Current Developments on Labelling of AI-Generated Content: Comments on the AI Convention and the EU AI Act

Letrône, William

Hayashi, Mika

(Citation)
国際協力論集,33:59-67

(Issue Date)
2025-12

(Resource Type)
departmental bulletin paper

(Version)
Version of Record

(JaLCDOI)
https://doi.org/10.24546/0100498910

(URL)
https://hdl.handle.net/20.500.14094/0100498910

［論　説］

Current Developments on Labelling of AI-Generated Content: Comments on the AI Convention and the EU AI Act

William Letrône＊1
Mika Hayashi＊2

Abstract

The note addresses the growing concern over disinformation fueled by AI-generated content（AIGC）, particularly images and videos, and makes a snap-shot analysis of the two legal instruments on AI: the Council of Europe's AI Convention, an international treaty, and the AI Act, a regulation of the EU, both of which were adopted in 2024. As a background to this recent development, the note identifies a broad consensus among governments and international organisations that content provenance and authentication, understood as the identification of AIGC, are the key strategy to combat misinformation and disinformation of AIGC. There is also common acknowledgment that this is a matter of human rights, especially freedom of expression, according to the official discourse in international organisations. In this regard, the AI Convention promotes common principles rooted in human rights but leaves all the details to States Parties. It does not propose any labelling obligation for states or for AI businesses. In contrast, the EU AI Act establishes regulations regarding labelling of AIGC, and has specific rules regarding deep fakes, with clear obligations for AI providers and deployers. Other regulatory frameworks within the EU are likely to complement and support the AI Act, too. In the AI Act and indeed in any other instrument, while labeling AIGC enhances transparency, its effectiveness is reduced if different techniques are used on different platforms. The standardisation of these different techniques of labelling is needed. Its effectiveness also depends on the user AI literacy. There is also a need to heed the impact of labelling on privacy, and more generally, free speech.

I. Introduction

Israel launched strikes on Iran on 13 June 2025, and the following days saw several rounds of Iranian missile attacks on Israel. In a very short period of time, visuals generated by artificial intelligence（AI）tools, falsely claimed as real, started to emerge, and soon, there was a veritable "torrent" of

＊1　Postdoctoral researcher, CNRS, at DCS（Nantes University-France）, IPoP project.
＊2　Professor, Graduate School of International Cooperation Studies, Kobe University.

disinformation using generative AI.[1] For example, destruction of buildings in Tel-Aviv, allegedly as a result of Iranian attacks, was shown in a realistic-looking video produced by AI.[2] It was shared on multiple social media platforms. At the time of this writing, the Teheran Times still maintained the video as "Doomsday in Tel-Aviv" on its X account,[3] even though it did not actually happen. While false content can be produced by both humans and generative AI, the presence of generative AI today brought the problem of online disinformation to a new level. Realistic-looking images and videos can now be produced with ease, at low-to-no-cost, and at an extremely high rate. They are viewed, followed and circulated infinite number of times.

In the sphere of international law, solutions to this particular problem of AI are not immediately clear. What emerges from the examination is that content provenance and authentication, understood as the identification of AI-generated content (AIGC), are deemed important, and that labelling is a preferred technical solution. Moreover, the concern is framed as that of international human rights law. What that concretely means in terms of obligations for states and AI businesses remains, in contrast, unclear (Section II). Against that backdrop, there are currently two binding instruments of international or supra-national calibre regarding AI: the Council of Europe's Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (the AI Convention)[4] and the Artificial Intelligence Act (the AI Act),[5] a regulation of the European Union (EU). As seen in their titles, both instruments affirm human rights as a framework in tackling challenges of AI. Therefore, their approaches to the labelling of AIGC as a means to bring transparency about content provenance, and possibly a remedy to the problem of disinformation, are examined (Section III). While every aspect of AI technology falls within the scope of both the AI Convention and the EU AI Act, the focus of this note is on their approaches to content provenance and authentication. Other human rights issues of AI, such as privacy and data protection, are not dealt with by this note.

Disinformation in this AI context means purposefully passing off, or falsely claiming, AI-generated content as real events. The term misinformation is used in a broader sense, where the mental state of the disseminator of the AIGC is irrelevant.[6] While AIGC can be texts, and not just images and videos, a major problem of disinformation, illustrated by the example of Iran-Israel conflict above, is that of images and videos. The rules regarding "deep fake" in the EU AI Act are also about image, audio or video content. For these reasons, AIGC, when used in this note by the present authors, refer to AI-generated images, videos and audio content. In many, but not all, documents cited in this note, the same term includes AI-generated texts.

