UDC 341.171
DOI https://doi.org/10.32782/2524-0374/2024-4/23

INFRINGEMENT OF INTELLECTUAL PROPERTY RIGHTS IN THE COURSE OF ADVERSARIAL ATTACKS ON ARTIFICIAL INTELLIGENCE SYSTEMS FROM THE PERSPECTIVE OF EUROPEAN UNION LAW[1]

ПОРУШЕННЯ ПРАВ ІНТЕЛЕКТУАЛЬНОЇ ВЛАСНОСТІ ПРИ ЗДІЙСНЕННІ ЗМАГАЛЬНИХ АТАК (ADVERSARIAL ATTACKS) НА СИСТЕМИ ШТУЧНОГО ІНТЕЛЕКТУ З ТОЧКИ ЗОРУ ПРАВА ЄВРОПЕЙСЬКОГО СОЮЗУ

Bohatchuk D.P., PhD in Law,
Senior Lecturer at the Department of General Theoretical Jurisprudence and Public Law
National University of "Kyiv-Mohyla Academy"

Artificial intelligence is having an increasing impact in many areas. At the same time, artificial intelligence systems can be targeted by so-called "adversarial attacks". Appropriate measures are needed to counter adversarial attacks on artificial intelligence, including with the help of law.

The paper aims to analyze whether adversarial attacks on artificial intelligence systems constitute infringement of intellectual property rights to such systems or elements thereof. The paper focuses on copyright law, patent law, and *sui generis* database protection and examines how intellectual property rights can be infringed in the course of conducting adversarial attacks.

It is concluded that the copying of artificial intelligence system or its parts and components (if they are copyrighted) in the course of an adversarial attack may qualify as reproduction and constitute copyright infringement if certain conditions are met. In this context, it is important to take into account that the act of reproduction committed for adversarial attack generally cannot be considered as accidental, but on the contrary, may be considered as intentional.

When considering possible patent infringements in the context of adversarial attacks, in particular, the doctrine of equivalents and the principle of "exhaustion of rights" must be taken into account.

The *sui generis* database right, which is an additional protection possibility for databases in the European Union, may also be infringed within the adversarial attacks under certain conditions.

The need to include provisions on the unlawful character of adversarial attacks in the relevant legislation is determined. The provisions prohibiting adversarial attacks on artificial intelligence in the license agreements can also increase the level of legal protection of artificial intelligence systems.

Key words: artificial intelligence, adversarial attacks, copyright law, patent law, sui generis database right, intellectual property law.

Штучний інтелект посилює свій вплив у різних сферах. Водночас, системи штучного інтелекту можуть зазнавати так званих «змагальних атак» («adversarial attacks»). Є потреба у відповідних заходах для протидії «змагальним атакам» (adversarial attacks) на штучний інтелект, у тому числі за допомогою права.

Стаття має на меті проаналізувати, чи є змагальні атаки (adversarial attacks) на системи штучного інтелекту порушенням прав інтелектуальної власності на такі системи або їх складові. Стаття зосереджується на авторському праві, патентному праві та захисті баз даних *sui generis* та розглядає, яким чином права інтелектуальної власності можуть бути порушені під час здійснення змагальних атак (adversarial attacks).

Зроблено висновок, що копіювання системи штучного інтелекту або її частин і компонентів (якщо вони захищені авторським правом) під час змагальної атаки (adversarial attack) може розглядатися як відтворення і становити порушення авторського права за певних умов. У цьому контексті важливо враховувати, що відтворення, яке здійснюється для змагальної атаки (adversarial attack), як правило, не може розглядатися як випадкове, а навпаки, може вважатися умисним.

При розгляді можливих порушень прав на винахід у контексті змагальних атак (adversarial attacks), необхідно враховувати, зокрема, доктрину еквівалентів та принцип «вичерпання прав».

Право *sui generis* на базу даних, яке є додатковою можливістю захисту баз даних в Європейському Союзі, також може бути порушене в ході змагальних атак (adversarial attacks) за певних умов.

Встановлено необхідність закріплення у відповідному законодавстві положень про протиправний характер змагальних атак (adversarial attacks). Положення про заборону змагальних атак (adversarial attacks) на штучний інтелект у ліцензійних договорах можуть підвищити рівень правового захисту систем штучного інтелекту.

