward).

Type: String

Length: 1 character

Code/Format: P=Prime Award
S= Subaward

Purpose: The distinction between prime awards and subawards helps to: 1) better identify unique award counts in university reporting and for research purposes; 2) trace how many prime awards and how much research expenses out of prime awards are directed towards subawards; 3) characterize research collaborations across universities through subawards.

Examples: P
S

'DirectSponsor'

Full Name: Direct Sponsor

Description: If an award is a subaward, this field includes the name of the pass-through entity (e.g., university, foundation or private company) that subcontracted to your university. So, for example, if the National Science Foundation funded an award for the University of Michigan, and the University of Michigan subcontracted funding to your institution, this field would contain the value "University of Michigan".

To be used when Primeflag=S; otherwise, NA.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: This field helps minimize inconsistency in the way the Funding Source Name field in the Award field is used, and provides better information on direct sponsors regardless of whether awards are federal, nonfederal, prime awards, or subawards.

Examples: Nova University
NA

'InternalAwardTitle'

Full Name: Internal Award Title

Description: The award name used in your grants management system. If the award was a subcontract from another university, this could be a very different title than the External Award Title. Keep university-specific information in this field, including any additional identifiers commonly used at each university.

Type: String

Length: 100

Code/Format: Text field using any characters.

Purpose: It helps to avoid truncation of external award titles (full titles) if any internally used titles or internally used award-related information are entered in this field, separate from the external award title.

Examples: Gamma Ray measurement

'Dept'

Full Name: Department

Description: Department administering the award for your university.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: This will help to increase accuracy in recording the award-administering departments for each award.

Examples: Physics Department
Sociology
Nuclear Engineering

'TotalAwardAmt'

Full Name: Total Award Amount

Description: Total estimated awarded amount across all years of this award. For multi-year awards, this should be the total amount across all years that is or was expected to be received by your university. At different institutions this may be called "total expected amount", "awarded amount", or "anticipated funding".

Type: Number

Code/Format: Dollar value - numeric without $ sign or commas

Purpose: This helps to better characterize award totals in terms of value of awards (in addition to number of awards as well as provide a benchmark and/or reference point in university reporting in comparison to actual spending records available from our current UMETRICS dataset)

Examples: 1222.31

'AwardBeginDate'

Full Name: Award Begin Date

Description: First day of sponsored funding period of this award across all years.

Type: Date

Code/Format: YEAR, MONTH, DAY (YYYY-MM-DD)
All dates must appear in the following manner: DD must be the number of a day between 01 and 31. YYYY represents the year. MM corresponds to the month (01=January, 12=December).

Purpose: This field helps to more accurately characterize awards and match to federal awards for additional information based on date and fiscal year information. Since the transactional files contain temporal information directly associated with transaction records, there is often a gap between dates of payments and award data on project begin/end dates and budget begin/end dates.

Examples: 2013-10-01

'AwardEndDate'

Full Name: Award End Date

Description: Last day of sponsored funding period of this award across all years.

Type: Date

Code/Format: YEAR, MONTH, DAY (YYYY-MM-DD)
All dates must appear in the following manner: DD must be the number of a day between 01 and 31. YYYY represents the year. MM corresponds to the month (01=January, 12=December).

Purpose: As with the AwardStartDate field, this field helps to more accurately characterize awards and match to federal awards for additional information based on date and fiscal year information.

Examples: 2019-09-01

'PrimePIFirstName'

Full Name: Prime PI First Name

Description: The first name of the Contact PI or Project Director on the original prime award. Leave blank (null) if not available.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Providing PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Reed
Ann

'PrimePILastName'

Full Name: Prime PI Last Name

Description: The last name of the Contact PI or Project Director on the original prime award. Leave blank (null) if not available.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Providing PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Richards
Schatz

'SubPIFirstName'

Full Name: Sub PI First Name

Description: The first name of the principal investigator at your university for awards received from prime awardees. Leave blank (null) if not available or if the award is a prime award.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Providing PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Anthony

'SubPILastName'

Full Name: Sub PI Last Name

Description: The last name of the principal investigator at your university for awards received from prime awardees. Leave blank (null) if not available or if the award is a prime award.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Providing PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Stark

'OrcidID'

Full Name: ORCID iD

Description: The ORCID iD of the award PI. An ORCID iD is a nonproprietary alphanumeric code used by the nonprofit organization ORCID to uniquely identify scientific and other academic authors and contributors.

If the award is a prime award, then the ORCID iD of the prime award PI. If the award is a subaward then the ORCID iD of the subaward PI at your university.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Providing an ORCID iD helps to track individuals who move across institutions over years.

