

## **Advisory Committee on Enforcement**

**Eighteenth Session**  
**Geneva, June 2 to 4, 2026**

### **ARTIFICIAL INTELLIGENCE TOOLS FOR DEALING WITH COPYRIGHT PIRACY: TECHNOLOGICAL, LEGAL AND POLICY DIMENSIONS\***

*Document prepared by Scott Martin, Aspen IP Consulting (former Deputy General Counsel and Executive Vice-President Intellectual Property at Paramount Pictures)\*\**

#### **ABSTRACT**

This study provides an overview of the use of artificial intelligence (AI) tools to address copyright piracy from a technical, legal and policy perspective. This will include an assessment of the potential value of AI tools in the context of copyright piracy, together with a description of traditional techniques for identifying and responding to copyright piracy (including those embedded in content distribution, monitoring tools and responsive tactics). Changes in the piracy landscape which are increasing the need among copyright owners for new tools to respond to content piracy. The study then lays out the current and potential role of AI tools in this sector, before turning to potential challenges of using such tools, including technological, legal and policy challenges, and their implications for rights holders, distribution platforms and consumers. The study concludes with a discussion of a few practical recommendations and good practices.

---

\* The study was conducted with funds provided by the Ministry of Culture, Sports and Tourism of the Republic of Korea. A copy is available in English on the World Intellectual Property Organization Advisory Committee of Enforcement website, available at: [https://www.wipo.int/meetings/en/details.jsp?meeting\\_id=90608](https://www.wipo.int/meetings/en/details.jsp?meeting_id=90608).

\*\* The views express in this document are those of the author and not necessarily those of the Secretariat or of the Member States of WIPO.

I.	INTRODUCTION AND SCOPE OF THE STUDY .....	4
II.	OVERVIEW OF TRADITIONAL TACTICS FOR IDENTIFYING AND RESPONDING TO COPYRIGHT PIRACY.....	5
	(a) Embedded protections .....	5
	(b) Monitoring and detection tools.....	5
	(c) Responsive tactics .....	6
	(i) <i>Graduated response leading to the termination of Internet access</i> .....	7
	(ii) <i>Notice and take down</i> .....	8
	(iii) <i>Site blocking</i> .....	9
	(iv) <i>Copyright infringement litigation</i> .....	10
III.	CHANGES IN THE PIRACY LANDSCAPE THAT ARE INCREASING THE NEED AMONG COPYRIGHT OWNERS FOR NEW TOOLS TO COMBAT PIRACY.....	12
	(a) Overview .....	12
	(b) Aspects unique to various types of content .....	12
	(i) Audiovisual content .....	12
	(ii) Live sports.....	13
	(iii) Music.....	14
	(iv) Print .....	15
	(v) Journalism and news.....	15
	(vi) Games .....	15
	(c) Impact on platforms and search engines.....	16
	(d) Risks to consumers.....	16
IV.	OVERVIEW OF THE ROLE OF AI TOOLS IN COMBATING COPYRIGHT PIRACY.....	16
	(a) Roles for AI tools in combating piracy .....	16
	(b) Summary of functionality of AI-enhanced tools .....	16
	(i) Machine learning.....	16
	(ii) Computer vision .....	17
	(iii) Natural language processing .....	17
	(iv) Watermarking and fingerprinting.....	17
	(v) Automated AI monitoring .....	18
	(vi) Automated AI take downs.....	18
	(vii) AI data processing.....	18
	(viii) AI tools for platforms .....	18
	(ix) Platform music tools .....	18
V.	BENEFITS OF AI IN COMBATING PIRACY .....	19
	(a) Scalability and efficiency.....	19
	(b) Reduced false positives.....	19
	(c) Proactive protection .....	19
	(d) Enhancement of enforcement efforts .....	19
	(e) Tools for user-generated content platforms .....	20

VI.	POTENTIAL CHALLENGES RELATED TO AI TOOLS .....	20
(a)	Technological challenges .....	20
(i)	Efficacy of AI tools.....	20
(ii)	False positives .....	20
(iii)	Fair use determinations.....	21
(iv)	Embedded licensed works.....	21
(v)	Crowdsourcing .....	21
(b)	Legal and policy challenges.....	21
(c)	Recommendations and good practices.....	21
VII.	CONCLUSION .....	22

## I. INTRODUCTION AND SCOPE OF THE STUDY

1. The scope of the study will cover the use of AI tools to tackle copyright piracy, together with related factors and challenges. Focus will be given to the distribution of copyright infringing work, including exact copies of a copyrighted work, excerpts from a work and unauthorized derivative works. Various types of copyrighted content will be considered, including audiovisual works, print works, journalism (including scientific journals), music and games.
2. The study does not address the concerns of content owners regarding the unauthorized training of AI foundation models using copyrighted content, nor does it address AI-enhanced tools for the detection of works used to train AI foundation models, or the interplay between input to AI systems and the output from such systems. AI-generated content will be considered in terms of how output infringes copyrights, regardless of how the content was created. The focus of the study is on copyrighted content; as such, the complex issues surrounding infringement of the names, voices and likenesses of individuals (notably performers and celebrities) are not addressed as those indications of identity are generally not protected by copyright.<sup>1</sup>
3. AI tools have been developed to assist with determining whether content, notably images, are AI-generated fakes.<sup>2</sup> Tools of this type have great potential value for news reporting services and for other types of users and platforms working to identify real versus fake images. However, given that these tools are used to determine whether original images are indeed what they appear to be, rather than being used to analyze infringing versions of copyrightable works, they are outside the scope of this study.
4. AI tools are also being developed to assist with identifying plagiarism, particularly in academic settings.<sup>3</sup> Even though plagiarism can involve acts of copyright infringement, these tools tend to focus on lack of originality rather than on copyright infringement. It is also possible, indeed likely, that use of excerpts from a work in an academic context may qualify as fair use/fair dealing, rather than as an act of piracy or copyright infringement.
5. Finally, AI tools are being developed to enable and assist with the monetization of content when it is distributed online. Since these tools do not aim to combat infringing content, but rather are designed to assist with the licensing and monetization of the distribution of such content, they fall outside the scope of this study.

---

<sup>1</sup> One of the first deep fake images to capture the public attention was a fake photo of Pope Francis wearing a puffy coat. See: <https://www.cbsnews.com/news/pope-francis-puffer-jacket-fake-photos-deepfake-power-peril-of-ai/>. In February 2026, fake videos featuring Brad Pitt fighting Tom Cruise, created by Seedance 2.0, spread across the Internet. See also: <https://www.theguardian.com/film/2026/feb/13/new-ai-video-generator-seedance-tom-cruise-brad-pitt>. AI-enabled tools are being developed and deployed to assist performers in detecting and responding to deep fakes using their likeness. For example, YouTube offers a detection tool through which a celebrity can upload their likeness and an AI-enabled system flags potentially infringing content – for example, detecting an AI-version of an actor playing a role in AI-generated deep fake footage – which may then be taken down. See: <https://www.hollywoodreporter.com/business/digital/youtube-ai-deepfake-detection-tool-1236569593/>.

<sup>2</sup> Identifying AI-generated fake images is of particular importance for journalists and news reporting organizations. Research indicates that people correctly identify high-quality deepfakes only about 24 per cent of the time, while the best AI-enabled detection tools achieve 85–94% per cent accuracy and no detector is 100 per cent accurate. News platforms routinely use multiple AI detectors on a single image in efforts to determine whether the image is legitimate, altered or fake. For example, see “Best AI Image Detectors in 2026”: <https://ddiy.co/ai-image-detection-tools/>.

<sup>3</sup> Some examples of tools used in the academic context to check originality and plagiarism in text include GPTZero, which combines AI detection with plagiarism checking, Turnitin with AI writing detection identifying content generated by GPT-4, Gemini and Claude (available only via academic institutional licensing), and Scribbr, which allows individual users to upload their own unpublished documents for comparison, helping to catch unintended examples of plagiarism. See also: <https://ddiy.co/best-plagiarism-ai-tools/>.