Ключові слова: штучний інтелект, змагальні атаки (adversarial attacks), авторське право, патентне право, право sui generis на базу даних, право інтелектуальної власності.

Introduction. Artificial intelligence (AI) is having an increasing impact on almost every sphere of our lives, and the development of AI is opening up new horizons. The AI systems are used in many areas, at the same time the AI systems can be targeted by attacks on them, which can negatively affect and damage their technical functioning. The so-called "adversarial attacks" on AI systems require appropriate countermeasures, in particular with the help of law.

The purpose of this paper is to analyze whether adversarial attacks on the AI systems may constitute infringement of intellectual property (IP) rights to such systems or parts thereof, namely whether copyright, patent rights, or *sui generis* database right may be infringed. This is not intended to be a comprehensive analysis, but rather an introductory overview in the author's attempt to begin to study the subject.

The scope of this paper does not cover the assessment of the use of the IP-protected data to create the so-called "adversarial examples" [1, p. 2, 4, 20] from the point of view of the Berne Convention for the Protection of Literary and Artistic Works [2].

IP protection of AI systems has been the subject of research by leading scholars and experts in the field of IP, in particular Jean-Marc Deltorn, Josef Drexl, Reto M. Hilty, Alfred Früh, Peter R. Slowinski, Matt Hervey, Virginia Driver, Tom Woodhouse, and others. Adversarial attacks on AI have been studied by many scholars in the field of technology. The issue of possible IP infringement in the course of conducting adversarial attacks on AI systems requires extensive research.

[1] The author thanks Prof. Dr. Alfred Früh for the great support during the author's research at the Center for Life Sciences Law (ZLSR), Faculty of Law, University of Basel. This article represents the views of the author alone.

Main material. This paper further attempts to analyze the IP rights that may be infringed in the course of conducting adversarial attacks against AI systems, in particular, copyright, patent rights and *sui generis* database rights.

Copyright Law. The AI system and its components, may be protected by copyright, provided that the respective requirements for copyright protection are met [7, p. 196–197].

Articles 2–4 of the InfoSoc Directive prohibit the use of the copyrighted works by their reproduction, communication to the public or distribution without the authorization of the copyright holders [3]. The actions committed within the adversarial attacks can hardly be considered as "communication to the public" or "distribution" in the meaning of copyright law. However, it is necessary to consider whether such actions can be classified as "reproduction".

Thus, Article 2 of the InfoSoc Directive establishes that "Member States shall provide for the exclusive right to authorise or prohibit direct or indirect, temporary or permanent reproduction by any means and in any form, in whole or in part: (a) for authors, of their work (…)" [3]. At the same time, under Article 5(1) of the InfoSoc Directive, an act of reproduction is exempted from the reproduction right envisaged by Article 2 thereof provided that it fulfils five conditions, namely, where (i) the act is temporary; (ii) it is transient or incidental; (iii) it is an integral and essential part of a technological process; (iv) its sole purpose is to enable a transmission in a network between third parties by an intermediary or a lawful use of a work or protected subject-matter; and (v) the act has no independent economic significance [3; 4, para 25]. The mentioned conditions are cumulative in the sense that non-compliance with any one of them will lead to the act of reproduction not being exempted pursuant to Article 5(1) of Infosoc Directive from the reproduction right provided for in Article 2 of this Directive [5, para 55].

For a work to be infringing, it must be copied and derived from the copyright work [6, p. 43]. For example, copying of algorithms or sequences of operations in a computer program might constitute an infringement of copyright in the program [6, p. 73], as well as copying the computer program's architecture in case it falls under the protection by copyright [6, p. 73]. Therefore, copying an AI system or its parts and components (if they are copyrighted) conducted in the course of an adversarial attack may qualify as reproduction and constitute copyright infringement in case certain conditions are met.