Examples: 0000-0001-5000-0007

'CoIFlag'

Full Name: Co-Investigator Flag

Description: This field indicates whether or not there are co-investigators listed on the associated award. If it is flagged 'Y', the associated co-investigators should be included in the **Award Co-Investigators** file.

Type: String

Length: 1 Character

Code/Format: Must be 'Y' if there are associated co-investigators on the associated award, or 'N' if there are no associated co-investigators on an award.

Purpose: This provides validation for the **Award Co-Investigators** file.

Examples: N

Y

Award Co-Investigators File

Data Fields
InternalAwardID
InternalAwardTitle
CoFIrstName
CoLastName
CoNumber
CoRole
CoAffiliation

'InternalAwardTitle'

Full Name: Internal Award Title

Description: The award name used in your grants management system. If the award was a subcontract from another university, this could be a very different title than the External Award Title. Keep university-specific information in this field, including any additional identifiers commonly used at each university.

Type: String

Length: 100

Code/Format: Text field using any characters.

Purpose: It helps to avoid truncation of external award titles (full titles) if any internally used titles or internally used award-related information are entered in this field, separate from the external award title.

Examples: Gamma Ray measurement

'CoFirstName'

Full Name: Co-Investigator First Name

Description: First name of this Co-Investigator

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Any Co-PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Alex

'CoLastName'

Full Name: Co-Investigator Last Name

Description: Last name of this Co-Investigator

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: Any Co-PI names will help to associate input and output of each award, especially in linking to publication authors, patent inventors, etc.

Examples: Cortez

'CoNumber'

Full Name: Co-Investigator De-Identified Employee ID Number

Description: The unique Employee ID (not Social Security Number) of this investigator. This number should match the DeidentifiedEmployeeIdNumber found in the **Employee Name** file if a co-investigator is from your own institution. If the co-investigator is from another institution, but there is an identifier in your university's system for this non-employee, please provide it. Otherwise this field should be null for co-investigators who are not employees at your institution.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: This field can be used to match employees up with their details found in the **Employee Name** file, in cases where the investigator is at the same institution.

Examples: G911123

'CoRole'

Full Name: Co-Investigator Role

Description: The role of this person on this award.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: To describe the status of this person on this particular award. Are they the lead author/ Contact Principal Investigator, Co-Principal Investigator, or Co-Investigator? Some systems have these field values predefined in their award systems.

Examples: Contact Principal Investigator
Co-Principal Investigator
Participating Investigator

'CoAffiliation'

Full Name: Co-Investigator Affiliation

Description: This is the university with which the Co-Investigator is affiliated.

Type: String

Length: 100 Characters

Code/Format: No required format

Purpose: To assist in the identification of CoPIs with common names and to identify affiliate universities when collaborations are across more than one university.

Examples: University of Michigan
New York University

Employee File

'OccupationalClassification'

Full Name: Occupational Classification

Description: The occupation title (original text from HR system) of an employee. There are no defined occupation titles that are required. We recognize that every institution uses their own naming conventions and classification system, so this field is very open-ended.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
DeidentifiedEmployeeIdNumber
OccupationalClassification
JobTitle
ObjectCode
SOCCode
FteStatus
ProportionOfEarningsAllocated

Purpose: Providing the occupation title for the employee assists IRIS in classifying people using the same classifications that the university is using. If a university has their own classification system, we use that in our IRIS algorithm to guide how people are classified into the specific categories presented in IRIS university reports. Job titles can be ambiguous, so internal classifications can help clarify what type of work the person is actually doing.

Examples:
Faculty
Technician/ Staff Scientist
Research Analyst / Coordinator
Post Graduate Research
Undergraduate Students
Graduate Students
Clinician
Research Support

'JobTitle'

Full Name: Job Title

Description: Job title (original text from HR system) for an employee. This is the specific job title assigned to the employee used in your university's HR system for a particular position.

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: The purpose is to provide detailed information about the positions conducting research work. Additionally, this field is to provide guidance in the classification process used by IRIS to ensure that people are categorized appropriately into occupational groups.

Examples: Research Analyst
Lab Assistant
Senior Zamboni Administrator

'SOCCode'

Full Name: Standard Occupational Classification Code

Description: Standard Occupational Classification codes that are required for federal agency reporting. http://www.bls.gov/soc/

Type: String

Length: 30 Characters

Code/Format: Two digits, a hyphen, and four more digits

Purpose: Provides an alternative bucketing tool for analysis and can be matched with federal reporting data.

