

Interns and Residents Information System (IRIS) New XML File Format Summary

Overview

The existing IRIS DBF file format will be retired and replaced by a redesigned XML file format. The new format will capture a few additional fields and will organize the data inside a normalized data structure but otherwise generally captures the same fields as the legacy DBF format and maintains the same core data model built around resident records and assignment records.

As part of this change, the IRISv3 and IRISEDv3 applications will also be retired. For the new XML format CMS will publicly release all the technical specifications and documentation needed to create an IRIS file, but will not release a replacement for the IRISv3/IRISEDv3 applications for creating IRIS files. Instead, Providers will be encouraged to use IRIS vendor software in order to prepare their IRIS submissions.

The process by which IRIS files are submitted will stay the same. Providers will continue to submit their IRIS submission to their MAC alongside their cost report.

New Fields

Except for one field being removed (which is addressed in a subsequent section below), the new XML format will contain the same fields as the old DBF format plus the following new fields:

1. Assignment IPF Percentage (Psych): The percentage of the Intern/Resident(IR)'s rotational assignment time period the hospital provider is allowed to count in its total number of FTE residents for Psych in the 2552-10 Cost Report's Worksheet E-3 Part II.
2. Assignment IRF Percentage (Rehab): The percentage of the IR's rotational assignment time period the hospital provider is allowed to count in its total number of FTE residents for Rehab in the 2552-10 Cost Report's Worksheet E-3 Part III.
3. Assignment Non-Provider Site Percentage: The percentage of the IR's rotational assignment time that was spent in allowable non-provider site settings. See 2552-10 cost report worksheet S2 Lines 66 & 67.
4. Assignment Displaced Resident (True/False): Indicates whether the IR is an allowable displaced resident for which the hospital may receive a temporary cap adjustment. See 2552-10 worksheet E-4 line 16 (DGME) and worksheet E Part A line 17 (IME). Note that IRIS will track the raw number of displaced resident FTEs while what gets recorded in the cost report is an adjustment whose calculation, among other things, takes into account free cap slots. The displaced resident assignments recorded in IRIS do NOT directly sum to the displaced resident FTEs recorded in the cost report.

5. Assignment New Program (True/False): Indicates whether the resident is in the “initial years of a program that meets the exception to the rolling average rules” as per the cost report instructions. See 2552-10 worksheet E-4 Line 15 (DGME), worksheet E Part A Line 16 (IME), worksheet E-3 Part II Line 7 (Psych), and worksheet E-3 Part III Line 8 (Rehab).
6. Resident Non-IRP Year One Residency: For IRs that either participated in a preliminary/transitional year or a simultaneous match, this records the code for the residency type they were enrolled in during their first year as well as a ‘type’ value indicating whether it was a preliminary year or a simultaneous match.
7. Creation Software Name: Simple text field for recording the name of the software or vendor used to create the IRIS submission. This is meant to help CMS debug issues with specific files by identifying their source.

Removed Field

The XML format will not include an equivalent of the DBF master file Residency Years Completed (RESYEAR). This field was removed due to being redundant because the same value was already being tracked in a more granular and useful way at the assignment level (ARESYEAR in the assignment file).

New Format Technical Summary

- The new IRIS file format will consist of an XML file whose structure is defined by an XML Schema (XSD 1.0).
 - XSD Background: an XSD schema is a document that precisely defines the structure and constraints that an XML file must follow in order to be considered valid and which allows said validity check to be done programmatically by any system producing or consuming the XML file. All third-party IRIS software and CMS would rely on the same XSD file to validate the basic structure of IRIS submission. Normally end-users will never interact with the XSD schema directly; they are meant to be consumed by computer programs.
 - These technologies have extremely wide support. XML is as universally supported as any piece of technology can be while XSD is by far the most commonly supported XML schema language.
- XML IRIS submissions will consist of a single file, rather than the file pair scheme currently used.
- CMS will publish the list of all additional error checks that files must pass beyond the basic XSD validation.
- The IRIS data in the XML format will be stored using a normalized scheme relying on parent/child records. See accompanying XML Mockup.
- The XML files will be human readable and can be viewed in any text editor such as Notepad. However, the files will be large and verbose and are generally meant to be viewed and edited using specialized software.

Problems with old DBF Format:

- Built around 30-year old dBase III database format that is increasingly unsupported by modern software. For example, Microsoft Excel 2013 and Access 2013 have dropped DBF file support.
- No Unicode support, cannot capture many of the characters used in non-english names.
- dBase III has a number of characteristics that make it a poor choice for use as a data transmission format, leading to many otherwise preventable errors in the data received by CMS:
 - Very weak support for data constraints beyond basic datatypes
 - Format requires one file per table, forcing each IRIS submission to consist of two DBF files: master and assignment. This complicates the transmission and storage of IRIS files and leads to errors when one of the files gets separated from its matching pair.
 - Current format lacks foreign keys (referential integrity) and employs a non-normalized data structure. This leads to massive duplication of certain fields, which get repeated on every row.
- dBase III's strictly tabular nature and one file per table scheme make it difficult to extend the format in the future in order to capture additional data.