

Károli Gáspár University of the Reformed Church in Hungary

Doctoral School of Law

Head of the Doctoral School:

Prof. Dr. Éva Jakab PhD

Professor, Doctor of the Hungarian Academy of Sciences



Liability issues in the light of technological developments – criminal liability and liability for damages for artificial intelligence

Doctoral Dissertation Thesis

dr. Tamás Bicskei

Témavezetők:

Prof. Dr. Andrea Domokos PhD

Vice Dean for Talent Management

Head of Department, Professor

Prof. Dr. Ádám Boóc PhD

Vice Dean for International and Scientific Affairs

Head of Department, Professor

Budapest, 2024

Table of contents

I. Outline of the research objective	3
II. Description of the research, investigation and analysis carried out, description of the method and sources used, structure of the dissertation.....	4
III. Thesis summary of new scientific findings	7
IV. Full list of publications of the doctoral candidate	14

I. Outline of the research objective

The volume of artificial intelligence (AI) developments is on a waving trend. Their historical origins date back to the mid-20th century, but after a long hiatus, the development of these technologies has gained momentum in the early 21st century. High-level research is currently underway in a number of application areas, with deep learning AI systems in particular expected to deliver the best results.

But artificial intelligence is no longer just a technology that exists in the scientific world and has no impact on the daily lives of citizens. AI is permeating our daily lives without our noticing: machines based on self-learning recommend movies, calculate optimal transport routes, produce high-quality translations. Developments are also extending to robotic devices such as self-driving cars, autonomous drones and asset protection devices.

The development methods of AI systems represent a qualitative difference compared to traditionally built algorithms. They are able to process extremely large amounts of data, revealing many more patterns and correlations than human cognition or traditional software due to the larger sample size. It should be noted, however, that the use of the term intelligence is somewhat misleading when describing these tools: they are certainly not intelligent in the human sense, and their operation can best be described as highly efficient data processing.

AI systems make life easier for the citizens who use them, make society work more efficiently, and make some dangerous areas (e.g. road transport) safer. But as the famous saying goes, "what can go wrong, will go wrong." This is doubly true for AI systems: like any other product or service, AI systems can be developed incorrectly, resulting in damage and material loss during operation. From a legal point of view, however, it is the second type of defect that is of real interest. The black box effect is a phenomenon observed in deep learning systems that fundamentally orients liability thinking. Due to the nature of the technology, the developer only knows the results of the input and the output. The process by which the developed machine connects the desired output to the input, based on which method and logical operations, cannot be known. Therefore, the inner workings of the technology, the specific operations that the machine will perform in response to the instruction, cannot be known. In hindsight (e.g. if the operation of the machine has caused damage), it is possible to discover how the AI system got to the output. Unlike traditionally developed algorithms, however, no one - neither the developer of the system, much less the user - can predict with 100% certainty in advance what decision the machine developed using deep learning will make.

So there is the problem, which is discussed in international and Hungarian legal literature. In Hungarian, several monographs have been published that analyse the regulatory and liability aspects of AI, and there are more than a hundred studies published in Hungarian scientific journals that deal with similar issues. The Hungarian government has adopted a National Digitalisation Strategy and an Artificial Intelligence Strategy. European legislation is also actively addressing the issue. As the manuscript is being finalised, the European Parliament has adopted its negotiating position on the AI Regulation, which is expected to be adopted by the end of 2023. The European Commission has submitted proposals for two directives containing rules governing MI systems in the areas of non-contractual liability and product liability. In addition, the European Union is trying to dominate the discourse on the regulation of MI through a series of background papers, ethical guidelines, impact assessment studies, reports and other soft law instruments. Of course, the research area is also very popular in foreign legal literature, especially in the Anglo-Saxon language area, with numerous academic articles dealing specifically with the legal aspects of AI.

