

Consultation response form

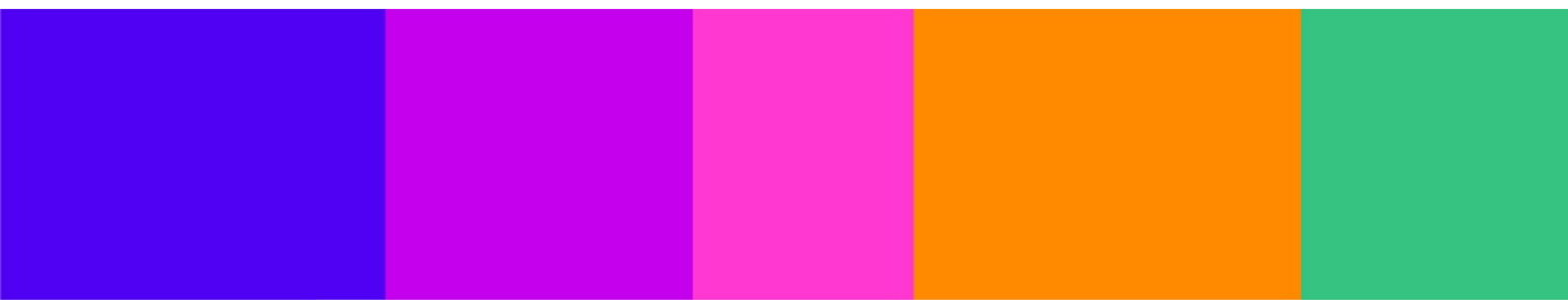
Please complete this form in full and return to cat1additionalduties@ofcom.org.uk.

Consultation title	Consultation: Giving users more choice, control and trust in their online experience
Full name	James Baker
Contact phone number	07817605162
Representing (delete as appropriate)	Organisation
Organisation name	Open Rights Group
Email address	James.baker@openrightsgroup.org

Confidentiality

We ask for your contact details along with your response so that we can engage with you on this consultation. For further information about how Ofcom handles your personal information and your corresponding rights, see [Ofcom's General Privacy Statement](#).

Your details: We will keep your contact number and email address confidential. Is there anything else you want to keep confidential? Delete as appropriate.	Nothing
Your response: Please indicate how much of your response you want to keep confidential. Delete as appropriate.	None
For confidential responses, can Ofcom publish a reference to the contents of your response?	Yes



Your response

Please tell us how you came across about this consultation.

- ☐ Email from Ofcom
- ☐ Saw it on social media
- ☐ Found it on Ofcom's website
- ☐ Found it on another website
- ☐ Heard about it on TV or radio
- ☐ Read about it in a newspaper or magazine
- ☐ Heard about it at an event
- ☐ Somebody told me or shared it with me
- ☐ Other (please specify)

Question	Your response
Question 1: Do you have any comments on our Equality Impact Assessment and Welsh Language Assessment at Annex 1?	Confidential? – N Ofcom acknowledges that some groups are less willing or able to verify and that these groups may disproportionately experience the consequences of being filtered out or receiving different treatment. We agree with this assessment, but think there are additional consequences that arise The impact assessment should consider in more detail the consequences of choosing not to verify. For examples platforms could implement proposals in a manner that includes lower visibility, reduced interaction, algorithmic downranking, perceived lower credibility (including in job recruitment processes) and when standing as candidates for elections. All of these could harm people's rights in a discriminatory fashion. Those consequences could be particularly significant for people who have legitimate reasons to participate pseudonymously or anonymously. The equality assessment should therefore examine whether apparently neutral product choices create a material disadvantage for such groups when they remain unverified. Furthermore platforms are increasingly having to comply with safety requirements from different legal jurisdictions. This increases the complexity of the equalities impact assessment because Ofcom's code of practice will start interacting with a wide