## II. OVERVIEW OF TRADITIONAL TACTICS FOR IDENTIFYING AND RESPONDING TO COPYRIGHT PIRACY

6. There are three general categories of traditional tactics for identifying and responding to piracy of copyrighted content: (i) protections that are embedded in the content or the distribution platform to limit unauthorized access; (ii) tactics for monitoring infringing content; and (iii) responsive tactics when infringing content is identified. AI-enhanced tools have implications for each of these categories.

### (a) Embedded protections

7. A first line of defense in protecting copyrighted content against piracy is the use of embedded protections through encryption, which can be implemented either on a content-based level or a platform-based level. The core of both types of encryptions are digital rights management (DRM) tools which safeguard digital content against unauthorized access and distribution. DRM tools encrypt digital content and then provide access to authorized users only. With content encryption a digital lock is placed on the content and only users who have the encryption key can unlock and view the content. These tools enable anti-piracy protections on smartphones, tablets, computers and smart televisions (TVs).<sup>4</sup> Tools of this type have traditionally included FairPlay,<sup>5</sup> PlayReady,<sup>6</sup> Widevine<sup>7</sup> and others.

8. DRM solutions can also provide for variable licensing scenarios, such as a time-restricted license (for example as a movie “rental” that expires after a few days) and other conditions, such as limiting content to be payable only on certain devices. At the platform level the DRM solutions can limit access, not only to credentialed users such as subscribers, but also to control the scope of subscriber access. For example, the number of devices that can simultaneously access the content can be limited, as can the number of concurrent points of access (such as two users in the same family accessing the platform at the same time).<sup>8</sup>

### (b) Monitoring and detection tools

9. The second category of traditional tactics is monitoring the Internet for pirated copies. There are two primary types of technology used to automate the process of detecting pirated content: watermarking and fingerprinting.

<sup>4</sup> A smart TV, sometimes also referred to as a “connected TV” (CTV), refers to a traditional television set which also features integrated Internet access enabling users to stream music and videos. Smart televisions represent a technological convergence of televisions and computers to add the functionality of digital media players. These devices can provide access to over-the-top media services such as streaming video, music and podcasts.

<sup>5</sup> FairPlay is a digital rights management tool developed by Apple Inc. for protecting videos, music, books and apps.

<sup>6</sup> PlayReady is a media file copy prevention technology from Microsoft, first launched in 2008, that includes encryption, output prevention and digital rights management.

<sup>7</sup> Widevine is a digital rights management system that is deployed in major web browsers and in the Android and iOS operating systems. It is also deployed by streaming services (including Netflix, Amazon Prime Video, Peacock and Hulu) to allow authorized users to view media while preventing them from creating unauthorized copies. Widevine was first developed in 1999 and was acquired by Google in 2010.

<sup>8</sup> Approaches to access control for content and content platforms include token-based authentication and multi-factor authentication. With token-based authentication each user is provided with a unique, time-limited token for their session. These can take the form of session tokens (an authorized user receives a token that serves as a digital key valid only for a period time or a session, limiting access or content sharing) and single-use tokens (the token works once and then becomes invalid – this prevents the token from being shared by multiple users). Multi-factor authentication adds an extra layer of use control by requiring both a username/password combination and a confirmation code sent to the authorized user via text or email. This limits both hacking access and unauthorized credential sharing.

10. Watermarks are unique information embedded into the content. A unique watermark can be applied to each copy of a work in order to trace the source of a pirated copy to the user who had possession of the original copy. This tactic has been used with screener copies sent to award voters, such as for the Academy Awards and Golden Globes, where additional levels of security may be required because screener copies are often distributed prior to the release date for the home entertainment copies of the film. Advanced watermarking technologies include dynamic watermarking where the watermark information changes during playback. This makes it more difficult for pirates to remove or alter the watermark without compromising the quality of the content.

11. The advantages of watermarking over fingerprinting include the ability to identify the specific source of the pirated content. This can have a deterrent effect if a user is aware that their illegal copying or sharing can be traced back to them. Disadvantages include the cost of creating multiple copies of a film, each containing a unique watermark; the possibility that the watermark may have a negative impact on the experience of watching the film; and the limitation that only watermarked frames of a film – and not the entire runtime of the film – will trigger a match.

12. Fingerprinting is a powerful technique used in video anti-piracy to identify and track unauthorized copies of content across the Internet. Unlike with watermarking, fingerprinting does not rely on markers being added to the content; instead it generates a unique digital “signature” based on the content itself. This signature can then be used to search for and detect unauthorized copies.<sup>9</sup> A content fingerprint is a unique code derived from images and audio in the video, including pixel patterns, frame sequences and audio tracks. It acts like digital DNA, enabling detection systems to recognize a video even when the pirate modifies the image through compression, cropping or color adjustment.

13. There are several stages to the fingerprinting process. The first stage is processing the video to create a fingerprint. Separate fingerprints may be created at various stages of film production and post-production in order to protect each version. The second stage is the inclusion of the digital fingerprint in a database, together with metadata like the title and description of the video. The third stage involves scanning websites, streaming platforms, file-sharing services and social media services to locate matches of the fingerprint. And the fourth stage is taking action based on the match, such as flagging the identified content for review or taking it down.

14. The advantages of fingerprinting include the lack of visible marks which can degrade viewer experience (watermarks are often visible). Moreover, fingerprinting can be more resistant than watermarking to thwarting tactics such as re-framing and re-encoding, and detection systems can be triggered by a differing length of match duration in order to address fair use concerns. Limitations include the inability to trace a leak to specific users.<sup>10</sup>

(c) Responsive tactics

15. When an infringing copy of a work is located, there are several approaches responding to the infringing copy. These include a graduated response to termination of Internet access for the infringing user, notice and take down the infringing copy, copyright infringement litigation and site blocking. During the pre-digital era when videotape and digital versatile discs (DVDs)

---

<sup>9</sup> One example of a platform-based application of fingerprinting is the YouTube Content ID system, which detects and blocks infringing uploads to YouTube. See “How Content ID Works”, available at: <https://support.google.com/youtube/answer/2797370?hl=en-GB>.

<sup>10</sup> Some motion picture studios, in an effort to determine the source of leaks of theatrical prints, have removed a series of arbitrary, non-noticeable frames from the print supplied to each theater in order to create and use unique watermarks to identify the exhibitor that is the source of piracy leaks.

were the primary formats for home distribution, and therefore the predominant form of content piracy, the responsive tactics focused on disrupting hard goods distribution. Those efforts included police raids on pirate manufacturing facilities and customs enforcement at borders.<sup>11</sup>

16. While the disruption of manufacturing facilities and customs enforcement remain key tools for combating the flow of infringing consumer products, the challenges for audiovisual content, music and software have largely shifted to focusing on online digital distribution. The following is a summary of traditional responsive tactics for addressing the digital distribution of pirated content. These summaries will provide a context for the need for and potential role of AI-enabled tools.

(i) *Graduated response leading to the termination of Internet access*

17. One of the earliest tactics for combating the online distribution of pirated copies focused on identifying infringing users and terminating their Internet access. This technique was based on the nature of early online sources of piracy, which largely consisted of file-sharing by individuals. Until revenue sources developed for piracy (including subscription-based pirate websites and advertising that supported pirate websites), the primary source for pirated copies of content was individual users “sharing” their copies of content. The graduated response applied not only to the individual who uploaded the pirated content, but also to end users who downloaded copies.

18. The premise of the graduated response was the following: when pirated traffic was identified originating from a specific internet service provider (ISP) address, notice would be sent to the individual to whom that ISP address was assigned. The account holder would be informed of the infringing activity and of the consequences of continued infringement, and that additional activity would result in a termination of Internet access.<sup>12</sup>

19. The graduated response tactic was not ideal for several reasons. From a technical perspective, some ISPs allocate a pool of IP addresses as needed, rather than assigning a unique and static IP address to each computer. This made it more difficult to identify exactly which subscriber was responsible for the infringing traffic. From a practical perspective, some ISP addresses are shared by multiple users, for example, on college campuses, in hotels and at coffee shops. There were concerns about terminating ISP access for colleges based on infringing uses of that ISP address by a group of students. In addition, some ISPs were more cooperative than others in sharing traffic information and subscriber identities where there were allegations of copyright infringing activity on the account. There was also variation in the extent to which service providers were willing to slow the Internet speed or terminate the accounts of users. The lack of cooperation could often be traced to the business reality that, for ISPs, subscribers who engaged in file sharing piracy activities were among their most profitable customers because of their willingness to pay for high-speed access and high-data-volume

<sup>11</sup> In the early 2000s two DVD-sniffing dogs named Flo and Lucky were deployed to locate shipments of pirate DVDs in Malaysia by using their trained sense of smell to detect the polycarbonates used in DVDs. The press reported that one deployment of the dogs resulted in 26 arrests and seizures of illegal discs worth over US\$ 6 million. See: <https://www.nbcnews.com/id/wbna20355981>.