The dissertation partly connects to the academic discourse described above, but partly takes a different approach to the issues raised. It is not the aim of the thesis to present the regulatory issues, processes and tools related to AI. As such, the paper aims to provide a dogmatic analysis of the issues raised by legislation and scientific developments, and to present the related regulatory issues only to the extent necessary. The novelty of the thesis lies, on the one hand, in the fact that a monograph focusing on liability dogmatics and dealing exclusively with artificial intelligence has not yet been published on the subject. The novelty of the study is also due to the fact that the dissertation discusses in parallel the issues of criminal and civil liability, which are interrelated. However, the AI aspects do not justify treating the area as a separate branch of law or as a separate area of legal doctrine requiring a specific approach.

The aim of the research is thus to integrate AI into existing criminal and civil law doctrines with as little confusion and friction as possible, and to present useful doctrinal solutions for both the law practitioner and the legislator.

II. Description of the research, investigation and analysis carried out, description of the method and sources used, structure of the dissertation

The starting point was to identify the problem posed by the technology. First, the exact nature of AI needs to be described and appropriate legal concepts for each technological system need

5 to be established. By defining at the highest possible level of abstraction, the AI system, which otherwise exists as a technological phenomenon, can be transformed into a legal concept, after which the legal concept can be embedded in the system of liability doctrine and further dogmatic statements can be made on this basis.

The paper is thus the first to attempt to define artificial intelligence as a legal concept. For the sake of dogmatic usability, the conceptualisation is responsibility-oriented, approaching AI systems not from a technological point of view, but on the basis of their functionality. In this way, three categories of AI can be defined, highlighting that in fact only one of these categories needs to be examined in more detail.

Following the legal transformation of AI, it is also necessary to define the operation of AI as a legal fact from a liability perspective. The different aspects of the functioning of AI deserve different liability assessments and therefore need to be categorised in terms of autonomy and unpredictability of the functioning.

The definition of the concept of AI and the functioning of AI is dealt with in Chapter II. Chapter II also deals with the legal personality and autonomous responsibility of MI, with the legal doctrinal analysis of this issue also being dealt with in Chapters IV and V

Chapter III, after defining the definitions, undertakes a closer look at some AI systems, proving the hypothesis that the abstract concept of an AI system can be applied regardless of the field of application of the technology. The core of the liability problem is also the same for all such technologies: it can be captured in the uncertainty that arises from unpredictability. Finally, three technologies (drones, stock market algorithms, medical diagnostic tools) are analysed in this thesis. The study does not cover additional technologies due to scope limitations, and the author has tried to select tools with a lower level of literature coverage compared to e.g. selfdriving cars or text generator AI, and the author has investigated these systems in more depth in his research at doctoral school.

Finally, Chapters IV and V seek to situate AI in the doctrine of criminal and civil liability. The author's aim is to make explicit doctrinal statements that are of use to both law enforcement and legislation. Of course, the mass emergence of the liability problems raised is not yet expected in the near future, but the author argues that the confrontation of the legal practitioner with liability problems related to the unpredictable AI system will be inevitable.

The investigation has led to different results as regards criminal liability and liability for damages. The author argues that there is a case for revising the criminal law rules in order to reconcile the operation of the unpredictable AI with the objectives of criminal law. To this end, Chapter IV contains a *de lege ferenda* proposal to resolve the discrepancy.

With regard to liability for damages, it did not seem appropriate to legislate to resolve the doctrinal issues raised, which could be resolved by means of case law. It is, however, justified to reconsider the doctrinal framework in the area of causation and foreseeability, and to broaden it slightly. Chapter V therefore examines the relevant parts of the liability rules in detail, bringing them into line with trends in European legislation and providing guidance for the legal practitioners in dealing with the problems.

The main research method was the dogmatic method, as described above. In addition, a legislative proposal was drafted, preceded by a public policy analysis, both in the area of criminal and civil liability. Finally, the thesis also draws on an empirical research method, which is reflected in an interview with a practitioner with in-depth knowledge of the development of a deep learning MI system. This research method was based on a qualitative methodology and was used to provide the thesis with a more precise understanding of the characteristics of MI systems.