Adversarial attacks in black-box settings can involve building a copy of the target model for further performing adversarial techniques on that copy [8]. Within the model extraction attacks an attacker uses information gleaned from queries to the inference API (via inference APIs machine learning models are typically made available to potential client users) of a victim model to build a surrogate model with comparable functionality [9, p. 1]. Recent research has shown that an adversary can successfully extract functional surrogate models by querying a victim model using data from the same domain as the training data for the victim model and without any information about it beyond its intended task [9, p. 1]. However, adversary is not trying to obtain the exact replica of the victim AI model, but rather to achieve a level of performance comparable to the attacked AI [9, p. 6]. Thus, such surrogate model should be perceived as a different model compared to the attacked one [10].

Reproduction of the AI system or its components could happen in white-box attacks, when the attacker knows (and, therefore, can copy) everything about the victim model including the learning algorithm, model topology, defense mechanism, and model/defender parameters [11, p. 2] or in other settings. The white-box attacks are usually not a realistic setting and black-box attacks are more common in reality [12, p. 4].

For establishing the copyright infringement by reproduction, 5-step test on exemption should apply. Against this background, it should be observed that the act of reproduction committed for adversarial attack generally can't be considered as accidental, but on the contrary, may be considered as intentional. In addition, it can not be concluded that the sole purpose of such act is to enable a transmission in a network between third parties by an intermediary or a lawful use of a work or protected subject-matter. Therefore, the reproduction of the AI system or its components, if carried out in the course of adversarial attacks, generally does not meet the criteria required for being exempted from the reproduction right pursuant to Article 5(1) of Infosoc Directive.

It should also be noted that if the second work is substantially the same as the first work, but it has been created independently, the second work will not infringe the first to be created [6, p. 43–44]. However, it is not the case of adversarial attacks, as similarity between the attacked AI model and the model that can be used for such attack is not the mere coincidence.

At the same time, the following legal observations can be applicable to the AI systems in the context of IP rights infringement. Thus, the infringing work does not have to be copied directly from the original work to constitute an infringement [6, p. 44]. For instance, in Plix Products Ltd v Winstone (Merchants) (1986) the New Zealand Court of Appeal established that it was possible to copy indirectly through the intermediary verbal instructions about the substantial features of the original [6, p. 44]. Therefore, direct access to the copyrighted object, "seeing" its substantial features is not a compulsory precondition for infringing copying, provided that such features of the original copyrighted object can be foreseen by the copyist. In addition, a copyrighted work or a part of it can be infringed even when the infringing work is not an exact copy [6, p. 47]. Here, when literal copying is not at issue, the court should decide whether what is copied is the idea behind the program or its expression [6, p. 72]. Despite the peculiarities of legal regulation varying from jurisdiction to jurisdiction, finding of infringement of copyright turns on whether a substantial part of a copyright work is essentially reproduced or adapted [13, p. 121]. Whether the copied part is substantial or not is defined by courts largely discretionary, based on the qualitative rather than quantitative characteristics [13, p. 121].

Countermeasures against copyright infringement in AI are also supported by technical means. Today anti-plagiarism software tools for easy detection of copyright violations in respect of AI are being developed [14, p. 35]. Such techniques as code clone detectors, watermarking and birthmarking schemes allow detecting code fragments which are considered to be equal, determine whether one program is likely to be a copy of another and identify other forms of software plagiarism in AI systems [15, p. 2]. There are also techniques for preventing unauthorized copying source code of AI systems such as obfuscation, symmetric encryption and cryptographic hash functions [15, p. 2, 3].

In general, however, from a point of view of law, reverse engineering a computer program to gain access to its functionalities may be considered an authorized act, as the ideas underlying the computer program are free [16, p. 94]. The ECJ ruled: "… a person who has obtained a copy of a computer program under a license is entitled, without the authorisation of the owner of the copyright, to observe, study or test the functioning of that program so as to determine the ideas and principles which underlie any element of the program, in the case where that person carries out acts covered by that license and acts of loading and running necessary for the use of the computer program, and on condition that that person does not infringe the exclusive rights of the owner of the copyright in that program" (para 62) [17]. Here, it should be mentioned that the goal of an adversarial attack on the AI system goes beyond the mere determination of the underlying ideas and principles.

