

The Assistive Exemption

Content provenance marking and the compelled disclosure of disability

Europe told AI companies to mark what their models write. The same sentence exempts tools that only help a person edit their own words. The first companies to build it marked everything anyway.

THE SHORT VERSION

Somebody wrote a rule to protect people who write with help

Europe now requires AI companies to hide a mark inside the text their models produce, so a machine can later tell the text came from a machine.

The same sentence that creates that duty says it does not apply when the tool is only helping a person edit their own writing.

That sentence is the reason this paper exists. The first companies to build the mark did not use it.

THE CLAIM

The mark cannot tell writing from editing, and the companies say so themselves

What the mark proves

Only that a particular company's model touched the text at some point. The company's own help page says a hit means the text may have been processed by its model, and that the model may not be the author.

What it cannot separate

A model that wrote the paragraph from a model that fixed the grammar in a paragraph a person wrote. Two of the three uses the company lists as indistinguishable are uses Europe exempted.

Who pays for that

People who use these tools because they cannot produce the document otherwise. For them the tool is not a shortcut. It is how the writing happens at all.



The Rule

What Europe actually asked for, and what it excused

Article 50(2) creates the marking duty and limits it in the same sentence. The limit has now been defined by the European Commission, and it is broader than anyone implemented.

THE DUTY

Mark it so a machine can find it later

Companies that build text-generating models have to leave something inside the output that software can detect. A watermark, hidden metadata, a cryptographic signature, or something similar. The rule does not say which.

It started on 2 August 2026. Models already on sale get until 2 December 2026.

It reaches companies outside Europe only when the output is used in Europe.

THE EXCEPTION

The duty stops where the tool is only helping you edit

"shall not apply to the extent the AI systems perform an assistive function for standard editing or do not substantially alter the input data provided by the deployer or the semantics thereof"

Article 50(2), quoted because the exact words are what the argument turns on

Two things to notice. It says to the extent, which means it can switch on and off depending on what the tool just did. And it asks what the tool did, not what the user is. Nobody has to claim it or prove anything.

WHAT THE COMMISSION SAYS COUNTS

Fixing your grammar is not writing your paper

Inside the exception

Grammar correction and spellchecking. Small stylistic polish that does not change what you meant. Machine translation of your text. Transcription of speech.

Named explicitly

Tools that convert a disabled person's input into speech or text so they can communicate, including communication boards and synthesised voices, because they do not change the meaning.

Switches per use

Where one tool does both jobs, the duty applies only to the parts it actually generated. The Commission says this in as many words.

WHERE THE LINE FALLS

It covers people who cannot speak. It does not cover people who cannot organise.

The exception protects assistive technology that carries what you already meant. A speech device. A synthesised voice. Translation.

Summarising a document, or reworking its structure, sits outside the exception on the Commission's own examples. So does substantive rewriting.

If your disability affects speech, the rule shelters you. If it affects attention, memory, or the ability to organise a thought into a paragraph, it does not. The line runs through the middle of the same population, and nothing in the public record suggests anyone noticed where it landed.



What Got Built

Marking everything, everywhere, for everyone

The rule is a floor. Companies may mark more than they must, and the first one to ship did.

AUGUST 10, 2026

The first frontier lab shipped it across everything at once

Anthropic announced that its Claude models now weave a mark into generated text. It survives copying and pasting. Users cannot see it and cannot remove it.

It covers the products people actually use to work: the chat interface, the coding tool, the developer platform, and the models sold through the big cloud providers. Images and vector files get signed provenance data instead.

The detection tool is not out yet. The company says it is building it for businesses, educators, and third parties.

THREE GAPS

Further than the law asked, in three separate directions

- 1 What it marks**

The Commission exempts grammar correction and translation. The company lists proofreading and translation among the things its mark cannot rule out. It marks the exempt work and the non-exempt work identically, and then says it cannot tell them apart.
- 2 Where it marks**

The law reaches non-European companies only when the output is used in Europe. The mark applies wherever the product is sold. A subscriber in Illinois whose writing never touches Europe is marked under a European rule that does not reach him.
- 3 Whose writing it marks**

The law leaves personal, non-commercial activity out of scope entirely. The mark does not distinguish a court filing from a birthday card.



Who Reads It

The document that decides who gets the key

The Code of Practice concedes the text signal is unreliable, then guarantees access to it. The person whose writing gets checked is not mentioned anywhere in it.