<sup>12</sup> France was among the first countries to implement a graduated response law, adopting three strikes policy in its High Authority for the distribution of works and the protection of rights on the Internet (HADOPI) law in 2009. In 2013, the portion of the HADOPI law that allowed for the suspension of Internet access for a repeat infringer was revoked by the French Government because that penalty was considered to be disproportionate. The power to impose fines or other sanctions on repeat infringers remained in effect ([www.wipo.int/en/web/wipolex/w/news/2013/article\\_0022](http://www.wipo.int/en/web/wipolex/w/news/2013/article_0022)). Since then, HADOPI has been replaced by the French regulatory authority for audiovisual and digital communication (ARCOM). By a decision of 30 April 2026, the French Council of State decided that the graduated response scheme, in its current form, must be modified as it is not in conformity with EU law due to the way in which personal data is being processed ([www.conseil-etat.fr/actualites/protection-des-droits-d-auteur-contre-le-piratage-le-traitement-de-donnees-personnelles-doit-etre-revu](http://www.conseil-etat.fr/actualites/protection-des-droits-d-auteur-contre-le-piratage-le-traitement-de-donnees-personnelles-doit-etre-revu)).

service. These shortcomings led graduated response systems to be viewed as relatively ineffective at stemming the flow of pirated content.<sup>13</sup>

(ii) *Notice and take down*

20. Notice and take down is a tactic used so that pirated copies of a work will be removed from an online platform. This tactic involves the copyright owner giving notice to an online host of illegal content being present on the host service, followed by the online host removing the content.

21. Under United States and European Union law, notice and take down are mandated as a condition for statutory limited liability, or safe harbor protections for online services. In the United States this limitation on the liability of the service provider was established by the Digital Millennium Copyright Act (DMCA) in 1998.<sup>14</sup> In the European Union, the limitations were set out in the Electronic Commerce Directive in 2000,<sup>15</sup> and are now covered in the Digital Services Act (DSA).<sup>16</sup> Both laws contain provisions requiring, as a condition for limited liability, that online service providers expeditiously remove or disable access to content they are hosting upon notification of alleged illegality.

22. Using the DMCA as an example, the take-down process involves the copyright owner (or a vendor engaged by the copyright owner) sending a take-down notice to a service provider requesting the provider to remove material that is infringing on their copyrights. Service providers include ISPs, website operators (i.e. YouTube or eBay), search engines (i.e. Google or Bing), web hosts (i.e. Amazon Web Services or Google Cloud), domain registrants (i.e. GoDaddy or CloudFlare) and other types of online sites and service providers.

23. The DMCA specifies elements required for an effective take-down notice. If required elements are not included in the notice, the service provider may decline to take down the targeted content. Even if a takedown notice meets all the legal requirements, the service provider still may refuse to take down the material. However, if they fail to do so, they forfeit their protections (i.e., a safe harbor) against claims of secondary liability for assisting with copyright infringement.

24. Registration of the copyright with the United States Copyright Office is not required for use of the DMCA take-down process, but the process is limited to copyright infringement claims<sup>17</sup> and cannot be used for trademark claims, for example.<sup>18</sup>

25. Many service providers offer online tools for the direct submission of claims to the provider through an online DMCA take-down portal or form. After a take-down notice is sent to a service provider, the provider notifies the user, subscriber or other person responsible for uploading the infringing content. If that person does not think the activity is infringing, they can send a counter-notice to the service provider explaining why they disagree with the copyright

<sup>13</sup> Giblin, R., "Evaluating Graduated Response". *The Columbia Journal of Law & the Arts*. Available at: <https://journals.library.columbia.edu/index.php/lawandarts/article/view/2142>, January 2014.

<sup>14</sup> For a description of the Digital Millennium Copyright Act, see the United States Copyright Office summary available at: <https://www.copyright.gov/dmca/>.

<sup>15</sup> <https://eur-lex.europa.eu/eli/dir/2000/31/oj/eng>. A description of the Electronic Commerce Directive is available at: [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648797/IPOL\\_STU\(2020\)648797\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/648797/IPOL_STU(2020)648797_EN.pdf).

<sup>16</sup> European Union Regulation 2022/2065 of the European Parliament and of the Council of October 19, 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act). Available at: <https://eur-lex.europa.eu/eli/reg/2022/2065/oj/eng>.

<sup>17</sup> The DMCA states that "A service provider shall not be liable [...] for infringement of copyright [...]". Available at: [17 U.S.C. § 512\(a\)](https://www.copyright.gov/dmca/).

<sup>18</sup> Most platforms, including Google, Facebook, Amazon and YouTube, have their own non-statutory trademark infringement reporting tools.



owner. After receiving a counter-notice, the service provider is obligated to forward that counter-notice to the person who sent the original take-down notice.

26. Once the service provider has received a valid DMCA counter-notice from the alleged infringer, the copyright owner has a limited period of time in which to initiate copyright infringement litigation against the alleged infringer. If the copyright owner files litigation within that time frame, the service provider is obligated to keep the content inaccessible until the litigation is finally resolved. If, however, litigation is not filed within that time frame, the service provider must reactivate or allow access to the alleged infringing content.

27. While the DMCA notice and takedown process remains a valuable tool for content owners, the sheer volume of uploads of infringing content limits the effectiveness of this tactic in reducing online piracy. There have also been complications for content owners arising from malicious actors issuing unauthorized take-down notices.<sup>19</sup>

### (iii) *Site blocking*

28. Site blocking requires ISPs, based on a court or administrative order, to prevent users from accessing specific websites that provide access to websites hosting infringing content. While site blocking is a tactic available to content owners in the European Union, the United Kingdom, Australia, Brazil, India, and numerous other countries,<sup>20</sup> the United States has yet to enact site blocking legislation.

29. The United States came close to enacting the first site blocking law in 2011. The Stop Online Piracy Act (SOPA)<sup>21</sup> would have required online service providers and internet search engines to block access to sites hosting infringing copyrighted material. This proposed legislation spurred opposition by open Internet advocates claiming that these measures would “break the Internet”.<sup>22</sup> The claim was that site blocking would damage the Domain Name Service (DNS) Internet security infrastructure.<sup>23</sup> In the end, the existence of such threats resulted in the bill never being enacted into law.

30. The implementation of site blocking in other territories not only did not break the Internet,<sup>24</sup> it also proved to be an effective anti-piracy tool for content owners. For example, in 2012, the United Kingdom blocked The Pirate Bay, one of the most notorious pirate websites.<sup>25</sup> After the initial block, there was only a minimal impact on the piracy market in the United Kingdom, as most users simply pivoted to other piracy websites, resulting in no increase in visits

<sup>19</sup> One example of this type of activity involved Nintendo, where DMCA takedown notices were sent from a non-Nintendo domain and signed by someone who did not exist. False take-down requests of this type can negatively impact the reputation of a copyright holder. Similarly, the video game company Bungie sued an individual who created email accounts posing as a Bungie employee and sent 96 take-down requests to YouTube to remove content related to its Destiny 2 video game, including videos on the Bungie channel itself. Bungie claimed that the individual harmed its reputation and caused economic damage by angering the Destiny 2 community through confusion over whether they could continue to create derivative works to post on YouTube. Available at: <https://torrentfreak.com/8-1m-damages-agreed-by-youtuber-bungie-for-96-bogus-dmca-notices-240627/>. See also: <https://www.tuckerellis.com/ip-tip-of-the-month-blog/the-power-and-perils-of-dmca-takedown-notices-protecting-copyrights-or-silencing-creators/>.

