



**Science and Technology in America's
Reinvestment:**

**Measuring the Effects of Research on Innovation, Competitiveness and
Science**

*As of May 4th, 2015, STAR METRICS is no longer accepting
new institutions and will be discontinuing Level I data
collections and reports generation on January 1, 2016.*

**Technical Specification for Research Institution
Data Submissions for Level I**

Introduction

This document has been created to provide guidance on the technical specifications for data files to be sent as a part of Level I of the STAR METRICS® project.

If the research institution is using XML to send these files, this guide together with the XSD code provided (See Appendix A) outlines how to produce a readable XML file. If the research institution is sending data via CSV files this guide outlines the minimum requirements needed for the data system as well as suggested formatting.

As always, the STAR METRICS® approach is flexible. The STAR METRICS® team will work individually with every research institution to ensure that data is being sent in the easiest manner for the research institution.

Process

Once a file is submitted to the STAR METRICS® data system, each submission is examined at a macro and micro level to validate that the correct data is being submitted. At the macro level, the system verifies that the file sent has been sent in a readable format (CSV or XML) with the correct number of columns. It also verifies how many of the requested data elements have been provided. At the micro level, the system verifies whether each data field provided matches a certain set of pre-defined characteristics (detailed below in the Data Characteristics section). The system will provide feedback to help ensure that data is correctly formatted.

The STAR METRICS® team will analyze each file that is submitted, either manually or through the data system. Reports will be generated that provide feedback to research institutions on how their data were examined by the data system and any possible action items.

Please use this document in conjunction with our Participation Guide and Frequently Asked Questions posted on the STAR METRICS® website. The Data Dictionary on the website also provides an overview of the information listed below.

Requested Data Files

Data should be submitted for the following four submission types:

1. Award
2. Employee
3. Sub-Award
4. Vendor

Please note that an Indirect Cost Proposal Ratio is also requested. However, this document does not need to be sent quarterly and has unique specifications. The details below pertain only to the four files listed above. See the Indirect Cost Proposal Specifications section for further details.

Each file type can be submitted in one of the following two data file formats:

1. XML
2. CSV

When submitting files, please use the following format for the file names:

UniversityAbbreviation_FileType_DateSubmitted using underscores and YYYY_MM_DD for the date

Ex. UnivX_Award_2010_12_30

Data Element Names

The following data fields are included in the data files. For a more detailed description of the meaning of the fields, please see the Data Dictionary.

All Fields

<i>Data Field Name</i>	<i>XML/CSV Element/Field Name</i>
Period Start Date	PeriodStartDate
Period End Date	PeriodEndDate
Unique Award Number	UniqueAwardNumber
Recipient Account Number	RecipientAccountNumber
Overhead Charged	OverheadCharged
De-identified Employee ID Number	DeidentifiedEmployeeIdNumber
Occupational Classification	OccupationalClassification
FTE Status	FteStatus
Proportion of Earnings Allocated to Award	ProportionOfEarningsAllocatedToAward <i>OR</i> ProportionOfEarningsAllocated
Sub-Award Recipient DUNS Number	SubAwardRecipientDunsNumber
Sub-Award Payment Amount	SubAwardPaymentAmount
Vendor DUNS Number	VendorDunsNumber
Vendor Payment Amount	VendorPaymentAmount

Data Elements by Data File

<i>Data File Type</i>	<i>Element/Field Name</i>
Award	PeriodStartDate
	PeriodEndDate
	UniqueAwardNumber
	RecipientAccountNumber
	OverheadCharged
Employee	PeriodStartDate
	PeriodEndDate
	UniqueAwardNumber
	RecipientAccountNumber
	DeidentifiedEmployeeIdNumber
	OccupationalClassification
	FteStatus
	ProportionOfEarningsAllocatedToAward <i>OR</i> ProportionOfEarningsAllocated
Sub-Award	PeriodStartDate
	PeriodEndDate
	UniqueAwardNumber
	RecipientAccountNumber
	SubAwardRecipientDunsNumber
	SubAwardPaymentAmount