Question	Your response
	range of other legal duties that may increase and decrease the likelihood of particular design decisions not foreseen as likely by Ofcom that have higher discriminatory outcomes.
Question 2: Do you have any comments on our overarching approach to implementing Category 1 services' additional duties, including the guiding principles for our approach, and framework for the Additional Duties Code?	Confidential? – N We believe Ofcom's conception of user empowerment should extend beyond controls designed and operated by the provider of the underlying service. Genuine user empowerment could be facilitated through interoperability, allowing users to choose third-party tools which filter, label, rank or recommend content according to their own preferences. Our November 2025 report 'Making platforms accountable – Empowering users & creating safety' sets out the case for this. (https://www.openrightsgroup.org/app/uploads/2025/11/Making_platforms_accountable.pdf). One possibility is a modular model where vertical interoperability allows for independent moderation or recommendation service chosen by the user which operates between the underlying platform and the interface through which the user experiences it. Ofcom expressly recognises this modular model in its 2023 paper 'Mandated interoperability in digital markets' (https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/economic-discussion-papers-/discussion-paper-mandated-interoperability-in-digital-markets). At paragraph 5.20 it says greater vertical interoperability can allow competition at a more granular level over individual "services, components or functionalities", producing greater variety, innovation, flexibility for users and benefits from specialisation. It describes this as facilitating a "mix-and-match approach". A user could, for example, choose a filtering service designed for particular accessibility requirements, a family-oriented moderation service operated by a children's charity or a public service broadcasters, a civil-society-developed hate-

Question	Your response
	speech filter, or a recommender system designed not to optimise for engagement. Additionally Ofcom expressly connects interoperability with protection from harm in the same paper - Paragraph 6.24 says users might value: “privacy, security, resilience, protection from harm” This would transfer some control over online safety away from both the platform and the state and towards both the individual user and organisations with an interest in online safety. It could also create competition between different approaches to safety rather than requiring a single provider to determine the appropriate experience for every user. This would force platforms to compete economically on the safety and user-experiences rather than relying on their large network effects to retain their userbase. Our clear recommendation here is that Ofcom should ensure that its Code and guidance do not assume that user empowerment features must always be developed and controlled exclusively by the platform itself, and that vertical interoperability should be promoted as a way to improve safety outcomes. Ofcom should consider how providers could achieve supporting users with interoperable interfaces could give users more control to select alternative filtering, moderation and recommendation tools. This would fit incredibly well as Section 15 requires Category 1 providers, where proportionate, to give adults features they can choose to use to increase their control over relevant content.
Question 3: Do you have any comments on our combined impact assessment for the Additional Duties Code of Practice at Annex 6?	Confidential? – N The combined impact assessment should place

Question	Your response
	greater weight on the costs of erroneous classification and the cumulative effects of different measures on lawful expression. Relevant impacts include false-positive classifications, loss of reach, chilling effects, differential impacts on particular communities, additional collection of personal information, pressure to verify identity, and the interaction between user filters and recommender systems. Impact assessments should also compare proposed measures against credible less intrusive alternatives, rather than comparing intervention against doing nothing. For example, a warning, content-level control or user-selected recommender setting may achieve a significant proportion of the desired outcome without excluding content entirely. This is particularly important for identity verification. Ofcom's own supporting assessment concludes that there is currently "limited direct evidence" of the effectiveness of identity verification in addressing illegal harm and that the scale of any beneficial effects remains unclear. Where there is a weak evidence base for the effectiveness of an intervention, but privacy and expression impacts are foreseeable that should form part of the proportionality assessment.
Question 4: Do you agree with our proposed approach to explain the content categories the same way as the categories in the Guidance for Content Harmful to Children? Please provide supporting reasoning and evidence for your views.	Confidential? – N We partly agree with Ofcom's proposed alignment. We think that consistency can reduce compliance burdens, which can have a harmful impact on SME service providers. We also accept that the statutory descriptions of the categories substantially overlap. But we do have a concern that the regulatory purpose is different. Children's duties concern protec-

Question	Your response
	tion of children from content considered harmful to them, whereas the user empowerment duty concerns giving adults choices about lawful content. It would therefore be inappropriate if alignment resulted in a children's-protection model being carried across wholesale into the treatment of adult speech. In the adult context, context, purpose and user choice should have considerably greater weight. We welcome the draft Guidance's recognition that political, journalistic, artistic and educational content requires careful consideration and its examples excluding counter-speech, personal testimony, satire, reclaimed language, reporting of hate incidents and robust disagreement with political candidates from the hate and abuse category. These safeguards should be strengthened rather than treated as peripheral examples. International human-rights standards demonstrate why context matters. The Rabat Plan of Action (https://www.ohchr.org/en/documents/outcome-documents/rabat-plan-action) assesses potentially unlawful incitement by reference to factors including context, speaker, intent, content and form, extent and likelihood of harm. This illustrates the danger of reducing difficult speech classifications to isolated words or phrases without adequate context.
Question 5: Do you have any further comments on the draft Guidance on Relevant Content, published as part of this consultation (including on our assessment on the causes and impacts of relevant content, specificity of examples, and inclusion of contextual information)?	We are pleased that Ofcom recognises relevant-content determinations can be inherently subjective and that providers should regard context where classification is unclear. This does however create a procedural issue which the Guidance currently does not adequately address, that is 'What happens when a provider gets the classification wrong'. Even where content remains online, being classified as hate, abuse, self-harm or eating-disorder

