Fraudulent Advertising Codes Consultation

Volume 2: Risk, governance and control

Consultation

Published 10 July 2026
Closing date for responses: 02 October 2026

For more information on this publication, please visit ofcom.org.uk

Contents

Section

1. Risk, governance and control (Volume 2) – Introduction

1.1 In this volume, we explain our proposals on how providers can more effectively identify material risks on their services and seek to mitigate them effectively and proportionately. In particular, our governance proposals aim to ensure there is senior oversight and accountability for the implementation and operation of the measures set out in the Codes.

1.2 This volume is structured as follows:

- **Section 2, 'Advertising intermediaries'** explains our proposals for how providers who may be unable to apply one or more of the proposed Fraudulent Advertising Codes measures (due to degree of control) can comply with their fraudulent advertising duties.
- **Section 3, 'Fraud indicator assessment'** explains our proposals for how providers should assess relevant evidence to understand how fraudulent advertising manifests on their service and what indicators suggest that individual advertisements and accounts pose a high risk of fraud. These insights can then be used to apply mitigations to tackle fraudulent advertising more effectively and proportionately.
- **Section 4, 'Governance and accountability'** explains our proposals for ensuring providers have appropriate governance arrangements in place, including senior oversight and accountability, for compliance with the fraudulent advertising duties.
- **Section 5, 'Testing advertisement generation tools'** explains our proposals for how providers should test their advertisement generation tools to identify vulnerabilities that fraudsters could look to exploit and how they should take steps to address identified vulnerabilities.

2. Advertising intermediaries

What is this section about?

Where providers use advertising intermediaries to place paid-for advertisements on their service, the degree of control providers have over decisions around the placement of advertisements and the application of safety measures may vary. In some circumstances, providers that use advertising intermediaries in an open-display supply chain may have insufficient control to apply a proposed measure.

This section explains our proposed measure for how providers can apply safety measures and comply with their fraudulent advertising duties in circumstances where advertising intermediaries play a role in the placement of paid-for advertisements.

Our proposal

Number in our Codes	Proposed measure *Applicable to providers of Category 1 and 2A services that have insufficient control in relation to the placement of paid-for advertisements on their service to apply a measure in the draft Fraudulent Advertising Codes*
FAU K1 and FAS K1	Where a provider cannot implement a measure recommended in the draft Fraudulent Advertising Codes, in whole or in part, due to the degree of control it has in relation to the placement of paid-for advertisements on the categorised service, that provider should **use all reasonable endeavours to implement a version of the measure(s)**, or the relevant part of it, that is as similar to the measure as possible.

Why are we proposing this?

Our expectations for safety mitigations against fraudulent advertising for providers of Category 1 and 2A services are high. The pathway a provider uses to place paid-for advertisements on its service(s) may affect the degree of control the provider has over the placement of paid-for advertisements and the application of safety measures. We are aware that providers might use multiple pathways to place paid-for advertisements on a single service. However, the pathway(s) that a provider uses to place paid-for advertisements on a service should not result in different safety outcomes; providers need to take steps to protect users from fraudulent advertisements.

Consultation questions

- Do you agree with our proposal? Please provide any arguments and supporting evidence.
- We have included two illustrative examples of when and how this measure may be applied, and what 'all reasonable endeavours' might look like. Do you have any comments on these examples, and do you have any additional examples of how providers could work with or require intermediaries to implement safety measures in circumstances where they have insufficient control? Please provide any arguments and supporting evidence.

Introduction

2.1 Paid-for advertisements can be placed on categorised services[1] through different pathways.[2] These pathways are dynamic, complex and intricate, and will often affect the degree of control a service provider has over the application of safety measures on a service.[3] For more information on how we understand advertising pathways, see Volume 1, Section 3, 'Online advertising ecosystem'.

2.2 Industry evidence indicates that the complexity of the open-display market can lead to a lack of standardisation and transparency.[4] Existing initiatives aim to address this by encouraging best practice across the supply chain. The measure we are proposing to recommend draws on these initiatives.

2.3 Our initial analysis suggests that most providers of Category 1 and Category 2A services use owned-and-operated supply chains (sometimes known as 'walled gardens') to place paid-for advertisements on their services.[5] Some providers also use advertising intermediaries in the open-display market to place paid-for advertisements on their services.[6] Decisions around the placement of advertisements on categorised services that use the open-display market are likely to be shared among several advertising intermediaries.

2.4 Providers can use a combination of pathways to place paid-for advertisements on a single service: we refer to this as a 'hybrid arrangement'.[7] Hybrid arrangements might mean that the degree of control a provider has over the placement of paid-for advertisements and the application of safety measures could differ on a single service. However, the pathway(s) that a provider uses to place paid-for advertisements on its service – and the degree of control a provider has over the placement of advertisements and application of safety measures – should not result in different safety outcomes for users.

2.5 Generally, we refer to providers having either 'sufficient control' or 'insufficient control' to apply the measures we are proposing to recommend as part of our draft Fraudulent Advertising Codes of Practice. We have not defined 'sufficient control' or 'insufficient

---

[1] For more information on categorisation, see Volume 1, Section 2, 'Introduction', paragraphs 2.14 to 2.17.
[2] We use the term 'pathways' to describe how providers can place an advertisement on a service. These pathways are generally categorised by the type of supply chain(s) that a provider uses. See Volume 1, Section 3, 'Online advertising ecosystem', paragraphs 3.16 to 3.25.
[3] For an explanation of how we have approached 'degree of control', see Volume 1, Section 5, 'Approach to Codes', paragraphs 5.43 to 5.49.
[4] Beruku and Which?, 2022. Fraud in the open-display advertising market, p.6. [accessed 22 June 2026]; Incorporated Society of British Advertisers (ISBA), 2020. Programmatic Supply Chain Transparency Study [accessed 22 June 2026].
[5] We sent formal and informal information requests to many of the most widely used user-to-user and search services to gather information about their advertising systems. The large majority of those services used some form of owned-and-operated supply chains to place paid-for advertisements on their service(s). Some used advertising intermediaries as well, but these tended to be found in hybrid arrangements. Source: [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 30 January 2026; [✂] response to our informal information request issued 30 July 2025; [✂] response to our informal information request issued 16 September 2025; [✂] response to our formal information request issued 30 January 2026; and [✂] response to our formal information request issued 30 January 2026.
[6] An 'intermediary' is an actor (usually in an open-display supply chain) that is involved in the automatic buying, selling and serving of online advertisements. A third-party advertising intermediary refers to an intermediary that is separate to the service provider.
[7] See Volume 1, Section 3, 'Online advertising ecosystem', paragraph 3.25.

control' because we know that this will likely be context-dependent and is not static. In this section (see Example 2.1 and Example 2.2), we provide some illustrative examples of 'all reasonable endeavours' that providers might use to implement a version of a proposed measure (or the relevant part of that measure) that is as similar to the measure as possible: in those examples, we explain why or where a provider might have 'insufficient' control.

Advertising intermediaries

Explanation of the proposed measure

2.6 Where service providers conclude that they are unable to apply a measure recommended in the draft Fraudulent Advertising Codes, in whole or in part, due to the degree of control they have in relation to the placement of paid-for advertisements on their service, they should:

i) use all reasonable endeavours to implement a version of the measure(s), or the relevant part of it, that is as similar to the measure as possible; and
ii) keep a record of which measure(s) they have been unable to implement, why they have been unable to implement the measure(s), and what steps they have taken instead to implement a version of that measure(s) (or the relevant part of it) that is as similar to the measure as possible. This should also include detail of governance processes that have been followed.

2.7 We understand that there is variation in the way providers place paid-for advertisements on their services: a provider of two categorised services could set up pathways for placing paid-for advertisements differently on each of those services; it might also use a hybrid arrangement on one categorised service.[8] Variations can also occur around arrangements with advertising intermediaries, which can affect the degree of control.

2.8 To account for these variations, this proposed measure only applies in relation to paid-for advertisements that are placed on services through pathways where a provider has insufficient control over the placement of paid-for advertisements and application of safety measures. A provider should not apply this measure to paid-for advertisements that have been placed via pathways that are integrated with the service.[9] We expect this proposed measure to apply only in limited circumstances because we understand that there is a greater prevalence of owned-and-operated supply chains among categorised services (see paragraph 2.3).

2.9 Our proposed wording for other measures in the draft Codes reflect this. Where we have determined that degree of control is a relevant factor in the application of a recommended measure, we have drafted the 'application' section of that measure to acknowledge that.[10] We set out in relevant proposed measures that the intermediaries measure applies where the provider cannot implement the proposed measure in question (either in full or in part) due to the degree of control the provider has over the placement of paid-for advertisements on its service. An assessment of the application of the intermediaries measure should be

[8] For example, [✂], and Meta is the provider of Facebook and Instagram. [✂] told us that [✂] Source: [✂] response to our formal information request issued 26 June 2025, [✂]; [✂] explained that [✂] Source: [✂] response to our formal information request issued 26 June 2025, [✂].
[9] For more detail on how we consider 'degree of control' in relation to advertising pathways, see Volume 1, Section 5, 'Approach to Codes', paragraphs 5.43 to 5.49.
[10] In our draft Codes, we set out that the intermediaries measure would not apply to: the proposed Terms of service/Publicly available statements measures, the Governance measures, and the Testing measure.

made on a measure-by-measure basis (with regard to ensuring a coherent approach, as described in paragraphs 2.18 to 2.19) and in relation to all advertisements from a given pathway. Applying the intermediaries measure in respect of one measure (either in full or in part) does not mean that other measures will not also apply – they do apply where providers have sufficient control over the placement of paid-for advertisements.[11]

2.10 We consider that this proposed measure acts as an 'enabling' measure. It would enable a provider to implement a version of the proposed measures in the draft Fraudulent Advertising Codes in a way that is suited to the degree of control the provider has over the placement of paid-for advertisements and application of safety measures on its service. It also enables a provider to ensure that any arrangements it has with advertising intermediaries contribute to safety mitigations more broadly and the provider's ability to meet the fraudulent advertising duties. As such, we broadly consider this proposed measure to be a core part of a provider's governance activities. The way a provider works with advertising intermediaries should complement other governance activities, be supported by appropriate governance structures, and be subject to appropriate principles and guidelines.

2.11 We have not been prescriptive about how providers should apply this proposed measure. We consider that this offers sufficient flexibility to ensure that the proposed measure works for the range of advertising arrangements that providers may have.

Using 'all reasonable endeavours' to implement a version of the measure(s) not applied

2.12 There may be some circumstances where a provider is unable to apply a measure recommended in the draft Fraudulent Advertising Codes (in whole or in part) due to the degree of control it has over the placement of paid-for advertisements. In such circumstances, the provider should use all reasonable endeavours to implement a version of the measure(s), or the relevant part of the measure(s), that is as similar to the measure as possible. 'All reasonable endeavours' should include, where applicable:

- requiring advertising intermediaries involved in the placement of paid-for advertisements on the service to implement a version of the measure or the relevant part of it; and
- working with advertising intermediaries involved in the placement of paid-for advertisements on the service to implement a version of the measure or the relevant part of it.