<sup>20</sup> More than 50 countries have laws that enable website blocking of sites hosing pirated content and estimates are that 39 countries are actively block pirate sites. See: <https://itif.org/publications/2025/06/09/blocking-access-to-foreign-pirate-sites-a-long-overdue-task-for-congress/>. See also: [www.wipo.int/edocs/mdocs/enforcement/en/wipo\\_ace\\_17/wipo\\_ace\\_17\\_13.pdf](http://www.wipo.int/edocs/mdocs/enforcement/en/wipo_ace_17/wipo_ace_17_13.pdf).

<sup>21</sup> Available at: <https://www.congress.gov/bills/112/house-bills/3261>.

<sup>22</sup> See: <https://www.eff.org/deeplinks/2011/10/sopa-hollywood-finally-gets-chance-break-internet>.

<sup>23</sup> See: <https://www.extremetech.com/defense/109533-how-sopa-could-actually-break-the-internet>.

<sup>24</sup> See: <https://www2.itif.org/2016-website-blocking.pdf>.

<sup>25</sup> It was estimated that The Pirate Bay had an estimated 3.7 million users and made approximately US\$ 3 million per month. See: <https://www2.itif.org/2016-website-blocking.pdf>.

to legitimate websites. However, when widespread site blocking took place in 2013, data showed a 12 percent increase in traffic to legal content platforms.<sup>26</sup>

31. Site blocking also enables content owners to respond to the use of proxies by pirate websites. After the ineffective single block of Pirate Bay in the United Kingdom, a subsequent court order in 2015 instructed ISPs to block all proxy websites without the content owners having to undertake the same time-consuming review process for each subsequent ISP address, providing a faster and far more effective tool, not only for taking infringing content offline and but also keeping it offline.<sup>27</sup> This approach is often referred to as dynamic site blocking.<sup>28</sup>

32. Site blocking has tangential benefits related to the distribution of malware. Consumers are estimated to be 30 times more likely to be exposed to phishing, malware, scams and spam on piracy sites than on legitimate websites. Studies have also found that 80 per cent of piracy sites serve malware-infected ads to their visitors and that one in four sites expose users to malicious content.<sup>29</sup>

33. As mentioned above, the United States remains an outlier in terms of the lack of provision of site blocking to content owners as a tool for combatting content piracy, although Congresswoman Zoe Lofgren introduced the Foreign Anti-Digital Piracy Act in 2025.<sup>30</sup> The proposed legislation would enable courts to issue site blocking orders against sites hosting pirated content, which would apply to both ISPs and DNS resolvers.<sup>31</sup>

(iv) *Copyright infringement litigation*

34. In addition to the tools of graduated response, notice and take down, and site blocking, content owners also have the option of initiating copyright infringement litigation processes, both against pirate sites and against the online platforms that enable piracy. There are several drawbacks to initiating copyright infringement litigation processes against individual pirate sites and services, including practical complications, jurisdictional issues and cost.

35. The practical complications include the difficulty of identifying and locating the individual or entity responsible for the infringing activity. This is the “whack-a-mole”<sup>32</sup> nature of online digital piracy; where a webpage is taken down and another online listing pops up under a different uniform resource locator (URL) almost instantly. Another complication for content owners is the sheer volume of online content and the speed at which pirated content is made available online.

<sup>26</sup> See: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2612063](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2612063); see also: <https://www.heinz.cmu.edu/media/2019/October/anti-piracy-blocking-multiple-websites>.

<sup>27</sup> See: <https://www.bbc.com/news/technology-31832137>; see also: <https://techpolicyinstitute.org/publications/intellectual-property/the-effectiveness-of-site-blocking-as-an-anti-piracy-strategy-evidence-from-the-u-k/>.

<sup>28</sup> See description and discussion of dynamic site blocking in the report of the World Intellectual Property Organization Advisory Committee on Enforcement 17th Session, February 4–6, 2025 entitled “Sharing Experiences and Best Practices on Site Blocking/No-Fault Injunctions”, available at: [https://www.wipo.int/edocs/mdocs/enforcement/en/wipo\\_ace\\_17/wipo\\_ace\\_17\\_14\\_prov.pdf](https://www.wipo.int/edocs/mdocs/enforcement/en/wipo_ace_17/wipo_ace_17_14_prov.pdf).

<sup>29</sup> The original version of The Pirate Bay was located at “thepiratebay.org”, and the 2015 order extended the site blocking order to proxy sites including “piratebayproxy.co.uk”, “piratebayproxylist.com” and “ukbay.org”. <https://www.digitalcitizensalliance.org/clientuploads/directory/Reports/Unholy-Triangle-Report.pdf>.

<sup>30</sup> See: <https://torrentfreak.com/new-bill-aims-to-block-foreign-pirate-sites-in-the-u-s-250129/>.

<sup>31</sup> The inclusion of DNS resolvers is significant. Major tech companies such as Google and Cloudflare offer DNS services internationally, increasing the possibility of blocking orders having a worldwide impact.

<sup>32</sup> Whack-a-Mole is a classic arcade game that has also become a popular cultural metaphor for problems that keep reappearing as fast as they are solved, for fixing one issue only to see another to pop up elsewhere, and for challenges that seem endless and repetitive. See: <https://en.wikipedia.org/wiki/Whac-A-Mole>.

36. The jurisdictional complications arise from the borderless nature of the Internet, which creates three categories of challenges. First, pirate services are often intentionally located in countries that lack robust IP laws or lack effective judicial systems. Second, where a site is located in a country which is not impacted by piracy,<sup>33</sup> local law enforcement and local courts may not be motivated to assist with efforts to combat the site. And third, the cross-border nature of piracy raises issues of international jurisdiction and the enforcement of foreign judgments. However, perhaps the biggest impediment to copyright infringement litigation against specific pirate sites is the cost involved in suing individual sites, as recovery of damages and of the litigation costs is unlikely, and impact is limited when the pirated material shifts from the targeted website to other websites.

37. The abovementioned challenges may have less of an impact on litigation against service providers such as ISPs and social media platforms. Nonetheless, those litigations involve their own unique set of challenges. For example, under United States law there are the hurdles of the DMCA safe harbor immunity provisions, as well as the need to prove secondary liability.

38. The DMCA safe harbor (Section 512) shields online platforms from copyright liability for content posted by users, as long as the platform meets certain conditions.<sup>34</sup> To qualify, a platform must have no actual knowledge of infringing content and, upon gaining knowledge, it must act "expeditiously" to remove it; it must not receive a direct financial benefit from infringement while having the ability to control it; it must designate a registered DMCA agent with the United States Copyright Office to receive takedown notices; it must have and follow a repeat infringer policy (terminating accounts of repeat offenders); and it must comply with the take-down/counter-notice process.

39. Section 512 creates numerous challenges for content owners. Platforms are obligated to act only in response to specific notices with no duty to monitor for infringing content uploaded by users; courts have grappled with determining how much knowledge a platform must have before it is deemed to have actual knowledge of infringing content;<sup>35</sup> and Section 512 does not address the issue of infringing content being endlessly re-uploaded, since the law requires only take-down and not stay-down for infringing content.

40. The second hurdle under United States law in holding platforms responsible for the infringing content they host is the newly narrowed concept of secondary liability based on the March 2026 Supreme Court ruling in the Cox Communications vs. Sony Music case.<sup>36</sup> The case dealt with the issue of whether an ISP can be held liable for copyright infringement committed by its subscribers when the ISP receives repeated notices of infringement but does not terminate user accounts. Sony Music and other record labels sued Cox, which is one of the largest ISPs in the United States, after sending hundreds of thousands of notices alleging that subscribers repeatedly shared copyrighted music via peer-to-peer networks.<sup>37</sup> In 2019, a Virginia court found Cox liable for willful contributory infringement and awarded \$1 billion in

<sup>33</sup> For example, some pirate websites apply geofilters that prevent the service from being accessed in the hosting country.