Question	Your response
	content may result in warnings, or worst an invisible 'shadow ban' with heavily reduced distribution or exclusion from users' feeds. The person who created the content therefore has a legitimate interest in understanding and challenging a classification that materially affects its reach. We recommend that Ofcom encourage providers to: 1. notify users where classification materially restricts the distribution or accessibility of their content; 2. identify the category applied and provide an intelligible reason; 3. provide an accessible mechanism to challenge erroneous classification and obtain meaningful human review; 4. reverse downstream restrictions, including recommender or account-level signals, when a challenge succeeds; and 5. measure and publish meaningful information about false positives and successful appeals to measure for any discrimination within these processes. 6. evaluate whether classification systems perform differently across languages, dialects, reclaimed terminology and communities. This reflects recommendations Open Rights Group has previously made for stronger transparency, accuracy, proportionality and appeal safeguards in our May 2025 report 'How to fix the online safety act: a rights based approach' (https://www.openrightsgroup.org/app/uploads/2025/05/How to fix the Online Safety Act A Rights first approach.pdf) The need for such safeguards is not theoretical. ORG and other civil-society organisations working in the human rights space have previously raised concerns about political and journalistic material being incorrectly classified through automated

Question	Your response
	moderation, including the risks of confusing discussion, reporting or criticism relating to Palestine Action with unlawful support. That work highlighted the importance of monitoring, appeals and restoration or “put-backs”. (https://www.openrightsgroup.org/press-releases/palestine-action-ruling-human-rights-organisations-call-for-ofcom-to-issue-guidance-on-content-takedowns/)
Question 6: Do you have any evidence to support our draft Guidance on Relevant Content?	Confidential? – Y / N Ofcom's own recent research provides useful evidence for an approach based on nuanced user control rather than indiscriminate suppression. Research published on 17 September found a consistent desire for greater control while retaining choice. Some participants actively did not want broad filtering because they wished to remain informed, support people experiencing abuse or encounter different perspectives. Some preferred click-through warnings to whole categories being hidden. The research also found that unwanted material is frequently encountered through recommendations, personalised feeds and comments rather than deliberately sought out. It's important for us to the prevalence and amplification of harmful content should also be considered in the context of platforms' underlying business models. Academic research describes an advertising-funded model in which platforms have incentives to maximise users' attention and engagement, while recommender systems use data about users' past behaviour to predict which content will keep them engaged (Lasser and Poechhacker, 2025). The New York Academy of Sciences (https://nyaspubs.onlinelibrary.wiley.com/doi/full/10.1111/nyas.15359). Personal data is an important input into this model: behavioural advertising relies on monitoring users' online behaviour in or-

Question	Your response
	der to target advertising (Boerman, Kruijkemeier and Zuiderveen Borgesius, 2017) (https://www.tandfonline.com/doi/full/10.1080/0913367.2017.1339368). There is also empirical evidence of digital services shifting towards greater data extraction: Affeldt and Kesler found that apps retained following acquisition by major technology companies tended to become free while requesting more privacy-sensitive permissions. (https://reinholdkesler.github.io/publication/ieclap_21) Ofcom should therefore consider whether given users more control over how much of their personal data is obtainable by recommender systems and used for advertising purposes might also be a form of user control that can drive down harms.
Question 7: Do you agree with our proposed approach for the draft User Empowerment Assessment Guidance? In particular, do you agree with our proposal to align our guidance with the Illegal Content and Children's Risk Assessment Guidance? Please provide any supporting reasoning and evidence for your comments.	Confidential? – N We support aligning the structure of the User Empowerment Assessment Guidance with Ofcom's existing risk-assessment methodologies where this improves consistency. However, Ofcom should maintain a clear distinction between a risk assessment and a user empowerment assessment. The purpose of these duties is not to require a Category 1 service to minimise adults' exposure to all relevant content. It is to determine where user-controlled features should be available so adults can exercise their own choices. The assessment should therefore measure both the risk of unwanted exposure and users' legitimate reasons for wanting access to the same content. We welcome the use of evidence such as complaints and reports, but caution against encouraging additional behavioural or demographic profiling solely for the purpose of carrying out the as-