2.13 We consider 'all reasonable endeavours' to mean providers are required to take every reasonable step within their power to implement a version of the measure(s), or the relevant part(s) of it, that is as similar as possible to the measure they are unable to apply: they are not required to take steps which would be commercially or financially unreasonable. However, given that our objective and the purpose of the fraudulent advertising duties is to reduce harm from fraudulent advertising, steps are likely to be reasonable even if they result in a reduction in a provider's revenue.

2.14 This proposed threshold is context-dependent, allowing for variations in agreements between providers and advertising intermediaries. In this way, 'all reasonable endeavours'

[11] For example, on a single service, a provider with a hybrid arrangement may apply a moderation measure to advertisements that have been placed via an owned-and-operated supply chain, and also apply the intermediaries measure on advertisements that have been placed via an open-display supply chain (see Example 2.1).

offers flexibility for providers while remaining consistent with the obligations placed on providers to operate a service using proportionate systems and processes designed to meet the fraudulent advertising duties.

2.15 The measure we are proposing sets out that, where a provider cannot implement a measure, in whole or in part, they should use all reasonable endeavours to implement a version of that measure, or the relevant part of that measure, that is as similar to the measure as possible. 'All reasonable endeavours' should include working with or requiring advertising intermediaries they have arrangements with to implement a version of the measure (or the relevant part of it) they are unable to apply, that is as similar to the measure as possible.

2.16 The intermediaries measure is not intended to limit providers' commercial freedom to contract with industry service providers or third parties to source, deploy or operate the systems and processes needed to implement the other measures proposed in the draft Codes on their behalf, such as content moderation or account verification. However, in those circumstances, providers remain fully responsible for securing the outcome of the measures in full in order to benefit from the safe harbour, as they do have sufficient control to implement the measure themselves if they wish to do so.

2.17 To help providers understand how this proposed measure works, and how it should enable them to implement a version of the measure(s), or the relevant part of it, that is as similar as possible to the measure they are unable to apply, we provide two examples of the circumstances in which this measure might be used and what 'all reasonable endeavours' might look like. These examples are illustrative only: the systems and processes a provider will use to place paid-for advertisements on their service will evolve, and there are likely to be other ways to implement a version of the measure, or the relevant part of it, that is as similar to the measure the provider is unable to apply as possible.[12]

Example 2.1: Moderation

This example is about implementing a version of the proposed advertising moderation measure that recommends that providers have systems and processes in place to review and assess paid-for advertisements they have reason to suspect may be fraudulent.

It considers a scenario where the moderation measure would be used on a paid-for advertisement that has been reported as a suspected fraudulent advertisement. Where a provider is using an open-display supply chain to place paid-for advertisements on its service – and a range of advertising intermediaries are involved in sourcing and determining the paid-for advertisement – the provider might not have access to the necessary information to moderate the reported paid-for advertisement.

In this scenario, the advertisement slot is hosted on the service that the user is visiting, but the provider of that service may not be able to see the paid-for advertisement, particularly if it is using intermediaries that are not owned by the provider or the provider's parent company. This could mean that it is difficult for the provider to apply the advertising moderation measure, even when it has reason to suspect that a paid-for advertisement is fraudulent.

[12] These examples only relate to paid-for advertisements that are served to users through pathways where a provider has insufficient control over the placement of paid-for advertisements and the application of safety measures. They are not intended to illustrate situations where paid-for advertisements have been served to users through pathways that are integrated with the service.

As such, the provider should work with – or require – intermediaries to ensure that moderation processes are carried out on suspected fraudulent paid-for advertisements. If found to be fraudulent, the provider or intermediary should take steps to prevent the paid-for advertisement from being displayed on the service in the future, and take appropriate action. This can be achieved through activity such as:

- the provider setting clear parameters consistent with its advertising policy for intermediaries to apply where they carry out moderation on the provider's behalf; and
- the provider requiring an advertising intermediary to assess the reported paid-for advertisement against the contract it has with the provider or against the provider's terms of service or publicly available statement, and to take steps to prevent the paid-for advertisement from being displayed on the service again if it is identified to be fraudulent.

When considering how to use all reasonable endeavours to apply the moderation measure more generally, providers could also:

- put arrangements in place to ensure that the provider has visibility of all paid-for advertisements placed on its service, so that it can carry out moderation activity or have moderation activity carried out on its behalf;
- ensure that, where placements are made on the basis of a bidding process, a winning bid is not passed to the publisher ad server unless it meets the advertisement serving parameters set by the provider;
- put in place a mechanism, either at an intermediary or provider level, to prevent a fraudulent advertisement from being displayed on the service in the future, following the moderation outcome; and
- ensure that there is effective senior oversight of the development, operation and effectiveness of any processes, and that the processes are monitored and regularly reviewed and audited, and any weaknesses identified are addressed.

Note: the above lists are non-exhaustive and there may be other ways that providers can work with or require intermediaries to implement a version of the proposed moderation measure.

Example 2.2: Account checks and actions

This example is about implementing an aspect of the proposed account checks and actions measure that recommends that providers should have and apply an account checks and actions policy that, among other things, verifies that advertising account holders work for, or on behalf of, the individual or organisation they are advertising for.

It considers a scenario where the provider is unable to determine who the advertiser is ahead of the paid-for advertisement being placed on its service. This may also be the case once the paid-for advertisement goes live because, in some scenarios, the provider might not be able to see or access the paid-for advertisement (or the ad inventory where such information may be stored). This scenario may arise where a provider is using an open-display supply chain to place paid-for advertisements on its service and a range of advertising intermediaries are involved in sourcing and determining the paid-for advertisement that the provider will place.

In this scenario, checks would likely happen at the start or on the demand-side of the supply chain where an advertiser registers with media agencies, advertiser ad servers, or demand-

side platforms (DSPs). A service provider is more likely to have direct relationships with supply-side intermediaries. This could mean that the provider has insufficient control to verify that advertising account holders work for or on behalf of the individual or organisation they are advertising for.

As such, the provider should work with – or require – intermediaries to ensure that account checks are carried out on all new accounts. This can be achieved through activity such as:

- requiring that the supply-side platform (SSP) or publisher ad server that it has an arrangement with only accepts paid-for advertisements from advertisers that have been verified as working for or on behalf of the individual or organisation they are advertising for; and
- working with and via intermediaries it has arrangements with to clearly specify the checks that need to be carried out – this could include the provider requiring that SSPs or publisher ad servers, in turn, make arrangements with a DSP or advertiser ad server that they have relationships with and requiring clear performance standards and targets that are met and subject to regular audit;

In such a scenario, it is important to ensure:

- ongoing senior oversight of the development, operation and effectiveness of the processes; and
- that the effectiveness of the process is regularly reviewed to identify and address any weaknesses.

Note: the above list is non-exhaustive and there may be other ways that providers can work with or require intermediaries to implement a version of this aspect of the proposed account checks and actions measure.

A coherent approach to 'all reasonable endeavours'

2.18 When determining what 'all reasonable endeavours' might entail, providers should consider the measure(s) they have been unable to apply – and the steps they are taking to implement a version of that measure – holistically. This will help ensure that actions they take to reach 'all reasonable endeavours' for one proposed measure do not undermine the implementation of another measure.

2.19 Many of the measures we are proposing to recommend as part of the draft Fraudulent Advertising Codes are interlinked with other measures in the Codes. Even though they appear in this consultation as standalone measures, some form part of a process. For example, our proposed advertising complaints measures are closely linked (with some relating to the user experience of making an advertising complaint, and some relating to the back-end processes of how an advertising complaint is dealt with) and, in many cases, build on one another. Taken together, they form a coherent set of recommendations that would result in strong advertising complaints systems and processes when implemented. They also cross-refer to our proposed advertising moderation measures. Where a provider is unable to apply a particular advertising complaints measure and instead relies on this intermediaries measure to implement a version of that measure, it should consider how the steps it takes may affect the operation and effectiveness of the other advertising complaints measures.

Record-keeping

2.20 This proposed measure includes a record-keeping element. For each draft Fraudulent Advertising Codes measure the provider has been unable to apply, the provider should, at minimum, record the following:

a) which proposed measure(s) in the draft Codes it has been unable to implement;
b) why it has been unable to implement the measure(s);[13] and
c) what steps it has taken instead to implement a version of the measure(s), or the relevant part of it, that is as similar to the measure as possible, and the outcome of those steps. This should include detail of the governance processes that have been followed for those steps.

2.21 Providers could also record the following:

a) the broader effect of the steps considered and taken to implement a version of the measure(s), or the relevant part of it, that is as similar to the measure as possible; and
b) why those steps amount to 'all reasonable endeavours'.

2.22 This record should form part of the annual review that we are proposing to recommend providers undertake each year. For more details on the annual review measure, see Volume 2, Section 4, 'Governance and accountability'.

2.23 We consider that such records should be kept for a minimum of three years (consistent with Ofcom's Record-Keeping and Review Guidance), or in accordance with the organisation's record retention policies, if longer.[14]

2.24 A record would enable providers to clearly outline any expectations they have of advertising intermediaries to help them implement a version of any measures (in whole or in part) the provider has been unable to apply. Recording which intermediaries they have made arrangements with, and what steps they have taken, would help providers assess whether expectations have been met. This would also allow providers to reassess their arrangements over time and adjust to any innovations or changes in the market. Additionally, records would enable providers to evaluate what steps they can take when using all reasonable endeavours to implement a version of the measure(s) they are unable to apply, and whether the steps they have taken achieve similar safety outcomes.

2.25 The application of this proposed measure will likely necessitate some engagement with advertising intermediaries and may involve some adjustments of pre-existing or new arrangements with advertising intermediaries. For this reason, we consider it proportionate to recommend that the record for this measure includes details of the governance steps taken for the decisions made under this measure. For more detail on the governance measures we are proposing, see Volume 2, Section 4, 'Governance and accountability'.

Existing industry and government initiatives

2.26 As we explain in paragraph 2.2, the measure we are proposing draws on existing industry and government initiatives aimed at the open-display market. Such initiatives demonstrate

[13] Providers should explain why they do not have sufficient operational or technical control over the system components required to apply the proposed measure. This should include technical details describing how those components are configured within the architecture of the categorised service and should clearly identify which party exercises control over each component. If a provider uses multiple pathways to place paid-for advertisements on its service, this step should occur for all pathways a provider uses.
[14] Ofcom, 2025. Record-Keeping and Review Guidance, p.5.

that actors operating in the open-display supply chain are able, and willing, to implement processes to detect and remove harmful advertisements before they are encountered by users, and can potentially influence other actors in the chain.