In this context, the lawfulness of adversarial attacks needs to be separately considered. Thus, in the judgement as of 17 January 2012 in case C-302/10 (Infopaq International A/S v Danske Dagblades Forening) the ECJ made the following observations: "In respect of the lawful or unlawful character of the use, it is not disputed that the drafting of a summary of newspaper articles is not, in the present case, authorised by the holders of the copyright over these articles. However, it should be noted that such an activity is not restricted by European Union legislation. Furthermore, it is apparent from the statements of both Infopaq and the DDF that the drafting of that summary is not an activity which is restricted by Danish legislation. (…) In those circumstances, that use cannot be considered to be unlawful" (paras 44, 45) [4]. In view of this, it should be mentioned that unlawfulness of adversarial attacks against AI systems should be directly established by the legislation. The respective legal provisions determining that adversarial attacks against AI systems constitute infringement must be adopted. It could be also advised to include provisions on prohibition of adversarial attacks against AI systems to the relevant license agreements.

Patent Law. The AI system can receive patent protection as a computer-implemented invention or a part thereof, if the relevant requirements are met [7, p. 198].

The EPC establishes that any infringement of a European patent shall be dealt with by national law (Article 64 (3)) [18].

The rights attributed to a patent owner depend on whether the patent is for a product, a process, or a product obtained directly from a process [19, p. 540]. Thus, Article 28 of the TRIPS Agreement confers a patent owner with the following exclusive rights depending on the subject matter of a patent: "(a) where the subject matter of a patent is a product, to prevent third parties not having the owner's consent from the acts of: making, using, offering for sale, selling, or importing for these purposes that product; (b) where the subject matter of a patent is a process, to prevent third parties not having the owner's consent from the act of using the process, and from the acts of: using, offering for sale, selling, or importing for these purposes at least the product obtained directly by that process" [20].

It should be noted in this context that there is the private use exception from liability for patent infringement, which exception is usually explained on the basis that private uses may increase scientific knowledge and do not pose serious threat to the patent monopoly [19, p. 563–564]. If the infringer had non-commercial subjective purposes, the immunity to liability for infringement could apply [19, p. 564]. At the same time, in case the infringer was motivated by commercial interests, such defense would not apply [19, p. 564].

However, the private use exception can generally be considered inapplicable to adversarial attacks on AI systems, which are typically developed by experts in the field who perform such attacks as part of their practice based on technical knowledge of the functioning of AI.

The AI systems may be reverse engineered within the adversarial attacks. By observing how the AI systems operate, many of such systems may be reverse engineered and thereby at least partially replicated [21, p. 391, 392, 397]. It is difficult to protect AI systems from black-box reverse engineering unless they are isolated from the public [21, p. 399]. At the same time, a slight modification of the parameters of the AI model, that form the essential characteristic of the invention, would avoid falling under the scope of patent protection without significant impact on the underlying technical effect [16, p. 109].

The doctrine of equivalents could be considered as a solution, as it could extend the protection on a particular set of weights of trained AI model to variants of the same network [16, p. 110]. According to the Protocol on the interpretation of Article 69 of the EPC, "[f]or the purpose of determining the extent of protection conferred by a European patent, due account shall be taken of any element which is equivalent to an element specified in the claims" [22]. In the Text of the Basic Proposal for the Treaty and the Regulations as Submitted to the Diplomatic Conference for the Conclusion of a Treaty Supplementing the Paris Convention (Proposal to the Treaty Supplementing the Paris Convention), WIPO envisaged that "a claim shall be considered to cover not only all the elements as expressed in the claim but also equivalents" (Article 21(2)(a)) [23]. However, there are no precise definitions of "the equivalent elements" in general and of an 'equivalence class' in the context of machine learning models, in particular [16, p. 110]. Thus, the mentioned Proposal to the Treaty Supplementing the Paris Convention, contains the following definition: "An element ("the equivalent element") shall generally be considered as being equivalent to an element as expressed in a claim if, at the time of any alleged infringement, either of the following conditions is fulfilled in regard to the invention as claimed: (i) the equivalent element performs substantially the same function in substantially the same way and produces substantially the same result as the element as expressed in the claim, or (ii) it is obvious to a person skilled in the art that the same result as that achieved by means of the element as expressed in the claim can be achieved by means of the equivalent element" (Article 21(2)(b)) [23]. However, different other tests to evaluate equivalents were also developed in the case law by European national jurisdictions [16, p. 110]. In case a second AI model is trained on the knowledge of a first trained AI model to perform substantially the same functions, it is necessary to evaluate, in particular, whether the two models would achieve the same effect "in substantially the same way" for any category of AI models [16, p. 110]. Thus, there is a need in the development of the metric for evaluation and quantification of the similarities between two arbitrary AI models (assessment of 'equivalents'), including the values of their parameters and their architectures as well as their link to the associated function, otherwise practical application of the doctrine of equivalents to AI models will remain elusive [16, p. 110, 112]. Although such comparison between AI models is non-trivial (in particular, due to the fact that the infringing model may not be white-box accessible and/or may serve different tasks), there are a plenty of developed technical approaches and methods for AI models similarity comparison [24, p. 1].

