Q1 2023 AI PTF Teleconference Notes

A set of five teleconferences were organized in the first quarter of 2023 to make progress, with the following suggested topics:

- AI taxonomy
- A model of AI risk
- A synthesis paper of what the AI PTF has done since its inception

---

16 January 2023 Meeting

This meeting was inadvertently convened on a U.S. holiday, preventing at least one person from attending. The participants were:

- Claude Baudoin (co-chair)
- Ademola Adejokun (Lockheed Martin)
- Nick Stavros (Jackrabbit Consulting)
- Karl Gosejacob (GOSEJACOB)
- Mike Abramson (present but did not intervene)

Taxonomy Content

The Dokuwiki implementation of the current draft taxonomy within the private pages of this AI PTF wiki was demonstrated to Ademola, who expressed interest in working on this. Claude also showed the current state of the IEEE P3123 working group's AI terminology and Ademola asked for access. Claude e-mailed the Secretary of the P3123 WG to request those privileges for Ademola.

Taxonomy Representation

We discussed (in her absence) Elisa's intent to use this taxonomy as a use case for the new Multiple Vocabulary Facility (MVF) specification. Nick said that he would like to know what the desired format is, so he might write the JavaScript code to export the wiki content as an MVF file.

We also discussed the need for users who are not ontology/taxonomy experts to edit the taxonomy. This was partially based on finding that moving a taxon from one place to another is not obvious, even in the wiki. Nick should create a simple explanation of how to do that in the “Administrivia” section of the taxonomy pages.

Claude reminded others that he had created a graphical visualization of an earlier version of the taxonomy (shown as a hyperbolic tree by the Unilexicon tool), and said he could revive and update it, but that tool only has limited capabilities. This led Nick to suggest that we should capture the requirements for a good vocabulary management tool. Claude said that this is an interesting idea, but it is not specific to AI so that it would be a deliverable (discussion paper?) for the Ontology PSIG.

Risk Model

Karl said that he is interested in working on an AI risk model. This work should take into account the classification contained in the draft EU AI Act, as well as the NIST AI Risk Management Framework.

(Note that IEEE had started a study group in Q3 2022 about “AI implementation risk tiers” to respond to the NIST framework, but that effort was shut down because the study group was too small, couldn't agree on the scope, or even on what “implementation risk tiers” really meant. While this should serve as a cautionary note, it also means that there is no overlap with or competition from IEEE if we do this.)

---

30 January 2023 Meeting

The participants were:

- Claude Baudoin (co-chair)
- Ademola Adejokun (Lockheed Martin)
- Nick Stavros (Jackrabbit Consulting)
- Karl Gosejacob (GOSEJACOB)
- Alan Johnston (MIMOSA)
- Arnaud Billion (IBM France)
- Elisa Kendall (Thematix Partners)

We went through a round of introductions since we had a couple of new participants.

Arnaud Billion (LinkedIn) is a PhD in Intellectual Property Law who is an ethics advisor at IBM France. He also collaborates with Responsible Computing, an initiative launched by IBM Germany which became a managed program of OMG in 2021. Some of his work bears on whether the productions of AI systems should be copyrighted, while another area of work is captured in his book “Governed by Machines” (this refers to our reliance on prescriptive rules of governance, as opposed to natural law, not to computing machines).

Discussion on Legal Issues

Arnaud's introduction led to a discussion with Karl, who commented on his work on the IP aspects of AI, including the difference between EU and US laws, or the fact that in most countries a copyrighted work can only be licensed, not sold: the author remains the owner of the work. An exception is Switzerland, where the copyright to a work can be sold.

Arnaud thinks that OMG can play a role in “explaining to lawyers and the outside world that AI is just another form of data transformation.”

Alan said that he is regularly meeting with academics who have projects in data science and analytics that “creep” into the AI area, and they increasingly ask what can be patented. The university patent attorneys, whose job is to constantly watch for inventions that can be protected and monetized, are

struggling to find answers. This input, combined with Arnaud's and Karl's work, supports the idea of developing an OMG discussion paper on “AI and Intellectual Property” or a similar title.

Taxonomy

Claude reported that in his work with the IEEE P3123 Working Group, he added a taxonomy dimension to the terminology table under development. This simply consisted of adding a column called “Kind of” and, for now, making all forms of machine learning (supervised, unsupervised, transfer, etc.) “kinds of” machine learning – a term that was missing and that he added, together with a definition from the author who apparently invented the term.