Question	Your response
	assessment. Ofcom notes that platforms may already collect or infer demographic characteristics from users' behaviour. Providers should not infer sensitive personal information merely because doing so would make an assessment administratively easier as this could drive up data-driven online harms.
Question 8: Do you have any further comments on the draft User Empowerment Assessment Guidance, published as part of this consultation?	Confidential? – N Ofcom should encourage providers, when considering how to empower users, to assess whether interoperability could give users greater control over their online experience. This could include standardised interfaces through which a user-selected third-party service could provide content labels, filtering rules, recommendation signals or moderation preferences. Ofcom could also discourage platform terms and conditions that prohibit people developing their own tools to interact with content and force users to use apps that aggressively harvest personal data for the purposes of targeted advertising and attention maximisation on platforms. Such an architecture would allow safety interventions to reflect different users' needs without requiring the underlying platform to make a single substantive judgment for its entire user base. A person who wishes to avoid particular categories of lawful content could select a service that filters them aggressively, while another user could choose warnings, different classifications or no additional filtering. Again this type of software isn't theoretical, it exists but would often break large platform's terms and services if someone were to distribute it. If this approach is of interest then we would be happy to demonstrate the safety potential for it with Ofcom. This approach would be particularly valuable both for creating safe 'white lists' of the type of content

Question	Your response
	an adult is happy with their child seeing, and also where judgments about harmful but lawful material are contextual or contested. Rather than requiring a dominant platform to become the arbiter of those judgments, interoperable moderation could permit trusted third parties, civil-society organisations or communities to develop their own judgements and filter systems that users were free to choose from.
Question 9: Do you agree with the proposed easy to access and defaults measure? Please provide supporting reasoning and evidence for your views.	Confidential? – N We support making user-control tools easy to find and easy to change. However, Ofcom should recognise (as is widely understood with data protection consent opt-ins) that defaults themselves exercise considerable behavioural power. Ofcom's research illustrates this clearly. In one experiment, 42% retained "Reduced sensitive content" when it was pre-selected, compared with 24% choosing it when neither option was selected. Among children, nearly seven in ten retained "Don't recommend harmful content" where it was defaulted, compared with around five in ten with no default. As such adult users should be required to actively choose between options rather than defaulting to any particular option. The choice should be informed with consequences explained. Users should also be prompted to review choices periodically.
Question 10: Do you agree with the inclusion of these four proposed content-specific measures: (1) content determination, (2) not encountering, (3) alerting, and (4) reporting? Please provide supporting reasoning and	Confidential? – Y / N We broadly support these functions, subject to stronger safeguards around content determination.

Question	Your response
evidence for your views.	“Not encountering” and alerting can be valuable precisely because the decision remains with the user. A click-through content warning as is common on the Fediverse and Mastodon for example, can prepare a user for potentially distressing material without removing that material from public discussion. The significant rights risk arises earlier: either a user or the provider must decide that material belongs in one of the statutory categories before these tools can work. A mistaken determination may consequently reduce distribution of lawful political, journalistic, educational or personal speech
Question 11: Do you agree that, for the purposes of these content-specific measures, providers should apply the measures to the following three categories of relevant content: (i) suicide and self-harm content (ii) eating disorder content, and (iii) hate and abuse content? Please provide supporting reasoning and evidence for your views.	Confidential? – Y / N We do not object to organising the Code around the three statutory groups identified by Ofcom: suicide and self-harm; eating disorders; and hate and abuse. However, we should like to highlight that if users had more freedom to use their own software to access interoperable platforms then users could pick categories of speech themselves. We do think caution is required around hate and abuse because political argument, documentation of discrimination, counter-speech, satire and reclaimed language may use the same vocabulary as the material being criticised. Ofcom's own draft Guidance correctly identifies examples where apparently hostile or offensive language should not be treated as hate or abuse, including counter-speech, testimony from survivors, artistic material and strong disagreement with a political candidate. It is already the case that content creators use ‘algo-speak’ phrases such as ‘grape for rape’ or ‘un-alive’ instead of ‘suicide’ or ‘kill’ to try and avoid algorithmic censorship

