

CFPS 17

(Call for Papers Submission number 17)

Proposal to Accommodate Gregorian Dates using a Modified ISO 8601

Submitted by: Proctor, Tony

Type: Technical proposal

Created: 2013-03-06

Last updated: 2013-04-20

URL: Most recent version: <http://fhiso.org/files/cfp/cfps17.pdf>
This version: http://fhiso.org/files/cfp/cfps17_v1-1.pdf

Description: Proposed adoption of a modified ISO 8601 subset to accommodate Gregorian dates, and yearly quarters thereof

Keywords: Dates, ISO-8601, Quarters

Contents

1. Abstract	3
2. Proposal.....	3
3. Not Covered or Not Required.....	3
4. Illustration.....	4
5. Use Cases	4
6. Recommendation	4
7. References	4

1. Abstract

Proposal to adopt a modified subset of the ISO 8601 date standard for the encoding of Gregorian date values in the data model. The proposal defines an extension of that ISO standard in order to accommodate a representation of yearly quarters.

2. Proposal

The ISO 8601 date standard was designed to provide for unambiguous exchange of Gregorian date and time values. All date field values are encoded numerically to avoid locale dependencies related to the use of month or day names. These fields are also arranged most-to-least significant in order to provide a natural sorting capability by using the textual value as a key. A representation may be truncated to support different date/time granularities, such as a year, a day, or an hour. For instance:

yyyy-mm-dd	(for a known day, e.g. 1956-06-09)
yyyy-mm	(for a known month)
yyyy	(for a known year)

The standard actually includes many separate features which were each covered by separate previous standards. For instance, numeric date/time representations (ISO 2014), week dates (ISO 2015), ordinal dates (ISO 2711), and a number of time-related standards. It also includes some support for date/time ranges, and for offsets from UTC. It was created in 1988 but revised in 2000, and again in 2004.

The W3CDTF discussion note [NOTE-datetime](#) examines the need for a subset of the overloaded ISO 8601 standard that can be used on the Internet. The US Library of Congress Extended Date Time Format ([EDTF](#)) proposes a similar sub-setting of the ISO 8601 standard. These subsets amount to the following syntax: yyyy[-mm[--dd]] when only dates are concerned.

The use of such an encoding would ensure Gregorian date values can be exchanged unambiguously between different genealogical products. Note that this does not affect the display format for such dates, which can be generated according to the locale and preferences of any particular end-user, or the recording of an actual date transcriptions as part of historical evidence. In other words, these values are purely designed to be machine-readable interpretations of dates found in evidence.

The proposal includes a modification to use a single letter 'Q' to introduce a yearly quarter number (1-4), i.e. yyyy-Qq. This choice was selected because it is unambiguous and follows a precedent already set in ISO 8601 for week dates (i.e. yyyy-Www).

3. Not Covered or Not Required

The ISO 8601 implementation of durations is not required. Any expression of a date range should be done within the data model rather than within just one datum of the model.

Ordinal dates and week dates are part of ISO 8601 but are unlikely to be of any use for family history. The aforementioned subsets do not include them.

Although full date/time specifications will be applicable to data creation dates, modification dates, etc., they are unlikely to be relevant to historical dates within the data itself.

ISO time-zone designators are unlikely to be relevant to historical dates. All historical dates should be a representation of local-time at the corresponding historical place.

There is no known requirement to address days relative to each quarter in the way that the existing standard addresses days relative to each week in the 'Week dates' format (e.g. 2009-W53-7).

This proposal does not consider alternative calendars such as Hindu or Julian. ISO 8601 is completely specific to the Gregorian calendar and so this modified subset will be similarly specific. A separate proposal has been made for generalising the date syntax to accommodate any number of alternative world calendars.

This proposal does not consider negative years, e.g. for addressing dates BC. Although the ISO standard suggests these can be supported by agreement between two parties, it breaks the goal of the fixed field widths. Also, such dates would not be Gregorian, and possibly not even Julian.

This proposal does not address dual dates (i.e. Gregorian/Julian pairs). These are the subject of a separate FHISO proposal

This modified subset caters for date granularity, just as ISO 8601 does, but it does not cover date imprecision. This must be the subject of a separate proposal.

4. Illustration

A date of birth recorded as “13th March 1901” would be encoded as follows:

1901-03-13

NB: the model must allow a transcription of the original text to be associated with the machine-readable ISO-encoded value.

A registration occurring in quarter 2 (i.e. April-June) of 1956 would be encoded as follows:

1956-Q2

5. Use Cases

For Gregorian dates, the ISO 8601 standard provides an unambiguous, machine-readable encoding. Unfortunately, the standard is overloaded and this has already caused other organisations to adopt a working subset, as suggested here.

Many registrations occur on a quarterly basis. Examples include the GRO index of civil registrations of births, marriages, and deaths in England & Wales. Without this extension it is not possible to record the dates of those registrations from the index accurately.

6. Recommendation

Since the proposed extension of the ISO 8601 is both straightforward and unambiguous, it is recommended that FHISO liaise with the [ISO TC 154](#) that handles data standards to have it incorporated into a future revision of the international standard.

The absence of this support is a glaring omission within the existing ISO standard, and other areas will benefit from its inclusion. This would be preferable to FHISO using a custom extension that is not recognised elsewhere.

7. References

ISO 8601:2004(E) – Data elements and interchange formats — Information interchange — Representation of dates and times.

STEMMA Discussion. <http://www.familyhistorydata.parallaxview.co/research-notes/dates-calendars> (‘STEMMA Dates’ section).