II. Current landscape

1. General consensus: content provenance matters

Social media companies acknowledge that their users want to know the provenance of the content, in particular, the images. Thus, the Community Guidelines of TikTok require its users to label AI-generated visual content.[7] Meta, the current owner of Instagram, Facebook and WhatsApp, confirms that

"As the difference between human and synthetic content gets blurred, people want to know where the boundary lies."[8] Users of these social media platforms are therefore asked to "disclose when they share AI-generated video or audio so we can add a label to it."[9] Since May 2025, Google's own generative AI model, Veo 3, adds visible watermark to videos it generates. There is also an initiative called the Coalition for Content Provenance and Authenticity (C2PA), an initiative joined by Google, Microsoft, Intel, the BBC and others to develop interoperable provenance labels in the form of Content Credentials.[10]

States have expressed increasingly strong concerns regarding content provenance and authentication in the context of generative AI, too. These concerns are shown in their national regulatory efforts,[11] and are further confirmed in international fora. The UN General Assembly in its Resolution in 2024 entitled "Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development" thus calls upon the UN Member States to encourage the development and deployment of "reliable content authentication and provenance mechanisms – such as watermarking or labelling [...]."[12] The United Nations' Global Principles for Information Integrity (2024) also show a serious concern for risks introduced by generative AI.[13] Seeing content provenance and authentication as a way forward for AI, this report recommends, *inter alia*, labelling, too.[14] A year earlier, the members of the Hiroshima AI process, a G7 initiative, similarly encouraged the development and the deployment of "reliable content authentication and provenance mechanisms, where technically feasible, such as watermarking or other techniques to enable users to identify AI-generated content."[15] The OECD and UNESCO had also expressed similar views on the importance of identification of AIGC.[16]

To sum up, there is currently general consensus that content provenance matters. Moreover, labelling AIGC is seen as an answer, or at least a realistic remedy, to the problem of disinformation using AIGC, in many of the international documents previously mentioned. Furthermore, the frequent coupling of the two terms, content provenance and content authentication, indicates that the interest in content provenance stems from a concern for misinformation or disinformation involving AIGC.

2. Common acknowledgment: content provenance is a matter for human rights

There also appears to be common and consistent acknowledgment that content provenance of AIGC is an issue for human rights and international human rights law, because of the problem of disinformation using AIGC. The UN Human Rights Council squarely sees disinformation, including the one by digital technology, as an issue of human rights.[17] So did a number of regional organisations and UN Special Rapporteurs, as early as in 2017.[18] More recent UN reports and documents also view disinformation involving digital technology, including AI, as a human rights issue.[19] The UN Global Principles for Information Integrity (2024), previously mentioned, are also "grounded in international law, including international human rights law," and affirm "an unwavering commitment to human rights."[20]

Why is content provenance, and as a consequence, labelling of AIGC, important in light of this human rights concern? Because users as individuals must be capable of identifying and distinguishing "the authentic digital content and artificial intelligence-generated or manipulated digital content."[21] Why is it important that individuals can make this identification or distinction? Because individuals' self-

determination and self-fulfilment can only be realized upon certain conditions, and the right to freely form an opinion is one of such indispensable conditions,[22] whereas if one is fed with lies, this aspect of self-determination is harmed. A long line of national and international regulations of false speech in the context of human rights law attest to this thinking.[23] Therefore, labelling of AIGC is a matter for, in a traditional human rights language, freedom of expression and the right to freely form an opinion: passing off AIGC as real is a human-rights matter.

To sum up, there is common acknowledgement that states must be guided by international human rights standards, especially freedom of expression, in dealing with AIGC. There is also common acknowledgement that the fight against AI-generated disinformation must be a multi-stakeholder effort, and that the AI businesses should be guided by international human rights standards, too. What is not so clear is the consequences of these, seemingly uncontroversial ideas. Beyond the formal discourse referring to human rights in the UN and beyond, what these human rights standards actually mean in terms of obligations of states and AI businesses is not easy to ascertain. Does this common discourse mean that freedom of expression sets a positive obligation for states to take positive measures vis-à-vis AI businesses, so that the individuals are protected from disinformation using AIGC? Does "the recognition of the right to be informed translate[s] into providing for the right to correct information"?[24] In academic writings, such ideas in the European context are explored.[25]

3. Advantages of labelling as a solution

As previously mentioned in Section II.1, labelling of AIGC is regularly referred to as a solution, and is obviously an attractive answer to the problem of disinformation using AIGC. There are two reasons for this popularity.