Question	Your response
	around these categories of speech. The Code should make clear that these are categories of lawful content over which adults are being given individual control, not categories that providers should seek to eliminate from the service.
Question 12: Do you agree with the flexibility we are offering providers over the technical actions in how they meet the not encountering and alerting measures? Please provide supporting reasoning and evidence for your views.	Confidential? – N We support technical flexibility and considers that Ofcom should expressly leave space for interoperable and user-selected approaches. “Not encountering” need not necessarily mean that the underlying platform itself suppresses content. A provider could enable a user to select a third-party filtering or moderation service and apply that service's signals when constructing the user's feed. Such services could even be vetted by the provider or Ofcom in a manner that ensured a degree of quality. Similarly, an independent service such as a browser plug-in could provide labels or contextual warnings which the underlying platform displays at the user's request if platforms were designed in a manner to allow the integration of such software tools. This would preserve the statutory objective while potentially reducing the need for platforms to make universal decisions about lawful expression.
Question 13: Do you agree with the inclusion of these two proposed measures: (1) blocking/muting, and (2) disabling comments? Please provide supporting reasoning and evidence for your views.	Confidential? – N We supports both measures. Blocking and muting are particularly strong examples of genuine user empowerment because they give individuals direct control over interac-

Question	Your response
	tions without requiring the platform to decide that another person's lawful speech should be removed for everyone. Likewise, the ability to limit or disable comments can provide users with control over their own space without restricting the other user's ability to speak elsewhere. Ofcom should ensure that use of these functions does not itself become an undisclosed moderation signal used to penalise or downrank the blocked user across the wider service. A user's individual decision to avoid another user should not automatically become a platform-wide enforcement judgment. This is an important defence against brigading.
Question 14: Do you agree with our approach to the proposed filter tool measure? Please provide supporting reasoning and evidence for your views.	Confidential? – N We have significant concerns about how this statutory duty is implemented. The filter must remain genuinely user initiated with the consequences of the choice being clearly explained to users. It should not evolve into a platform-wide distinction between a first class of verified users and a second class of pseudonymous users. Ofcom's own proposals illustrate the potential reach implications. Its consultation gives an example in which content generated by non-verified users could be downranked within recommender feeds for users who have switched on the filter. Ofcom also acknowledges that blanket application across some services could be disproportionate. Open Rights Group recommends that the Code distinguish between the unavoidable consequence of an individual's decision to filter non-verified users and platform-created disadvantages imposed on non-verified users irrespective of indi-

Question	Your response
	vidual choice.
Question 15: Do you agree with the proposed objectives that providers should meet in determining which functionalities to apply the filter tool to? Please provide supporting reasoning and evidence for your views.	Confidential? – N We agree that providers need flexibility to avoid applying the filter in ways that fundamentally undermine the service. However, the objectives should expressly include protection of the rights and interests of people who choose not to verify. We recommend adding an objective along the following lines: Providers should seek to avoid materially degrading the ability of non-verified users to receive and impart information or participate in the ordinary functions of the service, except to the extent that this is the direct and proportionate consequence of another user's decision to apply the filter. Ofcom already recognises this issue in its draft identity guidance, stating in 3.48 that people with valid reasons not to verify should not receive an “unduly degraded experience”. https://www.ofcom.org.uk/siteassets/resources/documents/consultations/giving-users-more-choice-control-and-trust-in-their-online-experience/documents/draft-user-identity-verification-guidance.pdf The same principle should appear directly in the filter-tool framework. Providers should also monitor whether the implementation of filtering produces substantially different reach or participation outcomes for groups who are less likely or able to verify.
Question 16: Do you agree with our proposed approach to the draft User Identity Verification Guidance? Please provide supporting reasoning and evidence for your	Confidential? – N a) Ofcom should distinguish more clearly between verifying a person's real-world identity, verifying a particular attribute, establishing control of an ac-

Question	Your response
views. In particular: a) Do you agree with our definitions relating to identity verification? Please provide supporting reasoning and evidence for your views. b) Do you agree that the identity verification offered by providers should be consistent with the proposed four principles? Please provide supporting reasoning and evidence for your views. Question 17: Do you have any further comments on the draft User Identity Verification Guidance?	count, and establishing that an account is operated by a person rather than an automated system. A service does not necessarily need to know who someone is in order to establish the information needed for a particular purpose. Ofcom should therefore draw on the Government's Identity Assurance Principles, originally developed by the Privacy and Consumer Advisory Group (PCAG https://www.gov.uk/government/groups/privacy-and-consumer-advisory-group). These include the Multiplicity Principle: "I can use and choose as many different identifiers as I want to" — and the Data Minimisation Principle: "My interactions only use the minimum data necessary to meet my needs". The current guidance further states that identity assurance should only be used where a need has been established and only to the minimum appropriate level of assurance. Ofcom should therefore distinguish identity assurance from identity disclosure: users should be able to prove what is necessary about themselves without necessarily having to prove who they are. b) Relevance, reliability, inclusivity and clarity are useful principles. However, without an explicit principle covering privacy, data minimisation and security this is incomplete. Relevance partially addresses unnecessary attributes, but due to the way that personal data drives harms online privacy is a distinct issue that requires its own principle. Providers should minimise information collected and retained, avoid disclosure of unnecessary identity information to other users, prevent secondary use for advertising or unrelated profiling, and provide privacy-preserving verification routes wherever possible. Reliability should also encompass effective correction and appeal mechanisms where verification fails or produces an erroneous result. Finally, verified status must not be presented as