<sup>34</sup> Section 512 provides four types of safe harbor covering different functions served by online platforms: storing user content (hosting), transmitting content (conduit), caching and linking/search. The most relevant practice here is the hosting safe harbor which is addressed in Section §512(c).

<sup>35</sup> See *Viacom vs. YouTube*, 676 F.3d 19 (2d Cir. 2012); *UMG vs. Shelter Capital*, 718 F.3d 1006 (9th Cir. 2013); and *Capitol Records vs. Vimeo*, 826 F.3d 78 (2d Cir. 2016).

<sup>36</sup> See *Cox Communications, Inc. vs. Sony Music Entertainment*, No. 24-171, 607 U.S. (2026).

<sup>37</sup> The Supreme Court decision described Cox as follows: "Cox Communications, Inc., is an Internet service provider serving approximately six million subscribers, each associated with a unique Internet Protocol address. Internet service providers like Cox have limited knowledge about how their services are used; they know which IP address corresponds to which subscriber account but cannot distinguish individual users or directly control how services are used. Cox contractually prohibits subscribers from using their connection to post, copy, transmit, or disseminate content that infringes copyrights." *Cox Communications, Inc. vs. Sony Music Entertainment*, No. 24-171, 607 U.S. (2026).

damages. The Supreme Court reversed the ruling of the lower court that Cox was vicariously and contributorily liable for music piracy by users, holding that an ISP is contributorily liable “only if it intended” that the service be used for infringement, and that intent can be shown “only if the party induced the infringement or the provided service is tailored to that infringement”.

### III. CHANGES IN THE PIRACY LANDSCAPE THAT ARE INCREASING THE NEED AMONG COPYRIGHT OWNERS FOR NEW TOOLS TO COMBAT PIRACY

#### (a) Overview

41. In addition to the abovementioned challenges faced by content owners in dealing with piracy of their works, a primary driver of demand for new tools to combat piracy of audiovisual copyrighted content, including AI-enhanced tools, has been the shift from physical media (DVDs and compact discs (CDs)) to online distribution, including digital downloads and on-demand streaming. The shift of consumer demand for content from physical media to online content has made it easier for pirates to create and distribute unauthorized copies of works, resulting in a explosive growth of pirated copies of content.<sup>38</sup> At the same time, the borderless nature of the Internet has created new challenges for effective enforcement. The demand for pirated online content coupled with the complexities of cross-border enforcement has created an environment in which organized crime is flourishing, further compounding the challenges faced by content owners.<sup>39</sup> These changes have an impact on all forms of content, however, there are aspects unique to each form of content including audiovisual works, live sports, print and electronic versions of books journalism, music and games.

#### (b) Aspects unique to various types of content

42. Below is a brief summary of the unique impact of piracy in the AI era on various types of content.

##### (i) Audiovisual content

43. In addition to the tremendous growth of online pirated copies of works, the quality of pirated copies has improved exponentially, with the era of shaky hand-held camcorder film copies giving way to the current era of perfect digital copies.<sup>40</sup> At the same time the pirate

<sup>38</sup> The growth of online piracy is reminiscent of Moore's law, which is the observation that the number of transistors on a microchip doubles approximately every two years, while the cost of computing roughly halves over the same period. That observation was made in 1965 by Gordon Moore, co-founder of Intel.

<sup>39</sup> IP House and the Digital Citizens Alliance recently released an investigative report examining the convergence between global piracy networks and organized crime. The report, entitled “*Organized. Piracy. Crime.: How Global Piracy Networks Became Organized Crime Syndicates – And What Needs to be Done About It*”, provides an analysis of how large-scale digital piracy has evolved into a multibillion-dollar criminal enterprise. See <https://ip-house.com/resources/organized-piracy-crime/>; <https://ip-house.com/assets/Uploads/resources/DCA-IPH-Organized-Piracy.-Crime..pdf>.

<sup>40</sup> Digital copies are sourced from multiple different sources including DVD and Blu-ray rips, screener leaks and telecines (high quality frame-by-frame film transfers). However, the issue of camcording has not disappeared, as it still constitutes a threat during the pre-release window for digital copies. The 2026 Special 301 Report of the United States Trade Representative notes: “The proliferation of camcords continues to be a significant trade problem. Unauthorized camcording is the primary source of infringing copies found online of newly released movies. The recordings made in movie theaters today are very different from those by a single person sitting in a theater with a bulky videotape recorder. The results are not shaky, inaudible recordings. It is now easy for a surreptitious recording in a movie theater to result in a clean digital copy of a movie with perfect audio that can be quickly distributed online. The pirated version of the newly released movie may be available online while it is still showing in theaters. The economic damage is magnified because movies may be released in different markets at different times. Thus, a camcord of a movie released in one market can be made available unlawfully in another market before the movie enters the theaters there. In addition to theater owners who lose revenue, legitimate digital platforms, which often negotiate for a certain period of exclusivity after the theatrical run, cannot fairly compete in the

[Footnote continued on next page]

platforms have become more professional in appearance and operation, making it more difficult for consumers to distinguish between pirate and legitimate services. Early pirate sites were cluttered with pop-ups, broken links, offensive advertisements and other red flags of illegitimacy. Today, sites have layouts that mimic legitimate streaming service with clean interfaces, recommendation engines and professional-looking logos. Some even feature advertising by well-known consumer brands.<sup>41</sup>

44. These challenges are compounded by the growth in high quality AI-generated infringing derivative works. In the past, content owners had to address the threat of pirated copies of their works; now they have the challenge of addressing AI-generated infringing derivative works. For example, the challenge of detecting and responding to pirated copies of Star Wars films now also includes the challenge of dealing with unauthorized new AI-generated live action Star Wars films or episodes created without the consent of the copyright owner.<sup>42</sup>

(ii) *Live sports*

45. The owners of live sports broadcasting rights face the similar challenges of pirate websites offering competing illegal feeds of live events using platform interfaces that may appear legitimate to consumers, potentially due to the appearance of advertisements for well-known brands on the pirate site. Unauthorized streaming services, illegal retransmission through social media platforms and illicit Internet Protocol television (IPTV) operations increasingly undermine legitimate markets.

46. The worldwide sports broadcast rights market was valued at approximately \$62 billion in 2024, continuing a rapid upward trend over the past decade. The number of United States viewers who stream a sports event at least once a month is estimated to have risen steeply from 57 million in 2021 to more than 90 million today. Major sporting events, including the Olympic Games, the International Federation of Association Football (FIFA) World Cup, and championship finals for professional leagues, attract billions of viewers worldwide, generating enormous commercial value for broadcasters, leagues, teams and sponsors. Broadcasting revenue has become the dominant financial pillar for most major sporting events. For example, media rights licenses for the United States National Football League are valued at approximately \$110 billion for the period from 2023 to 2033. The domestic and international broadcast rights of the English Premier League for 2022 to 2025 exceeded \$13 billion.<sup>43</sup> That

---

market due to unauthorized camcording". See:

<https://ustr.gov/sites/default/files/files/Press/Releases/2026/2026%20Special%20301%20Report.pdf>.

<sup>41</sup> Consumer brands often unknowingly advertise on pirate sites through automated advertisement networks which may not distinguish between legitimate platforms and reputable-looking pirate platforms which generate significant user traffic. The WIPO ALERT platform is an excellent tool for advertisers and advertising agencies to avoid the placement of advertisements on infringing websites. The service maintains lists of infringing websites and those lists can be used by advertisers and advertising placement agencies in their automated advertising systems to avoid placing advertisements on those infringing sites. This assists advertisers in avoiding the subsidization of copyright infringement while also protecting their brands from the negative reputational effect that can arise from association with illegal activities, including infringing sites that use pirated content as bait for the distribution of malware. See <https://www.wipo.int/en/web/wipo-alert>.

<sup>42</sup> Forbes published an article entitled "*The Best Star Wars Movie in Years Is Made With AI*", available at: <https://www.forbes.com/sites/charliefink/2025/02/12/the-best-star-wars-movie-in-years-is-made-with-ai/>.