Existing government work

2.27 The National Cyber Security Centre (NCSC) has issued guidance to help brands work safely with actors in the advertising ecosystem. The NCSC recommends that advertisers ask their partners to adhere to eight principles, covering some areas that the draft Fraudulent Advertising Codes will target.[15] From this guidance, it is clear there are instances where providers can reasonably expect, and work with, advertising intermediaries to implement a version of the measure(s), or the relevant part of it, that is as similar as possible to the measure they are unable to apply.

2.28 We also note that the UK Government's Online Advertising Taskforce is working with actors in the online advertising ecosystem to improve evidence around advertising harms and address issues around transparency. Since 2025, the Taskforce's working groups have focused on specific challenges in the advertising sector including, but not limited to, information sharing across the supply chain and implementation of initiatives such as the Interactive Advertising Bureau UK's (IAB UK) Gold Standard (see Example 2.4) and the Advertising Standards Authority's (ASA) Intermediaries and Platform Principles (see Example 2.3). In November 2025, the Taskforce agreed to establish a new working group on 'ad fraud and standards', with a specific focus on fraudulent advertising and understanding existing best practice across the advertising ecosystem.[16]

2.29 As the Online Advertising Taskforce's working groups progress, there may be complementary findings which support providers to work with intermediaries in the supply chain and identify steps which can be taken to implement measures in the draft Fraudulent Advertising Codes.

Existing industry initiatives

2.30 In the example boxes, we describe some existing industry initiatives that seek to encourage safety- and transparency-minded change in open-display supply chains.

Example 2.3: The ASA's Intermediary and Platform Principles pilot

The ASA's Intermediary and Platform Principles pilot explored how the UK advertising self-regulatory system could encourage more transparency and broader accountability in online spaces.[17]

The pilot found that, where appropriate, participating companies (which included advertising intermediaries) were willing to implement the principles, and proactively engage with the ASA to share information on their implementation.

[15] NCSC, 2024. Guidance for brands to help advertising partners counter malvertising. [accessed 22 June 2026].
[16] The Online Advertising Taskforce is a Department for Culture, Media and Sport (DCMS)-led programme, which aims to "deliver a programme of work to help improve transparency and accountability in online advertising". For more information on the Taskforce's working groups and their current status see: Department of Culture, Media and Sport, 2026. Online Advertising Taskforce – progress report 2025. [accessed 22 June 2026].
[17] ASA, 2023. Intermediary and Platform Principles Pilot: Final Report. [accessed 22 June 2026].

It noted improvements relating to reporting processes and the removal of advertisements which did not comply with the UK Code of Non-broadcast Advertising and Direct & Promotional Marketing (CAP Code).

Some participating companies introduced dedicated and user-friendly reporting channels so that non-compliant advertisements could quickly and easily be reported.[18]

Example 2.4: The IAB UK Gold Standard

The IAB UK Gold Standard is a certification for actors across the advertising supply chain, including providers, agency groups, media owners and advertising intermediaries that perform adtech functions.[19]

Two of its primary aims are to reduce ad fraud[20] and strengthen transparency in the supply chain.

To become accredited, companies must implement a variety of standards and processes to improve their operations, including filtering their advertisement inventory, and working closely with their downstream and upstream partners.[21]

2.31 These initiatives demonstrate that there is already work being done to improve transparency and accountability across the open-display supply chain. The measure we are proposing complements these initiatives.

Benefits and effectiveness

2.32 We are proposing this measure to ensure that, in circumstances where service providers do not have sufficient control to apply other draft Fraudulent Advertising Codes measures, they nonetheless take steps to implement a version of the measure(s), or the relevant part of it, that is as similar to the measure as possible to protect their users from fraudulent advertising. We consider that, in general, in applying this proposed measure the provider would derive benefits similar to the proposed measures it was unable to apply.

2.33 For the reasons set out in the following paragraphs, we consider the proposed measure would be an effective means of addressing the risk of fraudulent advertising in open-display and hybrid settings.

Ensuring similar safety outcomes for users

2.34 As we explain in Volume 1, Section 5, 'Approach to Codes', we consider that providers using owned-and-operated supply chains to place advertisements on their services should be able to apply all the proposed measures in these draft Codes.[22] In Volume 1, Section 3, 'Online advertising ecosystem', paragraph 3.25, we explain that many of the largest providers of

[18] *Ibid.* p.20.
[19] For more information, see Interactive Advertising Bureau. The Gold Standard. [accessed 22 June 2026].
[20] For more on the difference between fraudulent advertising and ad fraud, see Volume 1, Section 4, 'Causes and impacts of fraudulent advertising', paragraph 4.7.
[21] For more information, see Interactive Advertising Bureau, 2025. Compliance Grid Master. [accessed 22 June 2026].
[22] As we explain in Volume 1, Section 3, 'Online advertising ecosystem', paragraph 3.27, a supply chain is 'owned and operated' when it is integrated with the service; it is integrated with the service because the functions of the supply chain are handled by the provider itself rather than by separate actors within the market.

categorised services will sometimes use a combination of pathways to place advertisements on their service. We therefore do not expect that this proposed measure will need to be widely relied upon. However, providers using an open-display supply chain may not be able to implement all the measures proposed in these draft Codes fully because the intermediaries involved in that supply chain may be separate to the provider. Not applying safety measures in circumstances where advertising intermediaries are involved (particularly in hybrid arrangements) could create a disparity in safety outcomes across a single service. This, in turn, could have a negative effect on user safety and trust.

2.35 We therefore think that this proposed measure will increase protection to users. If we did not propose this measure, service providers who have insufficient control to implement a measure could reach the safe harbour without doing anything to protect users. As we explain in Volume 1, Section 5, 'Approach to Codes', we do not consider this scenario to be in keeping with the duties in the Online Safety Act 2023 (the Act).

2.36 The fraudulent advertising duties require Category 1 and 2A services to use proportionate systems and processes designed to (in summary) address fraudulent advertising. In addition to the nature of potential harm, providers should consider the degree of control they have in relation to the placement of paid-for advertisements on a categorised service. Even if a provider has insufficient control over the placement of paid-for advertisements and application of a given safety measure, they still must meet their fraudulent advertising duties in a way that is proportionate. Not doing so in instances of insufficient control may result in disparity in safety outcomes and in a provider not meeting its duties.

2.37 As an 'enabling' measure, this proposed measure is intended to ensure providers achieve appropriate standards of safety across all paid-for advertisements placed on a service, regardless of the pathways they have come from, but with consideration of the control awarded to the provider by the type of advertising supply chain used. As such, this proposed measure also accounts for situations in which a provider uses multiple pathways to place paid-for advertisements on a single service, ensuring there is no disparity in safety outcomes for users of that service.

Addressing risk of displacement

2.38 This proposed measure addresses the risk of displacement. Bad actors are opportunistic, and the harm from fraudulent advertising evolves rapidly.[23] A disparity in safety measures and outcomes could create a risk of displacement as fraudulent actors seek less obstructive opportunities to defraud users. By proposing a measure that applies to paid-for advertisements that are served through open-display supply chains, we are mitigating against the risk that fraudulent advertisements are displaced from categorised services that use owned-and-operated supply chains to categorised services that use open-display supply chains. We recognise that we cannot protect against all displacement risks but consider that efforts on Category 1 and 2A services could have positive trickle-down effects, for example, by indirectly creating incentives for intermediaries to filter out bad actors.

Complementing wider government and industry developments

2.39 Our proposed measure can complement wider government work and guidance that seeks to improve transparency and promote user safety along the supply chain. Examples 2.3 and 2.4 demonstrate that there is a precedent of actors in the open-display supply chain taking

[23] See Volume 1, Section 4, 'Causes and impacts of fraudulent advertising' for more detail on how fraudulent advertising manifests and evolves.

action to reduce the risk of harm to users. If such initiatives can improve safety standards, we consider that proposing providers use 'all reasonable endeavours' to implement a version of the measure(s), or the relevant part of it, that is as similar as possible to the measure the provider is unable to apply would complement these initiatives as they could enable providers to meet their fraudulent advertising duties.

Record-keeping

2.40 We consider that the record-keeping element of this proposed measure would also help providers to meet the fraudulent advertising duties. This is because it would enable a provider to evaluate whether the alternative steps it has taken achieve an appropriate standard of safety and to have the evidence base to enable it to make adjustments or take further steps to better protect users.

2.41 We set out more information on the benefits of record-keeping under the subheading 'Approach to record-keeping' in Volume 1, Section 5 'Approach to Codes'.

Impacts and costs on service providers

2.42 The impacts and costs of this proposed measure will depend on a variety of factors, including:

- the number of proposed Fraudulent Advertising Codes measures a provider is unable to apply;
- the number of advertising intermediaries a provider works with;
- what existing safety mitigations are in place; and
- the specific arrangements that a provider has with any advertising intermediaries.

2.43 For providers that have hybrid arrangements, this proposed measure may result in some operational complexity. This is because, depending on the specific advertising arrangements that a provider has, different safety measures may need to be applied to paid-for advertisements placed through different pathways to achieve a similar standard of safety for users.

2.44 However, the threshold that this proposed measure sets should reduce any unnecessary negative impact or costs on the provider. 'All reasonable endeavours' encourages providers to do what they feasibly can to implement a version of the measure(s), or the relevant part of it, that is as similar as possible to the measure they are unable to apply. Therefore, providers have sufficient flexibility to choose the most appropriate and cost-effective solution.

2.45 In addition, if providers or intermediaries already participate in existing industry initiatives as described in Examples 2.3 and 2.4, this may reduce some of the impacts and costs of our proposed measure and could increase the efficiency of implementation.

Costs

2.46 In applying this proposed measure, providers may change how they engage with advertising intermediaries they have arrangements with. For example, they may meet more regularly with these advertising intermediaries, seek safety- or regulatory-minded adjustments to their arrangements, or work with different advertising intermediaries. This could increase the time and resources required to maintain these arrangements.

2.47 There would likely be resourcing costs associated with identifying any steps that a provider could take to implement a version of the measure(s), or the relevant part of it, that is as

similar as possible to the measure they were unable to apply, and in determining what is feasible for a service and any specific advertising arrangements it has. This is likely to involve technical and legal expertise.

2.48 There may also be costs associated with the implementation of any mitigation strategies or safety measures that a provider determines are reasonable and proportionate for implementing a version of measures, or the relevant part of it, that is as similar to the measure they were unable to apply as possible.

2.49 There are also likely to be costs associated with ensuring appropriate records are kept.

2.50 The costs will vary by provider and will depend on the number of measures it cannot apply, the number of advertising intermediaries it works with, its existing arrangements with those intermediaries and the mitigations already in place. We consider that the overall costs are likely to be limited and proportionate. This is because the proposed measure is expected to apply only in limited circumstances, and the 'all reasonable endeavours' threshold gives providers flexibility to take steps that are feasible and proportionate to their existing arrangements.