Examples: 11-1000
12-1000

'FTEStatus'

Full Name: Full Time Employment Status

Description: Designation of the status (percent) of the funded personnel (full time = 1.0, half time = 0.5) during the given reporting period. This describes whether an employee is a full time or part time employee of the institution. This is not specific to an award, this is specific to an employee, so the FTE status is the employee's status at the institution, not just their status on the defined award. Keep in mind that FTE is a sliding scale between 1 (full time) and 0 (not employed). An employee FTE status can fall between either of those two end points. A temporary student employee may have an FTE status of 0.1, for example. Typically, an employee that works 40+ hours per week is considered full time with an FTE of 1 (100% Full Time Equivalent). Someone who works 15 hours per week would have an FTE of .375 (37.5% Full Time Equivalent).

Type: Number

Code/Format: Text field using any characters.

Purpose: FTE Status allows us to distinguish full time, part time, and temporary employees.

Examples: 0.5
1
0.75

'ProportionofEarningsAllocated'

Full Name: Proportion of Earnings Allocated

Description: Calculated portion of earnings charged by funded personnel to the award in the specified period. This field is a simple calculation:

$$Proportion\ of\ Earnings\ Allocated = \frac{Earnings\ from\ Specified\ Award}{Total\ Earnings\ for\ Employee}$$

The 'Total Earnings for Employee' is the total gross amount that an employee earned in a given transactional period.

The 'Earnings from Specified Award' are those earnings that came exclusively from work or association with the given award for the given employee in the given time period.

Remember that an employee can work on more than one award, so it is possible that an employee could have multiple entries in the **Employee** file during a given time period. In that case, each one would be for a specific award/account number, and have its own separate Proportion of Earnings Allocated.

Type: Number

Code/Format: Text field using any characters.

Purpose: This allows us to determine how much of an employee's salary is derived from the specified award.

Examples: 0.33
0.50
0.75

Employee Name File

Data Fields

DeidentifiedEmployeeIdNumber
LastName
MiddleName
FirstName
BirthYear
BirthMonth

'LastName'

Full Name: Last Name

Description: This is the employee's last name. The most up-to-date name should be used in instances where an employee has changed names over time.

Type: String

Length: 50

Code/Format: Text field using any characters.

Purpose: This helps with the matching of data to other external datasets that are used for report generation and research.

Examples: Smith
Truex
Bjorne

'MiddleName'

Full Name: Middle Name

Description: This is the employee's middle name. If the full middle name is not available, middle initial can be put in this field. The most up-to-date name should be used in instances where an employee has changed names over time.

Type: String

Length: 50

Code/Format: Text field using any characters.

Purpose: This helps with the matching of data to other external datasets that are used for report generation and research.

Examples: Gregory
G
Alan

'FirstName'

Full Name: First Name

Description: This is the employee's first name. It should not include middle initials, which should instead go in the 'MiddleName' field. Suffixes such as "Jr." should be appended to this field. The most up-to-date name should be used in instances where an employee has changed names over time.

Type: String

Length: 50

Code/Format: Text field using any characters.

Purpose: This helps with the matching of data to other external datasets that are used for report generation and research.

Examples: Kevin Jr.
Rob
Amy

'BirthYear'

Full Name: Birth Year

Description: This is only the birth year for the specific employee.

Type: Numeric

Length: 4

Code/Format: 4 digit Integer

Purpose: Providing month and year of birth is essential for matching these data with dataset providers such as the US Census Bureau that can provide information about the work status of former employees when they leave your university.

Examples: 1999
1974
2009

'BirthMonth'

Full Name: Birth Month

Description: This is only the birth month for the specific employee.

Type: Numeric

Length: 2

Code/Format: 2 digit Integer. Each MM must be one of the following:

01 January	05 May	09 September
02 February	06 June	10 October

03 March	07 July	11 November
04 April	08 August	12 December

Purpose: Providing birth month in addition to birth year adds significantly to IRIS' ability to match university employee records with their post-university work status and earnings.

Examples: 01
02
03

Vendor File

'OrganizationID'

Full Name: Organization ID

Description: Identifier used within your university specifying the organization or institution of the vendor. This identifier should be at the most granular level possible. For example, there should be a different number tied to each McDonald's establishment rather than a single number for all McDonald's. If this is not possible, still provide the establishment-level location information in the **Vendor** file rather than the firm headquarters. There is no required naming convention for this data element, so any combination of characters and numbers that represents your institution's id for the organization is acceptable.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field is helpful for tracking vendors used by an institution and reporting on these organizations across different awards.

Examples: BB48566
1233955-8912
660587119.CAM

'EIN'

Full Name: Employer Identification Number

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
ObjectCode
OrganizationID
EIN
DUNS
VendorPaymentAmount
OrgName
POBox
BldgNum
StrNumber
StrName
Address
City
State
DomesticZipCode
ForeignZipCode
Country

Description: The vendor's 9 digit EIN number. This is the government-issued EIN that most businesses should have assigned to them. Each EIN should be unique to the organization. Leave this field blank if there is no EIN for the vendor. If the vendor is a person, do *not* include their social security number in this field.