<sup>43</sup> For sample data on the value of sports broadcasting rights, see the following: <https://www.sportbusiness.com/news/global-value-of-sports-media-rights-tops-60bn/>; <https://www.spglobal.com/market-intelligence/en/news-insights/research/2026/04/global-sports-rights-climb-to-over-67-billion-in-2026>; <https://www.sportspro.com/news/broadcast-ott/sports-media-rights-global-spending-ampere-analysis-november-2025/>.



revenue stream, both for broadcasters and for the sports leagues, is under threat from the growth of live-stream piracy.<sup>44</sup>

47. Online piracy has had a uniquely detrimental impact on broadcasts of sports and live entertainment because, unlike movies or television series which retain their revenue-generating value for extended periods, sports content derives almost all its economic value from real-time or near real-time viewing. A soccer match generates minimal viewing interest hours after its conclusion. This compressed value window means that by the time enforcement mechanisms detect and remove unauthorized streams, much or all of the commercial value has been lost. The live, time-sensitive nature of sports creates an asymmetry that heavily favors pirate operations over rights holders and enforcement authorities and increases the need for AI-enabled tools to combat piracy.

(iii) *Music*

48. Music content owners face much the same new threats from piracy as do audiovisual works, but they also encounters challenges that are uniquely associated with music.

49. The advent of AI in music is difficult to pin down because computer assisted machine learning has long been applied to sound, including tools for pitch correction and editing, and the use of electronic and synthesized instruments. However, the role of AI in the creation of infringing music became headline news in 2023 with the release of a song falsely purporting to be by the artist Drake, widely referred to as the “Fake Drake song”. A TikTok user with the pseudonym Ghostwriter created a viral sensation with the song “Heart on My Sleeve” which used AI to deepfake vocal performances of the music creators/performers Drake and The Weekend.<sup>45</sup> The song invoked issues triggered by the use of AI to create sound-alike and deep fake works based on the vocals of well-known artists. Several music platforms subsequently announced policies aimed at combating AI-generated soundalike recordings.<sup>46</sup>

50. Another AI-related challenge is the use of AI-generated music to fraudulently claim streaming royalty payments. For example, in the United States a North Carolina musician was indicted by federal prosecutors over allegations that he had used AI to create hundreds of thousands of songs and then used those AI tracks to earn more than \$10 million in fraudulent

<sup>44</sup> The 2025 Special 301 Report of the United States Trade Representative focused on the impact of piracy on live sports broadcasts. See: <https://ustr.gov/sites/default/files/files/Press/Releases/2026/2026%20Special%20301%20Report.pdf>. The report notes that “piracy operations siphon revenues from right holders, devalue official broadcasting licenses, and weaken the economic incentives for continued investment in sports production and infrastructure. As streaming technology becomes increasingly accessible and piracy operations grow more sophisticated, the challenge intensifies” [https://ustr.gov/sites/default/files/files/Press/Releases/2026/2025%20Notorious%20Markets%20List%20\(final\).pdf](https://ustr.gov/sites/default/files/files/Press/Releases/2026/2025%20Notorious%20Markets%20List%20(final).pdf).

<sup>45</sup> See: <https://www.billboard.com/music/pop/ghostwriter-heart-on-my-sleeve-drake-ai-grammy-exclusive-interview-1235434099/>.

This scope of this study is focused on AI tools for content protection and does not address legislative responses. However, it is worth noting that there have been efforts to address these concerns through legislation, including through the Tennessee state law titled the Enduring Likeness Voice and Image Security (ELVIS) Act, which is designed to protect artists from AI deepfakes (<https://legiscan.com/TN/text/HB2091/id/2900923>). The ELVIS Act replaces the State’s previous right of publicity law, which only included protections for “name, photograph, or likeness”, expanding protections to also address voice and AI concerns.

<sup>46</sup> For example, iHeartRadio announced a policy that it calls “Guaranteed Human”, pledging that the company will not play AI music that features synthetic vocalists pretending to be human. See: <https://www.billboard.com/pro/iheartradio-bans-ai-music-podcasts-radio-djs-new-program/>. Similarly, Spotify announced policies aimed at addressing the impersonation frauds that come with AI-generated music. While the new policies of Spotify do not specifically relate to AI-generated content – any form of impersonation would be penalized – the goal is to limit malicious uses of AI while still allowing for the creative applications. See: <https://www.billboard.com/pro/spotify-updates-ai-music-policies-spam-tracks-removed/>. Another online music platform, Bandcamp, also bans the use of any AI tools “to impersonate other artists or styles”. See: <https://www.billboard.com/pro/bandcamp-bans-ai-generated-music-songs-policy/>.

streaming royalty payments. One monitoring service has estimated that nearly 40 per cent of daily music uploads are generated entirely with AI.<sup>47</sup>

(iv) *Print*

51. Print media faces its own unique challenges in the AI era, in addition to the challenges created by the global, decentralized nature of the Internet and the challenges of cross-border enforcement. Digital e-book formats have made piracy significantly easier. E-books can be stripped of DRM protections and spread across torrent sites, forums and file-hosting services within hours of release, sometimes before the official launch date. Sites hosting pirated copies of e-books (such as Z-Library, which has repeatedly been shut down only to reappear using new URLs), have hosted millions of infringing e-book titles.<sup>48</sup>

(v) *Journalism and news*

52. Journalism faces unique challenges because the core of reporting is facts, which are not subject to copyright protection. Nonetheless, without financial revenue, reporting services cannot exist. There are three primary online threats to journalism during the AI era. The first is paywall circumvention. Sites or browser extensions that enable users to bypass paywalls undermine the subscription model on which many news outlets depend. When users can easily access premium content for free, the incentive for subscriptions weakens, negatively impacting newsroom budgets.<sup>49</sup>

53. The second threat is content theft and re-use. Scrapers and aggregators copy articles republish them without permission or attribution. This diverts traffic away from original publishers, costing them ad revenue and subscription conversions. Even mainstream aggregators, such as Google News and Apple News, have been accused of "good enough" summarization: giving readers the gist without sending them to the journalistic source. Generative AI trained on journalistic content, including summarization of paywalled articles, represent a new form of the problem of copyrighted journalism being reused without payment or attribution and without directing user traffic to the original source of the reporting.<sup>50</sup>

54. The third threat to journalism during the AI era is the amplification of distorted reporting. Pirated or repackaged journalism is often stripped of context, corrections or nuance, which can damage public trust and the reputational value of the original reporting. Finally, the impact of online piracy on academic and technical publishing has been particularly significant, undermining the subscription and per-article revenue models on which researchers and academic publishers depend.

(vi) *Games*

<sup>47</sup> See: <https://www.billboard.com/pro/deezer-execs-ai-music-how-fighting-interview/>.

<sup>48</sup> See: <https://techhq.com/news/how-is-z-library-down-again-alternatives-ebooks/>.

<sup>49</sup> The News/Media Alliance has successfully taken down paywall bypass websites including 12ft.io. See: <https://www.newsmediaalliance.org/takedown-of-12ftio/>.

<sup>50</sup> One of the tools for responding to the negative economic impact of AI on journalism is legislative solutions. For example, legislation has been introduced in the United States Congress to create an anti-trust exemption enabling digital journalism providers to form joint negotiation entities to collectively negotiate with online platforms over the pricing and the terms and conditions under which the platform can access digital news content. It would also require online platforms to negotiate in good faith with news organizations, while imposing for mandatory arbitration where negotiations prove unsuccessful. See the proposed Journalism Competition and Preservation Act: <https://www.klobuchar.senate.gov/public/index.cfm/2023/3/klobuchar-kennedy-introduce-bipartisan-legislation-to-save-local-journalism>.

55. Game companies face significant and unique challenges as a result of online piracy in the AI era. Companies invest heavily in DRM systems,<sup>51</sup> always-online requirements, or hardware-locked licenses. These systems are expensive to implement and can be frustrating to legitimate players due to performance issues or the requirement for constant Internet connections. Popular games cost millions of dollars to develop and piracy makes it harder to recoup that investment. The global nature of popular games creates regional pricing challenges that can be difficult to resolve. Piracy rates tend to be higher in regions where games are expensive relative to local incomes. The global nature of gaming makes it difficult for companies to price games appropriately for emerging markets, sometimes causing a reluctance to enter those markets which, ironically, can increase demand for pirated access.