Question	Your response
	meaning that a user is truthful, safe or authoritative. Ofcom's own research says that verification does not guarantee that content is genuine or true. Its experimental research also found that verification had little effect for general content but could increase willingness to follow verified accounts in the case of political content. 17: We welcome Ofcom's recognition of the importance of anonymity and pseudonymity. The draft notes explicitly that people discussing hidden disabilities or experiencing domestic abuse may wish to participate without disclosing personal information. Providers should also be discouraged from retaining identity evidence once verification has been completed where retention is unnecessary, and from reusing verification data for unrelated commercial purposes. There are sadly already examples such as Reddit using verified age data data being used to target people with advertising by platforms (https://support.reddithelp.com/hc/en-us/articles/36429514849428-Why-is-Reddit-asking-for-my-age)
Question 18: Do you agree with our proposed approach to the draft News Publisher Guidance? Please provide your reasoning and any supporting evidence. In particular: a) Do you agree with the proposed technical approaches for identifying recognised news publishers and news publisher content? Are you aware of alternative methods that have not been explored in the draft guidance? b) Are you aware of any technical means of identifying news publisher content when it is uploaded to or shared on a service by users who are not themselves recognised news publishers? c) Are there any additional factors you consider providers should take into account	Confidential? – Y / N

Question	Your response
when setting or applying relevant time-frames?	
Question 19: Do you have any further comments on the draft News Publisher Content Guidance, published as part of this consultation?	Confidential? – N Our concern is that enhanced protection for recognised news publishers should not imply a lower standard of protection for other forms of journalism or public-interest speech. We have previously recommended in ‘How to fix the online safety act: a rights based approach’ (https://www.openrightsgroup.org/publications/how-to-fix-the-online-safety-act-a-rights-first-approach/) that protections should attach as far as possible to the nature of the speech rather than solely to the institutional identity of the speaker. Citizen journalists, freelance reporters, campaign groups, candidates, civil-society organisations and ordinary users may all disseminate journalism or evidence of matters of public interest. Having a pathway for wider journalistic-content, democratic-content, freedom-of-expression and complaints therefore remain essential even where content falls outside the recognised-news-publisher regime.
Question 20: Do you agree the inclusion of these seven proposed measures: (1) internal content policies (2) user sanctions policies (3) training for individuals working in content moderation (4) materials for volunteers in content moderation (5) compliance training for operational staff (6) dedicated and expedited complaints for journalistic content (7) review of upheld complaints for content of democratic importance? Please provide supporting reasoning and evidence for your views.	Confidential? – N Internal policies, training and guidance for moderators are useful steps towards reducing both under- and over-moderation. We support a dedicated and expedited complaints process for journalistic content and the requirement for swift reinstatement or reversal of sanctions where complaints are upheld. However, restoration should be understood broadly. Where content has been removed or downranked incorrectly, simply making it technically accessible again may not restore the harm caused. Where technically feasible, successful appeals should reverse associated account sanctions,

Question	Your response
	distribution restrictions and negative recommendation signals. Ofcom's Measure B7 concentrates on reviewing upheld complaints. But systematic failure to recognise democratic content may result precisely in complaints not being upheld. Providers should therefore also analyse samples of rejected complaints and cases where users asserted that material was democratic or journalistic content.
Question 21: Do you agree with our approach to the draft Indicative Factors Guidance? Please provide supporting reasoning and evidence for your views.	Confidential? – N We broadly support Ofcom's approach, particularly its statement that any type of user can generate, upload or share journalistic content or content of democratic importance, and its warning that excessive reliance on account identity may lead to content being wrongly excluded or misclassified. However, we recommend expressly expanding the examples of accounts likely to generate content of democratic importance. At present, Ofcom lists elected officials, candidates for public office, campaign or issue groups and political activists. We recommend adding: civil-society organisations, charities, companies limited by guarantee with political or social objects of incorporation, collectives, co-operatives, NGOs, trade unions, community organisations and public-interest advocacy organisations. Such organisations routinely contribute to democratic political debate through policy analysis, consultation responses, parliamentary evidence, litigation, research, campaigning and scrutiny of public institutions. This should not become a whitelist or privileged category. Account identity should remain only a supporting signal. An ordinary individual can also

