Gmail by Google

Eric Haas <ehaas@healthedatainc.com>

US Language Requirements

8 messages

Brett Marquard <brett@riverrockassociates.com> Mon, Feb 13, 2017 at 7:22 AM
To: Robert McClure MD <rmcclure@mdpartners.com>, "grahame@healthintersections.com.au" <grahame@healthintersections.com.au>
Cc: "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Good day,

I hope everyone is doing well - snowy wonderland here, it's unbelievable!

Grahame, What is the best way to create the following value set for language (cc to Rob since I know he thought about this for VSAC...) (gforge 11919<http://gforge.hl7.org/gf/project/fhir/tracker/?action=TrackerItemEdit&tracker_item_id=11919&start=0>):

Consistent with the RFC 5646 the shortest alpha code for the language should be used. Testing will only test the primary language tag and not test for extension components specified in RFC 5646 such as extended language sub-tags, script tag, nor region tag. [see also 80 FR 16817]
Specifically:

* use alpha 2 character code if one exists (ISO 639-1);

* use alpha 3 character code if an alpha 2 character code does not exist (ISO 639-2); and

* region extensions (ISO 3166-1) are permitted but not required (however, if a region extension is used, it will be verified for accuracy as part of testing and must be correct).

Full text from ONC clarification...
[cid:image001.png@01D285E3.06CAA560]

Brett Marquard
Principal
River Rock Associates
t: 413.549.6886

image001.png
71K

Robert McClure MD <rmcclure@mdpartners.com> Mon, Feb 13, 2017 at 10:35 AM
To: Brett Marquard <brett@riverrockassociates.com>
Cc: Grahame Grieve <grahame@healthintersections.com.au>, "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Brett,
First a question on content. I suspect you will grimace, as you so often do when I start one of these...

Regarding the use of ISO3166-1: I don't know what the "region extensions" are and I don't see any mention of them here: http://www.iso.org/iso/home/standards/country_codes.htm or see them here https://www.iso.org/obp/ui/#search.

As fun as that is, I suspect you'll be even more thrilled to see that the ISO3166-1 two letter (and I suspect 3-letter) codes **collide** with the ISO639-1 and ISO639-2 codes - look at AF for example: for 639-1 = Afrikaans *language,* but in ISO 3166-1 = Afghanistan *country*. That is actually **okay** because RFC5646 uses these different codes in different places (to represent different things.)
I want to note that in RFC5646 the 3166-1 two letter code is used to represent *the region* and **that is it** so far as I can tell.

So I'm confused a to how :

> region extensions (ISO 3166-1) are permitted but not required (however, if a region extension is used, it will be verified for accuracy as part of testing and must be correct).

is to be implemented because I can't tell what it's suppose to mean.

I also wonder how once that is figured out, how this:

> Testing will only test the primary language tag and not test for extension components specified in RFC 5646 such as extended language sub-tags, script tag, nor region tag. [see also 80 FR 16817]

can also be true given that the "region extensions" are not "the primary language tag."

As for setting up value sets for everything **but "region extensions"**, I think we'd have a value set for the ISO 639-1 codes, and another value set for the 639-2 codes **that do not have** a corresponding 639-1 code. Then we group those value sets into a grouper that contains all the allowed codes, but yo can access the subtype value sets as needed. BTW, this set of codes **does change** on occasion.

So what about the "region extensions"? I've looked again at the FR80 sections regarding RFC5646 and I don't see this 3166-1 stuff mentioned. I also don't see it **at all** in the FR 80 text but it can be hard to search.
Does anyone on this thread know?
. . .
Consultant to ONC / NLM on terminology
All content represents my view only
Robert McClure MD : 303.926.6771 : rmcclure@mdpartners.com

On Feb 13, 2017, at 8:22 AM, Brett Marquard <brett@riverrockassociates.com> wrote:

Good day,

I hope everyone is doing well – snowy wonderland here, it's unbelievable!

Grahame, What is the best way to create the following value set for language (cc to Rob since I know he thought about this for VSAC…) (gforge 11919):

> Consistent with the RFC 5646 the shortest alpha code for the language should be used.
> Testing will only test the primary language tag and not test for extension components specified in RFC 5646 such as extended language sub-tags, script tag, nor region tag. [see also 80 FR 16817]
> Specifically:
> - use alpha 2 character code if one exists (ISO 639-1);
> - use alpha 3 character code if an alpha 2 character code does not exist (ISO 639-2); and

- region extensions (ISO 3166-1) are permitted but not required (however, if a region extension is used, it will be verified for accuracy as part of testing and must be correct).

Full text from ONC clarification…
<image001.png>

Brett Marquard
Principal
River Rock Associates
t: 413.549.6886

Brett Marquard <brett@riverrockassociates.com> Mon, Feb 13, 2017 at 10:53 AM
To: Robert McClure MD <rmcclure@mdpartners.com>
Cc: Grahame Grieve <grahame@healthintersections.com.au>, "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Hi Rob,

Thanks, I do love this response.

-Did you ever attempt to build this for C-CDA R2.1?

…just curious before I start digging up some history.