(c) Impact on platforms and search engines

56. While this study is focused on aspects of AI tools for content owners to deal with piracy of their copyrighted material, it should be noted that piracy also has a significant impact on legitimate content hosting platforms and search engines. Google, for example, needs new AI-enhanced tools in order to better identify and de-index infringing URLs. When pirated content ranks at the top of search responses it degrades the relevance and trustworthiness of results for users and opens the service to claims from rights owners. Google alone has processed billions of copyright removal requests under the DMCA. Legislation like the European Union Directive on Copyright, which requires hosting platforms to implement proactive filtering systems, creates demand for new AI-enhanced tools to assist with that filtering process.

(d) Risks to consumers

57. It should be noted that the risks created by pirated content in the AI era are not limited to content owners and to platforms; there are also risks to consumers. Pirated files like cracked software, games, and film/episodic content are common vehicles for trojans, ransomware, and spyware. Attackers bundle malware directly into the installer tool. Even visiting a pirate website can trigger automatic downloads of malicious scripts without clicking on any links by exploiting unpatched browser or plug-in vulnerabilities. Pirate sites also push through browser hijackers and adware that redirect searches and harvest browser and log-in data.<sup>52</sup>

#### **IV. OVERVIEW OF THE ROLE OF AI TOOLS IN COMBATING COPYRIGHT PIRACY**

(a) Roles for AI tools in combating piracy

58. In the light of the increasing piracy challenges described above, content owners are looking at ways in which AI-enhanced tools can assist with the process of detecting and responding to online piracy of content. Applications for AI-enhanced tools can include content recognition and tracking, automated takedowns, watermarking, digital fingerprinting, anti-piracy measures for software and games and dark web monitoring.

(b) Summary of functionality of AI-enhanced tools

59. AI-enhanced tools for addressing content piracy can include machine learning, natural language processing and computer vision functionality.

(i) Machine learning

<sup>51</sup> One example of a DRM system for games is Denuvo. See: <https://irdeto.com/video-games>.

<sup>52</sup> See this summary by Fact UK of the dangers of illegal streaming: <https://www.fact-uk.org.uk/consumer-advice/dangers-of-illegal-streaming/>.



60. Machine learning algorithms can enhance detection accuracy over time by learning from past piracy patterns, decreasing false positives, and enabling automated review for fair use and fair dealing scenarios (the term “fair use” as used in the study encompasses similar concepts including fair dealing and the three-step test). While forms of online piracy continue to evolve rapidly, machine learning algorithms have the ability to refine their detection methods and to develop more effective response strategies.

(ii) Computer vision

61. Computer vision can detect pirated videos by analyzing visual frames, even where the image in the pirated copy has been cropped, mirrored or otherwise altered to avoid detection by traditional monitoring systems. By comparing video frames pixel by pixel, AI systems are able to detect even subtle modifications in pirated versions that have enabled pirated content to elude traditional forms of detection. Computer vision technologies are also being deployed to protect images and graphics. Stock photo companies, for example, are deploying AI-powered image recognition to locate and identify unauthorized use of their copyrighted images.

(iii) Natural language processing

62. Natural language processing enables AI-assisted scanning of metadata and file descriptions for piracy-related keywords. This adds another detection method to that of scanning for content itself. For example, natural language processing can detect and react to suspicious descriptions such as “free movie download” or “unlocked access”.

63. That functionality can also assist with monitoring online forums and other sites where pirated content is promoted and unauthorized access links are offered. This is a particularly valuable tool for detecting pirated live event streams, where the descriptive metadata that enables users to locate the service also enables the content owner to locate and disable the service during the live event.

(iv) Watermarking and fingerprinting

64. AI-assisted functionality can improve the effectiveness of watermarking and fingerprinting, while AI recognition capacity – including computer vision functionality – can also help to override the manipulation of watermarks and fingerprinted images by pirates, thus improving detection capacities for content owners and distribution platforms. Video fingerprinting has grown in sophistication, incorporating advances in the signal processing, machine learning and big data analytics that constitute the very foundations of AI.

65. This evolution has significantly improved the accuracy, speed and scalability of video fingerprinting. One key aspect of the value of AI-enhanced video fingerprinting arises from its capability for real-time content identification. This allows for automated takedown processes, where pirated content can be flagged and removed almost immediately upon detection, significantly reducing the window during which pirated content is available to consumers. Given the significant and rising economic threat of piracy to sports and live events, the ability to create and implement fingerprint detection and reactive strategies in real time during a sports or live event broadcast is crucial to protect the narrow window of economic value for such works.

66. One of the challenges with using fingerprinting as tool for video content protection is the wide range of content formats, resolutions and codecs required by the multiple platforms and markets in which content is distributed. This further complicates the fingerprinting process, making it necessary to have systems that are capable of handling a wide variety of formats for the same program. A strength of AI-enabled systems is the ability to process and analyze that

diversity of formats. However, the physical logistics and cost of inputting various formats into a fingerprinting system should be taken into account.

67. AI-enhanced fingerprinting is also being used for music. For example, one content protection vendor is promoting a product that uses fingerprints of song lyrics to detect even partial use of copyrighted lyrics, enabling copyright owners and platforms to distinguish between original, copyrighted and public-domain content.<sup>53</sup>

(v) Automated AI monitoring

68. Automation of monitoring is another key value that is inherent to AI-assisted anti-piracy tools. Expanding on traditional methods of manual scanning and monitoring, AI tools can automatically monitor vast amounts of data across multiple platforms, including websites, forums, torrent platforms and social media platforms. AI-assisted automation has the potential to detect and respond to pirated copies before they spread widely across the piracy ecosystem. That expanded scalability of monitoring is a key source of value for AI-enhanced anti-piracy tools.

(vi) Automated AI take downs

69. AI-enabled automation of the take-down process, including AI-generated take-down notices and monitoring compliance, have the ability to significantly expand the scope and accelerate the process of removing pirated content from platforms, while reducing false-positive identifications (such as a take down mistakenly issued based on a licensed third party musical composition contained in the soundtrack of a film), while not interfering with appropriate fair-use activity.

(vii) AI data processing

70. AI-enhanced content protection tools can supplement analytical guidance on content protection strategies by analyzing data trends to predict high-risk periods and territories for specific content and by identifying and targeting the platforms where pirated content first appears as a priority. Information and predictive data of this type can assist content owners with developing prevention strategies to ensure their content enforcement efforts are proactive as well as reactive. Data analysis can also assist with piercing the shield of anonymity that many pirate sites use to thwart enforcement efforts by making it difficult to determine the location of a site and the identities of its operators.

(viii) AI tools for platforms

71. AI-enabled anti-piracy tools have great potential value for content owners, but they can also help websites and platforms that host user-uploaded content to comply with notice and take-down obligations and to improve blocking of user uploads of pirated or infringing content. As noted above, the enormous volume of content being uploaded to sites hosting user-generated or user-supplied content requires automated tools with AI-learning capacity to effectively combat the seeding of infringing pirated works.

(ix) Platform music tools

---

<sup>53</sup> See: <https://www.musicbusinessworldwide.com/musixmatch-launches-sentinel-service-to-detect-when-copyrighted-music-and-lyrics-are-used-in-ai-and-user-generated-content/>.

72. One example of an AI tool that can assist platforms hosting user uploaded original content is an AI-powered tool launched by YouTube that enables creators facing copyright claims to replace contested music in their videos with royalty-free instrumental tracks, rather than merely muting the soundtrack or removing content altogether. The tool is integrated into the “replace song” feature on YouTube Studio, which enables users to replace existing music with non-infringing music as a path to resolving content identification disputes while keeping their videos posted.<sup>54</sup>

## **V. BENEFITS OF AI IN COMBATING PIRACY**

73. While AI enhanced tools for combating content piracy are still in the early stages of development, efficacy testing and deployment, the benefits of such tools are already evident. They include the following:

### **(a) Scalability and efficiency**

74. Unlike manual monitoring, AI systems have the capacity to process and analyze vast amounts of data. This scalability enables content owners to monitor both the use and misuse of their content across large numbers of streaming platforms and social media uploads, far beyond the capacity of current monitoring and tracking tactics. With automated detection enabled by AI, monitoring systems can operate with greater efficiency and scope, resulting in lower costs and improved accuracy.

