The Bar Council

Law reform essay competition 2025: First Place

'Invisible Violence: Restoring the Sight of the Online Safety Act 2023'

1. Introduction — The Boot Without Oversight.

"If you want a picture of the future, imagine a boot stamping on a human face—forever"[1]

Orwell's image of power crushing the individual has found a new expression in the digital age. Today, the "boot" is algorithmic: opaque moderation systems that decide whose speech is amplified, whose pain is ignored, and who is quietly pushed out of public space. These systems mediate the modern public sphere, yet remain shielded from scrutiny.

As digital feminist scholar Emma A. Jane observes, "Men have turned on women online. The place that was supposed to be radically inclusive ... is now delivering female users a blunt message: GTFO."[2]

For women, girls, gender-diverse people, and those at intersecting margins, that message is lived reality. Gender-based abuse is sustained, targeted, and structurally entrenched. It thrives in the blind spots of automated moderation systems that escape oversight.

The Online Safety Act 2023 (OSA) was intended to make the UK "the safest place to be online."[3] Yet under section 9, regulated platforms may discharge their statutory duties almost entirely through opaque algorithmic systems, without

---

[1] Orwell, G. (1949) *Nineteen Eighty-Four*. Available at: https://www.clarkchargers.org/ourpages/auto/2015/3/10/50720556/1984.pdf (Accessed: 21 October 2025), p. 155.

[2] Jane, E.A. (2016) *Misogyny Online: A Short (and Brutish) History*. London: SAGE Publications. Available at: https://www.researchgate.net/publication/328252988_Misogyny_Online_A_Short_and_Brutish_History (Accessed: 22 October 2025), p. 1, para. 1.

[3] UK Government, 'UK children and adults to be safer online as world-leading Bill becomes law' (19 October 2023) https://www.gov.uk/government/news/uk-children-and-adults-to-be-safer-online-as-world-leading-bill-becomes-law (Accessed: 21 October 2025).

any legal requirement to disclose how those systems operate, how often they fail, or whose voices they fail to protect. Parliament has legislated for safety but left its enforcement to mechanisms the law cannot see.

This paper argues for the insertion of a new section 9A into the OSA: a targeted transparency and audit duty that would make algorithmic enforcement visible, measurable, and accountable. Particularly in relation to gender-based violence.

This would not expand the scope of harmful content or censor speech, but would ensure that the tools platforms use to comply with their existing obligations are subject to the same democratic oversight as the duties themselves.

This reform is desirable, practical, and useful. It advances the UK's human rights and equality obligations, aligns with emerging international standards such as the EU Digital Services Act[4], and provides Ofcom with the visibility necessary to regulate effectively. Above all, it returns controlled oversight to the law — where it belongs.

2. The Harm: Gender-Based Violence and Algorithmic Invisibility

Gender-based violence (GBV) online is not a marginal problem; it is a structural harm with democratic consequences. It silences voices, forces self-censorship, and drives women, girls, and marginalised communities out of public spaces.

Globally, 38% of women personally experienced online violence, and 85% witnessed it[5]. 41% fear for their physical safety after abuse, and more than half experience lasting psychological harm.[6] One in ten women experiences online sexual harassment by the age of fifteen.[7]

---

[4] European Union (2022) *Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act)*. Official Journal L 277/1. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065 (Accessed: 21 October 2025).

[5] UN Women (2020) *Online and ICT-facilitated violence against women and girls during COVID-19*. New York: UN Women. Available at: https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2020/Brief-Online-and-ICT-facilitated-violence-against-women-and-girls-during-COVID-19-en.pdf

[6] Amnesty International (2018) *Toxic Twitter: Triggers of violence and abuse against women on Twitter.* London: Amnesty International. Available at: https://www.amnesty.org/en/latest/news/2018/03/online-violence-against-women-chapter-2 (Accessed: 21 October 2025).

[7] European Union Agency for Fundamental Rights (2014) *Violence against women: an EU-wide survey — Main results report.* Luxembourg: Publications Office of the European Union. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2014-vaw-survey-main-results-apr14_en.pdf (Accessed: 22 October 2025).