The first reason is that labelling AIGC is technically feasible.[26] For example, as mentioned previously, videos generated by Veo 3 already carry visible watermarks at the bottom corner. Asking a user to add a sticker to an AIGC, which social media companies do, is also certainly feasible. The issue of evaders that do not label the AIGC, as well as the issue of how robust the watermarking can be, of course remain, but at least labelling AI-generated images and videos is a feasible option.

The second reason why this can be an attractive solution relates to the businesses that offer social media platforms. It is easy to say that these companies should be guided by human rights standards related to freedom of expression, but combating online disinformation would require a drawing of a line between what is disinformation and what is not. This is a huge and complex task for private businesses. There is also a risk that the companies take this task too seriously and take off too many content, perhaps offensive but not disinformation.[27] Compared to the daunting responsibility of being a truth arbiter in this way, the responsibility of requiring labelling AIGC is relatively straightforward. If an image is generated by AI, then it must be labelled. Labelling addresses only a small portion of the problem of online disinformation and has its own challenges, but all things considered, it is an attractive solution worth pursuing.

III. The AI Framework Convention and the EU AI Act

The summary above in Section II. 3 is confirmed by two current international or supra-national binding instruments. These instruments are the AI Convention and the EU AI Act, both of which were adopted in 2024. Both of them look to labelling as a way to ensure transparency of AIGC, though there are also marked differences.

The AI Convention will come into force after five states including three member states of the Council of Europe ratify it, and the EU AI Act will be fully in force in 2026. Both of them, as their official names testify, approach AI with human rights framework.

1. AI Convention

While the AI Convention is a product of the Council of Europe,[28] it has already been signed by states such as the US, Canada and Japan since its adoption on 17 May 2024. The AI Convention is a framework convention, which means this is not a treaty that establishes specific regulations to be imposed on states or AI deployers, developers and social media companies. Instead, the treaty establishes obligations of general nature and common principles. Thus, the treaty brings no immediate change to AIGC regulations.

From the perspective of this note, what it does nonetheless is that it takes "human rights, democracy and the rule of law" very purposefully as frameworks in formulating these common principles and general obligations regarding AI, as the treaty name confirms. Indeed, the first principle the treaty identifies is that of human dignity and individual autonomy in relation to AI (Art. 7). Given this particular perspective, the problem of disinformation using AIGC is an acknowledged concern.[29] Concretely, there is a need for identifying AIGC "in order to avoid the risk of deception and enable distinction between authentic, human-generated content and AI-generated content as it becomes increasingly hard for people to identify."[30] An important general obligation of States Parties in this regard is to have measures in place for individuals' "ability to freely form opinions" (Art. 5(2)). In order to freely form opinions and act on these opinions, the access to information is indispensable. Disinformation and misinformation via AIGC can undermine information integrity and the right of access to information, in particular, in the context of elections.[31]

As strong as the concern may be, what is actually required of States Parties in the AI Convention is to have measures that ensure adequate transparency requirements, without a further specification (Art. 8). Though measures "with regard to the identification of content generated by artificial intelligence systems," thus, labelling of AIGC, is included in the scope of this article, the exact measure of identification and the nature of the requirement are entirely left in the hands of each State Party. As such, the AI Convention only confirms the importance of identifying AIGC, as many non-binding documents in the past surveyed in Section II did.

There is also a structural limitation in the AI Convention: its target by default is public authorities and private actors acting on their behalf (Art. 3(1)(a)). Obviously, the identification of AIGC as a measure to ensure transparency does little to limit disinformation if the scope of the treaty does not extend to

companies developing and deploying AI, social media companies, and other news outlets. This limited scope of the treaty was reportedly a compromise to keep the US interested, which signed the treaty subsequently. Each State Party is asked to decide, either at the time of signature or ratification, how the obligations apply to businesses that do not act on behalf of the public authorities. It can either accept the Convention as directly applicable to the private actors or to prepare other measures that ensure the application of the treaty's relevant principles to private actors (Art. 3(1)(b)). Upon signature, Norway notified that it would apply them to the private actors, and Ukraine notified that it would go the other way. Other states that signed so far did not express their preferences upon signature.

