VCoI Call Notes

12 April 2021

# Attendees

* Bobbin Teegarden
* Claude Baudoin (Cébé)
* Mike Bennett
* Lars Toomre
* Robert Nehmer

# Agenda

* Set objectives for Q2

# Meeting Notes

## Current Status

### MB summarizes.

Seems we have the idea of a small-s standard for each of:

* Structural (meta-model): how to maintain concepts, terms, abbrevs etc. for a TF
* Annotations (metadata): textual annotations for a TF’s resource
* Process guidelines

Presume the idea is a kind of resource (using maybe SKOS, RDF, SBVR) that TFs can be encouraged to use.

### Process:

Define the workflow: How does a TF:

* Initiate
* Conduct
* Get approval for
* Publish

a Vocabulary.

Ideally a simple BPMN diagram.

This would help TF leads understand what they might be doing, who needs to be consulted, who it goes to in order to be published and so on.

Also whether this recommendation has the same standing as a Specification, a Discussion Paper and so on. What is the process between someone identifying the need for a vocabulary for their community, to the point where there is a formal thing that the OMG has ‘set’ as the way to do this.

#### Related question:

This may be a simple guideline for people to use or not use, or it could be a more formal thing with some backing from OMG or the AB, as ‘the’ of using the available standards in the ways defined here.

Could be something like a standard – may have e.g. UML Conceptual mode, OWL file, SKOS file and so on. Should such a thing be submitted (by the TF) via the RFC process?

* A wrinkle: if SIG, they would have to go to a TF.
* Would propose the vocabulary resource to the TF and recommend for adoption
* Might be >1 SIG liaising with a given TF

Again it would be a matter of making that SIG’s Vocabulary resource an RFC of that TF.

* 2 swim-lanes in the process definition already

Do we need something new similar to an RFC or a White Paper or somewhere in between?

Disclaimer ‘Discussion paper is not an official product of the OMG’

* Would we want that?

If we were wanting to tell a future submitter to comply with a Vocabulary it would need some sort of normativity – would need to be adopted at the same level as an OMG Technology.

That’s one of two possible outcomes already:

* Something for a TF or a SIG to produce (e.g. RFP/RFI)
  + As optional to help TFs and SIGs be able to create consistent terms, refs etc.
  + As mandatory for the TFs and SIGs
* Something for a responder to an RFP (or a submitter of a RFC) to produce
  + Then we would need to have something normalized in order to include e it in instructions for submitters

If there is an informal thing that TFs and SIGs can use, they can’t enforce submitters to use it but they can still link to it in terms of guidance.

### Summary

So there is:

* Process
* Roadmap
  + i.e. The perceived route(s) forward for this

### Actions / Next Steps

Map out the first parts of roadmap:

1. A recommendation for TFs and SIGs to use
2. A formal thing for TFs and SIGs to follow
3. A thing that submitters can be required to follow

On Step 1, what it looks like seems to be:

* Wishlist for model resource to manage terms, concepts (ontology), abbrev etc. and their outputs as tables, wikistuff, doc inserts etc. etc.
  + How to use what parts of SKOS, MVF, SBVR, OWL, RDF etc.
* Wishlist for annotations

### Process:

On (1) an initial draft

On (2) you add more formal requirements on roles and responsibilities

On (3) you publish a formal document of some sort that enshrines the process and the formal roles and responsibilities (new Roles and Responsibilities that would not be in (1) and (2))

## Q2 Ambitions:

Get (1) in shape – have something concrete and useful

We are not delivering the vocabularies to the TFs but rather the ability to define vocabularies in a suitable way for use in different contexts.

That is to decentralize the production of vocabularies.

By having these in a common format and a common repository then everyone can have access to that.

BT: Did we lose the terms that Lars initially gave us?

LT: At the June 2019 meeting in Amsterdam, the FERM WG came to FDTF with 13 terms they had been asked by the Data Foundation to define. Some of these were incorporated into the draft FTA. Among other terms were ‘Taxonomy’ and ‘Ontology’. FDTF approved definitions we felt comfortable with at the June 2019 QM. Subsequent work here (VCoI) and in AI, BC-PSIG etc. have look at this.

LT: A 4th point is ‘discoverability’ – e.g. when someone new to the OMG wants to know what is meant by ‘ontology’ for example – t know at least what definitions we are comfortable.

May be:

* ‘What does the OMG mean by X?’
* What does this SIG/TF mean by Y?’

Including cross discoverability – so e.g. Gov can refer to concepts from FDTF and vice versa.

BT: Can we include thesauri (in the Roget sense not the technical sense)?

* By implication yes.

RN: One of the things to avoid is coming up with all of the terminology. One of the motivators for being more structural in terms of how this is built out such that those who want to put in their terminology are able to do that.

BT: But would not be restrictive? Would not be required?

RN: Need a description of the work product that would come out of this activity.

NOT: Proposing to come up with the list of words.