This harm is intensified at the intersections of race, disability, sexuality, and religion, where abuse is layered, targeted, and often linguistically disguised. Women of colour, disabled women, trans women, and gender-diverse people are systematically overexposed to harassment yet under-protected by moderation systems.[8]

These failures are baked into the infrastructure of moderation. Most platforms rely on natural language processing (NLP) models trained on narrow, standardised datasets. They are good at spotting overt slurs in standard English. They are not good at recognising the fluid, coded, multilingual, and context-dependent abuse that targets women and marginalised groups.[9]

For example:

- **"S h e d e s e r v e d i t"** may bypass word filters entirely.
- Dialectal or patois insults are routinely unrecognised.
- Slurs embedded in memes, acronyms, or phonetic distortions **(e.g. "b a c k 2 t h e k i t c h e n")** are often invisible to classifiers.
- Feminist advocacy or survivor testimony is frequently misclassified as abuse, silencing those resisting harm.

Abusers know this and adapt quickly. Linguistic evasion strategies exploit structural weaknesses in moderation systems. The result is predictable: the abuse that most urgently needs to be moderated is precisely the abuse most likely to be missed.

This is not a glitch. It is a structural injustice. By outsourcing safety to black-box systems, platforms reproduce and deepen existing inequalities. Those already marginalised are rendered invisible not only socially, but also algorithmically. When Parliament delegates statutory safety duties without ensuring their effectiveness, the law itself becomes complicit in the harm.

---

[8] Marshall, B. (2021) *Algorithmic misogynoir in content moderation practice.* Heinrich-Böll-Stiftung European Union, June. Available at: https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-Algorithmic-Misogynoir-in-Content-Moderation-Practice-200621_FINAL.pdf (Accessed: 21 October 2025).

[9] Center for Democracy & Technology (2024) *Intersectional disparities within automated hate-speech detection across US-centered social-media content*. Available at: https://cdt.org/insights/intersectional-disparities-within-automated-hate-speech-detection-across-us-centered-social-media-content/ (Accessed: 23 October 2025).

The state cannot ignore this asymmetry. Where the state imposes duties on platforms to address harmful content, it must also ensure that those duties are effective, proportionate, and accountable. A system that consistently misses the most harmful abuse while over-policing vulnerable voices does neither.[10]

As Dorn *et al*. have shown, even the most advanced language models remain largely incapable of accurately identifying harm when the content is written by those within targeted communities, whereas they perform better when outgroup speakers use the same language. This asymmetry is foreseeable, measurable, and remediable but yet currently unregulated.

This undermines the very legitimacy of the regulatory framework. It also risks placing the United Kingdom in breach of its positive obligations under Articles 8 and 10 of the European Convention on Human Rights,[11] which require the state to protect individuals' rights to private life and freedom of expression.

3. The Legal Gap: Section 9 and the Black Box

The Online Safety Act 2023 was presented as a landmark in online regulation, introducing statutory duties of care on platforms to address illegal content. Section 9 requires regulated services to take proportionate steps to mitigate and manage the risk of illegal content. In practice, however, most platforms discharge this duty through automated moderation systems — machine learning models that flag or remove harmful material at scale.

Yet the Act does not require platforms to:

- disclose how these algorithmic systems operate,
- reveal their detection thresholds or false negative rates,
- explain their handling of coded, dialectal or intersectional abuse, or
- submit their moderation systems to independent audit.

[10] Dorn, R., Morstatter, F., Kezar, L. and Lerman, K. (2024) Harmful Speech Detection by Language Models Exhibits Gender-Queer Dialect Bias. arXiv:2406.00020v2 [cs.CL], 21 June. University of Southern California, ISI, Marina del Rey, California, USA. Available at: https://arxiv.org/abs/2406.00020 (Accessed: 21 October 2025).

[11] Council of Europe (1950) *European Convention for the Protection of Human Rights and Fundamental Freedoms* (ECHR), Articles 8 and 10. Available at: https://www.echr.coe.int/documents/convention_eng.pdf (Accessed: 21 October 2025).

Ofcom may issue information notices, but in the absence of a specific statutory transparency duty, platforms retain broad discretion over what, if anything, they disclose.[12] Victims of gender-based violence have no practical or legal route to challenge invisible failures.

