BRIEFING NOTES

BN-59-The role of AI-May2021

REQUIREMENTS FOR LIABILITIES OF AI SYSTEMS

Authors: **:** Mohamadreza Nematollahi[1] and Kash Khorasani[2]

1 Graduate student, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada
2 Professor, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

SUMMARY

- Liabilities of designers and agencies that use Artificial Intelligence (AI) systems should be properly addressed by law makers.
- This should be addressed as soon as possible, since any ambiguity in it impedes the path toward next generation of AI systems, and stifle investments in this field.
- Design failure lawsuits in case of AI systems are not the appropriate mechanisms and path to manage the concerns and issues. Treating the AI systems in the same way as humans that make mistakes more fits the context, along with considering the similarities and differences between humans and AI systems.

CONTEXT

- Notwithstanding designers and agencies that use AI systems are subject to judgments by law because of their products or services and lawsuits, there should nevertheless at least minimally be a path in testing and rolling out at least limited versions of the technology.
- Hallevy's article [1] addressed the cases where the AI systems are directly liable. To hold an AI entity liable by law, one needs to determine the existence of the requirements, actus reus as the necessary condition, and mens rea if applicable based on the context.
- Usually, attributing the required mens rea to AI systems is much more challenging. AI systems may resemble most of the human cognitive capabilities and also creativity even in a more efficient form, but those are neither sufficient nor excluding for direct criminal liability of the entity. To hold an entity liable, one needs capabilities to acquire the related form of knowledge, generate criminal intent or negligence, or other mentally related capacities.
- The knowledge acquisition capability of an AI entity should be related to the criminal context. It should contain the minimum number of required sensory systems and required inference capabilities, which in most of the cases these capabilities well resemble the human counterpart while regarding the intent generation, AI systems may lack major human capabilities, and hence it is not an easy and trivial task to attribute AI systems, a specific criminal intent or negligence, at the same time, there is no reason to not being able to do so; hence a more detailed analysis is required.
- The author in [1] also compared the case of AI systems with those class of humans exempted from being liable with respect to criminal laws and concluded that in the cases where the primary requirements of the criminal liability exist in an AI entity, none of those exemptions are applicable.
- In cases where humans are liable, determining the punishments is trivial and is based on the currently available methodologies and is not subject to this study; however, when one holds an AI system liable, what are the punishments?

CONSIDERATIONS

- Design failure lawsuits, which are common due to high costs, ambiguities, and that can further negatively impact companies represent as companies' nightmare and discourage them in investing on certain new technology, since it is not easy to ensure and protect against all the responsibilities for the product.

- Due to those considerable impacts, liability models should be addressed as soon as possible, since any ambiguity in it would result in a lag in the path towards the next generation of AI systems and attraction of new investments.

- Most of the AI systems issues are patchable, and can be readily solved even without recalling the products. This gives another perspective on inappropriateness of impact of design failure lawsuits, which usually forces the companies to recall their products.

- Hallevy believes punishments are adaptable to the case of AI systems, in the same way as we have adaptations of the human criminal punishments for liable agencies or companies, and proposed some of the possible adaptations. However, we believe the proposed model has some major deficiencies since, in contrast with human criminals, which will be considered free creatures responsible for their own faith, the AI system is machines owned by the users.

- Hence applying some punishment is not related in this case, since they may result in financial loss. Simultaneously, although the quality of cognitive properties is not a matter in determining liabilities, they are one of the parameters that should be considered in defining proportional punishments, such that it effectively prevents future offenses while also not to cause others harm.

- Going from design failure lawsuits to Hallevy's model of criminal liabilities is quite a significant step. However, Hallevy's article looked at the subject mostly from a lawmaker view point based on a general understanding of the AI systems and lacked the capability to handle some technical issues. One such problem has been addressed in [2], is related to the cybersecurity of AI systems; what if a software virus causes an AI system to commit a crime?

- The author in [2] also did not address the required steps in implementing this framework, which should be addressed by policymakers in a way that it guarantees a smooth path toward the maturity of both the liability framework and AI technologies.

- Note that addressing the criminal liabilities alone cannot ensure safe interaction of AI systems with human society. Hallevy's article is open concerning some technical standards required by the definition of safe operation of the AI systems and defining the developers' and users' negligence, and as the first step, one needs to properly define the required standards of operation and training process of AI systems. This step itself requires to provide an AI friendly environment in which developers could release the early versions of their products for real-world experiments.

- At the same time, policymakers should facilitate their path in training AI systems, since in order to achieve the required level of standards, developers should access to informative

enough datasets, which otherwise are only accessible to some limited number of big tech companies.

NEXT STEPS

- Treating AI systems in the same way as treating humans making mistakes, more fits the context of this issue. Hence, instead of design failure lawsuits, negligence civil laws and criminal law models should be properly adapted to fit this context.
- In doing so, limitation and also differences of AI systems as compared with humans such as lack of general knowledge about the contexts in AI systems should be addressed by lawmakers and policy makers. These limitations have to be properly communicated to the users, otherwise the designers are responsible.
- In case of liabilities, AI systems should be considered as an innocent person, and responsibilities should be properly divided among all the stakeholders and agencies that are liable.

REFERENCES

[1] Gabriel Hallevy. The criminal liability of artificial intelligence entities-from science fiction to legal social control. *Akron Intell. Prop. J.*, 4:171, 2010.

[2] Kingston, John KC. "Artificial intelligence and legal liability." *International Conference on Innovative Techniques and Applications of Artificial Intelligence*. Springer, Cham, 2016.