

April 2019

Follow @Paul_Hastings



Is your AI system “trustworthy”?

By [Sarah Pearce](#) and [Ashley Webber](#)

This week, the European Commission’s High-Level Expert Group on AI (“HLEG”) released the final version of its *“Ethics Guidelines for Trustworthy AI”* (the “Guidelines”). The Guidelines were first released in 2018 and this final version follows their open consultation which attracted more than 500 responses.

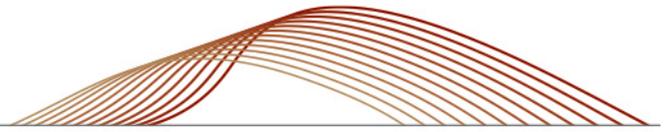
The Guidelines should be considered by organisations currently operating and managing AI tools, organisations which are in the process of creating AI tools, and those which are at the drawing board stage for their AI tool. The Guidelines are intended to apply to AI generally and as such, should be viewed as overarching principles. We say this because the reality of AI is that it covers such an incredibly broad remit of technologies, known and unknown, that one piece of guidance would not be specific enough to address the “trustworthy-ness” of each application of AI technology. It is therefore important – and indeed, best practice – that businesses should, in addition to framing their AI technology around the Guidelines, establish its own set of guiding principles which are specific to the AI it controls. Without specificity, nuances between different applications of AI and the potential risks they pose will not be appropriately handled.

The Guidelines are clear that, whilst HLEG accepts the “positive impact that AI systems already have and will continue having, both commercially and socially”, it is concerned with the “risks and other adverse impacts with which these technologies are associated”. To combat these risks, the HLEG, through the Guidelines, intends to promote the building of AI systems that are worthy of trust and states that trustworthy AI has the 3 components:

1. It should be lawful, complying with all applicable laws and regulations;
2. It should be ethical, ensuring adherence to ethical principles and values; and
3. It should be robust, both from a technical and social perspective, since, even with good instructions, AI systems can cause unintentional harm.

According to the Guidelines, the 3 components are all necessary to achieve trustworthy AI and ideally they will all work in harmony and overlap in their operation. However of course this could be problematic in practice in certain legal and social environments and HLEG seek to address this issue by stating it is “our individual and collective responsibility as a society” to try to make the components work together.

While the Guidelines are in their final form, they do include a pilot AI assessment test entitled the “Trustworthy AI Assessment List”. The HLEG invites stakeholders to take part in the test and to



provide their feedback on its implementability and completeness, amongst other things. Based on the feedback it receives, HLEG will then finalise the assessment. From an initial review of the pilot assessment, there is room for improvement. One issue will likely be raised by stakeholders: the assessment is prepared in question checklist format which, whilst many of the questions cover areas which every stakeholder definitely should take into account, as noted above, not all AI systems are the same and therefore not all questions will in fact have been, or should in have been, considered by the stakeholder when creating and/or operating the AI system. The result for such a stakeholder is that it could appear as though their AI system is not trustworthy when in fact for its particular circumstances, it is. We look forward to HLEG producing the final version of the assessment in 2020 to see whether and to what extent any substantial changes have been made.

◇ ◇ ◇

If you have any questions concerning these developing issues, please do not hesitate to contact any of the following Paul Hastings lawyers:

London

Sarah Pearce
44.020.3023.5168
sarahpearce@paulhastings.com

Ashley Webber
44.020.3023.5197
ashleywebber@paulhastings.com

Paul Hastings LLP

Stay Current is published solely for the interests of friends and clients of Paul Hastings LLP and should in no way be relied upon or construed as legal advice. The views expressed in this publication reflect those of the authors and not necessarily the views of Paul Hastings. For specific information on recent developments or particular factual situations, the opinion of legal counsel should be sought. These materials may be considered ATTORNEY ADVERTISING in some jurisdictions. Paul Hastings is a limited liability partnership. Copyright © 2019 Paul Hastings LLP.