THE COMPLIANCE DOCUMENT

It admits the signal is unreliable, then hands it to schools

The Code of Practice is what companies sign to show they are complying. On text watermarking it says, in its operative text, that detection has a lower level of reliability and robustness and may produce misleading or low-confidence results.

In the very next breath it says the results are still valuable to verified expert users, and it names them: market surveillance authorities, regulators, law enforcement, media, fact-checkers, trusted flaggers, independent researchers, civil society organisations, and educational and research organisations.

Universities are the largest deployer of AI detection in the world, and they already have a disciplinary process pointed at this exact question.

THE RECEIPT

One party gets a signed record. The other gets nothing.

38

pages in the Code of Practice. It requires that a detection result be downloadable as a digitally signed file carrying a hash of the text checked, an identifier for the tool, and a timestamp. That is a cryptographic receipt, built for whoever ran the check.

Search those pages for the person whose writing was checked. There is no requirement to tell her a check happened. No right to the receipt. No route to contest the result. She is not in the document.

ACCESSIBILITY, ONE DIRECTION ONLY

The rules make sure a blind person can read the verdict, not that a disabled person can contest it

Where accessibility appears

Notices to people exposed to AI content must meet accessibility standards. The screens that display a detection result must be accessible too, under the European Accessibility Act.

Where it does not

Nothing constrains who gets marked, or what marking costs them. The Guidelines say plainly that Article 50 imposes no accessibility requirements of its own.

The tell

The Guidelines say marking is about how content was made, not who created it. Correct as data protection. It also describes a set of rules that has defined the writer out of its own picture.

SOURCES European Commission, Guidelines on the Implementation of the Transparency Obligations for Certain AI Systems Under Article 50 of the AI Act (July 20, 2026). European Commission, Code of Practice on Transparency of AI-Generated Content (June 10, 2026).

IV

What It Does To People

Detection as actually practised

The mark arrives into an enforcement culture that was already built, already deployed, and already documented as failing in a pattern.

THE FAILURE UNDERNEATH ALL OF IT

People cannot do this, and practice makes them worse

39.93%

How often readers correctly identified AI-written short fiction in a 2026 study. Worse than flipping a coin. In the same study they rated the machine-written stories more highly than the human ones.

One study of teachers put trainees at 45.1 percent and experienced teachers at 37.8 percent. Experience made them worse. And accuracy falls as models improve: 57.9 percent against an older model, 49.9 percent against a newer one.

THE LIST PEOPLE MADE UP INSTEAD

Every item on it is ordinary English

The dash

Dickinson wrote almost entirely in them. So did Melville and Woolf.

The list of three

Of the people, by the people, for the people.

Negation then affirmation

Antithesis. One of the oldest named devices in Western rhetoric, unbroken from the King James translators through Dickens, Lincoln, and Kennedy.

Faced with a task they cannot perform, people substitute a rule they can apply. The rule punishes formally correct prose, which means it punishes hardest the writers drilled hardest on the rules: second language writers, and people whose thinking applies rules literally. Detection software already over-flags the same group for a different reason. Two mechanisms, one population.

SOURCES Liang et al., GPT Detectors Are Biased Against Non-Native English Writers, 4 Patterns 100779 (2023); Weber-Wulff et al., Testing of Detection Tools for AI-Generated Text, 19 Int'l J. for Educ. Integrity art. 26 (2023); Gegg-Harrison & Quarterman, AI Detection's High False Positive Rates and the Psychological and Material Impacts on Students (IGI Global 2024).

THE BASELINE NOBODY MENTIONS

No assistive technology has ever carried a mark

Dictation software, screen readers, braille displays, word prediction, and text to speech have all sat between a disabled person and a document for decades. None of them has ever left an identifier in the file.

Courts have gone further than tolerating this. In 2011 a federal appeals court ordered a national bar examiner to let a blind applicant use specific screen-reading and magnification software, holding that the law demands administration that best ensures the score measures ability rather than disability, a higher bar than mere reasonableness. The same court said the list of permitted aids is not closed, because accommodations have to keep pace with technology.

Fifteen years later the newest member of that category ships with a signature, and the compliance code guarantees schools the ability to read it.

THE MECHANISM

It does not compete with disclosure. It goes around it.