The problem statements (and also the hypotheses) of the research can be defined as follows:

1. The main specification of artificial intelligence is the possibility of autonomous operation, independent of the creator and the user
2. The autonomous functioning of AI is a doctrinal problem to be addressed either by legislation or by legal theory, to increase legal certainty.
3. The legal-dogmatic concept of AI should be defined not by a nominal definition of technical description, but by a functional approach. Beyond the definition of AI, it is also necessary to define the functional definition of AI.
4. An unnecessarily disruptive solution to legal doctrine is to establish the autonomous legal personality of AI. The ultimate subject of liability is the human being or the legal person created by the human being, so that to sanction AI on its own would either be ineffective or would in fact indirectly penalise the human being or legal person behind it
5. A correctly constructed legal doctrinal concept of artificial intelligence can be applied to liability doctrine regardless of the areas of use of the technology. It is not necessary

to define artificial intelligence used in self-driving cars, autonomous drones or other technologies.

6. In criminal liability doctrine, the problem is the ascertainability of the guilt of the perpetrator. The contradiction between the principle of guilt and the unpredictability of the operation of the system is a tension that can be resolved by legislation.
7. With regard to liability for damages, it is not a question of fault, but of causation, to examine the contradiction between unpredictable operation and foreseeable damage. A partial rethinking of the legal framework, complemented by European legislation on the presumption of causation, provides a solution to this problem.

In terms of the use of sources, in Chapters II and III, mainly English-language studies and monographs in the field of jurisprudence have been used, with the aim of presenting and processing all significant works of domestic authors on AI systems. In these chapters, related non-legal, but also IT and technical articles are necessarily included.

In Chapters IV and V, in which the aim of the thesis was to transpose the findings on the MI systems into the domestic criminal law and tort liability systems, the text mainly includes legal studies and monographs in Hungarian dealing with liability and certain aspects of liability. In these chapters, the works of international authors, such as American, British, German and French authors on criminal law and tort liability, have also been used.

In addition to legal and technical academic sources, the thesis also draws on the European legislator's acts, preparatory documents, and academic materials prepared by expert groups commissioned by the European institutions, and the relevant case law is presented in the liability chapters.

III. Thesis summary of new scientific findings

The main findings and results of the dissertation can be summarised as follows, along the hypotheses set out.

1. The main specification of artificial intelligence is the possibility of autonomous operation, independent of the creator and the user

This autonomous operation is the main characteristic of AI systems. At a theoretical level, three categories of AI can be distinguished: general AI, pre-coded algorithms, and AI based on machine learning and deep learning. Only the third category is of particular relevance for

jurisprudence. General AI can only be assessed as a theoretical category and, since it is probably not possible to build such a system, it is not necessary to consider the implications of liability. The operation of precoded algorithms (expert systems) is determinate and knowable, and therefore they do not have the characteristic feature of AI, i.e. autonomous, unpredictable operation. Their dogmatic role therefore consists in defining their operation in relation to 'real' AI, i.e. AI based on machine learning.

Thus, legal doctrine must first and foremost create and apply a definition of AI as an autonomous system, which overlaps with deep learning systems, but is not fully compatible with them. However, whether or not the AI system in question can be classified as autonomous and unpredictable will be of primary importance for the resolution of a dispute. The dogmatic insights discussed in this thesis are therefore only relevant in the case of "real" AI based on deep learning, whereas in the case of systems that can be considered as precoded algorithms, the legal issues arising can be decided without any dogmatic additions to the AI.

2. The autonomous functioning of AI is a doctrinal problem to be addressed either by legislation or by legal theory, to increase legal certainty.

The doctrinal problem can simply be described as the unpredictable, unforeseeable way in which the AI system operates, but the fact that the person liable must be aware of the fact of his conduct and the consequences of that conduct in order to establish liability. The question can in fact be resolved in two ways in the case-law: either the court can refuse to impose liability on the grounds of lack of fault or causation, or it can impose it by broadening the doctrinal concepts. The legislator should consider whether to adapt positive law to the challenges posed by AI systems in order to provide clarity or to leave it to the legislator to resolve.