Vendor	PeriodStartDate
	PeriodEndDate
	UniqueAwardNumber
	RecipientAccountNumber
	VendorDunsNumber
	VendorPaymentAmount

Data File Validation

Various validations will be done on each submitted file based on the data file format and on the individual data elements. These validations will help ensure that the correct data is being sent and processed. Any data submitted will be analyzed based on the data characteristics specifications listed below. The Data Dictionary provides an overview of this information while the Data Characteristics section in this document details all possible exceptions and clarifications. Reports will be generated (either manually or automatically) to provide feedback to the research institution based on the analysis from the data file validation process. These reports will help guide research institutions in future submissions.

Note: When "standard XSD" data types are referenced below they are referring to the data types defined in the namespace <http://www.w3.org/2001/XMLSchema>.

XML Document Level Validation

XML files will be validated to ensure they match the structure and format specified in the STAR METRICS® Data Files XSD (StarMetricsDataFiles.xsd). See Appendix A or the website for the XSD code. Data will also be checked to ensure they pass the validations in the Data Characteristics section below.

Example XML files for each submission type are available on the STAR METRICS® website.

Please note that the parent element for Employee submissions should be <Individual> instead of <Employee>.

CSV Document Level Validation

CSV files will be validated to ensure they are in the proper CSV format. Please include a header row where the CSV column names are the same as the field names listed above, namely:

Award

PeriodStartDate,PeriodEndDate,UniqueAwardNumber,RecipientAccountNumber,OverheadChar
ged

Employee

PeriodStartDate,PeriodEndDate,UniqueAwardNumber,RecipientAccountNumber,DeidentifiedEmployeeIdNumber,OccupationalClassification,FteStatus,ProportionOfEarningsAllocatedToAward

OR

PeriodStartDate,PeriodEndDate,UniqueAwardNumber,RecipientAccountNumber,DeidentifiedEmployeeIdNumber,OccupationalClassification,FteStatus,ProportionOfEarningsAllocated

Vendor

PeriodStartDate,PeriodEndDate,UniqueAwardNumber,RecipientAccountNumber,VendorDunsNumber,VendorPaymentAmount

Sub-Award

PeriodStartDate,PeriodEndDate,UniqueAwardNumber,RecipientAccountNumber,SubAwardRecipientDunsNumber,SubAwardPaymentAmount

Data Characteristics

In addition to the document level validation, the individual fields will be validated. Each element in the XML files is described below.

PeriodStartDate

- The field is required and should be specified exactly once per parent element.
- The start date should reflect the beginning of the time period in which the transaction occurred, not the start date of the award.
- Should be in standard XSD date format (YYYY-MM-DD).

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
2009-12-31	Yes	
12/31/2009	No	Not in the standard XSD date format.

PeriodEndDate

- PeriodEndDate has the same data type and rules as PeriodStartDate with the following additional rule:
- Should be greater than or equal to the value of PeriodStartDate within the same parent element.
- The end date should reflect the ending of the time period in which the transaction occurred, not the end date of the award.

UniqueAwardNumber

- The field is required and should be specified exactly once per parent element.
- The field is made up of a funding source code and an award identifier:
 - The funding source code portion is either the CFDA code (for federal awards) or the STAR Other Funding Source (OFS) code (for non-federal awards or those without CFDA codes).
 - The award identifier is either the actual federal award ID (for federal awards) or an internal award identifier (for non-federal awards).
- Both the funding source and award identifier components are required.
- The award identifier is an arbitrary string with no specific formatting requirements.
- The minimum length for the award identifier component is 1 character.
- The maximum length for the award identifier component is 50 characters.
- The funding source component is of the format: ##.### where # is a digit from 0 to 9.
- The funding source component should be prepended before the award identifier component and separated with a space as in:
 [Funding Source] [Award Identifier]
 00.000 [1- to 50-character award identifier]
- If neither a CFDA code nor OFS code is available or applicable, then the funding source component should be set to 00.000.
- If the award identifier component is unknown, then a zero should be inserted after a blank space, as in “47.123 0” rather than merely “47.123”.
- If the funding source component is a CFDA code and only the first two digits are available, three trailing zeros should be used to match the ##.### format (e.g. 33.000).
- A list of OFS codes is available on the [Resources page](#) of the STAR METRICS® website.