Under Article 14, those protections must be provided without discrimination. A regulatory system that systematically fails to protect women and marginalised groups against coded abuse risks breaching those obligations.[13]

Domestic equality law points the same way. While platforms are not themselves public authorities, the Public Sector Equality Duty under section 149 Equality Act 2010 demonstrates Parliament's expectation that public functions are exercised with due regard to equality impacts.[14] If platforms are entrusted to deliver online safety functions in the public interest, it is incoherent for those functions to remain opaque and un-auditable.

Parliament has imposed duties to protect against online harms but allowed those duties to be mediated through systems no one can meaningfully scrutinise. This is an accountability vacuum — and in that vacuum, gender-based violence is rendered algorithmically invisible. The law cannot credibly promise protection while tolerating opacity at the point of enforcement.[15] That is the gap the proposed Section 9A is designed to close.

4. The Reform Proposal — Section 9A

4.1 The Proposed Draft Clause

Section 9A — Algorithmic Transparency and Gender-Based Violence

---

[12] Ofcom, Statement: *Online Safety Transparency Reporting* (21 July 2025) 1.3. Available at: https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-draft-transparency-reporting-guidance/main-docs/statement-online-safety-transparency-reporting.pdf (Accessed: 21 October 2025).

[13] Council of Europe (1950) *European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR),* Articles B and 14. Available at: https://www.echr.coe.int/documents/convention_eng.pdf (Accessed: 21 October 2025).

[14] Equality Act 2010, s 149. Available at: https://www.legislation.gov.uk/ukpga/2010/15/section/149 (Accessed: 21 October 2025)

[15] Edwards, L. (2022) *Regulating AI in Europe: Four Problems and Four Solutions, Ada Lovelace Institute.* Available at: https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe (Accessed: 21 October 2025).

(1) A regulated service which uses automated or algorithmic systems to discharge its duties under section 9 must—

(a) provide Ofcom with documentation sufficient to enable an assessment of the design, training data composition, testing methodology, and operational thresholds of those systems;

(b) publish annual transparency reports setting out—

(i) the volume of content flagged and actioned by automated systems,

( ) disaggregated false negative and false positive rates, including for coded, dialectal and intersectional forms of abuse, and

( ) steps taken to identify, address and mitigate any discriminatory effects, including failures to detect coded or context-dependent forms of gender-based abuse;

(c) permit Ofcom, or an accredited third party acting on its behalf, to conduct independent audits of such systems; and

(d) take proportionate steps to improve the performance of algorithmic systems in detecting and addressing such harms, and to remedy any deficiencies identified through reports or audits.

(2) In exercising functions under subsection (1), Ofcom must have particular regard to—

(a) the need to protect individuals from gender-based violence and related harms online; and

(b) the need to ensure that algorithmic systems do not directly or indirectly discriminate on grounds of sex, gender reassignment, race, disability, sexual orientation, or any other protected characteristic within the meaning of the Equality Act 2010.

(3) Ofcom may issue codes of practice or guidance for the purpose of compliance with this section.

(4) A failure to comply with this section constitutes a breach of the duties imposed by section 9 and is enforceable accordingly.

(5) The Secretary of State may, by regulations made by statutory instrument, amend this section to include additional categories of harm or further transparency obligations as necessary.

4.2 Rationale

Section 9 already recognises the centrality of algorithms in risk assessments, but leaves these assessments locked inside private systems. Section 9A closes this gap by making those same assessments transparent, auditable, and accountable to the regulator.

Targeted transparency.

The purpose of Section 9A is simple: to make the systems that enforce online safety visible. It does not impose a new duty to remove content. It does not expand the scope of illegal content. It simply obliges platforms to disclose how their moderation works and where it fails, particularly in relation to gender-based violence.[16]

Mandatory disclosure.

Subsection (1)(a) ensures that Ofcom can understand how automated systems are designed and deployed. At present, platforms reveal only minimal information voluntarily, often in vague or technical terms that Tribunals may not understand.[17] Requiring structured documentation allows regulators to interrogate how well these systems capture coded abuse.[18]

Performance reporting.

---

[16] Ofcom, *Online Safety Transparency Reporting: Statement* (2023) https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-draft-transparency-reporting-guidance/main-docs/statement-online-safety-transparency-reporting.pdf (Accessed 21 October 2025).