Type: String

Length: 10 Characters

Code/Format: Text field using any characters.

Purpose: EIN helps matching organizations to external datasets and across institutions allowing for more robust reporting.

Examples: 123456789
112233445
987654321

'DUNS'

Full Name: Dun and Bradstreet Number

Description: The Vendor DUNS number. The DUNS (Dun and Bradstreet) number is not available for every organization or vendor so this field is not required.

Type: String

Length: 15 Characters

Code/Format: Text field using any characters.

Purpose: Allows for better matching to external datasets by connecting DUNS numbers between datasets.

Examples: 12-345-6789
19-587-6892
22-641-9731

'VendorPaymentAmount'

Full Name: Vendor Payment Amount

Description: The amount paid by the university to the vendor in the specified period from the specified award. Do not include commas or dollar signs in the dollar value but always include any of the decimal values (cents). As discussed in the period start date and period end date fields it is important to sum these data up to the monthly level. Negative values should be included and can be indicated either with a preceding hyphen ("-30.21") or parentheses ("(30.21)").

Type: Number

Length: N/A

Code/Format: Text field using any characters.

Purpose: Allows for reporting on the total dollar amount paid to a vendor related to the specified award. This amount should be the sum of all expenditures from a given grant to a given vendor in the time period reported.

Examples: 15975.45
(487296.57)
-234.96

'OrgName'

Full Name: Organization Name

Description: Name of the vendor. This name is used to match to external datasets, so values closest to a raw name (lacking account numbers, order numbers, etc.) are preferred.

Code/Format: Text field using any characters.

Type: String

Length: 200 Characters

Purpose: Allows for matching and tracking of vendors across awards, institutions, and external datasets.

Examples: Acme Rockets
LexCorp
Xavier's School for Gifted Children

'POBox'

Full Name: Post Office Box

Description: PO Box number of the vendor. It is not necessary to include the term "PO" or "PO Box" followed by the number. Any value entered into this field is assumed to be a PO Box number.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor.

Examples: 145
10

'BldgNum'

Full Name: Building Number

Description: Building number of vendor.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor.

Examples: 1234B
1122

'StrNumber'

Full Name: Street Number

Description: Street number of the vendor.

Code/Format: Text field using any characters.

Type: String

Length: 20 Characters

Purpose: Helps to identify the location or address of the vendor.

Examples: 1234B
1122

'StrName'

Full Name: Street Name

Description: Street name of vendor. This is the street name without additional address information (e.g. building number, PO Box, etc.).

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor.

Examples: Main Street
Packard Ave

'Address'

Full Name: Address

Description: Vendor address metadata (alternative to splitting the PO Box, Building Number, Street Number and Street Name into separate columns). This is the combined address made up of the PO Box, Building Number, Street Number, and Street Name. Ideally, IRIS would like the individual address components entered into their respective fields (e.g., PO Box in its own field, Street Name in its own field) but it is acceptable to provide the full street address here. Additionally, if you do provide the address components in their designated fields, you are still encouraged to concatenate them together and provide them here as well.

Addresses should be at the establishment level, rather than the firm level. For example, there should be a different address tied to each McDonald's establishment rather than a single address for McDonald's corporate headquarters.

Type: String

Length: 120 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor.

Examples: 1234 Main Street
608 Packard Street

'City'

Full Name: City

Description: City of the vendor. As with all location data, it should be determined by the establishment location rather than the corporate headquarters.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the vendor, which is used when linking to external datasets.

Examples: New York City
Hermosillo

'State'

Full Name: State

Description: State of the vendor. This is the two-letter state used by the United State post office. As with all location data, it should be determined by the establishment location rather than the corporate headquarters. For non-US vendors, leave this field blank.

Type: String

Length: 2 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the vendor, which is used when linking to external datasets.

Examples: MI
OH

'DomesticZipCode'

Full Name: Domestic Zip Code

Description: US Zip Code of vendor. If the associated vendor establishment is not in the US, this field should be left blank. You can use either the shorter five-digit zip code, or the full zip+4 code. As with all location data, it should be determined by the establishment location rather than the corporate headquarters.

Type: String

Length: 10 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the vendor, which is used when linking to external datasets.

Examples: 48130
49916-5741

'ForeignZipCode'

Full Name: Foreign Zip Code

Description: Foreign Zip Code of vendor organization. If the associated vendor establishment is in the US, this field should be left blank. As with all location data, it should be determined by the establishment location rather than the corporate headquarters.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the vendor.