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
12.345 abcde12345	Yes	CFDA is 12.345 and the federal award ID is abcde12345.
12.000 abcde12345	Yes	CFDA is 12 and the federal award ID is abcde12345. Trailing zeros are appended to two digit CFDA code.
00.000 abcde-12345	Yes	No applicable funding source code is available and the award identifier is abcde-12345.
00.200 State Award 1	Yes	OFS code (non-federal award) is 00.200 (home state funding) and the award identifier is “State Award 1”. Note that

		spaces are allowed in the award identifier component.
12.345abcde12345	No	CFDA is 12.345 and the federal award ID is abcde12345. There is no space between the funding source and award identifier components.
12.345	No	CFDA is 12.345 but no federal award ID was included. A blank space and then a zero should be added if the award ID is unknown (e.g. 12.345 0).
abcde12345	No	CFDA is not available and the award identifier is abcde12345. In this case, the funding source component is missing.

RecipientAccountNumber

- The field is required and should be specified exactly once per parent element.
- The value is an arbitrary string with no specific formatting requirements.
- The min length for the value is 1 character.
- The max length for the value is 255 characters.

OverheadCharged

- The field is required and should be specified exactly once per parent element.
- Should be in standard XSD decimal format.
- Any value with more than two decimal places will be rounded to two decimal places.

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
1000	Yes	
1000.1	Yes	
1000.12	Yes	
1000.123	Yes	
0	Yes	Zero is ok.

-1000	Yes	Negative numbers are ok.
\$1000	No	The dollar sign is not allowed by the XSD decimal data type.
1,000	No	Commas are not allowed by the XSD decimal data type.

DeidentifiedEmployeeIdNumber

- The field is required and should be specified exactly once per parent element.
- The value is an arbitrary string with no specific formatting requirements.
- The min length for the value is 1 character.
- The max length for the value is 50 characters.

OccupationalClassification

- The field is required and should be specified exactly once per parent element.
- The value is an arbitrary string with no specific formatting requirements.
- The min length for the value is 1 character.
- The max length for the value is 255 characters.
- Note: STAR METRICS® maintains a list of known occupational classifications internally. When a data file is received the value is mapped to one of the items in the list of known classifications.
- Note: There is no need for an institution to map their jobs to the STAR METRICS® job classifications. This will be done by the STAR METRICS® team and validated with the institution. However if the institution would like to change the mapped classifications, this can be done within the “Management” function on the website.

FteStatus

- The field is required and should be specified exactly once per parent element.
- Should be in standard XSD decimal format.
- The value should be greater than or equal to zero.
- The max value is 1.
- Any value with more than four decimal places will be rounded to four decimal places.
- If FTE status cannot be determined, please notify the STAR METRICS® team. Some institutions have chosen FTEs based on type of job held when exact calculations cannot be determined, for example using 0.25 FTE for undergraduate positions, 0.5 FTE for graduate student positions and 1.0 FTE for all other positions.

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
1	Yes	

1.0	Yes	
0.5	Yes	
0.1	Yes	
0.5678	Yes	
1.5	No	Numbers greater than 1 are not allowed.
0	Yes	
-0.5	No	Negative numbers are not allowed.

ProportionOfEarningsAllocatedToAward or ProportionOfEarningsAllocated

- The field is required and should be specified exactly once per parent element.
- Either name can be used for this field. The shorter name was introduced to address the character limit for column headings in some HR systems.
- Should be in standard XSD decimal format.
- The minimum value is -1000 (exclusive).
- The value can be zero (e.g. the Individual is on the contract but didn't charge any time).
- The max value is 1000 (exclusive).
- Any value with more than four decimal places will be rounded to four decimal places.