Impact on market

2.51 While the fraudulent advertising duties only apply to providers of Category 1 and 2A services, this proposed measure could have an indirect impact on other actors in the open-display supply chain, particularly advertising intermediaries serving paid-for advertisements to both categorised and non-categorised services. Depending on the steps providers take, some advertising intermediaries may incur additional operational costs.

2.52 Wider market effects are likely to depend on the intermediary's role, scale and bargaining position, as well as the nature of the provider's request. Some intermediaries may pass on additional costs through higher fees, particularly where they provide access to important demand or inventory infrastructure. Others may operate in more competitive or commoditised parts of the supply chain, where price pressure makes it more likely that they will absorb higher costs.

2.53 However, we consider that this proposed measure provides incentives for adopting best practices and reducing the risk of fraudulent advertising across the supply chain. The advertising intermediaries market for Category 1 and 2A services is small (relative to the advertising intermediaries market more broadly) and concentrated with larger intermediaries. These factors suggest that any wider market effects are likely to be limited, though the effect on a given intermediary may vary.

2.54 The existing industry initiatives outlined in Examples 2.3 and 2.4 suggest that transparency, accountability and safety are already features of the market, and that actors in the supply chain already dedicate resources to these. Consequently, we consider it unlikely that the proposed measure would result in substantial additional costs for these actors.

2.55 Overall, due to the adversarial nature of fraudulent advertising, and the significant harm it can cause to users, we conclude that the potential benefits for increasing best practice across the supply chain outweighs the risk of negative effects on the wider market.

Rights assessment

Freedom of expression

2.56 As explained in Volume 1, Section 5, 'Approach to Codes', Article 10 of the European Convention on Human Rights (ECHR) upholds the right to freedom of expression, which encompasses the right to hold opinions and to receive and impart information and ideas without unnecessary interference by a public authority. Article 10 is a qualified right, and we must exercise our duties under the Act in a way that does not restrict this right unless satisfied that is proportionate to the legitimate aim pursued. We start from the position that this proposed measure is prescribed by law, in pursuit of a legitimate aim and corresponds to a pressing social need.

2.57 As this is an 'enabling' measure, we consider that this proposed measure's impact will largely depend on how the provider has decided to implement a version of the measure, or the relevant part of that measure, that is as similar as possible to the measure it is unable to apply. We note that this proposed measure might lead to the application, or more effective application, of other measures which may have impacts on freedom of expression. Those impacts have been considered in the rights assessments for the relevant proposed measures. We therefore refer stakeholders to those assessments.

Data protection and privacy

2.58 As explained in Volume 1, Section 5, 'Approach to Codes', Article 8 of the ECHR confers the right to respect for an individual's private and family life. Article 8 is a qualified right, and we must exercise our duties under the Act in a way that does not restrict this right unless satisfied that is proportionate to the legitimate aim pursued. As noted in relation to freedom of expression, we start from the position that this proposed measure is prescribed by law, in pursuit of a legitimate aim and corresponds to a pressing social need. Article 8 underpins the data protection laws with which providers must comply.

2.59 As was the case for freedom of expression, we consider that this proposed measure's impact on data protection and privacy will largely depend on how the provider has decided to implement a version of the measure, or the relevant part of that measure, that is as similar as possible to the measure it is unable to apply. We note that the proposed measure may lead to the application, or more effective application, of other measures which have impacts on privacy and data protection. Those impacts have been considered in the rights assessments for the relevant measures. We do not consider that this proposed measure gives rise to any additional impacts beyond those already identified.

Provisional conclusion

2.60 Providers of Category 1 and 2A services must operate their services using systems and processes designed to meet the fraudulent advertising duties. We recognise that providers will place paid-for advertisements on these services through a variety of pathways. Our proposed measure reflects the importance of effectively applying safety mitigations to paid-for advertisements regardless of the pathway(s) providers use.

2.61 We consider that this proposed measure is the least intrusive means of ensuring that users are protected from fraudulent advertisements in circumstances where providers have insufficient control to implement the proposed measures in our draft Code. We do not consider that this proposed measure has any additional impacts in relation to freedom of expression rights, privacy rights and data protection beyond those impacts we have already considered in respect of the other proposed Fraudulent Advertising Codes measures.

2.62 The costs of the proposed measure will vary between providers. However, on balance we consider they are unlikely to be disproportionate:

- The threshold of 'all reasonable endeavours' does not require providers to take steps that would be unreasonable or infeasible.
- It is critical that the proposed Fraudulent Advertising Codes effectively protect users from fraudulent advertisements regardless of the pathway used to place paid-for advertisements on a service. We are not aware of any materially less costly measures which would achieve this objective as effectively.
- We have separately concluded that the costs that providers that operate owned-and-operated supply chains would incur from implementing our proposed measures are proportionate. It follows that the costs providers that use open-display supply chains would incur when using all reasonable endeavours to implement a version of the measure(s), or the relevant part of it, that is as similar as possible to the measure it is unable to apply should therefore be proportionate.

2.63 Based on this analysis, we consider this proposed measure to set out a proportionate and effective way for providers to meet their duties in circumstances where they may have insufficient control to implement the draft Fraudulent Advertising Codes measures.

2.64 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU K1 and FAS K1 respectively.

3. Fraud indicator assessment

What is this section about?

Fraudulent advertising is a diverse and adversarial harm. Ensuring safety measures are effective and tailored to address fraudster tactics requires a sophisticated understanding of what these tactics are and how they are evolving.

In this section we set out our proposed measure on how service providers should assess relevant evidence to understand how fraudulent advertising manifests, or is likely to manifest, on their service. This assessment should be used to more effectively and proportionately apply mitigations aimed at tackling fraudulent advertising.

Our proposal

Number in our Codes	Proposed measure applicable to providers of Category 1 and 2A services
FAU B1 and FAS B1	The provider should carry out a **fraud indicator assessment** to **identify characteristics that indicate,** either alone or in combination, **a material risk of fraudulent advertising**. The provider should review and update its assessment at least every 12 months or whenever the provider makes a significant change in relation to paid-for advertisements. The provider should **track** characteristics and groups of characteristics for the purpose of identifying **any new potential fraud indicators** and then assess, as soon as practicable, whether they indicate a material risk of fraudulent advertising.

Why are we proposing this?

The fraud indicator assessment will help service providers establish what indicators suggest there is a material risk that a paid-for advertisement on their service is fraudulent. Similarly, it will help service providers establish which indicators suggest there is a material risk of an advertising account posting fraudulent advertisements.

Service providers that have undertaken a fraud indicator assessment will therefore be better able to identify advertisements and advertising accounts that pose a material risk, and subject them to additional scrutiny. It will enable them to more effectively and proportionately apply a number of other proposed measures to address fraudulent advertising, and comply with their duties under the Act.

By extension, they will detect and remove more fraudulent advertisements than they would have done without undertaking this assessment. Given the scale and impact of fraudulent advertising, we consider the benefits of this proposed measure are likely to be significant.

Consultation questions

- Do you agree with our proposal? Please provide any arguments and supporting evidence.
- Do you think there are other information sources providers should consider in order to understand how fraudulent advertising manifests on their service? If so, what information and why would they need to consider it?

Introduction

3.1 This section sets out our proposed recommendations about how providers of Category 1 and 2A services (providers) should undertake an assessment to determine indicators of fraudulent advertising on their service.

3.2 By understanding which characteristics of paid-for advertisements and advertising accounts are indicative of fraudulent advertising, the 'fraud indicator assessment' will help a provider to understand how fraudulent advertising manifests on its service. This understanding will help providers to more effectively and proportionately apply safety measures, to help them comply with their fraudulent advertising duties under the Online Safety Act 2023 (the Act).[24]

3.3 We provisionally consider that an indicator of fraudulent advertising is a characteristic or group of characteristics that is linked to a material risk that a paid-for advertisement is a fraudulent advertisement or that an advertising account will post fraudulent advertisements.

3.4 Examples of indicators of fraudulent advertising could include:

- **Brand impersonation:** Where a well-known brand's name or logo is used to deceive users into thinking they are seeing an advertisement for a legitimate brand.[25]
- **Exaggerated, attention-grabbing and pressure-inducing content:** These advertisements rely on exaggerated, sensationalised or unrealistic claims designed to capture attention and prompt quick decisions.[26]
- **Suspicious behaviour:** This could include advertisers sharing the same or similar username, photos, bios and contact information as accounts previously known to post fraudulent advertisements.[27]

3.5 In this section we set out the explanation of the proposed measure, including:

- the characteristics that service providers should look for;
- the evidence that should inform the assessment;
- how providers can use the evidence in the assessment;
- how frequently providers should repeat the assessment; and
- how they should track evidence of new indicators.

3.6 We then set out the benefits, costs and rights implications of the proposed assessment, before explaining why we consider the proposed measure to be proportionate.

3.7 We acknowledge that a Category 1 or Category 2A service may be serving paid-for advertisements to its users through different advertising pathways. Where relevant, the proposed intermediaries measure would apply. The proposed intermediaries measure recommends that a provider should use all reasonable endeavours to implement a version of any measures it has been unable to apply. For more information, see Volume 2, Section 2, 'Advertising intermediaries'.

---

[24] See sections 38 and 39 of the Act.
[25] See paragraph 4.21 and 'Impersonation through content' in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.
[26] See 'Exaggerated, attention-grabbing and pressure-inducing content' in in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.
[27] See 'Suspicious behaviour' in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.

Fraud indicator assessment

Explanation of the measure

3.8 We propose that service providers should carry out an assessment which allows them to understand the characteristics of paid-for advertisements and advertising accounts which indicate a material risk of fraudulent advertising.[28]

3.9 We propose that service providers should then use the outputs of the fraud indicator assessment to inform the measures they take both in relation to advertising accounts and in relation to the paid-for advertisements themselves, to more effectively apply those measures to protect users from fraudulent advertising.

3.10 We propose that service providers should undertake an assessment that incorporates the following elements:

- The assessment should focus on characteristics, or combinations of characteristics, that are linked to fraudulent advertising on that service.[29]
- The assessment should be informed by relevant evidence.[30]
- The assessment should use appropriate analytical methods to determine whether the identified characteristics, alone or in combination, present a material risk of fraudulent advertising or of accounts that post fraudulent advertising.[31]

3.11 To ensure that this assessment stays up to date:

- Providers should review this assessment at least every 12 months or if there is a major change in their service design in relation to paid-for advertising.[32]
- Between assessments providers should track new indicators of fraudulent advertising and feed these indicators into trust and safety processes as appropriate.[33]

3.12 This assessment should enable providers to apply safety measures in a way that is tailored, effective and proportionate.