[17] Lilian Edwards, *Regulating AI in Europe: Four Problems and Four Solutions* (Ada Lovelace Institute, 2022) https://www.adalovelaceinstitute.org/wp-content/uploads/2022/03/Expert-opinion-Lilian-Edwards-Regulating-AI-in-Europe.pdf (Accessed 22 October 2025)

[18] Ofcom, *Online Safety Transparency Reporting: Statement* (2023) https://www.ofcom.org.uk/siteassets/resources/documents/consultations/category-1-10-weeks/consultation-draft-transparency-reporting-guidance/main-docs/statement-online-safety-transparency-reporting.pdf (Accessed 21 October 2025).

Subsection (1)(b) goes further by requiring publication of quantifiable metrics, including disaggregated error rates. This is critical because under-detection is often concentrated in dialectal and coded speech. Making those failures visible allows for evidence-based regulatory action.[19]

Independent audit.

Subsection (1)(c) gives Ofcom the power to require audits by accredited third parties. This mirrors the Digital Services Act's (DSA) Article 37 audit mechanism and draws on existing Ofcom enforcement powers.[20] Crucially, the audits are focused on discriminatory performance gaps, not business strategy.[21]

Regulatory teeth.

Subsection (1)(d) ensures transparency is not merely performative. Platforms must address any deficiencies identified. Because this is tied to section 9, Ofcom can enforce through existing penalties — including fines of up to 10 per cent of global turnover.[22]

Equality and non-discrimination.

Subsection (2) embeds equality considerations directly into the duty. This ensures Ofcom and platforms must explicitly consider whether algorithmic systems have discriminatory impacts on women, racialised groups, disabled users, or LGBTQ+ communities. This reflects the logic of Article 14 ECHR and the Public Sector Equality Duty.[23]

---

[19] Rebecca Dorn and others, 'H*armful Speech Detection by Language Models Exhibits Gender-Queer Dialect Bias*' (arXiv:2406.00020v2, 21 June 2024) 8 https://arxiv.org/pdf/2406.00020 (Accessed 22 October 2025).

[20] European Union (2022) *Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act)*. Official Journal L 277/1. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065 (Accessed: 21 October 2025).

[21] *Ibid*

[22] Online Safety Act 2023, sch 13 https://www.legislation.gov.uk/ukpga/2023/50/schedule/13 (Accessed 21 October 2025).

[23] Council of Europe (1950) *European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR),* Articles B and 14. Available at: https://www.echr.coe.int/documents/convention_eng.pdf (Accessed: 21 October 2025).
Equality Act 2010, s 149. Available at: https://www.legislation.gov.uk/ukpga/2010/15/section/149 (Accessed: 21 October 2025)

Guidance and proportionality.

Subsection (3) allows Ofcom to tailor guidance, recognising that not all services have the same scale or resources. This makes the reform practical and risk-based, focusing on the large platforms where harm is concentrated.[24]

Integration, not duplication.

Finally, subsection (4) integrates Section 9A with existing duties under section 9. This avoids legislative sprawl. Ofcom already enforces content-safety duties; this simply ensures those duties are auditable and accountable.[25]

4.3 Why This Reform Matters

The proposed reform is:

- **Desirable:** It addresses a legally and socially significant gap — the invisibility of coded gender-based violence online which has substantial public interest now that many of us have an online "life".
- **Practical:** It leverages existing Ofcom powers and follows the structure of the DSA's transparency and audit regime to maintain transparency.
- **Useful:** It creates measurable, reviewable standards for algorithmic performance, enabling effective oversight and enforcement.

Markedly, Section 9A does not tell platforms what speech to remove. It tells them to show their workings out, and to ensure that the legal promise of protection under the OSA is not lost inside a black box.

5. Freedom of Expression and Proportionality

No reform to platform governance can claim legitimacy if it silences the very voices it seeks to protect. Experience shows that when platforms face legal or regulatory pressure, they often over-enforce, removing lawful content to minimise risk. Automated systems are particularly prone to this: they flag

---

[24] Ofcom, *Online Nation 2024 Report* (2024) https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/online-research/online-nation/2024/online-nation-2024-report.pdf (Accessed 22 October 2025).

[25] Online Safety Act 2023, s 9 https://www.legislation.gov.uk/ukpga/2023/50/section/9 (Accessed 24 October 2025).

feminist speech, survivor testimony, satire, or political commentary as "harmful" because they cannot grasp context but can be trained to.