### **(b) Reduced false positives**

75. Traditional piracy detection methods can result in false positives, where legitimate content or fair use of third-party content is flagged as having been pirated. These false positives create complications and inconvenience for content owners, platforms and users. AI-enhanced tools have the capability to reduce false-positive errors by learning from historical data and refining detection methods over time. Improved accuracy also builds trust between content owners and the platforms that receive take-down notices.

### **(c) Proactive protection**

76. The predictive capabilities of AI allow content owners to anticipate piracy threats before they occur. This proactive protection data can assist content owners with decisions regarding release timing for problematic territories, as well as the need for additional security measures. For example, early release windows for premium video on demand (PVOD) in Asia have implications for enabling digital versions of films to be pirated ahead of the release date for standard (non-premium) priced releases. Data about early piracy leaks and the spread of those infringing copies helps studios to determine the optimal timing for PVOD and possible films to hold back from PVOD release where losses from piracy may exceed the profits from the PVOD release. In some cases, studios have been able to determine, based on predictive data, that adding burned-in subtitles in the local language makes the release a less-desirable piracy target for other territories.

### **(d) Enhancement of enforcement efforts**

---

<sup>54</sup> See: <https://www.musicbusinessworldwide.com/youtube-creators-hit-by-music-copyright-claims-can-now-replace-tracks-with-ai-at-the-touch-of-a-button/>.

77. Data gathered and analyzed by AI systems can be used to support legal actions, including identifying relationship between sites, tracking physical location of the operators of pirate services and identifying advertisers on sites that host copyright-infringing content.

(e) Tools for user-generated content platforms

78. Many of the benefits to content owners discussed above also apply to online platforms that host user-uploaded content. AI-enhanced tools improve the ability of hosting platforms to more efficiently and accurately review the vast scale of user-uploaded files in order to filter out infringing content. They also enable the use of data on those uploads to identify malicious actors with repeat offences, while decreasing the number of false-positive identifications. This may serve to reduce piracy on platforms while simultaneously reducing anti-piracy costs and erroneous take downs that alienate users.

## VI. POTENTIAL CHALLENGES RELATED TO AI TOOLS

(a) Technological challenges

79. Few revolutionary technologies are perfect during the early stages of their development, and AI is no exception. While AI-enhanced tools for content protection are in the early stages of development, additional challenges can be expected to arise, for example, complications created by false positives, the inability to appropriately assess possible fair-use activity, and the use of embedded license content owned by third parties.

(i) *Efficacy of AI tools*

80. An initial challenge with assessing the reliability of AI-enabled tools and their ability to contribute towards addressing the piracy of copyrighted works is an issue inherent in all new technology: accurately determining the efficacy of a service. Some computer-assisted tools that are branded as AI tools do not actually contain the types of AI functionality described in this study, while the purported efficiency and effectiveness of some tools have yet to be confirmed by way of proof-of-concept testing. As with all AI tools, it is also important to guard against AI hallucination,<sup>55</sup> in order to ensure that accurate results are being delivered.

(ii) *False positives*

81. While tools which incorporate AI functionality have the potential for greater accuracy and sophistication in detecting pirated content, false positives are unlikely to ever be completely eliminated. As a result, a key part of assessing any AI tool is to consider how it tackles and learns from false positives, including false positives resulting from AI hallucination. When an AI tool relies on crowd sourcing for monitoring (discussed below), the tool is only as reliable as the crowd that is sourcing it, therefore, extra scrutiny should be applied.

---

<sup>55</sup> AI hallucination refers to cases where an AI model generates information that appears plausible but is factually incorrect, fabricated or unsupported by its training data. Large language models do not research or look up facts, they predict the next most likely information based on patterns learned during training. There have been numerous examples in the litigations where attorneys filed briefs containing non-existent case law provided by AI research. Even purpose-built legal AI tools have shown cases of hallucination. A Stanford study found hallucinated responses from the AI-assisted research of Westlaw (inventing a statement in the Federal Rules of Bankruptcy Procedure that does not exist), Lexis+ AI of LexisNexis (citing overruled precedent as current law) and Ask Practical Law AI of Thomson Reuters (failing to correct a the mistaken premise of a user and instead adding more false information). See: <https://hai.stanford.edu/news/ai-trial-legal-models-hallucinate-1-out-6-or-more-benchmarking-queries>.

82. Achieving the right balance between sensitivity (the ability to detect pirated content) and accuracy (the ability to avoid false positives) can be a significant challenge for AI-enhanced tools. Overly sensitive systems may flag content incorrectly, leading to unjustified take downs, platform pushbacks, and potential exposure to legal claims. Conversely, systems that are not sensitive enough may fail to detect instances of piracy. Refining AI algorithms to minimize false positives without compromising detection capabilities remains a critical necessity.

(iii) *Fair use determinations*

83. A fair use determination requires the consideration of multiple factors, including purpose and character of the use, nature of the copyrighted work, amount and substantiality of the portion used in relation to the copyrighted work as a whole, and the effect of the use upon the potential market for or value of the copyrighted work. Those determinations cannot be automated easily. However, unlike non-AI systems, AI machine learning has the capacity, over time, to produce informed determinations.

(iv) *Embedded licensed works*

84. The use of embedded licensed works within an audiovisual work has sometimes been a challenge in terms of avoiding false positives when flagging pirated copies. For example, the licensed use of the recording of a popular song in the soundtrack of a motion picture could be mistakenly flagged as a pirated copy of the recording. AI systems have the capacity to learn and respond to allow lists and can identify content that needs to be flagged for human review.

(v) *Crowdsourcing*

85. There are content protection vendors who are promoting their use of crowdsourcing to monitor and locate infringing content that was trained on AI content. While crowdsourcing is not AI, where it forms part of the data training for an AI tool caution is advised. Any tool that relies on crowd sourcing is only as reliable as the crowd that sourced it.

(b) *Legal and policy challenges*

86. The legal and policy concerns that have been discussed with regard to existing anti-piracy tools apply equally to tools that deploy AI technology functionality. Those concerns include respect for fair use, exceptions and limitations, avoidance of false positives and possible suppression of free speech. Any forms of AI-enhanced automation of content take down, site blocking or other remedies that lacks direct human oversight risk triggering those concerns. One advantage of AI is its capacity to improve over time as the scope and variety of training data sets expand and become more refined. Nonetheless, it will be important that those legal and policy concerns are not sacrificed during the training period.

(c) *Recommendations and good practices*

87. Good practices with all AI tools are the same as those related to all other content protection services and vendors: know your vendor, rely on tools with demonstrated accuracy and reliability rather than accepting exciting new claims based on face value, consider the possible risks of mistakes by the AI tool (which can range from negative public relations and consumer alienation to possible legal exposure) and do not overlook the ongoing need for human review, control and oversight.

88. While the potential value of AI-enhanced tools for content protection is significant, we are still in the very early stages of the development, testing and implementation of AI tools that will fulfill that potential. As such tools are developed, they also should be explored, in line with good practices, to enable informed decision-making about whether they are appropriate and fit for purpose.

## **VII. CONCLUSION**

89. AI-enabled tools for combating copyright piracy are still in the early stages of development but they are already showing strong potential to assist copyright owners in detecting and responding to piracy of their works. Traditional and AI-enabled anti-piracy tools face many of the same challenges, including with regard to false-positives and respect for the fair use of copyrighted material. Nonetheless, AI tools have the capacity for far more sophisticated forms of automation that can address those concerns while keeping pace with the ever-expanding threat posed to copyright owners by digital piracy. This is particularly relevant in combating the piracy of sporting and other live event broadcasts, where the window of value is shorter and the need for effective and reliable anti-piracy tools is greater. Such tools must be able to evolve rapidly in response to the ongoing evolution of digital piracy, with a view to harnessing the potential of AI technology in content protection measures. That is the value and promise of AI technology for content protection tools.

[End of document]