Examples: CAD128554
188459672

'Country'

Full Name: Country

Description: Country of vendor organization. This is the two-digit code for a country. For example, US – United States, CA – Canada, MX – Mexico. The size of the field will accommodate larger country codes but the dataset should be limited to two letter country codes. The two letter country codes can be found here: http://www.worldatlas.com/aatlas/ctycodes.htm (The A2 column of country codes is what should be used.)

As with all location data, it should be determined by the establishment location rather than the corporate headquarters.

Type: String

Length: 2 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the vendor organization, which is used when linking to external datasets

Examples: US
MX

Subaward File

Data Fields

PeriodStartDate
PeriodEndDate
UniqueAwardNumber
RecipientAccountNumber
ObjectCode
OrganizationID
EIN
DUNS
SubAwardPaymentAmount
OrgName
POBox
BldgNum
StrNumber
StrName
Address
City
State
DomesticZipCode
ForeignZipCode
Country

'OrganizationID'

Full Name: Organization ID

Description: Identifier used within your university specifying the organization or institution of the subawardee (also commonly referred to as subrecipient) organization. This identifier should be at the most granular level possible. For example, there should be a different number tied to each McDonald's establishment rather than a single number for all McDonald's. If this is not possible, still provide the establishment-level location information in the **Vendor** and **Subaward** files rather than the firm headquarters. There is no naming convention for this data element so any combination of characters and numbers that represents your institution's id for the organization is acceptable.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field is helpful for tracking subawardees used by an institution and reporting on these organizations across different awards.

Examples: BB48566
1233955-8912
660587119.CAM

'EIN'

Full Name: Employer Identification Number

Description: The subawardee's 9-digit EIN number. This is the government-issued EIN that most businesses should have assigned to them. Each EIN should be unique to the organization. Leave this field blank if there is no EIN for the subawardee. People should not typically appear in subaward data, but if you do have a subaward to a person in your systems, do *not* include their social security number in this field.

Type: String

Length: 10 Characters

Code/Format: Text field using any characters.

Purpose: EIN helps matching organizations to external datasets and across institutions allowing for more robust reporting.

Examples: 123456789
112233445
987654321

'DUNS'

Full Name: Dun and Bradstreet Number

Description: The subawardee's DUNS number. The DUNS (Dun and Bradstreet) number is not available for every organization or vendor so this field is not required.

Type: String

Length: 15 Characters

Code/Format: Text field using any characters.

Purpose: Allows for better matching to external datasets by connecting DUNS numbers between datasets.

Examples: 12-345-6789
19-587-6892
22-641-9731

'SubAwardPaymentAmount'

Full Name: Subaward Payment Amount

Description: The amount paid by the university to the subwardee in the specified period from the specified award. Do not include commas or dollar signs in the dollar value, but always include any of the decimal values (cents). As discussed in the period start date and period end date fields, it is important to roll these data up to the monthly level. Negative values should be included and can be indicated either with a preceding hyphen ("-30.21") or parentheses ("(30.21)").

Type: Number

Length: N/A

Code/Format: Text field using any characters.

Purpose: Allows for reporting on the total dollar amount paid to a subawardee related to the specified award.

Examples: 15975.45
-487296.57
(234.96)

'OrgName'

Full Name: Organization Name

Description: Name of the subaward organization. This name is used to match to external datasets, so values closest to a raw name (lacking account numbers, order numbers, etc.) are preferred.

Code/Format: Text field using any characters.

Type: String

Length: 200 Characters

Purpose: Allows for matching and tracking of subaward organizations across awards, institutions, and external datasets.

Examples: University of Illinois Ubrana-Champaign
Ohio State University
University of Michigan

'POBox'

Full Name: Post Office Box

Description: PO Box number of subawardee. It is not necessary to include the term "PO" or "PO Box" followed by the number. Any value entered into this field is assumed to be a PO Box number.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor, subawardee, or organization.

Examples: 145
10

'BldgNum'

Full Name: Building Number

Description: Building number of subawardee.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the vendor, subawardee, or organization.

Examples: 1234B
1122

'StrNumber'

Full Name: Street Number

Description: Street number of the subawardee.

Code/Format: Text field using any characters.

Type: String

Length: 20 Characters

Purpose: Helps to identify the location or address of the subawardee.

Examples: 1234B
1122

'StrName'

Full Name: Street Name

Description: Street name of the subawardee. This is the street name without additional address information (e.g. building number, PO Box, etc.).

Type: String

Length: 100 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the subawardee.

Examples: Main Street
Packard Ave

'Address'

Full Name: Address

Description: Subawardee address metadata (alternative to splitting the PO Box, Building Number, Street Number and Street Name into separate columns). This is the combined address made up of the PO Box, Building Number, Street Number, and Street Name. If you do provide the address components in their designated fields, you are still encouraged to concatenate them together and provide them here as well.