3.13 We propose that service providers should use the fraud indicator assessment to inform actions they take at both the account level and the advertisement level. To support account-level actions (for example, account checks) providers should consider the indicators that an account may post fraudulent advertisements.[34] To support advertisement-level actions (such as advertising moderation) providers should consider the indicators that an advertisement may be fraudulent.[35]

The characteristics of fraudulent advertising

3.14 In identifying characteristics that indicate a material risk of fraudulent advertising, it is proportionate to focus on those relevant to the specific service. We propose that the

[28] We refer to this proposed assessment as a 'fraud indicator assessment'.
[29] See Table 3.2.
[30] See Table 3.3.
[31] See paragraph 3.20.
[32] See paragraph 3.22.
[33] See paragraphs 3.23 and 3.24, and Table 3.1.
[34] See 'Benefits of incorporating the findings of the assessment into other measures' in this section.
[35] Ibid.

assessment should focus on identifying characteristics for one or more of the following reasons:

- The characteristics have previously been found in fraudulent advertising on the service.
- The characteristics have been found in broader fraud trends which are relevant to the service.
- The provider has other evidence to suggest the characteristics indicate a material risk of fraudulent advertising on the service.[36]

3.15 Service providers should seek to assess characteristics relating to:

- content of the advertisement and relevant connected content;
- advertising accounts;
- account behaviour;
- targeted users; and
- service functionalities.

3.16 We consider that in some circumstances a single characteristic will be enough to indicate that an advertisement or an account poses a material risk of fraudulent advertising. For example, an account could be created from a device that was previously used by an account banned for fraudulent advertising. This single account characteristic could potentially be by itself an indicator of fraudulent advertising.

3.17 However, in other circumstances multiple characteristics may need to be combined to indicate a material risk of fraudulent advertising. For example, detecting brand impersonation may require multiple characteristics. While advertisements showcasing recognised brands are generally trustworthy, their placement by accounts linked to less reputable web domains may suggest an increased risk of fraudulent advertising.

3.18 We propose to recommend that service providers may choose to either assess whether certain indicators pose a material risk of fraudulent advertising directly or assess using a fraudulent advertising proxy.[37] Where a proxy is used, providers may identify advertisements and accounts that present a material risk of that proxy to inform the use of relevant systems and processes.

How providers should use evidence to identify indicators of fraudulent advertising

3.19 An assessment of indicators of fraudulent advertising should be informed by high-quality evidence. Sometimes relevant information will be held by third parties such as external experts and advertising intermediaries. Where a service provider is aware that a third party holds relevant information, it should request it from them in time to support the next planned assessment. Aside from information obtained from third parties, we propose that providers use the information they have already collected rather than generating new data

[36] A provider may have reason to believe that a characteristic of fraud is likely to occur on its service due to product testing (see Volume 2 Section 5, 'Testing advertisement generation tools') or other anticipatory safety activity.

[37] Where a provider assesses an advertisement that it suspects to be fraudulent against its own categories of prohibited advertisements (rather than making a fraudulent advertisement judgement per the draft guidance on fraudulent advertising judgements), this would be a judgement of 'fraudulent advertising proxy'. For more information see Volume 4, Section 2, 'Advertising moderation'.

for this assessment. We propose that providers should review relevant evidence relating to UK users from:

- results of advertising moderation systems;
- evidence drawn from existing controls including account-level action;
- relevant user data and advertiser data;
- user complaints, including user reports;
- engagement with intermediaries (to the extent they are used);
- trusted flaggers and views of independent experts in fraudulent advertising (including referrals from law enforcement);
- consultation with internal experts on fraudulent advertising risks and safety measures;
- retrospective analysis of incidents of fraudulent advertising;
- product safety testing;
- internal and external commissioned research;
- outcomes of internal audit, external audit or other risk assurance processes; and
- data sharing schemes.

3.20 We propose that service providers should have flexibility in how they use data from the listed sources to assess whether a characteristic, or combination of characteristics, presents a material risk of fraudulent advertising. We provisionally consider that an assessment of possible characteristics to find indicators could be done in several different ways, including:

- data-based analysis;
- testing using content detection technology;[38] or
- expert-led assessment of evidence.

3.21 We propose providers should use a minimum of 12 months of information to inform their assessment to the extent that it is relevant to determining whether an indicator exists.

Reviewing the assessment and tracking new indicators of fraudulent advertising

3.22 We propose that service providers should review and update their fraud indicator assessment, and share it with their most senior governance body, at least every 12 months.[39] We further consider that providers should update their assessment following any significant change to the design or functionality of their service in relation to paid-for advertising.

3.23 We propose that providers should track new characteristics linked to fraudulent advertising on their service. The latest fraud indicator assessment should serve as their baseline of characteristics that are known indicators of fraudulent advertising. Between assessments providers should look at the information they receive from outside sources and the data they collect from their trust and safety systems.[40] Providers should track new characteristics,

[38] By content detection technology we mean the same technologies that may be used as part of complying with our suggestions on advertising moderation in Volume 4, Section 2. We are not here proposing to recommend that providers should use this technology for this purpose, only acknowledging that this could be a way that a provider may choose to conduct the assessment.

[39] We consider that such records should be kept for a minimum of three years (consistent with our Record-Keeping and Review Guidance), or in accordance with the organisation's record retention policies, if longer. Source: Ofcom, 2025. Record-Keeping and Review Guidance.

[40] Trust and safety systems here refers to the systems and processes used to protect users from fraudulent advertising. This includes advertising moderation systems and account-level restrictions and sanctions.

and as soon as practicable, assess if they are new indicators of fraudulent advertising.[41] These new indicators should be fed into moderation processes.[42]

3.24 In tracking new indicators, we provisionally consider that service providers should use insights from, at least:

- referrals from law enforcement;
- information from trusted flaggers and any other expert group or body the provider considers appropriate;
- data sharing schemes; and
- their own trust and safety systems.

Benefits and effectiveness

3.25 The fraud indicator assessment will help service providers establish what indicators suggest there is a material risk that a paid-for advertisement is fraudulent. Similarly, it will help them establish what indicators suggest there is a material risk of an account posting fraudulent advertisements.

3.26 Service providers that have undertaken a fraud indicator assessment will therefore be better able to identify advertisements and accounts that pose a material risk, and subject them to additional scrutiny. It will enable them to more effectively and proportionately apply mitigations to address fraudulent advertising and comply with their fraudulent advertising duties.[43]

3.27 By extension, providers will detect and remove more fraudulent advertising than they would have done without undertaking this assessment. Given the harm fraudulent advertising causes both to individuals and the wider economy, we consider the benefits of this proposed measure are likely to be significant.[44]

Benefits of understanding the indicators of fraudulent advertising on a service

3.28 Fraudulent advertisements and the accounts that post them often share characteristics. We consider that service providers can use these characteristics to identify other advertisements and accounts that could be involved in fraudulent advertising. These accounts and advertisements can be directed for further targeted review.

3.29 We consider that the way that fraudulent advertising manifests depends on the service in question, due to differences in their features, functionalities and user base. As a result, the indicators of fraudulent advertising, and of the accounts posting it, will also vary across services. For example, [✂].[45] As fraudulent advertising presents differently across services, a service provider cannot rely entirely on generalised indicators of fraud to support its trust and safety systems and must base these decisions on indicators specific to its service. We consider that the findings from the fraud indicator assessment will support providers in

[41] New indicators meaning there are new characteristics or groups of characteristics that suggest there is a material risk that a paid-for advertisement is a fraudulent advertisement or that an account will post a fraudulent advertisement.
[42] See 'Benefits of incorporating the findings of the assessment into other measures' in this section.
[43] Sections 38 and 39 of the Act.
[44] For the wider benefits see Volume 1, Section 6, 'Combined impact assessment'.
[45] [✂] Sources: [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 24 November 2025.

tailoring their mitigations to ensure they are effective and proportionate in tackling fraudulent advertising as it manifests on their service.

Benefits of incorporating the findings of the assessment into other measures

3.30 Table 3.1 sets out how we provisionally consider that the assessment can help service providers more effectively and proportionately apply a number of our proposed measures, and comply with their fraudulent advertising duties.

Table 3.1 Benefits from the fraud indicator assessment to other measures[46]

Proposed measure	How it benefits from the assessment
Advertising moderation: Internal advertising policies[47]	The assessment can help providers to make higher-quality decisions about what to include in their internal advertising policies. This can help individuals working in moderation more easily identify fraudulent advertisements, and therefore increase the quality and certainty of their moderation decisions, as well as the speed with which staff can make them. Tracking new indicators can ensure internal content policies reflect current fraudster tactics.
Advertising moderation: Prioritisation[48]	Insights from the assessment can help providers to assess how likely it is that a suspected fraudulent advertisement is fraudulent, so they can prioritise moderation efforts on advertisements that are more likely to be fraudulent and deprioritise those with a lower likelihood (and lower potential impact). This supports more effective prioritisation of moderation actions, enabling providers to take down fraudulent advertising content more quickly. Tracking new indicators can ensure this understanding of likelihood is up to date, ensuring the right content is prioritised for moderation.
Advertising moderation: Training and materials[49]	This assessment will help providers tailor their training and materials for individuals working in moderation, to the risks that are present on their services. Similar to internal advertising policies, using insights from this assessment will help improve the quality and accuracy of moderation decisions. Tracking new indicators of fraudulent advertising can help ensure that individuals working in moderation are trained on new fraud indicators.

[46] We understand that many large service providers use proactive technology to detect fraudulent advertisements at scale. The assessment will give providers insights in how fraudulent advertising manifests on their service. As such, we expect it could be helpful context to inform what proactive technology a service should deploy. We intend to publish further proposals on proactive technology in autumn 2026 - please see Volume 4, Section 2, 'Advertising moderation' paragraphs 2.14 and 2.15 for more detail. These would set out how any proposed measures would work, and any benefits this assessment could give to those proposals.

[47] See Volume 4, Section 2, 'Advertising moderation'

[48] See Volume 4, Section 2, 'Advertising moderation'

[49] See Volume 4, Section 2, 'Advertising moderation'

Proposed measure	How it benefits from the assessment
Account checks[50]	Indicators identified through this assessment can help providers identify the types of checks that are effective in identifying accounts that have a material risk of posting fraudulent advertisements on their service. This should help them place appropriate restrictions on these accounts, and therefore help them comply with their duties, and better protect users from fraudulent advertising. Tracking new indicators can help ensure a provider's understanding of fraudster tactics is up to date, to help them more effectively and proportionately apply this mitigation.
Advertising bans[51]	The indicators identified through this assessment are likely to include common tactics to return to the service following an advertising ban. Therefore, using these indicators providers will be better able to tailor their mitigations to best address these tactics. Tracking new indicators can help to ensure a provider's understanding of how fraudsters may return to the service is up to date, to help them more effectively and proportionately apply this mitigation.
Governance and accountability: Annual review[52]	The assessment supports governance processes by giving providers insights into how fraudulent advertising manifests on their service. This is an important first step in the annual review to help the most senior governance body to assess whether measures are effectively addressing these identified concerns to comply with the fraudulent advertising duties.