Claude reported, before Nick Stavros joined the meeting, that work is going on to redesign the OMG wikis so that (a) there is a single installation of the Dokuwiki engine, (b) OMG members can sign in using their OMG credentials instead of us having to issue separately managed credentials by hand.

Elisa said that she wants examples of terms in order to apply the Multiple Vocabulary Facility (MVF) to them, after which she can demonstrate this. Claude said that in order to finalize enough terms, he needs to resume work on a long list of action items he has.

AI Risk and Trustworthiness

The discussion on taxonomy veered toward a discussion of risk, and this in turn led to AI trustworthiness. Claude showed a report which he just described a couple days earlier in an e-mail to the AI mailing list:

> Thanks to Clayton Pummill, who briefly attended our meetings when he was with Torch.ai and alerted me to this, I just looked through a white paper written by Jessica Newman, from UC Berkeley’s Center for Long-Term Cybersecurity (CLTC), which adds an extra dimension to the NIST AI Risk Management Framework.
>
> The report is entitled A Taxonomy of Trustworthiness for Artificial Intelligence and subtitled “Connecting Properties of Trustworthiness with Risk Management and the AI Lifecycle” (no paywall, no signing in – how refreshing!).
>
> As the subtitle indicates, the report creates a mapping between the concepts of the NIST AI RMF, in particular the lifecycle stages it defines (Plan and Design, Collect and Process Data, Build and Use Model, Verify and Validate, Deploy and Use, Operate and Monitor, Use or Impacted By) and the “characteristics of trustworthiness” (valid and reliable, safe, fair, secure and resilient, explainable and interpretable, privacy-enhanced, accountable and transparent, responsible practice and use). If you can imagine the resulting matrix of 7 stages by 8 characteristics, the author then goes on to define a set of properties within each cell of this matrix – sometimes just one property, often two to four, in one case 26 of them – for a grand total of 150 distinct properties.
>
> The report also lists (and used as inputs) a number of existing frameworks for AI trustworthiness, and specifically highlights these:
> * The “Ethics Guidelines for Trustworthy AI” from the High-Level Expert Group on Artificial Intelligence
> * The EU AI Act, which we’ve discussed several times in out OMG AI PTF meetings
> * The White House Blueprint for an AI Bill of Rights

> * The NIST AI Risk Management Framework
>
> This is not for the faint of heart (78 pages, 2 appendices, 69 footnotes…) but seems to be a really important piece of work for people interested in AI ethics and responsible computing.

---

13 February 2023 Meeting

The participants were:

- Claude Baudoin (co-chair)
- Ademola Adejokun (Lockheed Martin)
- Jürgen Boldt (OMG)
- Karl Gosejacob (GOSEJACOB)
- Elisa Kendall (Thematix)

This meeting was basically a review of ongoing action items.

Claude reported that he sent a message to the UC Berkeley Center for Long-Term Cybersecurity (CLTC) in order to invite Jessica Newman, the author of the report on a “taxonomy of AI trustworthiness” discussed last time, to speak at our March 21 meeting.

The wiki still needs work. There are four past meetings (March and September 2020 and 2021) that have not been documented, and this is still needed, even though this is in the relatively distant past, (TO BE CONTINUED)

Recap of Action Items

Several action items are mentioned in the above text.

- Claude: contact Jessica Newman at Berkeley and invite her to talk at our March 21 meeting
- Karl, Alan, Arnaud: discuss the content of a discussion paper on “AI and Intellectual Property”
- Claude: document the meetings (this page, done for the first two meetings)
- Claude: add definitions in the OMG AI taxonomy (see table of action items generated by Nick)
- Claude: update the main AI PTF wiki page to complete the calendar table with the missing meeting minutes (if they exist)
- Claude: ask Nikolay Tsanov to give Ademola access to the IEEE P3123 project workspace
- Claude: update the Unilexicon visual taxonomy to reflect the current state of our taxonomy
- Claude *and others*: brainstorm and capture requirements for a vocabulary management tool, submit to Ontology PSIG
- Nick: document how to move a taxon from one part of the tree to another
- Karl: start an AI Risk [meta?]model
- Claude: structure the wiki pages better, as the current page is starting to be too big. Create subpages for the reference architecture, the taxonomy, the risk model when it exists, and other future topics.

The next meeting is on 13 February 2023.