This over-enforcement is not a theoretical concern. It goes to the heart of Article 10 of the European Convention on Human Rights,[26] which protects not only comfortable expression but also speech that is satirical, critical, or unsettling. Excessive or poorly designed regulatory measures can produce a chilling effect, driving marginalised users out of public discourse.[27]

Section 9A is deliberately structured to avoid this. It does not mandate removal of content. It does not define new categories of illegality. It does not give Ofcom powers to censor. Instead, it requires transparency about how platforms already moderate and whether their systems are fair. It targets process, not expression.

The proposal also embeds safeguards. Platforms must have due regard to the protection of lawful expression, including survivor testimony and counter-speech. They must maintain accessible appeal mechanisms and human review for contested moderation decisions. These safeguards reflect Articles 17 and 20 of the EU Digital Services Act,[28] which aim to ensure procedural fairness and prevent the over-blocking of legitimate content.

In short, Section 9A is a proportionate, procedural reform. It strengthens accountability without handing regulators or platforms new censorship powers. Far from chilling speech, it creates the transparency and procedural safeguards necessary to protect it.

6. Comparative Models

[26] Council of Europe (1950) European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), Article 10. Available at: https://www.echr.coe.int/documents/convention_eng.pdf (Accessed: 21 October 2025).

[27] Victims' Commissioner for England and Wales and Domestic Abuse Commissioner for England and Wales (2025) *Joint response to the Office of Communications (Ofcom) consultation on draft Guidance: A safer online for women and girls*, 22 May. Available at: https://victimscommissioner.org.uk/document/ofcom-consultation-on-draft-guidance-a-safer-online-for-women-and-girls/ (Accessed: 21 October 2025).

[28] European Union (2022) *Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act)*. Official Journal L 277/1. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065 (Accessed: 21 October 2025).

Transparency in algorithmic decision-making is not an experimental idea. It is an emerging international legal norm, particularly in the regulation of large online platforms.[29]

6.1 The EU Digital Services Act[30]

The EU Digital Services Act (DSA) provides a clear comparative model.

- Article 34 requires very large online platforms to conduct systemic risk assessments, including risks related to gender-based violence and discrimination.
- **Article 37** mandates independent audits of their mitigation measures.
- **Article 40** provides for data access to regulators and vetted researchers to enable scrutiny of how moderation systems work.
- Article 42 requires detailed transparency reports.

Section 9A would mirror these obligations in a UK-specific way. By focusing on algorithmic performance and linguistic fairness, it translates the DSA's systemic transparency model into the UK's risk-based regulatory framework under Ofcom.

6.2 International Human Rights Standards

The UK has binding obligations under the Istanbul Convention[31] to take effective measures to protect women and girls from gender-based violence, including online. GREVIO, its monitoring body, has explicitly called for states to ensure that platforms adopt transparent and accountable moderation practices.

---

[29] Lepri, B., Oliver, N., Letouzé, E., Pentland, A. & Vinck, P. (2017) *Fair, Transparent and Accountable Algorithmic Decision-making Processes*: The Premise, the Proposed Solutions and the Open Challenges. Data-Pop Alliance White Paper. Available at: https://datapopalliance.org/wp-content/uploads/2020/09/Fair-Transparent-and-Accountable-Algorithmic-Decision-making-Processes.pdf (Accessed: 21 October 2025).

[30] European Union (2022) *Regulation (EU) 2022/2065 on a Single Market for Digital Services (Digital Services Act)*. Official Journal L 277/1. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065 (Accessed: 21 October 2025).

[31] Council of Europe Convention on Preventing and Combating Violence Against Women and Domestic Violence (known as the Istanbul Convention) (2011) Council of Europe Convention on Preventing and Combating Violence Against Women and Domestic Violence, opened for signature 11 May 2011, entered into force 1 August 2014. Available at: https://rm.coe.int/168008482e (Accessed: 21 October 2025).

The UN Special Rapporteur on Violence Against Women[32] has likewise urged governments to regulate online spaces in ways that are rights-based and survivor-centred, emphasising the need for transparency and accountability rather than censorship.

