The Challenges For Criminal Law in Facing The Passage From Technological Automation to Artificial Intelligence

We robot 2021 -

The possible and future criminal policies regarding the regulation of harms related to the use and functioning of AI systems

Beatrice Panattoni, University of Verona

Objective

AI, as part of the conceptual shift brought about by ICT, amplifies certain conceptual gaps. The distance between human actions and their consequences increases and responsibility models based on the category of "control" become out of date. The research aims to analyze and elaborate a set of possible answers to the conceptual gaps that the features of AI systems, such as "emergence", create to criminal law traditional categories and responsibility models.

Setting the background: "AI Crimes"

Since AI applications have already and will make it possible to realize new activities and to reach new goals, it is likely that new ways of realizing existing criminal offences will begin to occur and that also new forms of crimes will ask for the intervention of the legislator in criminalizing them.

Technically-oriented classification of "AI crime"

(1) **Malicious uses of AI**. Cases where the AI system is used by a criminal agent as the means to realize the crime

a. Present threats: AI malware; AI and social engineering (DeepPhis AI); AI and human impersonation (social bots); "criminal robot"

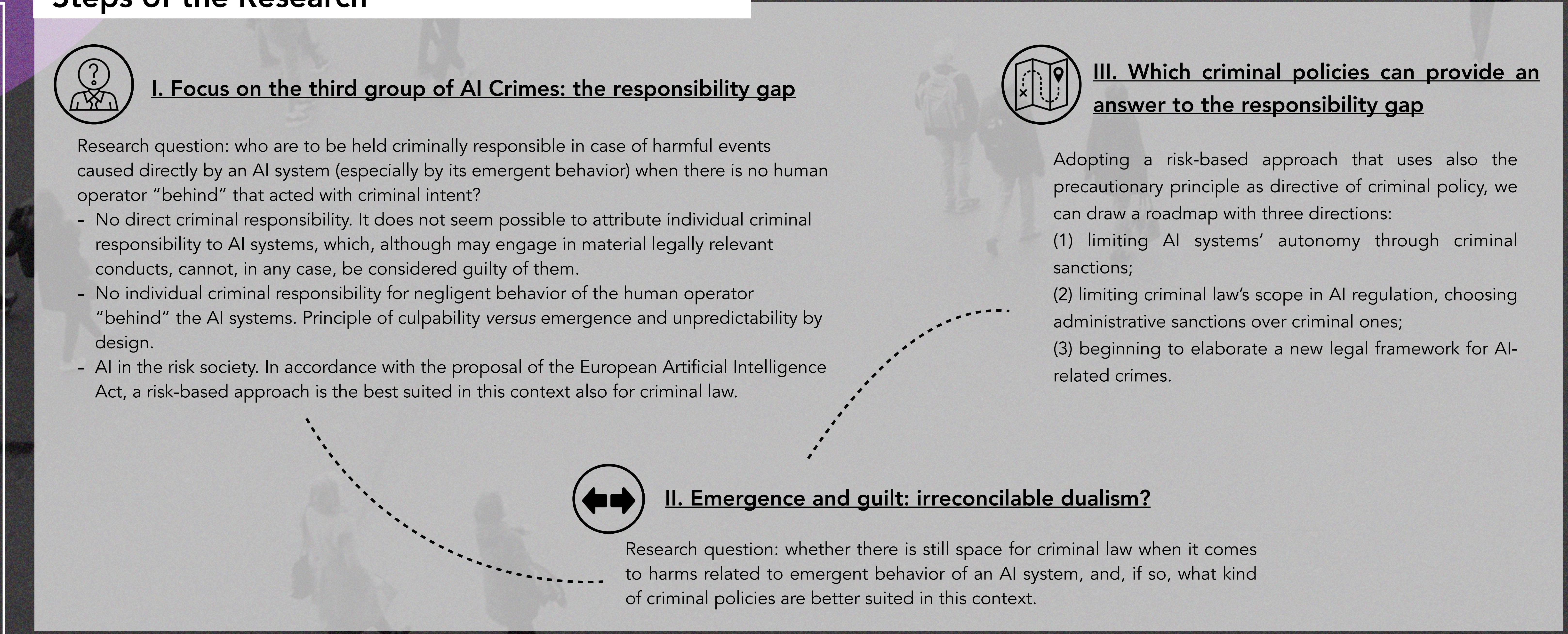
b. Future threats: social engineering at scale; data breach 2.0 (document-scraping malware); AI-aided stock market manipulation; deep fakes

(2) **Abuses against AI.** Cases where the AI system is the "object" against which is committed the crime

a. Present and future threats: abusing Smart Assistants; abuses against Image Recognition Systems ("tricking" AI systems)

(3) **AI crimes in a strict sense.** Cases where the harm is "committed" directly by an AI systems, including the case where the harmful event is caused by its emergent behavior

Steps of the Research



I. Focus on the third group of AI Crimes: the responsibility gap

Research question: who are to be held criminally responsible in case of harmful events caused directly by an AI system (especially by its emergent behavior) when there is no human operator "behind" that acted with criminal intent?

- No direct criminal responsibility. It does not seem possible to attribute individual criminal responsibility to AI systems, which, although may engage in material legally relevant conducts, cannot, in any case, be considered guilty of them.
- No individual criminal responsibility for negligent behavior of the human operator "behind" the AI systems. Principle of culpability *versus* emergence and unpredictability by design.
- AI in the risk society. In accordance with the proposal of the European Artificial Intelligence Act, a risk-based approach is the best suited in this context also for criminal law.

II. Emergence and guilt: irreconcilable dualism?

Research question: whether there is still space for criminal law when it comes to harms related to emergent behavior of an AI system, and, if so, what kind of criminal policies are better suited in this context.

III. Which criminal policies can provide an answer to the responsibility gap

Adopting a risk-based approach that uses also the precautionary principle as directive of criminal policy, we can draw a roadmap with three directions:
(1) limiting AI systems' autonomy through criminal sanctions;
(2) limiting criminal law's scope in AI regulation, choosing administrative sanctions over criminal ones;
(3) beginning to elaborate a new legal framework for AI-related crimes.

Conclusion

Forms of distributed legal responsibility, elaborated on a risk-based approach that keeps the principle of accountability on its core, represent the most suitable option in the context of AI. Proposal *de jure condendo*: resorting to corporate criminal laws in regulating potential threats posed by high risk AI systems to fundamental human rights and legal goods that need criminal law protection.