3.31 Conversely, we consider that without this understanding, service providers may fail to recognise the fraud threat on their own service and how it is evolving. This could lead them to rely on generic mitigations, deploying measures that are not applicable to their service or addressing fraud in an ad hoc manner. Applying the same generic verification checks for all advertisers, for example, can place unnecessary burden on low-risk advertisers, while failing to target the risks specific to higher-risk advertisers. Equally, attempting to address every potential threat without understanding if it is relevant to its service can materially overload a provider's trust and safety resources unnecessarily. The assessment will support providers to better target their resources and ensure they are incurring the costs of tackling fraudulent advertising where it is likely to be most impactful.

3.32 Without a structured way of recognising the current threats on their service and how they are evolving, providers may miss emerging or shifting patterns and fail to learn from previous incidents. Providers could end up relying on inconsistent or ad hoc approaches to identifying and removing fraudulent advertisements and accounts. This could mean that while individual advertisements may be removed, the same tactics are repeated undetected elsewhere on a service, ultimately resulting in a greater volume of fraudulent advertisements on the service, and worse outcomes for users.[53]

[50] See Volume 3, Section 2, 'Account checks and actions'
[51] See Volume 3, Section 5, 'Advertising bans'
[52] See Volume 2, Section 4, 'Governance and accountability'
[53] Gen Digital (Corrons, L., Karabeyli, E., Khmelnytskyi, D., Bühler, T. and Pachilakis, M.), 2026. The Scam Ad Machine. [accessed 6 March 2026].

Benefits and effectiveness of using the recommended evidence sources and characteristics of the fraud indicator assessment

3.33 The proposed fraud indicator assessment prompts service providers to focus on characteristics that are linked to fraudulent advertising on that service. The different types of characteristics are outlined in Table 3.2.

Table 3.2 Characteristics to assess

Characteristic type	Description
Content of the advertisement and relevant connected content	This refers to content that is linked to fraudulent advertisements on the service. For example, sensationalist and misleading claims promising unrealistic returns, or introducing artificial urgency, are aspects of content that could suggest that an advertisement is fraudulent.[54] Relevant connected content can include a phone number included in the advertisement, a URL, or the linked landing page.[55] For example, some bad actors use cloaking mechanisms where landing pages differ depending on the user or system accessing them.[56]
Advertising accounts	These include characteristics of advertising accounts that are linked to fraudulent advertising on the service. For example, accounts that have no followers or no non-paid-for activity could suggest that an account is fraudulent.[57]
Account behaviour	Providers should consider behaviours of advertising account holders that are linked to fraudulent advertisements on the service. For example, accounts created more rapidly than would normally be expected or where the advertiser frequently changes the image and text content of the advertisement could suggest that an account is fraudulent.[58]

[54] See 'Exaggerated, attention grabbing and pressure-inducing content' in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising'.
[55] See Volume 1, Section 2, 'Introduction' for further information on the meaning of paid-for advertisements and landing pages, including for Category 2A services how a paid-for advertisement includes the landing page when within one click of a paid-for search result. See Annex 9, 'Guidance on making fraudulent adverting judgement: updates to the Illegal Content Judgements Guidance' for proposed guidance on how information about the destination of an advert and use of URL-scanning technology may be relevant and reasonably available information for making judgements about advertisements.
[56] See 'Cloaking' in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising'. The Integrity Institute identifies landing page analysis alongside content and behaviour analysis as a part of the best practices for detecting fraud in advertising. Source: Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.
[57] See 'Suspicious behaviour' in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising'.
[58] See 'Suspicious behaviour' in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising'.

Characteristic type	Description
Targeted users	Though any user can become a victim of fraudulent advertisements, some users may be more likely to be targeted by fraudulent advertisements. This can be because they are directly targeted using advertising targeting functionality or because they are indirectly targeted.[59] For example, a bad actor could attempt to defraud people who have previously fallen victim to a scam.[60] The bad actor could use targeting systems to show advertisements to individual users who have previously clicked on fraudulent advertisements. Alternatively, the bad actor could indirectly target similar users by placing advertisements where they are likely to see them, such as on a sub-section of a service dedicated to supporting scam victims. The users encountering the advertisements (in these examples scam-victims) could indicate that an advertisement is fraudulent. A provider should consider the users that encounter fraudulent advertising as part of the characteristics of fraudulent advertising.
Service functionalities	Relevant service functionalities are those linked to the creation or dissemination of fraudulent advertising on the service. For example, automated advertisement generation tools can be used to impersonate famous individuals, which could suggest that advertisements using this feature may be fraudulent.[61]

3.34 Considering characteristics across all these areas should ensure that the service provider develops a fuller understanding of the techniques fraudsters use, including whether fraudsters are targeting users that are likely to be vulnerable to fraudulent advertising. For example, focusing only on content characteristics that indicate an advertisement has a material risk of being fraudulent may reveal what fraudulent advertisements look like to the user, but could miss important account-level characteristics that show fraudsters' tactics in committing fraudulent activity. We consider that this understanding of the different characteristics involved creates a stronger indicator of fraudulent advertising that can better inform a provider's safety systems.

3.35 We provisionally consider that using the proposed wide range of evidence sources will allow service providers to have a more comprehensive understanding of how fraudulent advertising manifests on their service. Internal sources of evidence should reveal the characteristics of fraudulent advertising they already detect on their service. Both internal and external sources, including expertise and research, should broaden this understanding by identifying characteristics that the provider may currently miss, as well as emerging characteristics likely to appear on their service soon. We have set out the proposed evidence in Table 3.3.

[59] The Integrity Institute identify small exposure to the "right" audience as a tactic used by fraudsters. Source: Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.
[60] Evidence suggestions that those who have already fallen victim to fraud are more susceptible to secondary attacks. See 'Vulnerability; in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising'.
[61] See 'Use of GenAI and deepfakes' in Volume 1 Section 4, 'Causes and impacts of fraudulent advertising' and Volume 2, Section 5, 'Testing advertisement generation tools'.

Table 3.3: Evidence to inform the assessment

Evidence type	Description
Results of advertising moderation systems	Providers should have an advertising moderation system in place to comply with their fraudulent advertising duties.[62] The relevant results from these systems will provide useful insights into the characteristics of the fraudulent advertisements the provider has detected.
Evidence drawn from existing controls including account-level action	Providers are highly likely to have existing controls for enacting sanctions against accounts used for fraudulent advertising. The relevant data from accounts that the provider has taken action against will provide useful insights into the characteristics of the accounts found to have posted fraudulent advertisements.
Relevant user data and advertiser data	It is highly likely that providers will have some data from users and advertising account holders.[63] Fraudulent advertisements or certain types of fraudulent advertisements may be more likely to target certain groups and providers should use relevant data to provide useful insights on who these groups are.[64]
User complaints, including user reports	Providers should have existing systems for users to report fraudulent advertisements and submit complaints relating to fraudulent advertisements.[65] The relevant data from user complaints and reports will include useful insights around the characteristics of fraudulent advertisements being complained about. This will help identify potential characteristics that a provider's systems are currently failing to detect.
Engagement with intermediaries (to the extent providers use them)	Some providers may use intermediaries as part of their advertising systems. These intermediaries may have controls that detect fraudulent advertising or detect accounts connected to fraudulent advertising. Intermediaries may also hold additional data about advertising and accounts that a service provider identifies as being fraudulent advertising or posting fraudulent advertising. Where this is the case, providers should request relevant information from the intermediaries they work with.

[62] See Volume 4, Section 2 for our proposals on advertising moderation. If a provider chooses to take alternative measures to those, we would still expect them to have some form of advertising moderation to meaningfully comply with the fraudulent advertising duties. It may be that this is achieved through intermediaries – see 'Engagement with intermediaries' in this table.

[63] By user data and advertiser data we mean data a provider holds that has been provided by users and advertisers including their personal data (for example, data provided when a user sets up an account), and data about users or advertisers that a provider has created, compiled or obtained (for example, data relating to when or where users access a service or how they use it).

[64] See 'Vulnerability' in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.

[65] See Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'. If a provider has taken alternative measures to those, we would still expect them to have some form of advertising reporting to meaningfully comply with the fraudulent advertising duties. It may be that this is achieved through intermediaries – see 'Engagement with intermediaries' in this table.

Evidence type	Description
Trusted flaggers and views of independent experts in fraudulent advertising (including referrals from law enforcement)	Trusted flaggers and other external experts can provide reports on specific incidents of fraudulent advertising on a service. They can also provide broader information on general characteristics of fraudulent advertising that they have observed, as well as industry trends and regulatory standards. Providers should proactively consult experts such as trusted flaggers to understand relevant fraudulent advertising trends to inform an assessment as well as receiving reports from them through dedicated reporting channels.
Consultation with internal experts on fraudulent advertising risks and safety measures	Internal fraud or advertising subject matter experts and relevant safety measure experts can highlight characteristics that they have directly observed or would expect to find in fraudulent advertising and related accounts on a service. For example, some providers may employ experts who investigate the behaviour of fraudsters occurring off-service.[66] This can highlight the characteristics associated with known tactics for fraud that might be applicable to their service.
Retrospective analysis of incidents of fraudulent advertising	Following significant incidents of harm providers often undertake retrospective 'lessons learned' exercises. These exercises would typically include understanding the underlying cause (including relevant characteristics) and consider what appropriate mitigations they could take.[67]
Product safety testing	Product testing can be used to identify vulnerabilities in any part of a product's functionalities that enables the creation, placement or targeting of fraudulent advertising.[68] This can identify service functionalities or associated content and behaviour that are characteristics of fraudulent advertising.
Internal and external commissioned research	Providers may choose to commission internal or external research to inform their approach to safety and moderation on the service. It should be used where it contains information and insights about the characteristics of fraudulent advertising.
Outcomes of internal, external audit or other risk assurance processes	Audits in relation to aspects of their service or other risk assurance processes may help providers identify characteristics of fraudulent advertising.
Data sharing schemes	Relevant information providers receive from data sharing schemes with other services or other industries can provide information on characteristics of fraudulent advertising that may be present on a provider's service. For example, in the Fraud Intelligence Reciprocal Exchange (FIRE) programme, Meta receives information from UK banks, and the Global Signal Exchange shares signals from multiple sectors about possible scams.[69]

[66] The Integrity Institute identify "threat intelligence" as a crucial part of combatting fraud. Source: Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.
[67] These insights could also be applied to various mitigations, for example, it could lead to an update in internal content moderation policies, changes to considerations on prioritisation of content for review or updates to the training provided to staff (see Volume 4, Section 2 where we set out proposals for advertising moderation).
[68] An example of product testing can be seen in Volume 2, Section 5 where we set out our proposals for testing of advertisement generators.
[69] Meta, 2024. Meta Partners with UK Banks to Combat Scams. [accessed 3 March 2026]; Global Signal Exchange, no date. About the Global Signal Exchange. [accessed 3 March 2026].