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
1	Yes	
1.0	Yes	
0.5	Yes	
0.1	Yes	
0.5678	Yes	
0	Yes	
-0.5	Yes	
-1	Yes	

1000	No	Numbers 1000 or greater are not allowed.
-1000	No	Numbers -1000 or less are not allowed.
999.9999	Yes	Number is less than 1000.
-999.9999	Yes	Number is greater than -1000.

SubAwardRecipientDunsNumber

- The field is required and should be specified exactly once per parent element.
- The field value should contain a DUNS number, a 5 or 9 digit ZIP code, or be an empty element (e.g. `<SubAwardRecipientDunsNumber></SubAwardRecipientDunsNumber>` or `<SubAwardRecipientDunsNumber />`).
- If the DUNS number is available it should be used.
- If the DUNS number is not available but the ZIP code is available the ZIP code should be used.
- If the DUNS number and ZIP code are not available an empty element should be included.
- If a DUNS number is specified, the format of the DUNS number should be ##### where # is a digit from 0 to 9.
- If a 5 digit ZIP code is specified, it should be in the format Z##### where # is a digit from 0 to 9.
- If a 9 digit ZIP code is specified, it should be in the format Z##### or Z#####-#### where # is a digit from 0 to 9.
- If a foreign postal code is being used, place an F before the postal code to indicate that it is foreign. Foreign postal codes do not need to follow the prescribed format above for 5 or 9 digit postal codes. Note that you may only include up to 9 characters (excluding the F prefix) for foreign postal codes.

Examples

<i>Value</i>	<i>Is Valid?</i>	<i>Notes</i>
123456789	Yes	This will be treated as a DUNS number.
Z12345	Yes	This will be treated as a 5 digit ZIP code.
Z123456789	Yes	This will be treated as a 9 digit ZIP code.
Z12345-6789	Yes	This will be treated as a 9 digit ZIP code.
FDK-1234	Yes	This will be treated as the foreign postal code 'DK-1234'.

1234567890	No	There can be a maximum of 9 digits.
12345-6789	No	DUNS numbers cannot have dashes and there is no "Z" present to indicate this is a ZIP code.
12345	No	DUNS numbers should be 9 digits and there is no "Z" present to indicate this is a ZIP code.
FABC123DEF	Yes	This will be treated as the foreign postal code 'ABC123DEF'.
FABC 123-DE	No	This exceeds the character limit. Spaces and hyphens are included in the 9-character limit (excluding the F prefix) for foreign postal codes.

SubAwardPaymentAmount

See rules for OverheadCharged.

VendorDunsNumber

See rules for SubAwardRecipientDunsNumber.

VendorPaymentAmount

See rules for OverheadCharged.

SubAwardPaymentAmount

See rules for OverheadCharged.

Indirect Cost Rate Proposal Specifications

Information on Overhead comes from your Indirect Cost Proposal. This requires the calculation of one number that should be provided initially and updated when there is a change. A subset of the STAR METRICS® FDP working group has determined that the least burdensome approach is as follows:

Retrieve your adjusted cost by cost pool. Then calculate the total salaries and fringe benefits, as well as the total non-labor costs. Sum each across cost pools, and calculate the proportion as (salaries and fringe benefits)/(salaries and fringe benefits plus non labor costs).

This number should be provided to the STAR METRICS® team during the initial data submission process using the [Submit Indirect Cost Proposal page](#). While the team only needs this calculated ratio, you may provide the entire Indirect Cost Proposal or Indirect Cost Proposal Spreadsheet if that is easier.

Note that these are intended to produce estimates of jobs in a statistical context, not an accounting context. So you can also provide us with a range of proportions, and we will generate separate estimates for each proportion you provide.

Data Submission Process

Data can be submitted through the Web Upload System on the STAR METRICS® website. Descriptions and instructions are detailed below.

Submitting Historical Data

For historical data submissions, please consider the following recommendations:

- Provide data at the transactional level if possible. This will allow for the most precise analysis.
- Multiple quarters worth of data can be provided in the same file as long as the Period Start Dates and Period End Dates would allow for the STAR METRICS® team to analyze the data at a quarterly level.
- If possible, provide data for Q1 2010 through Q4 2011.