According to the findings of the dissertation, legislation and the introduction of a new definition of criminal liability will probably be necessary, as the dogmatics, taking into account the constitutional requirements, either leaves behaviour dangerous to society unpunished or establishes guilt without taking into account the principle of guilt.

With regard to civil liability, the doctrine may be able to adapt to the challenges of the MI, but the legislation will provide presumptions to assist victims in pursuing their claims in the light of the MI Liability Directive, essentially for consumer protection reasons. Overall, the

challenge posed by the MI system cannot be addressed without changes in legal doctrine or legislation.

3. The legal-dogmatic concept of AI should be defined not by a nominal definition of technical description, but by a functional approach. Beyond the definition of AI, it is also necessary to define the functional definition of AI.

The functional approach means that it is not appropriate to define AI in terms of liability according to the technical and technical characteristics of the system - as defined in the unadopted version of the AI Regulation - but to do so explicitly along the lines of unpredictability and autonomy. Thus, in legal doctrine terms, a device is considered to be AI if it is capable of autonomous, unpredictable operation. This distinction may also be relevant in the context of a dispute. A given physical device usually contains several algorithms at the same time, such as machine learning based AI and pre-coded conventional algorithms. Under the concept of AI described as a nominal definition, the whole device cannot be considered as AI, only the part of it that is specifically based on machine learning. However, under the functional approach, the whole device can be considered as an AI system and no distinction is made between its subsystems operating as individual components.

The distinction between the way it operates is also important, because just because the AI system is capable of unpredictable operation does not mean that every decision it makes will be so. Therefore, from a liability point of view, a distinction is made between truly autonomous operation and the case where this is not the case, e.g. because the device has been given a specific instruction to perform a particular operation and has complied with it.

4. An unnecessarily disruptive solution to legal doctrine is to establish the autonomous legal personality of AI. The ultimate subject of liability is the human being or the legal person created by the human being, so that to sanction AI on its own would either be ineffective or would in fact indirectly penalise the human being or legal person behind it

A popular solution to the doctrinal challenges posed by MI in the legal literature is to recognise the legal personality of MI. Once legal personality is recognised, the MI system itself would be

subject to liability and could be subject to enforcement action, and the MI would be directly punished by law.

The criticisms of these views in the dissertation can be divided into two pillars. On the one hand, there are certain ethical considerations that hinder the recognition of the personality of the MI. Although these are not strictly speaking legal-dogmatic arguments, but rather legalpolitical ones, their weight should not be underestimated. Recognition of the legal personality of the MI would not reflect the reality of the situation, since, despite its unpredictability and autonomous way of functioning, it is far from being a free choice for the MI system. Unlike the human actor, the machine can only operate within the constraints imposed by its programming; unlike the human, it has no 'free' will in the legal-dogmatic sense, leaving aside the considerations of scientific theory. The 'perception' and 'consciousness' of the AI system is also far below that of primate animals, and a convincing argument could be made for the dangerousness of animals in terms of the potential for harm: before recognising the legal personality of AI, it might even be justified to recognise the legal personality of animals. Unlike legal partnerships, the recognition of the legal personality of an AI would be a mere liability convention and would not reflect the organisation and separation of its members. Furthermore, certain other ethical and religious considerations, stemming from western legal culture, militate against the recognition of legal personality.

The other argument against recognising legal personality is that it would ultimately become just another technical legal step in the chain of liability, and the sanction would end up being imposed on the owner or holder of the MI, or possibly its creator. If a pecuniary sanction were imposed on the machine as a legal entity, it would be the owner of the machine who would ultimately have to pay. The situation is similar in the case of insurance, in which case the owner of the MI would ultimately be liable to pay the insurance premium. Likewise, any criminal sanctions that might arise, such as the destruction, forced reprogramming or limitation of the operation of the machine, would necessarily disadvantage the owner. In view of the above, it is neither timely nor justified to recognise the autonomous subjection of the MI, as this would not effectively solve the liability problem.

5. A correctly constructed legal doctrinal concept of artificial intelligence can be applied to liability doctrine regardless of the areas of use of the technology. It is not necessary

to define artificial intelligence used in self-driving cars, autonomous drones or other technologies.