3.36 We provisionally recommend using relevant insights from a 12-month period to inform the assessment. A 12-month period provides a sufficiently robust and comprehensive evidence base, supporting the identification of links between characteristics and fraudulent advertising. It enables providers to capture both commonly observed and less frequently occurring characteristics, as well as to identify recurring or time-specific patterns in activity, including seasonal or periodic trends. This supports more reliable and well-informed assessments, enables providers to take account of when evidence arises within that period when designing mitigations, and is balanced against the cost of data storage and the risk of using out-of-date data.

3.37 Where a provider introduces paid-for advertising for the first time, we propose to recommend that they should conduct their initial assessment within three months. This will allow them time to accumulate initial evidence for how fraudulent advertising manifests on their service. This is also consistent with the period to carry out the illegal content and protection of children risk assessments as set out in Schedule 3 to the Act.

Benefits and effectiveness of reviewing the assessment and tracking new indicators

3.38 The fraudulent advertising space is adversarial, with indicators of fraudulent advertising changing at pace as fraudsters respond quickly to counter measures.[70] For this reason we provisionally propose that service providers should regularly assess the indicators on their service and do not propose to set out a comprehensive list of indicators of fraudulent advertising. This reflects the risk that any service-specific or centrally produced list of indicators may quickly become outdated.

3.39 To balance this risk of an assessment becoming out of date against the costs to service providers, we provisionally consider that providers should review and update their fraud indicator assessment at least every 12 months. This frequency would also allow alignment with other governance activities relating to fraudulent advertising,[71] user-generated content and search content under providers' illegal content duties.[72]

3.40 We also provisionally consider that providers should update their assessment following any significant change to the design or functionality of their service. Such changes may create new opportunities for exploitation or alter fraudsters' tactics. For example, introducing an advertisement generation tool that enables the use of celebrity likenesses could increase the risk of impersonation scams.[73]

3.41 We propose to recommend that service providers record the fraud indicator assessment and share it with their most senior governance body at least every 12 months. This will support informed decision-making and accountability for action to prevent fraudulent advertising in relation to the indicators identified.

---

[70] See 'How fraud manifests through online advertising' in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.
[71] See 'Annual review' in Volume 2 Section 4 'Governance and accountability' and 'Regular review and update of the policy' in Volume 3, Section 2 'Account checks and actions'.
[72] See the illegal content Risk Assessment Guidance. Providers may also wish to align with their protection of children duties where these apply.
[73] For more on this example see Volume 2, Section 5, 'Testing advertisement generation tools'.

3.42 Although we propose to recommend that service providers keep their fraud indicator assessments up to date through an annual review, we consider that monitoring real-time changes in fraudulent advertising will be important for ensuring the ongoing effectiveness of mitigations.[74] Fraudsters change tactics at pace to incorporate external news stories and events.[75]

3.43 The more limited sources of evidence we have proposed for tracking new indicators helps balance the need to stay informed of new trends in fraudulent advertising, and the costs involved in monitoring different forms of evidence. We note, there are significant costs to providers of not staying in front of the latest trends in fraudulent advertising, including the risk that fraudulent advertising becomes widespread on their service, and it becomes unattractive to both users and advertisers.

The fraud indicator assessment and existing industry best practice

3.44 We consider that our proposed approach aligns with recognised industry best practice for addressing fraud risks. For example, the Digital Trust and Safety Partnership's Safe Framework highlights the importance of "developing insight and analysis capabilities to understand patterns of abuse and identify preventive mitigations that can be integrated into products" for addressing content and conduct risks.[76] The Integrity Institute also identifies that effective fraud mitigation requires continuous monitoring.[77] These sources emphasise the value of maintaining a set of indicators that can inform the design of trust and safety interventions. As set out in paragraph 3.2, we consider that undertaking a fraud indicator assessment will help providers identify indicators, which will enable them to more effectively and proportionately apply other measures we propose to recommend in the draft Fraudulent Advertising Codes of Practice (or alternative measures, as relevant).

3.45 Our proposed approach is also in line with what many service providers are currently doing. Providers may use different language to describe these indicators of fraudulent advertising, such as 'risk factors' or 'signals'. However, the information we have received from several large service providers suggests that they currently monitor for the types of content that are likely to be fraudulent advertising and use them to effectively develop and deploy their trust and safety systems.[78] We consider that specifying how the assessment should be carried out

---

[74] We note that the International Chamber of Commerce and the Global Anti-Scam Alliance in their best practice guide have recognised the importance of maintaining an up to date understanding of the risk posed by fraudsters. Note that this refers to scams or fraudulent advertising more generally (not just paid-for fraudulent advertisements per the Act's definition) Source: International Chamber of Commerce and Global Anti-Scam Alliance, 2026. Best practices for combating scams in advertising. [accessed 16 June 2026].

[75] See 'Tools and techniques used in fraudulent advertising' in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.

[76] Digital Trust and Safety Partnership, 2025. The Safe Framework Specification, p.9. [accessed 3 March 2026].

[77] Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.

[78] [✂] explained that it monitors for 'indicators' that an advertisement or advertiser is violating its policies, considering the advertisement content and its destination, and behavioural signals of the account. [✂] explained that it analyses signals across ad content, advertiser characteristics and functional characteristics, with systems in place to detect misrepresentation of public entities and brands. [✂] explained that it identifies the common features in content found to be fraudulent and the common behaviours of advertiser accounts associated with fraudulent advertising by considering historic violative content. [✂] Sources: [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 24 November 2025; [✂] response to our formal information request issued 26 June 2025; [✂]

and used formalises existing practice and raises the baseline, rather than representing a significant change in how providers currently run their systems.

3.46 Overall, we consider that the proposed fraud indicator assessment provides a way for service providers to identify the characteristics of high-risk advertisements and accounts, allowing them to apply more effective and proportionate mitigations. This means that providers will detect and remove more fraudulent advertising than they would otherwise, improving user protection in a proportionate, evidence-based way. Given the large amount of harm that fraudulent advertising causes we consider that this should result in substantial benefits from this proposed measure.[79]

Impacts and costs on service providers

Direct costs for service providers

3.47 In this sub-section, we consider how service providers could implement the proposed measure and seek to estimate the respective costs (where possible) for each step of the proposed measure.

3.48 In our estimates we assume that service providers will incur the costs of implementing these steps from scratch. In practice services are already likely to have some of the relevant underlying systems needed to identify and assess risks and therefore to align with this proposed measure.[80] This means that the costs they incur would likely be lower than those set out in the rest of this sub-section.

Costs associated with collating and reviewing appropriate internal and external evidence sources to identify characteristics of fraudulent advertisements, or accounts that post fraudulent advertising

3.49 The proposed measure recommends that service providers use relevant internal and external evidence to identify the characteristics of fraudulent advertisements and accounts that have posted fraudulent advertisements.

3.50 The proposed measure specifies the type of evidence (in Table 3.3) generated by existing systems and processes that providers should use to inform their assessment. We therefore consider it is appropriate to focus our estimates on the incremental costs associated with collating and reviewing the relevant evidence in order to identify characteristics of fraudulent advertisements or accounts that post fraudulent advertisements rather than the costs providers would be expected to incur to set up and operate the systems and processes to generate this evidence.[81] To the extent that providers would be required to have in place such systems and processes to align with other proposed Fraudulent Advertising Codes measures (for example, advertising moderation and advertising complaints), we note that the respective costs have been considered and where appropriate estimated in the relevant sections.[82]

---

response to our formal information request issued 24 November 2025; [✂] response to our formal information request issued 26 June 2025.

[79] See our overall calculations of harm in Volume 1, Section 6, 'Combined impact assessment'.

[80] We consider this in further detail in paragraph 3.72.

[81] We note that there could be some costs associated with service providers requesting and accessing relevant information known to be held by third parties (for example, external experts), but we expect these costs to be small and again providers would not need to incur the costs of generating this evidence.

[82] See Volume 4, Section 2 for our proposals on moderation; and Volume, 4 Section 4 for our proposals on advertising complaints. We acknowledge that providers could take alternative measures to those that we

3.51 To effectively identify characteristics of fraudulent advertisements (or of accounts that post fraudulent advertising), we expect providers will need to create feedback loops between the systems generating the relevant evidence (for example, advertising moderation and advertising complaints) and the fraudulent indicator assessment workflows. This would involve the storing, collating and reviewing of relevant evidence on an ongoing basis.[83] We note that there may be additional costs to providers to the extent the proposed measure asks them to do more with the evidence and sources available to them than they currently do.

3.52 Our proposed measure allows service providers flexibility in how they choose to collate and review the relevant evidence. We expect that providers may choose to use analytical methods (incorporating approaches such as machine learning or statistical analysis), as we are aware that these approaches are consistent with the current practices of providers in relation to fraud more broadly.[84]

3.53 In the following paragraphs, we describe the costs service providers would be expected to incur if developing a machine-learning model in-house or alternatively if accessing a third-party platform for the purposes of synthesising and processing the relevant evidence.

Developing a machine-learning model in-house

3.54 There is limited publicly available information on how categorised services (in particular, the ones operating walled garden models) use machine learning to examine evidence in relation to fraudulent advertising, the amount of resourcing this takes or the likely costs involved. From formal requests for information to service providers, we understand that providers are often unable to identify the main inputs and costs solely dedicated towards dealing with fraudulent advertising.[85] We also understand that developing a machine-learning model is a continuous process and it can be difficult to distinguish between potential one-off and ongoing costs. Therefore, we have taken the approach of considering the potential costs involved at a high level.

3.55 To develop a machine-learning model in-house, we expect that service providers would go through the following high-level stages:

- Data collection and preparation: This would involve understanding, collating and storing the range of available data; cleaning and formatting the data; creating variables from the data (feature engineering); and so on. We expect these costs would scale with the amount of data and evidence sources available to providers.
- Model developing, training and testing: This would involve choosing an appropriate model, using the prepared data to train the model, assessing and

---

propose, but they would still need to comply with their fraudulent advertising duties, and therefore we think it is reasonable to assume for the purpose of estimating costs that they have done so by implementing our proposed recommended measures.

[83] We consider the potential storage costs further in paragraph 3.69 as part of our proposed recommendation for the ongoing tracking of evidence.

[84] Google indicated the use of machine-learning models to detect various fraud types, [✂] and LinkedIn use machine-learning models including account and member behavioural signals to identify for potential fraud. Sources: Google response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; LinkedIn response to our formal information request issued 26 June 2025.

[85] [✂] response to our formal information request issued 30 January 2026; [✂] response to our formal information request issued 30 January 2026; [✂] response to our formal information request issued 30 January 2026].

quality assuring the model, ensuring alignment with existing regulatory frameworks,[86] and so on.

- Model deployment and monitoring: This would involve integrating the model into existing infrastructures, monitoring to detect problems, fine-tuning the model, and so on.