Addresses should be at the establishment level, rather than the firm level. For example, there should be a different address tied to each McDonald's establishment rather than a single address for McDonald's corporate headquarters.

Type: String

Length: 120 Characters

Code/Format: Text field using any characters.

Purpose: Helps to identify the location or address of the subawardee.

Examples: 1234 Main Street
608 Packard Street

'City'

Full Name: City

Description: City of the subawardee organization. As with all location data, it should be determined by the establishment location rather than the firm headquarters.

Type: String

Length: 50 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the subawardee, which is used when linking to external datasets.

Examples: New York City
Hermosillo

'State'

Full Name: State

Description: State of subawardee organization. This is the two-letter state used by the United State post office. As with all location data, it should be determined by the establishment location rather than the firm headquarters. For non-US subawardees, leave this field blank.

Type: String

Length: 2 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the subawardee, which is used when linking to external datasets.

Examples: MI
OH

'DomesticZipCode'

Full Name: Domestic Zip Code

Description: US Zip Code of subawardee. If the associated subawardee establishment is not in the US, this field should be left blank. You can use either the shorter five-digit zip code, or the full zip+4 code. As with all location data, it should be determined by the establishment location rather than the headquarters.

Type: String

Length: 10 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the subawardee, which is used when linking to external datasets.

Examples: 48130
49916-5741

'ForeignZipCode'

Full Name: Foreign Zip Code

Description: Foreign Zip Code of subawardee. If the associated subawardee is in the US, this field should be left blank. As with all location data, it should be determined by the establishment location rather than the firm headquarters.

Type: String

Length: 20 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the subawardee.

Examples: CAD128554
188459672

'Country'

Full Name: Country

Description: Country of subawardee. This is the two-digit code for a country. For example, US – United States, CA – Canada, MX – Mexico. The size of the field will accommodate larger country codes, but the dataset should be limited to two letter country codes. The two letter country codes can be found here: http://www.worldatlas.com/aatlas/ctycodes.htm (The A2 column of country codes is what should be used.)

As with all location data, it should be determined by the establishment location rather than the firm headquarters.

Type: String

Length: 2 Characters

Code/Format: Text field using any characters.

Purpose: This field helps to identify the location or address of the subawardee, which is used when linking to external datasets

Examples: US
MX

Object Code File

'ObjectCodeText'

Full Name: Object Code Text

Description: Text that describes the submitted Object Codes found in the **Employee**, **Subaward**, and **Vendor** files. Every Object Code submitted should have a corresponding description included here.

Type: String

Length: 200 Characters

Code/Format: Text field using any characters.

Purpose: Object codes should be high-level accounting codes used to describe the type of purchase or expenditure that is being reported in these data. IRIS uses them to better understand the types of purchases of goods and services made from vendors and subawards, and the types of compensation paid to employees. This description field is also important for our algorithms to identify student status for employees found in these data, and for categorizing transactions in different ways to group and report on them.

Examples: Large infrastructure project
Moon Destroying Super Laser
Genetically Engineered Velociraptor Project

Data Fields
ObjectCode
ObjectCodeText

Organization Unit File

Data Fields

CampusId
SubOrgUnit
CampusName
SubOrgUnitName

'CampusName'

Full Name: Campus Name

Description: The name of the campus that is referenced in the 'CampusID'.

Type: String

Length: 200 Characters

Code/Format: Text field using any characters.

Purpose: Describes the campus that is referenced in Campus ID code found in the other files.

Examples: University of Minnesota - Duluth Campus
University of Michigan – Ann Arbor Campus
University of Michigan – Dearborn Campus

'SubOrgUnitName'

Full Name: Suborganizational Unit Name

Description: This is the text description of the suborganizational unit within the reporting institution.

Type: String

Length: 200 Characters

Code/Format: Text field using any characters.

Purpose: Describes the 'SubOrgUnit' that is found in the **Award** and **Award ID** files. This allows IRIS to examine data on research employees and expenditures at the sub-campus level.

Examples: School of Arts and Sciences
School of Physical Sciences
College of Natural Sciences

Appendix A: OFS (Other Funding Sources) Codes

Use these codes in place of an official CFDA code for nonfederal awards. If your systems do not provide enough information to distinguish between these groups, use the code 00.000. Where the code is 00.000, IRIS will impute these codes from award titles and funding source names.