Question	Your response
	generate content of democratic importance just as readily as an established organisation. Whistleblowers would be one example of this. Explicitly including civil society would nevertheless help prevent automated or institutional systems from interpreting “democratic importance” too narrowly as meaning primarily politicians, parties and formally registered political campaigners.
Question 22: Do you agree with our approach to the draft Freedom of Expression and Privacy Impact Assessment Guidance? Do you have any suggestions as to what further information might usefully be included in the guidance? Please provide supporting reasoning and evidence for your views.	Confidential? – N This is one of the areas where the proposals in the consultation can be significantly strengthened. We welcome Ofcom's proposed structured and evidence-based methodology and particularly its identification of reduced visibility, chilling effects, differential treatment, loss of access to information and privacy harms as relevant impacts. We also welcome the recommendation that providers draw on complaints, appeals, moderation data and external expertise including civil society. However, necessity and proportionality should form a central part of the assessment, rather than being presented as something providers “may wish” to conduct “where relevant”. Ofcom's draft currently says providers may conduct necessity and proportionality assessments, including considering whether less intrusive alternatives exist. We recommend stronger language requiring providers, wherever an identified measure has a material adverse impact on expression or privacy, to consider: whether the objective is sufficiently important; whether the measure is rationally connected to that objective; whether a less rights-intrusive intervention could achieve the objective without unacceptably compromising it; and whether the remaining interference strikes a fair balance. This tracks the established four-stage proportionality framework reaffirmed by the Supreme Court

Question	Your response
	in Shvidler (App) v Secretary of State for Foreign, Commonwealth and Development Affairs (Res); Dalston Projects Ltd and others (App) v Secretary of State for Transport (Res) https://www.supremecourt.uk/cases/judgments/uksc-2024-0045) This is particularly appropriate because Ofcom itself is subject to section 6 of the Human Rights Act when exercising its public functions. Section 6 provides that it is unlawful for a public authority to act incompatibly with Convention rights, subject to the statutory exceptions. Index on Censorship's commissioned legal analysis of Ofcom's earlier On-line Safety Act guidance has similarly emphasised the relevance of Ofcom's own Convention obligations when exercising its regulatory functions. The assessment should also explicitly consider cumulative effects. Ofcom usefully recommends periodic reviews capable of identifying incremental changes that become cumulatively significant. The same principle should be applied across different measures. Identity verification, filtering, warnings, moderation, recommender interventions and account sanctions may each appear proportionate in isolation while collectively having a substantial impact on a user's ability to participate. Providers should additionally be expected to publish enough substantive information to make these assessments capable of external scrutiny. This should include significant evidence relied upon, material uncertainties, alternatives considered, relevant error and appeal data, and the steps taken in response. The impact assessment process should not become a compliance document which records that freedom of expression was "considered"; it should demonstrate how rights considerations changed design or implementation.

Question	Your response
Question 23: Do you agree with the inclusion of these two proposed measures on: (1) substance and (2) clear and accessible terms of service. Please provide supporting reasoning and evidence for your views.	Confidential? – N Terms of service should enable a reasonably informed user to understand both what the rules are and what can happen to them. This should include not merely categories of prohibited or restricted content, but the consequences of moderation: removal, visibility restrictions, recommender downranking, warnings, account sanctions or other limitations. Where automated systems enforce these rules, platforms should test whether automated outcomes actually correspond to the published terms. We welcome Ofcom's draft Terms of Service Guidance in this respect: it recommends regular processes aimed at ensuring consistency in human moderation and states that automated content-moderation systems should be designed to reach outcomes consistent with the terms. https://www.ofcom.org.uk/siteassets/resources/documents/consultations/giving-users-more-choice-control-and-trust-in-their-online-experience/documents/draft-terms-of-service-guidance.pdf Clarity is particularly important where broad or subjective terminology is used. A platform should not be able to comply merely by saying that it may restrict anything it regards as “harmful” or “inappropriate” without providing users with meaningful information about how that discretion will operate. This also relates to the Article 10 requirement of foreseeability. As Graham Smith https://www.cyberleagle.com/2025/08/ofcoms-proactive-technology-measures.html) has highlighted in analysing Ofcom's wider regime, principles-based regulatory discretion still needs to be sufficiently predictable where its exercise affects lawful expression.