Web Upload

This method uses a secure web upload system on the STAR METRICS® website. This system will allow for research institutions to submit data and receive automatic validation on their file structure and data element characteristics.

1. In order to access this system, click on the “Log On” button at the top of any page on the STAR METRICS® website.
2. Log in using the instructions provided when your accounts were created after your signed Memorandum of Understanding (MOU) was received. If you need an account created to upload data, please contact us at starmetrics-it@nih.gov.
3. After logging in, navigate to the Upload page. This page will allow you to upload all files. Please follow the instructions on the page if uploading the Indirect Cost Proposal, as that has a separate upload page.
4. To upload a file:
 - a. Enter a label for the file that will correspond to the other files you are submitting for that time period. For example, if you are submitting files with data from the 3rd quarter of 2010 consider entering “Q3 2010” as a label.
 - b. Click “Choose File” to browse to your local submission file.
 - i. Note: You may only upload one file at a time.
 - c. Add a short note (optional).

- d. Click “Upload File.”
- e. The web upload system will examine the file to determine what type of file was submitted and whether it matches the expected structure and characteristics defined in this document. The system will provide one of three messages:
 - i. Submission Successful — The file matched the specified structure and characteristics. No further action required.
 - ii. Submission Successful (with Warnings) — The file matched the specified structure and characteristics. No further action required; however, there may be data quality concerns associated with the submission. You should review the warnings on the website and determine whether or not they should be addressed. If so, the file may be resubmitted. If you have questions regarding the issues, please contact us at starmetrics@nih.gov.
 - iii. Issues Found in Submission Data — The file was recognized by the system but contained issues with particular data fields. A detailed list of issues will be provided. If you would like the STAR METRICS® team to examine these issues further, please contact us at starmetrics@nih.gov.
 - iv. Unrecognized Format — The file was not recognized by the system. Please contact the STAR METRICS® team if you are unsure why you received this error at starmetrics@nih.gov.
5. Repeat step 4 for all award, employee, vendor, and subaward file submissions.
6. If you need to resubmit a file, click the Resubmit link in the Submission History next to the file you would like to resubmit. Please note that resubmitting a file completely replaces the previous submission; therefore, you should ensure that the new file for resubmission contains the same dataset as the previous submission.

Appendix A: XSD Code

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema id="StarMetricsDataFiles" targetNamespace="star-metrics-data-files" xmlns:star="star-
metrics-data-files" elementFormDefault="qualified" xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:simpleType name="UniqueAwardNumberElement">
    <xs:restriction base="xs:string">
      <xs:minLength value="1" />
      <xs:maxLength value="50" />
    </xs:restriction>
  </xs:simpleType>

  <xs:simpleType name="RecipientAccountNumberElement">
    <xs:restriction base="xs:string">
      <xs:minLength value="1" />
      <xs:maxLength value="255" />
    </xs:restriction>
  </xs:simpleType>

  <xs:simpleType name="DeidentifiedEmployeeIdNumberElement">
    <xs:restriction base="xs:string">
      <xs:minLength value="1" />
      <xs:maxLength value="50" />
    </xs:restriction>
  </xs:simpleType>

  <xs:simpleType name="OccupationalClassificationElement">
    <xs:restriction base="xs:string">
      <xs:minLength value="1" />
    </xs:restriction>
  </xs:simpleType>
</xs:schema>
```

```

        <xs:maxLength value="255" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="FteStatusElement">
    <xs:restriction base="xs:decimal">
        <xs:minInclusive value="0" />
        <xs:maxInclusive value="1" />
    </xs:restriction>
</xs:simpleType>

<xs:simpleType name="ProportionOfEarningsAllocatedToAwardElement">
    <xs:restriction base="xs:decimal">
        <xs:minExclusive value="-1000" />
        <xs:maxExclusive value="1000" />
    </xs:restriction>
</xs:simpleType>