Funding source	Other Funding Source (OFS) code	Definition	Examples of Funders included in this code
Unknown or Generic Nonfederal	**00.000**	Code for use when specific nonfederal funding type cannot be determined	
Federal - Other	**00.070**	Agencies or federal contracts that do not have CFDA numbers	CIA, FEMA, Veterans' Administration, etc.
CIA	00.071	Central Intelligence Agency	CIA
FEMA	00.072	Federal Emergency Management Administration	FEMA
Veterans' Administration	00.073	Veterans' Administration	VA
Insitutional Investment	**00.100**	Award provided internally by your own university to support research at your university	
State Funding - Home State	**00.200**	Funding from a state agency in the state where the university is based	For a university in New York, the New York State Division of Veterans' Services
State Funding - Nonresident State	**00.300**	Funding from a state agency in states other than the state where the university is based	For a university outside New York, the New York State Division of Veterans' Services
Specific Nonresident State	00.3xx	Funding from a state agency in a particular state other than the state where the university is based. See separate worksheet below for state-specific codes (xx is the FIPS state numeric code)	03.36 New York State Division of Veterans' Services 03.56 State of Wyoming Game and Fish Department
Local Funding (City/ County)	**00.400**	Funding provided by a local government body; that is, a government body smaller than a state.	City of Seattle Jefferson County Conservation District Menominee Indian Tribe of Wisconsin Cook County Health Department

Business / For Profit	**00.500**	Industrial / Commercial Funders that do business in the United States	American Steamship Company Rembrandt Foods
Nonprofit	**00.600**	Nonprofit funding; use a subcategory if desire, to be more specific	American Heart Association Annenberg Foundation
Philanthropic	00.610	Money given by an individual or corporation that sits in a University Foundation account and is expended to conduct research	Bill & Melinda Gates Foundation Ford Foundation
Public Foundations	00.620	Funding provided by a foundation that accepts public donations	The CDC Foundation The March of Dimes The Pediatric Cancer Research Foundation.
Private Foundations	00.630	Funding that is given by foundations that are privately funded.	the Annenberg Foundation the William & Flora Hewlett Foundation
Private Associations	00.640	Funding provided by professional medical societies, fraternal organizations, and other private associations	Indiana Elks Charities American College of Radiology Kiwanis International
Hospital / Medical Centers	00.700	Funding from a health care facility	Mayo Clinic Hospital for Sick Children Tufts Medical Center
Your university's hospital or medical center	00.710	Funding from your university's hospital or medical center	For the University of Pittsburgh, UPMC
A hospital or medical center other than your university's	00.720	Funding from a different university's hospital or medical center	For the University of Pennsylvania, UPMC
Universities / Colleges	00.800	Any award coming from universities and colleges including flow-through funding	The University of Maryland Black Hills State University London School of Economics
Foreign	00.900	Funding from non-US corporations, associations, or governments	Flax Canada Government of Pakistan
Foreign Government	00.910	Funding from governments outside the U.S.	Government of Pakistan
Foreign Non-Government	00.920	Funding from corporations, associations, etc outside the U.S.	Flax Canada

State Specific OFS Codes

State	OFS Code	FIPS State Numeric Code	FIPS State Alpha Code
Alabama	**00.301**	01	AL
Alaska	**00.302**	02	AK
Arizona	**00.304**	04	AZ
Arkansas	**00.305**	05	AR
California	**00.306**	06	CA
Colorado	**00.308**	08	CO
Connecticut	**00.309**	09	CT
Delaware	**00.310**	10	DE
District of Columbia	**00.311**	11	DC
Florida	**00.312**	12	FL
Georgia	**00.313**	13	GA
Hawaii	**00.315**	15	HI
Idaho	**00.316**	16	ID
Illinois	**00.317**	17	IL
Indiana	**00.318**	18	IN
Iowa	**00.319**	19	IA
Kansas	**00.320**	20	KS
Kentucky	**00.321**	21	KY
Louisiana	**00.322**	22	LA
Maine	**00.323**	23	ME
Maryland	**00.324**	24	MD
Massachusetts	**00.325**	25	MA
Michigan	**00.326**	26	MI
Minnesota	**00.327**	27	MN
Mississippi	**00.328**	28	MS
Missouri	**00.329**	29	MO
Montana	**00.330**	30	MT

State	OFS Code	FIPS State Numeric Code	FIPS State Alpha Code
Nebraska	**00.331**	31	NE
Nevada	**00.332**	32	NV
New Hampshire	**00.333**	33	NH
New Jersey	**00.334**	34	NJ
New Mexico	**00.335**	35	NM
New York	**00.336**	36	NY
North Carolina	**00.337**	37	NC
North Dakota	**00.338**	38	ND
Ohio	**00.339**	39	OH
Oklahoma	**00.340**	40	OK
Oregon	**00.341**	41	OR
Pennsylvania	**00.342**	42	PA
Rhode Island	**00.344**	44	RI
South Carolina	**00.345**	45	SC
South Dakota	**00.346**	46	SD
Tennessee	**00.347**	47	TN
Texas	**00.348**	48	TX
Utah	**00.349**	49	UT
Vermont	**00.350**	50	VT
Virginia	**00.351**	51	VA
Washington	**00.353**	53	WA
West Virginia	**00.354**	54	WV
Wisconsin	**00.355**	55	WI
Wyoming	**00.356**	56	WY