With regard to the systems presented in Chapter III, it can be concluded that, apart from certain specific features, their uniform assessment from a doctrinal point of view is justified. Chapter III is intended to demonstrate that, irrespective of the different areas of operation and regulatory environments, the problems of liability arise in the same place because of the identity of AI as a legal-dogmatic concept. Accordingly, the analysis in Chapters IV and V will be technologyneutral and will examine the relevant doctrinal issues solely in the light of the unpredictability that characterises all AI.

6. In criminal liability doctrine, the problem is the ascertainability of the guilt of the perpetrator. The contradiction between the principle of guilt and the unpredictability of the operation of the system is a tension that can be resolved by legislation.

From the point of view of the doctrine of the act, autonomous AI operation cannot be considered an act without further reduction of the concept of act, in agreement with the Hungarian legal literature. Therefore, according to the argumentation of this thesis, the legislator must declare not the machine operation, but the voluntary and effective human behaviour expressed before it to be punishable in order to establish liability without dogmatic problems.

In the area of culpability, in some cases, AI systems operate in such an unpredictable and unpredictable way that negligence cannot be established with regard to the offending outcome. Legal doctrine either recognises the impunity of certain socially dangerous behaviour, subject to the principle of culpability, or accepts the criminality of certain offences in the absence of the perpetrator's consciousness and culpability. It is up to the legislator to construct this, in which the offending mechanical operation would not be a result but an objective condition of criminal liability. This solution would not constitute an insurmountable doctrinal obstacle, since objective conditions of criminal liability have been part of criminal law for decades, as formulated in various factual situations. ¹² The theory of *actio libera in causa* would provide a solution to both the problem of the theory of acts and the problem of culpability as regards criminal responsibility for the illegal operation of AI systems. However, this can only be enforced by legislation on the basis of the principle of legality. The paper argues that in the case of the most serious offences of unpredictable AI, it is justified and necessary to establish the liability of the offender by breaking the principle of culpability to the extent that if the offender

is guilty of a breach of the rules on AI, he or she should be punishable for the serious events that occur thereafter.

7. With regard to liability for damages, it is not a question of fault, but of causation, to examine the contradiction between unpredictable operation and foreseeable damage. A partial rethinking of the legal framework, complemented by European legislation on the presumption of causation, provides a solution to this problem.

With reference to Chapter V, the legal-dogmatic analysis of liability for damages is not based on imputability, but on causation. The new Civil Code makes foreseeability one of the, but not the only, causal criterion in the causation context.

The standard of foreseeability is not to be assessed in relation to the specific tortfeasor, but in relation to an ideal person acting with due care. This standard also applies to the assessment of harm caused by an AI system, i.e. the court must assess whether either the developer or user who exercised due care or, in a slightly more formalised approach, the AI system exercising due care, could have foreseen the harm at the time the harm occurred.

Although foreseeability is theoretically a criterion to be examined in the context of liability for dangerous industrial activities, in the light of the relevant legal literature and judicial practice, damage is always considered foreseeable if the dangerous industrial nature of the conduct causing the damage is established. The same is true for claims for damages caused by an AIS of a dangerous operational nature, i.e. in fact the doctrinal problem arising from the unpredictable operation of the machine only appears at a theoretical level, and in practice it is presumably not to be examined by the legal practitioner.

On this basis, the dissertation argues that there is no need to change the existing legal framework by legislative means: legal doctrine is sufficiently flexible to allow the legal practitioner to assess claims for damages caused by AI systems.

The paper describes in detail the developments in European legislation on compensation, in particular the draft MI Liability Directive. The paper points out that the Directive, presumably driven by consumer protection considerations, will facilitate the enforcement of claims by victims by means of evidence, notably by introducing two rebuttable presumptions. As regards the wording and terminology of the Directive, it can be noted that transposition problems could

easily arise due to the different liability rules in Hungarian tort law compared to most Member States, in particular the provisions on exculpatory fault.