It is also important to pay a separate attention to the principle of "exhaustion of rights" in relation to patents which principle was established by the ECJ in Centrafarm BV v Sterling Drug Inc (1974) [6, p. 305] and is based on the idea that "…the specific subject matter of the industrial property is the guarantee that the patentee, to reward the creative effort of the inventor, has the exclusive right to use an invention with a view to manufacturing industrial products and putting them into circulation for the first time, either directly or by the grant of licenses to third parties, as well as the rights to oppose infringements" (para 9) [25]. However, the patent owner's exclusive rights of the patented object have been exhausted once the object has been sold by the patent owner and the purchaser can use the object, resell it without restriction and can't be sued by the patent owner for having an authorized copy of the patented object [26]. The ECJ confirmed the judgement in Centrafarm BV v Sterling Drug Inc (1974) in the judgement in Merck & Co Inc v Stephar BV (1981) where the following is stated: "… the substance of a patent right lies essentially in according the inventor an exclusive right of first placing the product on the market… That right of first placing a product on the market enables the inventor, by allowing him a monopoly in exploiting his product, to obtain the reward for his creative effort without, however, guaranteeing that he will obtain such a reward in all circumstances… It is for the proprietor of the patent to decide, in the light of all circumstances, under what conditions he will market his product… If he decides to do so he must accept the consequences of his choice as regards the free movement of the product within the common market, which is a fundamental principle forming part of the legal and economic circumstances which must be taken into account by the proprietor of the patent in determining the manner in which his exclusive right will be exercised" (para 9–11) [27].

The core issue of the mentioned decisions in Centrafarm and Merck v Stephar is the matter of consent to the first sale of the relevant goods [6, p. 306]. The "terms of use" on the relevant websites offering the AI for the public users or in the contracts for the commercial customers [21, p. 399, 401] should be considered in this context.

The significant fact to be established is whether the "terms of use" in respect of the AI systems include provisions prohibiting automated querying of the website and use of queries for reverse engineering of AI system exposed by the website [21, p. 399, 401], as well as whether these terms of use contain the reference to prohibition of the adversarial attacks against the AI system. Although the enforcement of the anti-reverse-engineering (and anti-adversarial-attacks) terms appears unsettled and not properly substantiated with court and business practice [21, p. 401–403], adding such terms to the relevant license agreements shall increase the level of legal protection against adversarial attacks on the AI systems.

The commercial strategies of the companies behind the development and use of the AI systems are important to consider. Thus, some companies adhere to the strategy of maximum protecting their AI by IP rights and, thus, implement restrictive terms under the proprietary licenses, whether other allow using their AI as open-source software (OSS) under the terms of the open-source licenses such as Apache 2.0, MIT, BSD 2 and 3 Clause [28, p. 224, 226, 234, 236]. A prevalent strategy among top AI developers today is combination of accumulating patents for AI and simultaneous sharing research via the open-source licenses [29, p. 2].

The terms and conditions of the proprietary licenses, open-source licenses and hybrid licensing forms need to be considered in the context of IP infringement in the course adversarial attacks. For example, such permissive licenses as the BSD 2 and 3 Clause licenses allow the "[r]edistribution and use in source and binary forms, with or without modification" [30], MIT license allows "to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software" [31], Apache 2.0 grants "a perpetual, worldwide, non-exclusive, no-charge, royalty-free, irrevocable" copyright and patent license "to reproduce, prepare Derivative Works of, publicly display, publicly perform, sublicense, and distribute the Work and such Derivative Works in Source or Object form", as well as "to make, have made, use, offer to sell, sell, import, and otherwise transfer the Work" [32]. However, for example, GPL licenses provide that derived works of the licensed software have to be licensed under the GPL and the source code of the modified version has to be provided [33, p. 36]. In general, proprietary licenses are, of course, more conducive to countering adversarial attacks with the help of IP law than open-source licenses.