Appendix B: Pre-Submission Checklist

IRIS uses a wide variety of data from your university to generate reports. Focusing on data quality and consistency in a few specific fields will allow us to provide the most accurate reports possible, will increase the match of your data to external data, and will minimize unexpected changes to reports over time. Included below are suggestions of what can be done prior to data file submission to optimize dataset quality. This is not a comprehensive list, but rather a summary of some common data issues we find within and across submissions and how to avoid them.

Table 12. Common data issues and how to avoid them

Verify …	… across all prior submissions (if applicable)	… in specific files (as listed)
Award number consistency • The 'UniqueAwardNumber' field must be consistent for each award across prior submissions as well as across the files listed • All award numbers present in any of the **Employee**, **Vendor**, or **Subaward** files should also be in the **Award** and **Award ID** files	√	√ **Award**, **Award ID**, **Employee**, **Vendor**, **Subaward**
Award number format • The 'UniqueAwardNumber' field matches the formatting laid out in the 'UniqueAwardNumber' Field Description		√ **Award**, **Award ID**,

		Employee, **Vendor**, **Subaward**
At least two calendar years of data are included ● To match employee data to federal data, IRIS needs at least two calendar years of data: one to mark an employee's presence at your university, and one to determine if they are still working there. The more years of data that are provided, the more robust reporting can be done. Because federal earnings data may not be available for recent years, data from earlier years is also valuable in providing a complete picture of your university's employees		√ **Employee**
Data going back to 2009 or earlier is provided (if possible) ● Data from earlier years allows us to generate a more robust Employee Report as we can build a more complete picture of career trajectories over time. We recommend going back to 2009 or earlier		√ **Employee**
Employee ID consistency ● The 'DeidentifiedEmployeeIdNumber' field must be consistent for each employee across submissions as well as across the **Employee** and **Employee Name** files ● **In particular, check for leading zero deletion.** One of the most common errors is differences in leading zero deletion, which can be caused by opening and saving files in Excel, which automatically deletes leading zeros unless fields are imported as text.	√	√ **Employee**, **Employee Name**

'FTEStatus' and 'ProportionsofEarningsAllocated' fields are proportions ● These fields should be proportions (data between 0 and 1), NOT sums of earnings		√ **Employee**
Transactions are rolled up to the monthly level		√ **Award**, **Employee**, **Vendor**, **Subaward**
Birth dates are included ● This dramatically improves our ability to match with external employee data sources.		√ **Employee Name**
Middle names or initials are in the 'MiddleName' field only ● Middle names or initials should not be included in the 'FirstName' or 'LastName' fields		√ **Employee Name**

Names are included for every transaction ● Employee names must be included for every transaction in the **Employee** file; otherwise, the unidentified employee transactions cannot be used to create earnings reports ● Transactions without vendor or subawardee names will not be included in the vendor report		√ **Employee**, **Employee Name**, **Vendor**, **Subaward**
Transactions are not duplicated ● A transaction should appear in either the **Vendor** or **Subaward** file, not both		√ **Vendor**, **Subaward**
None of your data is truncated ● Truncation is a common issue, especially in vendor names, and can damage our ability to match to Federal data		√ All files, but especially **Award**, **Vendor**, **Subaward**
Location information, especially zip code, is included for all transactions ● This is beneficial for matching to Federal data		√ **Vendor**, **Subaward**

Consistency of object codes ● The 'ObjectCode' field should be consistent between submissions (if applicable) ● **In particular, check for leading zero deletion.** Discrepancies are most often caused by choosing to include or remove leading zeros inconsistently	√	√ **Employee**, **Vendor**, **Subaward**, **Object Code**
Consistency of 'CampusId' and 'SubOrgUnit' fields ● These fields should be consistent across submissions (if applicable) ● 'CampusID' and 'SubOrgUnit' codes found in the **Organization Unit** file must be consistent with those in the **Award** and **Award ID** files ● **In particular, check for leading zero deletion.** Discrepancies are most often caused by choosing to include or remove leading zeros inconsistently	√	√ **Award**, **Award ID**, **Organization Unit**
'SubOrgUnit' field includes numbers not names ● The 'SubOrgUnit' field should be numbers or codes, not actual unit names. (Name corresponding to the number or code should be input 'SubOrgUnitName' field)		√ **Award**, **Award ID**, **Organization Unit**