<xs:complexType name="AwardsElement">
    <xs:sequence>
        <xs:choice minOccurs="0" maxOccurs="unbounded">
            <xs:element name="Award">
                <xs:complexType>
                    <xs:sequence>
                        <xs:element name="PeriodStartDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="PeriodEndDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="UniqueAwardNumber" type="star:UniqueAwardNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="RecipientAccountNumber" type="star:RecipientAccountNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="OverheadCharged" type="xs:decimal" minOccurs="1" maxOccurs="1" />
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
        </xs:choice>
    </xs:sequence>
</xs:complexType>

<xs:complexType name="IndividualsElement">
    <xs:sequence>
        <xs:choice minOccurs="0" maxOccurs="unbounded">
            <xs:element name="Individual">
                <xs:complexType>
                    <xs:sequence>
                        <xs:element name="PeriodStartDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="PeriodEndDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="UniqueAwardNumber" type="star:UniqueAwardNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="RecipientAccountNumber" type="star:RecipientAccountNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="DeidentifiedEmployeeIdNumber"
type="star:DeidentifiedEmployeeIdNumberElement" minOccurs="1" maxOccurs="1" />
                        <xs:element name="OccupationalClassification"
type="star:OccupationalClassificationElement" minOccurs="1" maxOccurs="1" />
                        <xs:element name="FteStatus" type="star:FteStatusElement" minOccurs="1"
maxOccurs="1" />
                        <xs:element name="ProportionOfEarningsAllocatedToAward"
type="star:ProportionOfEarningsAllocatedToAwardElement" minOccurs="1" maxOccurs="1" />
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
        </xs:choice>
    </xs:sequence>
</xs:complexType>

<xs:complexType name="SubAwardsElement">
    <xs:sequence>
        <xs:choice minOccurs="0" maxOccurs="unbounded">
            <xs:element name="SubAward">
                <xs:complexType>
                    <xs:sequence>

```

```

        <xs:element name="PeriodStartDate" type="xs:date" minOccurs="1" maxOccurs="1" />
        <xs:element name="PeriodEndDate" type="xs:date" minOccurs="1" maxOccurs="1" />
        <xs:element name="UniqueAwardNumber" type="star:UniqueAwardNumberElement"
minOccurs="1" maxOccurs="1" />
        <xs:element name="RecipientAccountNumber" type="star:RecipientAccountNumberElement"
minOccurs="1" maxOccurs="1" />
        <xs:element name="SubAwardRecipientDunsNumber" type="xs:string" minOccurs="1"
maxOccurs="1" />
        <xs:element name="SubAwardPaymentAmount" type="xs:decimal" minOccurs="1"
maxOccurs="1" />
    </xs:sequence>
</xs:complexType>
</xs:element>
</xs:choice>
</xs:sequence>
</xs:complexType>

<xs:complexType name="VendorsElement">
    <xs:sequence>
        <xs:choice minOccurs="0" maxOccurs="unbounded">
            <xs:element name="Vendor">
                <xs:complexType>
                    <xs:sequence>
                        <xs:element name="PeriodStartDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="PeriodEndDate" type="xs:date" minOccurs="1" maxOccurs="1" />
                        <xs:element name="UniqueAwardNumber" type="star:UniqueAwardNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="RecipientAccountNumber" type="star:RecipientAccountNumberElement"
minOccurs="1" maxOccurs="1" />
                        <xs:element name="VendorDunsNumber" type="xs:string" minOccurs="1" maxOccurs="1" />
                        <xs:element name="VendorPaymentAmount" type="xs:decimal" minOccurs="1"
maxOccurs="1" />
                    </xs:sequence>
                </xs:complexType>
            </xs:element>
        </xs:choice>
    </xs:sequence>
</xs:complexType>

<xs:element name="Awards" type="star:AwardsElement" />
<xs:element name="Individuals" type="star:IndividualsElement" />
    <xs:element name="SubAwards" type="star:SubAwardsElement" />
    <xs:element name="Vendors" type="star:VendorsElement" />
</xs:schema>

```