6.3 Domestic Compatibility

These international developments provide a coherent legal foundation for Section 9A. Far from creating a novel regulatory burden, it brings UK law into alignment with international best practice while respecting the UK's distinctive legal framework. It strengthens the state's compliance with Articles 8 and 14 ECHR[33] and demonstrates leadership in a field where the UK has historically lagged behind EU developments.[34]

7. Feasibility

Any credible law reform proposal must anticipate operational and legal objections. Section 9A is deliberately structured to work within the UK's existing regulatory architecture. It is both achievable and proportionate.

7.1 Enforcement

Ofcom already possesses robust enforcement powers under the Online Safety Act, including the ability to issue information notices, conduct investigations, and impose fines of up to 10% of global turnover for non-compliance. Section 9A would simply extend those powers to transparency failures. No new regulatory body or power structure would be required.

7.2 Trade Secrets and Security

[32] Office of the United Nations High Commissioner for Human Rights (2017) UN experts urge States and companies to address online gender-based abuse but warn against censorship, 8 March. Available at: https://www.ohchr.org/en/press-releases/2017/03/un-experts-urge-states-and-companies-address-online-gender-based-abuse-warn (Accessed: 21 October 2025).

[33] Council of Europe (1950) European Convention for the Protection of Human Rights and Fundamental Freedoms (ECHR), Article 8 & 14. Available at: https://www.echr.coe.int/documents/convention_eng.pdf (Accessed: 21 October 2025).

[34] The National (2024) *UK lagging behind Europe in AI regulation, SNP policy chief says*. Available at: https://www.thenational.scot/news/23589852.uk-lagging-behind-europe-ai-regulation-snp-policy-chief-says/ (Accessed: 21 October 2025).

Platforms may argue that mandatory disclosure would expose commercially sensitive information or enable adversarial actors to evade moderation.[35] Section 9A resolves this by requiring disclosure to Ofcom and accredited auditors, not full public disclosure of source code or proprietary data. Regulatory confidentiality and data protection provisions already govern Ofcom's handling of sensitive information. This mirrors the DSA audit regime, which has already been implemented across the EU without undermining platform integrity and security.

7.3 Proportionality and Scope

Concerns about compliance burden are met through Ofcom's code of practice powers.[36] The duty can be calibrated by platform size and risk, ensuring that the heaviest reporting requirements fall on the largest platforms, where harm is most concentrated. Smaller services would be subject to lighter obligations but in the same spirit. In terms of cost, a lot of the information would be information that the platforms would already have and utilise for both compliance and analysis for their own propriety algorithms such as those for marketing.

7.4 Effects

Some may suggest that transparency obligations could lead to the over-removal of content. But as set out in section 5 above, the duty targets process, not speech. It focuses on accountability, not censorship, with procedural safeguards to protect lawful expression.

Section 9A is neither intrusive nor unworkable. It is a measured, enforceable, regulatorily coherent reform that can be implemented with minimal legislative friction and significant public benefit.

8. Conclusion — Making the Boot Visible

Orwell's "boot on the human face" was never meant to describe technology. Yet in today's digital public sphere, it fits uncomfortably well. Algorithmic systems now shape the speech environment of billions. They decide whose pain

---

[35] MacCarthy, M. (2020) *Transparency Requirements for Digital Social Media Platforms: Recommendations for Policy Makers and Industry, Georgetown University*, 12 February. Available at: https://www.ivir.nl/publicaties/download/Transparency_MacCarthy_Feb_2020.pdf (Accessed: 21 October 2025).

[36] Home Office (2025) *Communications data code of practice (accessible)*, updated 6 June. Available at: https://www.gov.uk/government/publications/communications-data-code-of-practice/communications-data-code-of-practice-accessible--2 (Accessed: 21 October 2025).

is legible, whose abuse is actionable, and whose voices disappear beneath a flood of coded violence.

Law reform does not need to chase every evolving harm. It needs to make the systems that govern harm legible and accountable. That is what Section 9A achieves: it takes the invisible machinery of algorithmic exclusion and places it within the reach of law, scrutiny, and justice.

Orwell warned us about the boot. Law reform can ensure it no longer presses down unseen. This reform in particular will enable the State to be clear on what they expectfrom platforms, will empower Ofcom, re-invigorate public confidence and essentially contribute to making OSA both adaptable and future-focussed.