Question	Your response
Question 24: Do you agree with the inclusion of these three proposed measures on: (1) enabling complaints (2) easy to find, access and use complaints procedures (3) appropriate action for complaints. Please provide supporting reasoning and evidence for your views.	Confidential? – N We strongly support accessible complaints mechanisms but considers the proposed conception of “appropriate action” insufficient. In Volume terms of service and complaints duties for category 1 services https://www.ofcom.org.uk/siteassets/resources/documents/consultations/giving-users-more-choice-control-and-trust-in-their-online-experience/documents/volume-4-terms-of-service-complaints.pdf) Ofcom proposes that appropriate action can consist of directing a complaint to the relevant individual or team for processing. Routing a complaint correctly is an administrative prerequisite; it is not, by itself, meaningful redress. For complaints affecting expression, appropriate action should ordinarily involve an intelligible decision, including the relevant rule or classification, an opportunity for meaningful review, and an explanation of the outcome. Where an initial decision is found to have been wrong, providers should take reasonable steps to correct its effects. Depending on the case, this could include restoring content, removing account sanctions, correcting content classification and reversing associated restrictions on distribution or recommendation. This is important because timing matters online. A political, journalistic or public-interest post reinstated days after the discussion to which it related may have permanently lost its audience. Our previous work on the Online Safety Act has repeatedly identified meaningful appeal and correction as important safeguards against erroneous moderation. https://www.openrightsgroup.org/publications/o

Question	Your response
	pen-letter-implementation-of-palestine-action-proscription-on-social-media/ We think complaints procedures ultimately require access to independent redress. This is common in almost all industries and walks of life. Internal platform appeals leave the final decision with the same provider that made the original moderation decision. This is particularly problematic where decisions affect lawful expression, access to information or participation in democratic debate. Parliament anticipated this problem in sections 160 and 217 of the Online Safety Act. Section 160 requires Ofcom, when reporting on the operation of reporting and complaints procedures, to advise the Secretary of State on whether regulations should be made imposing an alternative dispute resolution duty on Category 1 services. Section 217 enables such regulations to require Category 1 providers to arrange and participate in an impartial ADR procedure for relevant complaints, potentially at the provider's expense. ORG believes Ofcom should recommend use of the section 217 power where evidence demonstrates that internal complaints processes do not provide effective redress. Independent adjudication should complement rather than replace access to the courts. Section 217 already requires that any ADR scheme must not restrict or exclude the availability of civil proceedings. The scheme should therefore be designed so that users do not waive legal rights by using ADR, and an ADR determination should not prevent a user from pursuing an available civil claim.

Question	Your response
Question 25: Do you agree with our approach to the draft Terms of Service Guidance? Please provide supporting reasoning and evidence for your views.	Confidential? – N We support most of the draft Guidance and particularly the emphasis on consistency between published terms and actual moderation practice. We welcome Ofcom's recognition in volume 4 terms of service and complaints duties for category 1 services https://www.ofcom.org.uk/siteassets/resources/documents/consultations/giving-users-more-choice-control-and-trust-in-their-online-experience/documents/volume-4-terms-of-service-complaints.pdf?v=421924) that automated moderation can struggle with context and nuance and that consistency requires attention both to human decisions and automated systems. The final Guidance should strengthen the expectation that providers test not only whether moderation systems technically implement their rules, but whether they produce unacceptable error rates in particular contexts or for particular communities. Providers should also be encouraged to explain clearly: how automated systems contribute to moderation decisions; the principal consequences of violating particular rules; how users can appeal; and what happens when an appeal succeeds. Material changes to terms should be clearly notified to users rather than incorporated silently into lengthy terms documents.
Question 26: Do you have any further comments on the draft Terms of Service Guidance, published as part of this consultation?	Confidential? – N We recommend that Ofcom encourage providers to maintain publicly accessible version histories or change logs for material moderation provisions. Users, civil society and researchers should be able to establish which rules applied at a particular

Question	Your response
	point in time and how they have subsequently changed. Consistency should also be understood as an outcome question, not merely a documentation question. A service can have carefully drafted terms while applying them inconsistently through automated tools, contractors, different languages or different regions. Terms of Service Guidance, complaints regime and freedom-of-expression impact assessments should operate together. Clear terms tell users what should happen; impact assessments identify whether those policies create disproportionate rights impacts; complaints and appeals establish whether the system worked correctly in individual cases. None of those mechanisms is sufficient by itself.

Please complete this form in full and return to Cat1additionalduties@ofcom.org.uk.