Sui Generis Database Right. The architecture of the AI system, the training data, and the entire AI system may in some cases be considered a database that requires substantial investment, and therefore protected by the *sui generis* database right provided by EU law, if certain conditions are met [7, p. 197, 199]. The maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents can prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database (Article 7(1) of the Database Directive) [34]. The 'extraction' means the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form (Article 7(2)(a) of the Database Directive) [34]. Thus, this term may be suitable to characterize certain acts conducted during the adversarial attacks on the AI systems (for example, the extraction of the training data [35, p. 1, 2] or a whole model [36, p. 1, 2]). Such actions may constitute an infringement of the *sui-generis* database rights (if applicable to the respective AI system), provided that the part of the AI system being extracted within the adversarial attack is deemed to be substantial.

In addition, it should be noted that under Article 7(5) of the Database Directive, the repeated and systematic extraction and/or re-utilization of insubstantial parts of the contents of the database implying acts which conflict with a normal exploitation of that database or which unreasonably prejudice the legitimate interests of the maker of the database shall not be permitted [34]. The repeated and systematic extraction for the adversarial attack on the AI system may fall within the framework of this provision. However, it is important to establish the unlawful nature of adversarial attacks that contradict a normal exploitation of the AI systems and unreasonably prejudice the legitimate interests of the makers or owners of the AI systems.

In this context, it is also worth mentioning the provision of Article 8(1) of the Database Directive, which provides that "[t]he maker of a database which is made available to the public in whatever manner may not prevent a lawful user of the database from extracting and/or re-utilizing insubstantial parts of its contents, evaluated qualitatively and/or quantitatively, for any purposes whatsoever" [34]. Moreover, Article 15 of the Database Directive stipulates that any contractual term contrary to this provision shall be null and void [34]. Therefore, a collection of data that is not protected under the Database Directive may be in some cases even better protected by the contractual terms limiting its use than by the *sui generis* right under the Database Directive [16, p. 98–99].

Conclusion. Adversarial attacks on AI systems may, under certain circumstances, infringe IP rights. Legislation must include a clear and unambiguous assessment of unlawfulness of adversarial attacks on AI systems.

REFERENCES

1. Alfred Früh and Dario Haux. Countermeasures against Adversarial Attacks on Computational Law. *Journal of Cross-disciplinary Research in Computational Law [CRCL]*. 2023. Volume 2. Issue 1. P. 1–30. URL: https://journalcrcl.org/crcl/article/view/31 (last accessed: 14.06.2023).
2. Berne Convention for the Protection of Literary and Artistic Works (as amended on September 28, 1979). URL: https://www.wipo.int/wipolex/en/text/283698 (last accessed: 20.04.2024).
3. Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society. 2001. OJ L167, 22/06/2001. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32001L0029 (last accessed: 20.04.2024).
4. Order of the Court (Third Chamber). 17 January 2012. Infopaq International A/S v Danske Dagblades Forening. ECJ case C-302/10. ECLI:EU:C:2012:16. URL: https://curia.europa.eu/juris/document/document.jsf?text=&docid=118441&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=3457456 (last accessed: 14.06.2023).
5. Judgment of the Court (Fourth Chamber) of 16 July 2009. Infopaq International A/S v Danske Dagblades Forening. Case C-5/08. ECLI:EU:C:2009:465. URL: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62008CJ0005 (last accessed: 14.06.2023).
6. Jennifer Davis. Intellectual Property Law. 3rd edn. Oxford: Oxford University Press, 2008.
7. Daria Bohatchuk. European Union Law Perspective on the Intellectual Property Protection of Artificial Intelligence Systems. *Uzhhorod National University Herald. Series: Law*. 2024. Volume 82. Part 3. P. 194–201.
8. Andrew Patel, Adversarial Attacks Against AI. *F-Secure Blog*. 11 July 2019. URL: https://blog.f-secure.com/adversarial-attacks-against-ai/ (last accessed: 14.06.2023).