* The ones we have looked at are pilots only, to identify what is needed to let this kind of thing happen and allow the sharing to happen
* We would not be the librarians of this system

We already determined that we want to provide the recommendations for the machinery for TFs to use. Last QM presentations were a chance to see what the reaction would be from some of our target users. Without buy-in from OMG as a whole this would be different to if there were.

Clearing house – unclear how that would happen in a consistent and reasonably manageable way.

Consistency and searchability – have something that, to the extent that each TF is using this approach, they are managing the Term / Context / Concept nexus in the same way.

That is, have a common metamodel and annotation scheme that each TF would use in the same way.

Then, rather than only being able to exchange at the level of words, be able to access other TFs stuff under the same metamodel

This depends on the search criteria – what would each TF need to do to query.

For example they would need to know the distinction between usage context and original context – that is, figure out the gotchas we already figured out.

## IIC Vocabulary Task Group

BT to CB: what about the IIC initiative Claude is involved in?

Vocabulary v 2.4 or 2.4 was published 2 months ago. Led by Eric Bourneville of Dell and Erin something of NIST, and Marcella Spughiet from weegle systems in Seattle. (spellings?)

* Got vocabulary
* Got graph of dependencies among terms
* Considering sub terms
* Meets every 2 weeks on Fridays
* CB is part of this

Can we use their model?

CB had initially thought so, but this VCoI has gone more in an ontology focused direction. Initially, this above group can offer a v simple model and a mechanism to deposit those contributions in a place, and be able to export as a SOS file.

* That’s his WG

The thing being described here is the IIC group. Has basically tables, with term, comments, source e.g. adopted from, adapted from, versus internal.

No attempt to cover any of the formality we are discussing here.

### IIC Vocabulary Status

Has been published multiple times – so it has found practical use. And there is a process where the groups that are introducing a term in their work can invoke the Vocabulary Task Group and say ‘we are adding this term in one of our papers and we would like that become part of the IIC Vocabulary’.

The IIC Vocabulary TF takes that in, and looks at certain criteria as to whether to include it

e.g. if it has no unique meaning in IIC that differs from a regular dictionary, then we would not include it. E.g. ‘Real World’ is not re-defined. Sensor might be – there is a v specific meaning for that in IIC.

So there is a Process in place.

Once they decide there should be a definition the next step is to look at the definition that may have been proposed by the stakeholder group that needed the term. Either we like it or we don’t, if not, we propose a counter proposal definition, followed by a formal review process. Then there is sign-off by the CT of the IIC.

#### Questions on the IIC Vocabulary TF:

1. Are a lot of the terms we are searching for already there
2. If not accessible an ontological approach, is that something we can do?

Answers

1. No
2. Yes

e.g. on (1) the terms are very specific to IIC.

For instance would they propose a definition for ‘Blockchain’?

* May consider for v 4.0

That suggests a 3rd thing:

* Cross-availability between IIC and the OMG TFs

Likewise subsidiary Blockchain terms would not be in there

CB has done a simple flow chart of the process that is used.

Interactions defined between:

* IIC members
* The TF
* IIC as a whole

#### On (2) the Task Force:

Would welcome input on how vocab can be represented in a formal structure that would allow people to do queries – e.g. whether to do SPARQL queries versus output vocabulary itself.

Would consider as a future version (would be Major revision e.g. 4.0). Can be 2nd half of next year if we have something ready. Can then consider how to manage it using some tools.

May not need OWL / Protégé support at this stage, but have plenty of terms, synonyms etc. that would lend themselves to a SKOS codification.

Last 2 meetings, they have resolved some things about abbreviations and acronyms. Currently there is one PDF with tables, no formalisms, and no segregation of acronyms v words. Have now proposed that as a separate table. CB has said no, it needs to be in a single table, for human alphabetical search of a single resource.

**Recall:** some people are searchers and some people are navigators.

VCoI needs to recognize that also.

Eg SaaS belongs in its own entry in the table alongside full terms. Then has

* SaaS: See ‘Software as a Service’

Definition associated with a full term

Abbreviation is a link to that, in its own place in the table

### BMI Terms / Vocabuary

LT: asks CB to think about: since many of us come out of Finance etc. not BMI, and BMI has come up with a bunch of useful things. Is there a good way to extract BMI terms out of 5 or 6 recent RFIs and RFPs.

* This is not limited to BMI but there has been a lot of activity recently that includes terms of wider interest such as occurrence, Provenance and so on.
* Other domains Finance, C4I, Space etc. are likely to have their own

Erin from Dell is maintaining a table in the Cloud in which every time a term is suggested or a modification requested, she makes an entry and labels the target version of the Vocabulary and a column to capture term discussions. This drives the meeting agenda at each meeting.

* Has a process
* We can use this!

Also do online research during a meeting to update these definitions.

Vocabulary Task Group (VTG) of the IIC. TGs are subsidiary to a WG, in this case marketing WG.

## Wrapping up:

### Next meeting (19 April)

* Look at process in more detail and see what drops out
* Draw lessons from the ICC activity
* See how our ‘output’ corresponds to IIC

## AoB?

No