A writer's disclosure lives in the document, because that is the only place it can live. Law reviews put it in a footnote on page one. Everything else puts it in acknowledgements at the back, where it satisfies the record and reaches almost nobody.

A detection result does not live in the document. It lives in a separate system, queried by somebody doing a different job, at a moment the writer does not choose and is never told about. The person running the check is not reading the paper. There is no place inside a document where a sentence reaches them.

So the finding gets made in a system the writer has no standing in, from a signal she cannot inspect, while her own statement on that exact subject sits in a file the decider has no reason to open. Writing a better disclosure cannot fix that. You cannot write your way into a channel you have no access to.

FOR EVERYONE WHO HAD NOT PLANNED TO DISCLOSE

Two options, and the only winning move costs you your medical privacy

Explain

The true and complete answer is that you use the tool as an accommodation for a disability. Saying it hands protected medical status to somebody who never asked for it and owes you nothing in keeping it.

Say nothing

The finding stands, and with it the inference the rule was written to produce: that you used a prohibited tool and hid it.

The injury

Being put to the choice. Everyone else caught by the same rule pays with a misconduct finding. This group pays with a civil rights interest.

ONE RULE, TWO PRICES

The law already has a name for this

Used it because it was faster

Detected, she faces an integrity finding. That is the loss the rule was designed to impose on her.

Identical treatment, applied identically, extracting two different things from two groups, and only one of those things is protected. Courts have recognised since 1985 that treating everyone the same can itself be the denial, when sameness is what produces the exclusion.

Used it because the document does not exist otherwise

Detected, she faces the same finding unless she spends something the first writer is never asked to spend.

THE OBVIOUS ANSWER, AND WHY IT FAILS

Register your accommodation in advance

- 1 There is no office**
Registration exists inside institutions. A journal has none. Nor does a conference committee, a preprint server, a newsletter, a grant panel, or a platform. That is where independent and disabled writers actually publish.
- 2 Registering is the disclosure**
Requiring it in advance moves the same choice earlier and makes it compulsory. Getting a legal entitlement means proving membership in a stigmatised group, and the access is renegotiated every term.
- 3 Accommodation expires, the mark does not**
A file at one institution in one year does nothing when the same paragraph surfaces four years later at a venue that did not exist when it was granted. Control over disclosure was never only about whether and to whom. It was about when.



The Law and the Gap

Who can be sued, and who cannot

The claims are weakest where the anger is strongest, and strongest where nobody has looked.

SUING THE COMPANY

The weakest case on the table

Disability law

Whether an online-only service is a public accommodation at all is split across the federal circuits, and in several of them it is not. Even where it is, no case has treated adding a feature as a denial of the service.

Consumer protection

More plausible and untested. Illinois asks whether a practice offends public policy, is oppressive, or causes substantial consumer injury, and says all three need not be met. There is no opt out and no unmarked tier, which is the strongest element.

Contract

The company assigns you the output, disclaims every warranty by name, and reserves the right to change the service without notice. You own a thing whose properties the other side may alter after the deal, warranted as to nothing.

SOURCES Robles v. Domino's Pizza, LLC, 913 F.3d 898, 905 (9th Cir. 2019); Carparts Distrib. Ctr., Inc. v. Automotive Wholesaler's Ass'n of New England, Inc., 37 F.3d 12, 19 (1st Cir. 1994); Robinson v. Toyota Motor Credit Corp., 201 Ill. 2d 403, 417-18 (Ill. 2002); 815 ILCS 505/2; 15 U.S.C. § 45(n); Anthropic, Consumer Terms of Service §§ 4, 11, 12 (2026).

SUING THE SCHOOL OR THE EMPLOYER

Considerably stronger, and almost nobody has tried

- 1 Public schools and anyone taking federal money**
Running detection and imposing consequences, with no route for an accommodated writer that avoids the disclosure, is a differential burden on meaningful access. Emotional distress damages are off the table; compensatory damages for lost opportunities are not.
- 2 Employers**
Treating a detection hit as misconduct, rebuttable in practice only by explaining an accommodation, builds a practice likely to elicit disability information. The employer never asks the forbidden question. It arranges the circumstances where the answer is volunteered.
- 3 Testing and credentialing bodies**
They carry their own duty to run examinations accessibly, measured by a best-ensures standard rather than reasonableness. No case has yet applied it to an integrity or proctoring dispute.