2. The EU Artificial Intelligence Act: labelling as an obligation

The EU AI Act, in contrast, is a set of regulations of direct applicability, intended for the Member States and businesses alike. Unlike the AI Convention, the AI Act does not give states the possibility to exclude the AI businesses from its scope. As a European regulation, it is directly applicable in the Member States.

The AI Act affirms the primary place of rights protected by the EU Charter of Fundamental Rights, and freedom of expression and information is of course one of these fundamental rights.[32] In contrast to the AI Convention's general principle of transparency which suggests labelling only as a possibility, the AI Act sets clear obligations of transparency, including labelling, for AIGC. The manner in which the AI Act addresses disinformation has already attracted considerable scholarly attention.[33]

The labelling of AIGC including AI-generated texts is outlined in Article 50. While the AI Act establishes a risk hierarchy which treats AI systems differently according to the gravity of risks they entail, Article 50 is applicable to AI systems of all risk categories, so long as the system is intended to interact directly with natural persons (Art. 50(1)), or the system generates synthetic audio, image, video or text (Art. 50(2)). The basic rule is that the outputs of the AI system must be "marked in a machine-readable format and detectable as artificially generated or manipulated" (Art. 50(2)). It is not clear whether marking AIGC in a machine-readable format, often invisible, can simultaneously satisfy the requirement that information under this article "shall be provided to the natural persons concerned in a clear and distinguishable manner" (Art. 50(5))[34] there may be a need to provide the information in a human-readable manner, too.[35] At any rate, the obligation of marking is clearly stated. The providers of an AI system generating AI content must in principle also ensure that their technical solutions for labelling are "effective, interoperable, robust and reliable as far as this is technically feasible" (Art. 50 (2)).[36]

In line with the focus on human rights and the concern for disinformation using AIGC, the AI Act also tackles deep fake explicitly. Deep fake in the AI Act is "AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places or other entities or events and would falsely appear to a person to be authentic or truthful" (Art. 3(60)). There is a specific transparency requirement for deep fake deployers, i.e., professional users: they must disclose that the content has been artificially generated or manipulated (Art. 50(4)). In this disclosure, deep fake deployers should "clearly

and distinguishably disclose that the content has been artificially created or manipulated by labelling the AI output accordingly and disclosing its artificial origin."[37]

In their current version, the transparency obligations contained in Article 50 only target the providers and deployers of AI systems, and thus, does not impose transparency obligations on social media companies. Nevertheless, the AI Act acknowledges the complementary relationship between its Article 50 and the responsibilities of providers of social media platforms under the Digital Services Act (DSA).[38] The DSA, adopted in 2022, sets out obligations to assess and mitigate systemic risks caused by online content for providers of very large platforms. Moreover, the Code of Conduct on Disinformation, which "aims to combat disinformation risks while fully upholding the freedom of speech and enhancing transparency under the Digital Services Act (DSA)," has been integrated into the DSA's regulatory framework.[39] The Code contains a specific reference to "transparency obligations for AI systems," and AI businesses as signatories are asked to take into consideration "the transparency obligations and the list of manipulative practices prohibited under the [AI Act]."[40] The signatories of the Code also recognize the importance of "the potential of provenance technology to empower users with tools to interrogate the provenance and authenticity of content [...]."[41] The labeling of AIGC in the AI Act is likely to be supported by these regulatory frameworks.

IV. Conclusion

The challenges that the rapid development of AI brought to our world are daunting. Regarding disinformation and AIGC, the best way forward in light of human rights concerns, chosen in the two binding international/supra-national instruments examined above, is labelling of AIGC. In particular, the EU AI Act goes as far as setting an obligation of labelling.

There is no doubt that labelling can, if well implemented, increase transparency, and can help individual users to identify AIGC. AI businesses are exploring and experimenting labelling technologies on different platforms.[42] Consequently, an immediate, practical challenge is the standardisation or interoperability of different labelling technologies. The AI Act is in fact mindful that techniques and methods of labelling should be interoperable.[43] AI businesses largely agree.[44] Labelling must be a user-friendly technique for individual users, preferably common across different platforms.