Finally, as a kind of "veterinary horse", I presented the liability for damages caused by the exercise of public authority combined with the liability for the AI system, describing the enforcement challenges that Hungarian and presumably other European law enforcers can expect in the near future.

IV. Full list of publications of the doctoral candidate

Publications on the subject of the thesis:

1. Bicskei Tamás: Liability problems arising from the nature of artificial intelligence. In: Miskolczi-Bodnár Péter (szerk.): *XVI. Jogász Doktoranduszok Szakmai Találkozója 2020*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2020. 35-50.
2. Bicskei Tamás: The ethical issues of artificial intelligence and criminal liability. In: Miskolczi-Bodnár Péter (szerk.): *XVII. Jogász Doktoranduszok Szakmai Találkozója 2020*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2020. 59-68.
3. Birher Nándor – Bicskei Tamás – Kovács Viktória: *Regulatory complexity*. *Glossa Iuridica*, 2020/7.1-2. 285-313.
4. Bicskei Tamás – Rideg Gergely: Artificial intelligence and drones: regulatory and liability issues. In: Miskolczi-Bodnár Péter (szerk.): *XVIII. Jogász Doktoranduszok Szakmai Találkozója 2020*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2020. 21-42.
5. Birher Nándor – Bicskei Tamás – Kovács Viktória: Law, morality, religion. In: Birher Nándor – Homicskó Árpád Olivér (szerk.): *Szabályozáskomplexitás*. Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2021. 23-56.
6. Bicskei Tamás – Kovács Viktória – Rideg Gergely: Regulatory complexity: extra-legal regulatory instruments and the rule of law. In: Birher Nándor – Homicskó Árpád Olivér (szerk.): *Szabályozáskomplexitás*. Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2021. 57-98.
7. Bicskei Tamás – Rideg Gergely: The legal subjectivity of artificial intelligence in the light of different cultures - the problem of independent criminal liability. In: Miskolczi-Bodnár Péter (szerk.): *XX. Jogász Doktoranduszok Szakmai Találkozója 2021*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2021. 9-18.
8. Bicskei Tamás – Rideg Gergely: Responsibility for and regulation of the use of artificial intelligence in financial markets in Europe. In: Miskolczi-Bodnár Péter (szerk.): *XXI. Jogász Doktoranduszok Szakmai Találkozója 2021*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2022. 9-16.

9. Bicskei Tamás: The Usability of Artificial Intelligence in the Judiciary. In: Miskolczi-Bodnár Péter (szerk.): *XXIII. Jogász Doktoranduszok Országos Konferenciája*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2022. 25-32.
10. Bicskei Tamás: *The damage caused by the use of artificial intelligence in public administration*. *Közigazgatás Tudomány*, 2023/1. 99-114.
11. Bicskei Tamás: Reflections on the harms caused by artificial intelligence in the administrative sphere in the light of European legislation. In: Jakab Éva – Miskolczi-Bodnár Péter (szerk.): *XXV. Jogász Doktoranduszok Országos Konferenciája*. Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2023. 45-52.
12. Bicskei Tamás – Rideg Gergely: Artificial Intelligence and Medicine, liability and regulatory issues. In: Rusinné Fedor Anita (szerk.): *Tudomány: út a világ megismeréséhez*, Debreceni Egyetemi Kiadó, 2024. 14-26.

Publications in other fields:

1. Bicskei Tamás: Fundamental Aspects of the Protection of the Right to Life in Civil Law. KRE-DIT, 2022/1. <https://www.kre-dit.hu/tanulmanyok/bicskei-tamas-az-elethez-valo-jog-polgari-jogi-vedelmenek-alapjogi-aspektusai/>
2. Bicskei Tamás: Usability of algorithms in criminal justice. In: Miskolczi-Bodnár Péter (szerk.): *XXII. Jogász Doktoranduszok Konferenciája*, Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar, Budapest, 2022. 21-30.
3. Bicskei Tamás: *The invalidity of foreign exchange contracts: the interests of the consumer in relation to the legal consequences of unfair contract terms in the light of case law decisions and DH law*. EUB 70 tanulmánykötet, OBH, 2023.