SOURCES 29 U.S.C. § 794(a); *Alexander v. Choate*, 469 U.S. 287, 301 (1985); *Cummings v. Premier Rehab Keller, P.L.L.C.*, 596 U.S. 212 (2022); *Payan v. L.A. Cmty. Coll. Dist.*, 169 F.4th 971, 977-79 (9th Cir. 2026); 42 U.S.C. § 12112(d)(4)(A); *Conroy v. N.Y. State Dep't of Corr. Servs.*, 333 F.3d 88, 94-95, 97 (2d Cir. 2003); 42 U.S.C. § 12189; *Enyart v. Nat'l Conference of Bar Examiners, Inc.*, 630 F.3d 1153, 1161-62 (9th Cir. 2011).

THE GAP

At the moment the harm lands, nobody is holding it

The company

Built the signal. Has no relationship with the school that read it, no duty to the person marked, and a European instrument to point at.

Accommodation law assumes two parties in a relationship: somebody asks, somebody grants or refuses, and the refusal can be reviewed. Marking introduces a third element that is not a party at all. It is a property of the file, made upstream by someone with no duty and read downstream by someone with guaranteed access and no idea where it came from. There is no interactive process with a watermark.

The school

Read the signal. Did not build it, cannot audit it, did not choose its scope, and has a rule of general application that says nothing about disability.

The writer

Injured by the combination, with a strong claim against neither.

FOUR REPAIRS, NO NEW LEGISLATION

Every one of these has a hook in a document that already exists

- 1 Mark by operation, not by product**

The Commission has already said the duty applies per use. A company can ship a mode where grammar correction, spellchecking, minor polish, and translation of the user's own text go unmarked. Summaries stay marked; that needs an amendment, and saying so is more honest than overclaiming.
- 2 Send the receipt both ways**

The signed, timestamped detection record already exists. Extending it to the account whose text was queried costs nothing that is not already built, and turns a hidden signal into an auditable one.
- 3 No finding without a route that avoids disclosure**

A respondent should be able to assert an accommodation without specifying it, and have the finding suspended pending review by somebody bound to confidentiality.
- 4 Offer an alternative that does not depend on the mark**

Where an institution's own policy creates the trap, providing another route to evaluation is a modification of policy, and existing law already governs that analysis.

CONCLUSION

The sentence exists. Nobody used it.

Somebody in the drafting understood that there is a difference between a machine that produces content and a machine that helps a person produce content, and wrote that difference into the law.

The line they drew reaches people who cannot speak and stops before people who cannot organise. Then the first companies to build it marked everything anyway, and published documentation conceding that the mark cannot separate the two cases the exception exists to separate. Then the compliance code conceded the signal is unreliable and guaranteed access to it to every university.

The people who pay already write with help. They already say so when they choose to. They now carry, in every paragraph they publish, a machine-readable statement about their own writing that they did not make and cannot unmake, readable on demand by parties named in a compliance document, with no notice to them, ever.

Travis Gilly, *The Assistive Exemption: Content Provenance Marking and the Compelled Disclosure of Disability* (working draft v0.8, August 2026)

Alexander v. Choate, 469 U.S. 287 (1985).

Americans with Disabilities Act of 1990, 42 U.S.C. §§ 12101–12213.

Andre Grenier v. Cyanamid Plastics, Inc., 70 F.3d 667 (1st Cir. 1995).

Anthropic. (2026). *Consumer terms of service*. <https://www.anthropic.com/legal/consumer-terms>.

Anthropic. (2026). *How Claude marks AI-generated content*. <https://support.claude.com/en/articles/16266773-how-claude-marks-ai-generated-content>.

Bobnar v. AstraZeneca Pharmaceuticals LP, 672 F. Supp. 3d 475 (N.D. Ohio 2023).

Carparts Distribution Center, Inc. v. Automotive Wholesaler's Association of New England, Inc., 37 F.3d 12 (1st Cir. 1994).

Cech, E. A. (2023). Engineering ableism: The exclusion and devaluation of engineering students and professionals with physical disabilities and chronic and mental illness. *J. Eng'g Educ.*, 112, 462. <https://doi.org/10.1002/jee.20522>.

Conroy v. New York State Department of Correctional Services, 333 F.3d 88 (2d Cir. 2003).

