fhiso

CFPS 103
(Call for Papers Submission number 103)

A trivial timestamp type

Submitted by: Smith, Richard

Created: 2014-01-29

URL: Most recent version: http://fhiso.org/files/cfp/cfps103.pdf
This version: http://fhiso.org/files/cfp/cfps103_v1-0.pdf

Description: This paper proposes a light-weight timestamp type for representing the creation and modification times of digital resources.

Keywords: dates, timestamps, time zones

Abstract

This paper considers a representation for timestamps suitable for recording the creation and modification times of digital resources. This format is compared with the `dateTime` type in XML Schema, and the 'simple date' used in GEDCOM X. It is concluded that certain details of the GEDCOM X representation are unfortunate, and that alignment with XML Schema's `dateTime` is desirable.

1   Introduction

There are a considerable number papers before the FHISO discussing how dates should be represented in a future FHISO standard. They offer the means of recording dates in multiple calendars, handling uncertainty in them, and recording other important characterists of the date as expressed in a source. None of the current papers propose a date format that includes time components. This paper does not comment on the desirability of being able to record the time of historical events.

However one context in which a time is important is when recording creation dates and modification times for digital resources. Adding a timestamp to a digital resource is considered good practice because it allows a user (or the user's software) to compare two versions of a document and determine which is the more recent. This paper proposes a simple timestamp data type that can be used for this purpose. It is not intended to be used for recording historical dates.

The timestamp format proposed here is compatible with ISO 8601 [1, §3.2.7], and the date part is the Gregorian date type proposed in CFPS 17, except that reduced representations (those in which certain components are omitted) are not permitted. The time part, which is separated from the date part by a literal T, must contain all three time components: hours, minutes and seconds. The format is fixed-width and naturally sorting. It is always interpretted as a date in the Gregorian calendar.

2   Comparison with XML Schema

The format proposed above very similar to XML Schema's `dateTime` type [3]. It differs only insofar that XML Schema optionally allows *(i)* an leading '-' sign on the year, *(ii)* sub-second resolution, and *(iii)* a time zone. Including a time zone is of genuine use when collaborating with researchers in different continents, or using remote web services. This paper therefore proposes following XML Schema in allowing a time zone to be recorded by suffixing it with either '±hh:mm' or a literal Z to denote UTC. In the former form, the minutes component are not optional, and therefore Central European Summer Time (CEST) must be denoted

+02:00. There is no default time zone. In the absense of an explicit one, and without knowledge of the creating system, it is not safe to assume any particular time zone.

Sub-second precision is neither particularly desired nor undesirable. As syntactical alignment with XML Schema's type seems sensible, this paper proposes allowing them, but it is suggested that the FHISO grants implementations licence to ignore the sub-second component if they so wish. Negative years are problematic in XML Schema because they are not ISO 8601 compliant, but fortunately they never arise in the uses for which this timestamp type is intended. Since the point is moot, this paper suggests that they are allowed solely for compatibility with XML Schema.

The representation of a timestamp must match the following POSIX extended regular expression:

```
-?[0-9]{4}-[0-9]{2}-[0-9]{2}T[0-9]{2}:[0-9]{2}:[0-9]{2}(\.[0-9]+)?
([+-][0-9]{2}:[0-9]{2}|Z)?
```

3   Comparison with GEDCOM X

The current GEDCOM X draft includes a 'simple date' representation very similar to the timestamp representation proposed here [4, §5.2]. In GEDCOM X, it is folded into its historical date format, and therefore reduced representations (such as pure dates without the time components) are permitted. This paper does not comment on the desirability of that, and a future FHISO paper may wish to do the same with this timestamp type. Aside from reduced representation, GEDCOM X's simple date differs from this proposal in the following respects: *(i)* a leading + is mandatory on years AD; *(ii)* sub-second parts are not permitted; and *(iii)* the minutes part of a time zone may be omitted.

The absense of sub-second components in GEDCOM X is inconsequential. The optionality of the minutes part of the time zone is unusual and may simply be an error in the current GEDCOM X draft. Time zones in computing are almost always denoted with both hours and minutes, even in areas such as mail or `http` headers which do not follow ISO 8601 and pre-date XML. The flexibility to omit the minutes part is simply not needed.

The requirement for a leading + is a more serious difference. It is far more common than not for the leading + sign to be omitted in ISO 8601 dates; by requiring it, GEDCOM X is running contrary with the standard XML practice, which is unfortunate for a language whose preferred serialisation is XML. This paper strongly urges GEDCOM X to reconsider this, especially if the proposal in CFPS 91 that the FHISO should ratify GEDCOM X is to be considered [5].

References

[1] International Organization for Standardization, 2004, *Data elements and interchange formats — Information interchange — Representation of dates and times* (ɪso 8601:2004)

[2] Tony Proctor, 2013, *Proposal to Accommodate Gregorian Dates using a Modified ɪso 8601* (ᴄғᴘs 17), `http://fhiso.org/files/cfp/cfps17.pdf`

[3] World Wide Web Consortium, 2004, *xmʟ Schema Part 2: Datatypes (Second Edition)*, `http://www.w3.org/TR/xmlschema-2/`

[4] Intellectual Reserve Inc., 2013, *The ɢᴇᴅᴄᴏᴍ x Standard Header Set*, `http://gedcomx.org/headers/v1`

[5] Ryan Heaton, 2013, *A substantive vision and proposal to validate and ratify the ɢᴇᴅᴄᴏᴍ x specification set*, `http://fhiso.org/files/cfp/cfps91.pdf`