To: aiframework
Subject: AI RMF: some review comments
Date: Tuesday, April 5, 2022 5:26:22 AM

Greetings from France,

I am a cybersecurity researcher at Thales, specialised on risk assessment. I have read through the initial draft AI Risk Management Framework dated March 17, 2022, and found it interesting, but I have some minor comments for your consideration:

- Page 5, lines 19-21. I would recommend sticking the ISO 27001's definition of risk, i.e., the effect of uncertainty on objectives. By contrast with yours, this definition has the advantage of remaining open to positive impacts, since this is your goal with this framework. Moreover, it is the Risk Level which is a function of Severity and Likelihood, rather than the risk itself. People often confuse risk and risk level, and your statement does not help.
- Page 6, lines 12-13: I am not sure what you mean by "Some AI risks may have a low probability in the short term but have a high likelihood for adverse impacts" and how that illustrates the temporal dimension of risk. This statement opposes "probability" and "likelihood", and opposes "short term" and "adverse impacts". I'm very confused.
- Page 7, line 23: Simple English issue: "The AI RMF is *neither* a checklist *nor* a compliance mechanism…"
- Page 7, lines 35 and following: You are proposing a taxonomy, but a taxonomy of what? It is not explicit in text, even if we finally guess what you are aiming at. The current text provides a blurry feeling. My understanding is that you present a taxonomy of "Attributes of trustworthy AI". I suggest you make that (more) explicit, and that you remove blurry statements like "Trustworthy AI: risks & characteristics".
- Page 16, lines 8-9: I suggest replacing "should" by "may" in "Both qualitative and quantitative methods should be used to track risks." This comment is supported by the fact that you wrote "qualitatively or quantitatively" in Table 2.
- Page 17, Table 2, ID 2: as mentioned above, risk is by construct related to uncertainty, thus the expression "potential risk" is a pleonasm and should be avoided.
- Page 17, Table 3, ID 1: risks are prioritized based on the *severity* of their impacts, not on their impacts.
- Page 18, line 4: Why do you add "and potential impacts". The latter are, by construction, already included in the risks. Again, the expression "risks and potential impacts" is a pleonasm and should be avoided.

That's all for me, wishing you a successful publication,
Best regards,
Stéphane PAUL.

[@@ OPEN @@]