9. Sebastian Szyller, Vasisht Duddu, Tommi Buder-Gröndahl, N. Asokan. Good Artists Copy, Great Artists Steal: Model Extraction Attacks Against Image Translation Models. *arXiv website*. 28 February 2023. P. 1–19. URL: https://arxiv.org/pdf/2104.12623.pdf (last accessed: 14.06.2023).

10. Joao Gomes. Adversarial Attacks and Defences for Convolutional Neural Networks. *Medium website*. 16 January 2018. URL: https://medium.com/onfido-tech/adversarial-attacks-and-defences-for-convolutional-neural-networks-66915ece52e7 (last accessed: 14.06.2023).

11. Bita Darvish Rouhani, Mohammad Samragh, Tara Javidi, and Farinaz Koushanfar. Safe Machine Learning and Defeating Adversarial Attacks. *Microsoft website*. P. 1–7. URL: https://www.microsoft.com/en-us/research/uploads/prod/2018/11/SP_Bita.pdf (last accessed: 14.06.2023).

12. Yinghua Zhang, Yangqiu Song, Jian Liang, Kun Bai, Qiang Yang. Two Sides of the Same Coin: White-box and Black-box Attacks for Transfer Learning. *arXiv website*. 25 August 2020. P. 1–9. URL: https://arxiv.org/pdf/2008.11089.pdf (last accessed: 14.06.2023).

13. Michael D. Pendleton. An abject failure of intelligence: intellectual property and artificial intelligence. In: Ryan Abbott (ed.). *Research Handbook on Intellectual Property and Artificial Intelligence*. Cheltenham: Edward Elgar Publishing Limited, 2022.

14. Till Jaeger. Enforcement of the GNU GPL in Germany and Europe. *JIPITEC*. 2010. No. 1. P. 34–39. URL: https://www.netfilter.org/documentation/licensing/dippadm1268746871.43.pdf (last accessed: 14.06.2023).

15. Fabrizio d'Amore and Lorenzo Zarfati. Source Code Anti-Plagiarism: a C# Implementation using the Routing Approach. *arXiv website*. 6 January 2022. P. 1–12. URL: https://arxiv.org/pdf/2201.02241 (last accessed: 14.06.2023).

16. Jean-Marc Deltorn. The elusive intellectual property protection of trained machine learning models: a European perspective. In: Ryan Abbott (ed.), *Research Handbook on Intellectual Property and Artificial Intelligence*. Cheltenham: Edward Elgar Publishing Limited, 2022.

17. Judgment of the Court (Grand Chamber). 2 May 2012. SAS Institute v World Programming Ltd. Case C-406/10. ECLI:EU:C:2012:259. URL: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:62010CJ0406 (last accessed: 20.04.2024).

18. Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973, as revised. OJ EPO 4/55. 2001. URL: https://www.epo.org/en/legal/epc/2020/convention.html (last accessed: 20.04.2024).

19. Lionel Bently and Brad Sherman. Intellectual Property Law. 3rd ed. Oxford: Oxford University Press, 2008.

20. Agreement on Trade-Related Aspects of Intellectual Property Rights as Amended by the 2005 Protocol Amending the TRIPS Agreement / World Trade Organization. URL: https://www.wto.org/english/docs_e/legal_e/trips_e.htm#art5 (last accessed: 14.06.2023).

21. Shawn Bayern. Reverse engineering (by) artificial intelligence. In: Ryan Abbott (ed.). *Research Handbook on Intellectual Property and Artificial Intelligence*. Cheltenham: Edward Elgar Publishing Limited, 2022.

22. Protocol on the Interpretation of Article 69 EPC, of 5 October 1973, as revised by the Act revising the EPC of 29 November 2000. OJ EPO 2001. URL: https://new.epo.org/en/legal/epc/2020/protinta69.html (last accessed: 14.06.2023).