At the same time, the downside of labelling must also be stressed from human rights perspective. Content provenance shown by labelling strengthens transparency, but also introduces new problems. A downside of labelling in light of human rights concerns is its impact on privacy. Labelling AIGC is likely to mean, in many cases, content traceability. In fact, as soon as labelling and watermarking started to be discussed, the concern was raised that they "could enable increased monitoring, recording, or disclosure of an individual's media purchases or usage."[45] Such disclosure, as well as authorship disclosure, could also have a chilling effect on freedom of speech, and is a serious side effect to be taken into account.

Finally, from the same human rights perspective, what should go hand in hand with labelling of AIGC is AI literacy of individual users. Labelling increases transparency but its utility is limited if individual

users did not know what to look for when presented with images and videos. Labelling may even be counter-productive if labels are perceived, or misunderstood, to guarantee anything other than the fact that the content is AI-generated.[46] The importance of AI literacy, together with targeted campaigns and training, is stressed in practically every comment and international document on this topic. While legal instruments such as the AI Act recognize the importance of AI literacy for individual users, concrete implementation measures are still lacking. An obligation of labelling or a policy of labelling for the sake of human rights is for the states to promote, and for the AI businesses to implement. Meanwhile, individual users need to remain vigil. To have trustworthy AI which is respectful of human rights is a multi-stakeholder effort, and that includes each of us.

Notes

1 Matt Murphy et al., "Israel-Iran conflict unleashes wave of AI disinformation", *BBC News* (21 June 2025) 〈https://www.bbc.com/news/articles/c0k78715enxo〉 accessed 22 August 2025.
2 AFP Middle East & North Africa, AFP Indonesia, "AI-generated video falsely shared as aftermath of Iran's attack on Israel", *AFP Fact Check* (23 June 2025) 〈https://factcheck.afp.com/doc.afp.com.62ZG3ZP〉 accessed 30 July 2025.
3 Tehran Times, "Doomsday in Tel Aviv", 〈https://x.com/TehranTimes79/status/1933850991599767749〉 accessed 28 August 2025.
4 Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, CETS No. 225.
5 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence.
6 Marko Milanovic and Philippa Webb, "False Speech", in Amal Clooney and David Neuberger (eds.), *Freedom of Speech in International Law, 1* (Online edn, Oxford Law Pro, 2024).
7 TikTok Community Guidelines (Released 14 August 2025, effective 13 September 2025), in the section on "Integrity and Authenticity" 〈https://www.tiktok.com/community-guidelines/en-GB〉 accessed 29 October 2025.
8 Nick Clegg (Meta), "Labeling AI-Generated Images on Facebook, Instagram and Threads" (6 February 2024) 〈https://about.fb.com/news/2024/02/labeling-ai-generated-images-on-facebook-instagram-and-threads/〉 accessed 31 July 2025.
9 *Ibid.*
10 For more information on social media platforms' policies regarding generative AI, Sandra Höltervennhoff et al., "Security Benefits and Side Effects of Labeling AI-Generated Images" *arXiv* (28 May 2025) 〈https://arxiv.org/abs/2505.22845〉 accessed 27 August 2025.
11 A few examples are described in Alexis Léautier,《Panorama et perspectives pour les solutions de détection de contenus artificiels [1/2]》*LINC* (27 octobre 2023) 〈https://linc.cnil.fr/panorama-et-perspectives-pour-les-solutions-de-detection-de-contenus-artificiels-12〉 accessed 30 July 2025; Höltervennhoff et al., *supra* note 10.
12 A/RES/78/265 (1 April 2024).
13 "The Global Principles for Information Integrity: Recommendations for Multi-stakeholder Action" (2024) p. 8. 〈https://www.un.org/sites/un2.un.org/files/un-global-principles-for-information-integrity-en.pdf〉 accessed 31 July 2025.
14 *Ibid.,* Recommendations, d.
15 "Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI Systems" (2023), para. 7 〈https://www.mofa.go.jp/files/100573471.pdf〉 accessed 31 July 2025.
16 Léautier, Panorama, *supra* note 11.
17 "Role of States in countering the negative impact of disinformation on the enjoyment and realization of human rights", A/HRC/RES/49/21 (8 April 2022).
18 UN Special Rapporteur on Freedom of Opinion and Expression, OSCE Representative on Freedom of the Media, OAS Special Rapporteur on Freedom of Expression and the African Commission on Human and Peoples' Rights Special Rapporteur on Freedom of Expression and Access to Information, "Joint declaration on freedom of expression and 'fake news', disinformation and propaganda", FOM.GAL/3/17 (3 March 2017) 〈https://www.osce.org/fom/302796〉 accessed 31 July 2025.
19 E.g. UNESCO Guidelines for the Governance of Digital Platforms (2023); Report of the Secretary-General, "Countering disinformation for the promotion and protection of human rights and fundamental freedoms", 2022

