EML - Bug #654

scope of the unit element

10/25/2002 10:02 AM - Peter McCartney

Status:	Resolved		**Start date:**	10/25/2002
Priority:	Immediate		**Due date:**	
Assignee:	Matt Jones		**% Done:**	0%
Category:	eml - general bugs		**Estimated time:**	0.00 hour
Target version:	EML2.0.0rc3		**Spent time:**	0.00 hour
Bugzilla-Id:	654			

Description

Discussion of stmml has revealed several isses, one of which is the fact that
units, as expressed by stmml, are applicable only to measurable quantities.
Many variables that ecologists put in an eml dataset and might intuitively
appear to have "unit's (geologic age, sex, or species names, for example) do
not have units and thus must be decleared "dimensionless" or "undefined"
because unit is required for all attributes. I dont think its intuitively
apparent to users that these are domains not units and that they should be
described as such.

in the required element <measurmentScale> we class all attributes as nominal,
ordinal, interval or ratio. Srictly speaking only interval scales have units,
the rest are dimensionless. In practice, there is still some value of knowing
the units of the denominator and/or numerator in ratios of two dimensions, so
we probably dont want to throw out the baby with the bath water there.

To help clarify this, we might consider merging units within measurementScale
so that things may be set requred when relevant. an example might be:

<measurementScale>
<interval>
<standardUnit>
metersPerSecond
</standardUnit>
</interval>
</measurementScale>
a variant does away with embedding custom units in additionalMetadata would be:
<measurementScale>
<interval>
<unit library="http://ecoinformatics.org/emlUnitDictionary.xml">
metersPerSecond
</unit>
</interval>
</measurementScale>

this would mean any custom unit definitions would need to be published online.

content model for measurement scale might look like:
element measurementScale(nominal | ordinal | interval | ratio)
element nominal
element ordinal
element interval (unit)
element ratio (i'm not sure what would go here - it seems like we're hacking
unit definitions in emlUnitDictionary for ratios already but maybe that should
be pulled out and we provide a structured ratio definition here that references
two (or more?) true dimensions)

all attributes would still have a domain element - the existing bug on that
still applies

History

#1 - 10/25/2002 10:36 AM - Matt Jones

Peter,
Thanks for the excellent summary. I think I agree with this approach, excpet
for one point: the "library" attribute. In general I think it is crucial that we
have the STMML unit definitions available with the metadata. Consequently, I
don't think it is appropriate to define units in stmml in an external location
-- they should either be 1) directly in the EML document (additionalMetadata or
elsewhere), or 2) should be well-known (meaning, shipped with and defined as
implicit to EML itself), as we are doing with the current unitDictionary.

The changes we've proposed in this bug affect the schemas and instance
documents, but I think it is critical to get this domain/unit/measurementScale
stuff worked out in a logical way for the EML 2 release. This is precisely why
we had a review of the RC1 and RC2 releases -- to catch fatal flaws like this.
Adn I do think this stuff is fatally flawed as it now stands. People would
complain once they tried to really use it.

#2 - 10/28/2002 01:07 PM - Matt Jones

Comments from Barbara Benson show that our previous definition of ratio scale is
wrong. It is, in fact, an extension of interval scale, and so always has units,
precision, and a domain. For a good summary of the definitions of the scales,
see: http://www.math.sfu.ca/~cschwarz/Stat-301/Handouts/node5.html

This will need to be reflected in our changes to EML2.

#3 - 10/29/2002 05:54 AM - Tim Bergsma

So... an attribute with celsius units has an interval measurement scale, but one
with kelvin units has a ratio measurement scale. I wouldn't have guessed that.
To limit guessing, perhaps we should officially identify the appropriate
measurement scale for each dictionary unit. I suspect the vast majority are
ratio.

#4 - 10/30/2002 12:25 PM - Matt Jones

DONE. Unit is now under measurement scale only where appropriate, and the
definition of ratio scale has been fixed.

#5 - 03/27/2013 02:14 PM - Redmine Admin

Original Bugzilla ID was 654