23. Text of the Basic Proposal for the Treaty and the Regulations as Submitted to the Diplomatic Conference for the Conclusion of a Treaty Supplementing the Paris Convention as far as Patents are Concerned / World Intellectual Property Organization, Standing Committee on the Law of Patents. Fourth session. Document prepared by the International Bureau. Geneva, November 6 to 10, 2000. SCP/4/3. URL: https://www.wipo.int/edocs/mdocs/scp/en/scp_4/scp_4_3.pdf (last accessed: 20.04.2024).

24. Yuanchun Li, Ziqi Zhang, Bingyan Liu, Ziyue Yang, Yunxin Liu. ModelDiff: Testing-Based DNN Similarity Comparison for Model Reuse Detection. ISSTA '21, July 11–17, 2021, Virtual, Denmark. *arXiv website*, 11 June 2021. P. 1–13. URL: https://arxiv.org/pdf/2106.08890.pdf (last accessed: 14.06.2023).

25. Judgment of the Court of 31 October 1974. Centrafarm BV et Adriaan de Peijper v Sterling Drug Inc. Case 15-74. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:61974CJ0015 (last accessed: 20.04.2024).

26. Exhaustion / Legal Information Institute. Wex. *Cornell Law School website*. URL: https://www.law.cornell.edu/wex/exhaustion#:~:text=Exhaustion%20refers%20to%20the%20doctrine,a%20sale%20has%20been%20made (last accessed: 14.06.2023).

27. Judgment of the Court of 14 July 1981. Merck & Co. Inc. v Stephar BV and Petrus Stephanus Exler. Case 187/8. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A61980CJ0187 (last accessed: 20.04.2024).

28. Carlos Muñoz Ferrandis and Marta Duque Lizarralde. Open sourcing AI: intellectual property at the service of platform leadership. *JIPITEC*. 2022. No. 13. P. 224–246. URL: https://www.jipitec.eu/issues/jipitec-13-3-2022/5557/ferrandis_13_3_2022.pdf (last accessed: 14.06.2023).

29. Nathan Calvin and Jade Leung. Who owns artificial intelligence? A preliminary analysis of corporate intellectual property strategies and why they matter / Centre for the Governance of AI, *Future of Humanity Institute, University of Oxford*. February 2020. P. 1–20. URL: https://www.fhi.ox.ac.uk/wp-content/uploads/Patents_-FHI-Working-Paper-Final-.pdf (last accessed: 14.06.2023).

30. The 3-Clause BSD License. *Opensource initiative website*. URL: https://opensource.org/license/bsd-3-clause/ (last accessed: 14.06.2023). The 2-Clause BSD License. *Opensource initiative website*. URL: https://opensource.org/license/bsd-2-clause/ (last accessed: 14.06.2023).

31. The MIT License. *Opensource initiative website*. URL: https://opensource.org/license/mit/ (last accessed: 14.06.2023).

32. Apache License, Version 2.0. *Opensource initiative website*. URL: https://opensource.org/license/apache-2-0/ (last accessed: 14.06.2023).

33. Till Jaeger. Enforcement of the GNU GPL in Germany and Europe. *JIPITEC*. 2010. No. 1. P. 34–39. URL: https://www.netfilter.org/documentation/licensing/dippadm1268746871.43.pdf. (last accessed: 14.06.2023).

34. Directive 96/9/EC of the European Parliament and of the Council dated 11 March 1996 on the legal protection of databases. OJ L 77, 27.3.1996. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31996L0009 (last accessed: 20.04.2024).

35. Nicholas Carlini, Florian Tramer, Eric Wallace, Matthew Jagielski, Ariel Herbert-Voss, Katherine Lee, Adam Roberts, Tom Brown, Dawn Song, Ulfar Erlingsson, Alina Oprea, Colin Raffel. Extracting Training Data from Large Language Models. *arXiv website*, 15 June 2021. P. 1–19. URL: https://arxiv.org/pdf/2012.07805 (last accessed: 14.06.2023).

36. Abdullah Caglar Oksuz, Anisa Halimi, Erman Ayday. AUTOLYCUS: Exploiting Explainable AI (XAI) for Model Extraction Attacks against White-Box Models. *arXiv website*, 7 May 2023. P. 1–14. URL: https://arxiv.org/abs/2302.02162 (last accessed: 14.06.2023).