Best

Brett

From: Robert McClure MD [mailto:rmcclure@mdpartners.com]
Sent: Monday, February 13, 2017 1:35 PM
To: Brett Marquard <brett@riverrockassociates.com>
Cc: Grahame Grieve <grahame@healthintersections.com.au>; Eric Haas (ehaas@healthedatainc.com) <ehaas@healthedatainc.com>; Nagesh Bashyam <nagesh.bashyam@drajer.com>
Subject: Re: US Language Requirements

[Quoted text hidden]

Robert McClure MD <rmcclure@mdpartners.com> Mon, Feb 13, 2017 at 11:00 AM
To: Brett Marquard <brett@riverrockassociates.com>
Cc: Grahame Grieve <grahame@healthintersections.com.au>, "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Nope, because we are waiting to get the code systems in. So far as I can tell, there is not an IP issue but that was not clear initially and it may be still confused along the way.

. . .
Consultant to ONC / NLM on terminology
All content represents my view only
Robert McClure MD : 303.926.6771 : rmcclure@mdpartners.com

[Quoted text hidden]

Grahame Grieve <grahame@healthintersections.com.au> Mon, Feb 13, 2017 at 2:06 PM
To: Brett Marquard <brett@riverrockassociates.com>
Cc: Robert McClure MD <rmcclure@mdpartners.com>, "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

I'm confused by this - is region tested or not?

generally, there's no way to represent this value set in a FHIR value set. only as a series of assertions, and even then, there's no way to make them complete, I think.

Grahame

[Quoted text hidden]
—
—
http://www.healthintersections.com.au / grahame@healthintersections.com.au / +61 411 867 065

Robert McClure MD <rmcclure@mdpartners.com> Thu, Feb 16, 2017 at 10:24 AM
To: Grahame Grieve <grahame@healthintersections.com.au>, Brett Marquard <brett@riverrockassociates.com>
Cc: "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Brett - where did you get the “ONC Clarification” text? Who is the author so I can track down what that last bullet is trying to convey.

Graham - We can create a value set that captures the first two bullets. I can give you two sets of data that will contain the 639-1 content and the allowed 639-2 content. Then we define a value set for this that combines the two 639-specific value set content. DO you want me to send you the current list? What I’m unclear on is how we want that to line up with what FHIR currently has as “Common Languages” http://build.fhir.org/valueset-languages.html. The CLD seems to infer that this is the set of tags described by whatever BCP 47 currently points to, RFC5646 at the moment. SO what we are talking about here is a subset of that because http://build.fhir.org/valueset-languages.html is much larger. Seems to me we need a different URI/OID for this more restricted set - agreed?

Brett - the current CCDA r2.1 Language value set (2.16.840.1.113883.1.11.11526) is tied to RFC4646. What do you want to do? We need to create a different value set that aligns with RFC5646 if we follow the approach attempted here. Is this what we plan to do? If yes, then how do we change CCDA to use this new value set? Did we already do this as an errata or something?

By the way, RFC5646 is consistent with the “use 639-1 when present” as the primary language code so it lines up with the “rule” described in the guidance below.

RFC5646 is a bit more complicated when options exist. All of this is resolved if we follow the content in the IANA registry. I don’t know if there is a single list of the current “real” languages - I’ve looked for this and it’s hard to find.

BTW, the best explanation of how the codes are created is here:
https://www.w3.org/International/articles/language-tags/index.en

Do we need a call to sort out our plan?

. . .
Consultant to ONC / NLM on terminology

All content represents my view only
Robert McClure MD : 303.926.6771 : rmcclure@mdpartners.com

[Quoted text hidden]

Brett Marquard <brett@riverrockassociates.com> Thu, Feb 16, 2017 at 10:27 AM
To: Robert McClure MD <rmcclure@mdpartners.com>, Grahame Grieve <grahame@healthintersections.com.au>
Cc: "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

Rob/I and will talk and get back everyone…

Best,

Brett

From: Robert McClure MD [mailto:rmcclure@mdpartners.com]
Sent: Thursday, February 16, 2017 1:25 PM
To: Grahame Grieve <grahame@healthintersections.com.au>; Brett Marquard <brett@riverrockassociates.com>
Cc: Eric Haas (ehaas@healthedatainc.com) <ehaas@healthedatainc.com>; Nagesh Bashyam <nagesh.bashyam@drajer.com>
Subject: Re: US Language Requirements

Brett - where did you get the "ONC Clarification" text? Who is the author so I can track down what that last bullet is trying to convey.

[Quoted text hidden]

Grahame Grieve <grahame@healthintersections.com.au> Thu, Feb 16, 2017 at 11:07 AM
To: Robert McClure MD <rmcclure@mdpartners.com>
Cc: Brett Marquard <brett@riverrockassociates.com>, "Eric Haas (ehaas@healthedatainc.com)" <ehaas@healthedatainc.com>, Nagesh Bashyam <nagesh.bashyam@drajer.com>

answering the FHIR questions:

> We can create a value set that captures the first two bullets. I can give you two sets of data that will contain the 639-1 content and the allowed 639-2 content.

ok

> Then we define a value set for this that combines the two 639-specific value set content. DO you want me to send you the current list? What I'm unclear on is how we want that to line up with what FHIR currently has as "Common Languages" http://build.fhir.org/valueset-languages.html. The CLD seems to infer that this is the set of tags described by whatever BCP 47 currently points to, RFC5646 at the moment. SO what we are talking about here is a subset of that because http://build.fhir.org/valueset-languages.html is much larger. Seems to me we need a different URI/OID for this more restricted set - agreed?

this is an extensible binding with a max binding. The extensible binding is a convenient subset defined intensionally, the max binding is defined extensionally. This pattern appears elsewhere in FHIR too.

We can do another enumerated value set for a US binding no problem. If we wanted to make it extensional rather than intensional - we haven't got language for that.
All the value sets get their own URI

Grahame