Cummings v. Premier Rehab Keller, P.L.L.C., 596 U.S. 212 (2022).

Dathathri, S., et al. (2024). Scalable watermarking for identifying large language model outputs. *Nature*, 634, 818.

Enyart v. National Conference of Bar Examiners, Inc., 630 F.3d 1153 (9th Cir. 2011).

European Commission. (2026). *Guidelines on transparency obligations for providers and deployers of certain AI systems*. <https://digital-strategy.ec.europa.eu/en/policies/guidelines-transparency-ai-generated-content>.

European Commission. (2026). *Transparency obligations under Article 50 of the AI Act*. <https://digital-strategy.ec.europa.eu/en/faqs/transparency-obligations-under-article-50-ai-act>.

Federal Trade Commission Act, 15 U.S.C. § 45.

Gaba, P., & De Cristofaro, E. (2026). *Group-differentiated discourse on generative AI in high school education*. arXiv:2603.24972.

Gegg-Harrison, W., & Quarterman, C. (2024). AI detection's high false positive rates and the psychological and material impacts on students. In *Academic integrity in the age of artificial intelligence* (p. 199). IGI Global. <https://doi.org/10.4018/979-8-3693-0240-8.ch011>.

Ibrahim, K. H. S., AlOtaibi, D., & Sibai, F. N. (2025). The robustness of AI-classifiers in the face of AI-assisted plagiarism. *Int'l J. Computer-Assisted Language Learning & Teaching*, 15, 1. <https://doi.org/10.4018/ijcallt.375428>.

Illinois Consumer Fraud and Deceptive Business Practices Act, 815 ILCS 505.

Kirchenbauer, J., et al. (2023). A watermark for large language models. In *Proceedings of the 40th International Conference on Machine Learning*.

Kroll v. White Lake Ambulance Authority, 691 F.3d 809 (6th Cir. 2012).

Liang, W., et al. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4, 100779. <https://doi.org/10.1016/j.patter.2023.100779>.

Nondiscrimination on the Basis of Disability by Public Accommodations, 28 C.F.R. § 36.309.

Payan v. Los Angeles Community College District, 11 F.4th 729 (9th Cir. 2021).

Payan v. Los Angeles Community College District, 169 F.4th 971 (9th Cir. 2026).

Regulation (EU) 2024/1689 (Artificial Intelligence Act), 2024 O.J. (L 1689).

Rehabilitation Act of 1973 § 504, 29 U.S.C. § 794.

Robinson v. Toyota Motor Credit Corp., 201 Ill. 2d 403 (Ill. 2002).

Robles v. Domino's Pizza, LLC, 913 F.3d 898 (9th Cir. 2019).

Sadasivan, V. S., et al. (2023). *Can AI-generated text be reliably detected?* arXiv:2303.11156.

Santuzzi, A. M., Waltz, P. R., Finkelstein, L. M., & Rupp, D. E. (2014). Invisible disabilities: Unique challenges for employees and organizations. *Indus. & Org. Psychol.*, 7, 204. <https://doi.org/10.1111/iops.12134>.

Shekhar, A. R., et al. (2026). *DoPE: Decoy oriented perturbation encapsulation*. arXiv:2601.12505.

Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltynsek, T., Guerrero-Dib, J., et al. (2023). Testing of detection tools for AI-generated text. *Int'l J. for Educ. Integrity*, 19, art. 26. <https://doi.org/10.1007/s40979-023-00146-z>.

Coley, M. (2023). *Guidance on AI detection and why we're disabling Turnitin's AI detector*. Vanderbilt University Brightspace. <https://www.vanderbilt.edu/brightspace/2023/08/16/guidance-on-ai-detection-and-why-were-disabling-turnitins-ai-detector/>.

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *Int'l J. for Educ. Integrity*, 19, art. 17. <https://doi.org/10.1007/s40979-023-00140-5>.

European Commission. (2026). *Code of practice on transparency of AI-generated content*. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-ai-generated-content>.

European Commission. (2026). *Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of the AI Act*. <https://digital-strategy.ec.europa.eu/en/library/guidelines-transparency-obligations-providers-and-deployers-ai-systems>.

Wechuli, Y. (2024). Passing as able-bodiedminded, disabled, or supercrip: Rethinking impression management strategies through a disability lens. *Symbolic Interaction*, 48, 293. <https://doi.org/10.1002/symb.711>.