(A/77/287); UNGA Resolution, *supra* note 12.
20 UN Principles, *supra* note 13, Recommendations, d.
21 UNGA Resolution, *supra* note 12.
22 HRC, "General Comment No 34: Freedoms of Opinion and Expression (Art 19)" (12 September 2011), UN Doc CCPR/C/GC/34, para 2.
23 Milanovic and Webb, *supra* note 6.
24 Giovanni De Gregorio and Pietro Dunn, "Artificial Intelligence and Freedom of Expression", in Jeroen Temperman and Alberto Quintavalla (eds.), *Artificial Intelligence and Human Rights* (Online edn, Oxford Law Pro, 2023), p. 77.
25 *Ibid.*, p. 77 and pp. 85–88. One of the principal architectures of this European context, the EU AI Act, will be examined in Section III.
26 Alexis Léautier, 《Le tatouage numérique, une mesure de transparence salutaire? [2/2]》 *LINC* (27 octobre 2023) 〈https://linc.cnil.fr/le-tatouage-numerique-une-mesure-de-transparence-salutaire-22〉 accessed 30 July 2025.
27 See also Milanovic and Webb, *supra* note 6.
28 For the drafting history, see Elżbieta Hanna Morawska, "Council of Europe Standards and Activities related to AI: towards a Framework Convention on AI and Human Rights?", in Michał Balcerzak and Julia Kapelańska-Pręgowska (eds.), *Artificial Intelligence and International Human Rights Law* (Edward Elgar Publishing, 2024).
29 Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, para. 43.
30 *Ibid.*, para. 59.
31 *Ibid.*, para. 43.
32 AI Act, the preamble (para. 48).
33 Matúš Mesarčík and Natália Slosiarová, "Regulating AI for a Truthful Tomorrow: Addressing Disinformation in the EU Artificial Intelligence Act", *International Journal of Law and Information Technology*, 33 (2025); Thomas Gils, "A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act", in Ceyhun Necati Pehlivan et al. (eds.), *The EU Artificial Intelligence (AI) Act: A Commentary* (Kluwer Law International, 2024); Mauro Fragale and Valentina Grilli, "Deepfake, Deep Trouble: The European AI Act and the Fight Against AI-Generated Misinformation" *CJEL: Preliminary Reference* (11 November 2024) 〈https://cjel.law.columbia.edu/preliminary-reference/2024/deepfake-deep-trouble-the-european-ai-act-and-the-fight-against-ai-generated-misinformation/〉 accessed 8 August 2025.
34 Gils, *supra* note 33.
35 Höltervennhoff et al., *supra* note 10.
36 See also AI Act, Recital 133.
37 AI Act, Recital 134.
38 AI Act, Recitals 120 and 136.
39 European Commission, The Code of Conduct on Disinformation (13 February 2025) 〈https://digital-strategy.ec.europa.eu/en/library/code-conduct-disinformation〉 accessed 27 August 2025.
40 *Ibid.*, Commitment 15.
41 *Ibid.*, Section V (Empowering Users), b.
42 Léautier, Panorama, *supra* note 11.
43 AI Act, Article 50(2); Recital 133.
44 Laurie Richardson (Google), "How we're increasing transparency for gen AI content with the C2PA" (17 September 2024) 〈https://blog.google/technology/ai/google-gen-ai-content-transparency-c2pa/〉 accessed 8 August 2025.
45 Center for Democracy and Technology, "Privacy Principles for Digital Watermarking-May 2008, Version 1.0" 〈https://cdt.org/wp-content/uploads/copyright/20080529watermarking.pdf〉 accessed 29 August 2025.
46 Höltervennhoff et al., *supra* note 10.