6 Item Data Specifications

OpenClinica supports user-defined CRF Item Data values according to Item Data properties defined during CRF and Study design. **OpenClinica enforces strict data typing** corresponding with the Item Datatype. Each data value must conform to the canonical format for its data type. If a value with an invalid data type is inserted or updated, it must be rejected and an exception thrown. This document describes data typing for user-defined CRF Item Data values in OpenClinica version 3.1.

This page is not approved for publication.

6.1 Item Properties

Each Item (specified in the item table) has a unique item_id and is associated with a single definition that does not change for the life of the Item* [[list exceptions to this such as ability to convert from date to pdate]]. The definition includes declaration of a datatype (item_data_type) as well as name, description and units, and Object Identifier (OID).

Each Item Data value is associated with exactly one Item definition and as such is bound to a specific datatype. The physical implementation of the OpenClinica data model (aka the database schema) stores all item data values in the item_data.value field, which is a text field of a maximum 4000 characters. Despite the fact that the DBMS treats all item data values as text, OpenClinica enforces strict data typing corresponding with the Items Datatype. Each data value must conform to the canonical format[1] for its data type. If a value with an invalid data type is inserted or updated, it must be rejected and an exception thrown.

An Item can be represented in one or more CRF Versions. Each Item has CRF version-specific properties (specified in the item_form_metadata and response_set tables) that control the allowed value domain and representation/formatting of the Item Data value in different interfaces of the system (e.g. CRF data entry input type, data import, data extract). The properties controlling these representations may change from CRF version to CRF version. The representations may be impacted by other system settings that could change over time or from user to user, such as locale settings in the case of dates, and by edit check constraints such as simple validation checks and OpenClinica Rules.

---

[1] A canonical form means that values of a particular type of resource can be described or represented in multiple ways, and one of those ways is chosen as the favored canonical form. (That form is canonized, like books that made it into the bible, and the other forms are not.) A classic

example of a canonical form is paths in a hierarchical file system, where a single file can be referenced in a number of ways:

myFile.txt # in current working dir

../conf/myFile.txt # relative to the CWD

/apps/tomcat/conf/myFile.txt # absolute path using symbolic links

/u1/local/apps/tomcat-5.5.1/conf/myFile.txt # absolute path with no symlinks

The classic definition of the canonical representation of that file would be the last path. With local or relative paths you cannot globally identify the resource without contextual information. With absolute paths you can identify the resource, but cannot tell if two paths refer to the same entity. With two or more paths converted to their canonical forms, you can do all the above, plus determine if two resources are the same or not, if that is important to your application (solve the aliasing problem).

Note that the canonical form of a resource is not a quality of that particular form itself; there can be multiple possible canonical forms for a given type like file paths (say, lexicographically first of all possible absolute paths). One form is just selected as the canonical form for a particular application reason, or maybe arbitrarily so that everyone speaks the same language. (taken from http://stackoverflow.com/questions/280107/what-does-the-term-canonical-form-or-canonical-representation-in-java-mean)

This page is not approved for publication.

6.2 Canonical Datatypes

The canonical formats are described below. They are derived from CDISC ODM 1.3 , which itself are derived from XML Schema datatypes. This is the basis for our specification, however only a subset of ODM formats are supported in OpenClinica.

OpenClinica Datatype Name	ODM Format Name	Schema Datatype	Canonical Format (Allowed String Pattern)	Notes
ST	text	xs:string	any sequence of characters	
INT	integer	xs:integer	-?digit+	
REAL	float	xs:decimal	-?digit+(.digit+)?	

DATE	date	xs:date	YYYY-MM-DD	An OpenClinica DATE is a string with year, month, day separated by a -which is the ISO8601 standard date format. [1] The OpenClinica user interface supports localization of dates, with the default date format in the user interface being DD-MMM-YYYY.[2] All dates must be translated by the application to the canonical format specified here prior to insert into the database. It is possible for OpenClinica instances upgraded from 3.0.x and older versions that some items of type DATE may have existing values that do not match the allowed string pattern. See OpenClinica Troika Item Data Specifications.doc and OpenClinica Date Format Change Specifications.doc for reference.
PDATE	partialDate	xs:date	YYYY-[[MM-]DD]	An OpenClinica partial date (PDATE) is a string with a year, optional month, and optional day separated by a -. No timezone is specified. The OpenClinica user interface supports localization of partial dates, with the default partial date format in the user interface being DD-MMM-YYYY. All partial dates must be translated by the application to the canonical format specified here prior to insert into the database. It is possible for OpenClinica instances upgraded from 3.0.x and older versions that some items of type PDATE may have existing values that do not match the allowed string pattern. See OpenClinica Troika Item Data Specifications.doc and OpenClinica Date Format Change Specifications.doc for reference.
FILE	URI	xs:anyURI	scheme ":" hier-part ["?" query] ["#" fragment]	See http://www.w3.org/Addressing/URL/uri-spec.txt for URI specification.

---

[1] http://www.iso.org/iso/date_and_time_format

[2] https://www.openclinica.org/dokuwiki/doku.php?id=developerwiki:i18n

This page is not approved for publication.