3.56 These stages would require input from a range of staff, including data scientists, data engineers or both, who are likely to play a crucial role. Based on our understanding of the salaries associated with these types of roles,[87] we consider that developing a machine-learning model could involve significant initial costs extending into the high hundreds of thousands of pounds, and annual ongoing costs in the low tens of thousands of pounds for maintaining the model (for example, updating and retraining the model as new data comes through).

3.57 We further expect these costs are likely to vary widely across individual services, depending on factors such as service size, existing infrastructure and expertise, volume of advertisements, type of advertising data that needs to be processed (for example, text, image, or video), signals provided by the data (for example, keywords, landing pages), the internal auditing processes of services, and so on.

Accessing a third-party solution

3.58 There are alternative options available in the market in the form of external third-party platforms.[88] These platforms can use deep learning to analyse images and videos, offering service providers an alternative to building machine-learning models from scratch. By doing so, these platforms can help providers to streamline development and therefore incur lower upfront costs.

3.59 Service providers can access these platforms through integrating the relevant application platform interfaces (APIs). These platforms tend to operate on a pay-as-you-go basis, meaning that customers only pay for what they choose to analyse, with the pricing being based on the type of analysis performed.[89]

3.60 We expect a service provider using third-party platforms would need to dedicate staff and engineering resources towards the following tasks:

- preparing storage of the relevant data (image and videos);

---

[86] For example, the International Organization for Standardization, 2023. IEC 42001:2023. Information technology artificial intelligence management system. [accessed 23 March 2026].

[87] As explained in Annex 8, 'Further detail on economic assumptions and analysis', we assume the earnings of a software engineer (£60,000 to £110,000 per year) from the 2025 Annual Survey of Hours and Earnings would be the closest match to the earnings of a data scientist or engineer in the UK. We expect service providers are overall likely to require several individuals from these professions (software engineers, data scientists, data engineers). However, we understand that the salaries for similar roles could be higher in some countries, such as the US. For example, TikTok indicates a US-based annual salary of around $202,160 to $368,220 for a machine-learning engineer. Source: TikTok. Machine Learning Engineer, Ecommerce Risk Control. [accessed 24 March 2026]; Similarly, [✂] estimates an average annual base salary (and performance bonus) for an engineer is [✂]. Source: [✂] response to our formal information request issued 30 January 2026.

[88] For example, Amazon Rekognition, Google Cloud Vision API and Microsoft Azure Computer Vision. [accessed 23 March 2026].

[89] For example, Amazon Web Services (AWS) Rekognition charges around $0.001 per image for the first 1 million images analysed and around $0.10 per minute of video being processed for label detection. Source: AWS. Amazon Rekognition pricing. [accessed 27 February 2026].

- ensuring the platform has continuous access to the relevant data (images and videos) to be analysed;
- integrating and testing the relevant APIs provided by the platform for analysis of the data; and
- ensuring compliance with data regulations, involving data governance processes and audits.

3.61 Based on the tasks set out in paragraph 3.60, we estimate the one-off costs associated with using a third-party platform to collate and review appropriate internal and external evidence sources to identify characteristics of fraudulent advertisements, or accounts that post fraudulent advertisements will be around £20,400 to £123,000.[90] There may also be ongoing costs in relation to monitoring and applying updates, as well as maintaining integration with the third-party platform over time. Consistent with our standard assumptions on maintenance costs, we estimate the maintenance costs to be around £5,100 to £30,700 per year.[91]

Costs to assess whether a characteristic, or group of characteristics, presents a material risk of fraudulent advertising

3.62 As explained in paragraph 3.20, service providers will have the flexibility to determine how they choose to undertake this assessment as a range of options varying in complexity and technical requirements are available to them.

3.63 Considering the volume of advertisements and range of characteristics service providers may need to assess to determine 'indicators of fraudulent advertising', we expect that some service providers may choose to develop an automated fraudulent advertisement risk-scoring tool using data-driven modelling approaches (for example, machine learning). This approach also appears consistent with the current practices of some providers.[92] There could be some cost synergies for providers choosing to use data-driven modelling approaches for collating and reviewing relevant evidence and also assessing for indicators of fraudulent advertising using this evidence.

3.64 To develop an automated risk-scoring tool in-house, we expect that the stages involved would be similar to those considered in paragraph 3.55. However, there would be further input needed in determining an appropriate risk-scoring mechanism (for example, to assess the relationship between characteristics and fraudulent advertising), which may involve the need to identify, test, and set thresholds and evaluation metrics.

3.65 We broadly expect a need for similar types of roles, including data scientists, data engineers or both. However, there could also be a greater reliance on external experts in identifying common fraud tactics. Based on the range of inputs likely to be necessary, we consider that service providers opting for a data-driven modelling approach could incur significant one-off costs in the high hundreds of thousands to millions of pounds, and ongoing maintenance costs in the mid-tens of thousands of pounds.

---

[90] We assume this would take around two to six months of full-time work by a software engineer and the equivalent time for a professional occupations staff. See our labour cost assumptions as set out in Annex 8, 'Further detail on economic assumptions and analysis'.

[91] We assume an ongoing annual maintenance cost of 25% of the initial one-off cost.

[92] [✂] response to our formal information request issued 30 January 2026; [✂] response to our formal information request issued 30 January 2026.

3.66 However, we note that these costs are based on service providers opting for the most expensive approach and developing it from scratch just for this draft Fraudulent Advertising Code. This is not the only option available to providers to align with this proposed measure. Providers could instead choose to adopt less technical, lower-cost methods, that are more proportionate to the size and nature of risk on their service. For example, providers could choose to do a qualitative assessment to determine whether a characteristic should be treated as an indicator of fraudulent advertising. This is likely to require some input from internal staff and external experts to identify, group and label potential characteristics, as well as engineering input to convert these labels into identifiable signals that can be detected by safety systems. We would expect the costs associated with this to be materially lower than the costs set out in paragraph 3.65.

Costs associated with reviewing the fraud indicator assessment at least every 12 months and tracking new indicators of fraudulent advertising between assessments

3.67 The proposed measure recommends that service providers should keep their understanding of potential indicators of fraudulent advertising up to date, through annual reviews of the fraud indicator assessment as well as the ongoing tracking of evidence for potential new indicators.

3.68 The proposed measure specifies the type of evidence that providers should refer to for the ongoing tracking of new indicators (see paragraph 3.23). We expect this evidence would mostly be generated by existing systems and processes, or the systems and processes service providers would need to implement to align with our other proposed Fraudulent Advertising Codes measures.[93] Therefore, we do not expect service providers to incur the costs of setting up these systems and processes. Instead, we consider the main costs will be in relation to the ongoing collation and reviewing of evidence generated from these systems and processes, and assessment of the characteristics or group of characteristics identified from this evidence.

3.69 There could also be storage and data management costs to service providers if having to retain more data from advertisements (compared to what they may have previously). To do this, providers may need to expand their storage capacity and maintain larger datasets over time. However, based on our internal expertise, we understand that data retention requirements for text and image content in particular are likely to be small.[94]

3.70 Where new characteristics are found, we also expect there may be additional costs in adding this to the list of indicators of fraudulent advertising and feeding them into the relevant measures to inform processes.

Overall costs

3.71 Our approach to identifying and where appropriate estimating the costs associated with implementing the different components of this proposed measure has mostly used the most expensive option available to providers as its basis (for example, machine-learning models).

[93] We note that the respective costs of these other proposed measures (for example, advertising moderation) have been considered and where appropriate estimated in the relevant sections. We acknowledge that providers could take alternative measures to those that we propose, but they would still need to comply with their fraudulent advertising duties, and therefore we think it is reasonable to assume for the purpose of estimating costs that they have done so by implementing our proposed recommended measures.
[94] We also think it would be possible for service providers to retain relevant data from videos in text or image format (for example, key frames from videos, transcribed audio).

Based on this assessment, we acknowledge that there could be significant costs if adopting certain methods to conduct a fraud indicator assessment from scratch.

3.72 However, in practice, we expect the costs service providers would incur to comply with our proposed measure would be lower and potentially materially lower, due to the following reasons:

- Providers would have the flexibility to choose the methods that are most cost-effective for them based on their existing infrastructure and capabilities, which we have not accounted for in our cost estimates. As explained in paragraph 3.66, these more cost-effective options are likely to be very materially less expensive than some of the more sophisticated approaches for which we have produced cost estimates.
- We understand that most providers have methods already in place to identify and assess risks (for example, machine-learning classifiers and risk-scoring models) using evidence from their core trust and safety systems,[95] which the provider could leverage in a cost-effective way to undertake a fraud indicator assessment, and therefore these providers may only incur incremental costs associated with adapting and building on their existing practices.

Rights assessment

Freedom of expression

3.73 As explained in 'Approach to human rights assessments' in Volume 1 Section 5 'Approach to Codes', Article 10 of the European Convention on Human Rights (ECHR) sets out the right to freedom of expression, which encompasses the right to hold opinions and to receive and impart information and ideas without unnecessary interference by a public authority. Article 10 is a qualified right and we must exercise our duties under the Act in a way that does not restrict this right unless we are satisfied that it is proportionate to the legitimate aim pursued. As noted in 'Approach to Codes', we start from the position that this proposed measure is prescribed by law, in pursuit of a legitimate aim and corresponds to a pressing social need.

3.74 We note that the proposed measure does not directly recommend service providers remove any kinds of advertisements, nor that providers take any action in relation to any kind of advertising accounts. As such, we provisionally consider that there would not be any direct interference with providers', advertisers' or users' rights to freedom of expression from this proposed measure.

3.75 However, as discussed in paragraph 3.9, the purpose of the proposed measure is to inform the actions that a service provider may take into relation to both paid-for advertisements and advertisers accounts on its service. Providers may therefore take actions based on the findings of the fraud indicator assessment. As noted in Volume 4 Section 2, 'Advertising moderation', Volume 3 Section 2, 'Advertiser checks and actions' and Volume 3 Section 5, 'Advertising bans', actions taken in accordance with the proposed measures set out therein may interfere with rights to freedom of expression, but we consider that such interference is justified and proportionate to the legitimate aims of those measures. We recognise that the way in which a provider implements the proposed fraud indicator assessment may affect

[95] For example, their moderation and reporting and complaints systems and processes. Service providers are already likely to store, collate and review relevant evidence coming from these systems and processes, to be able to identify fraudsters' activities and mitigate against these.

how actions are taken in accordance with the other proposed measures and therefore the potential interference by those measures with rights to freedom of expression of users and advertisers. This would include how a provider determines when an advertisement or account should be considered to have a material risk of being a fraudulent advertisement or an account which posts fraudulent advertisements. To the extent that there is any interference with the rights of service providers, we consider this is limited given that providers have flexibility in how they conduct the assessment and in how they implement other measures.

3.76 If a service provider were to set an unduly low threshold for when it considers there to be a material risk, this could lead to higher numbers of legitimate advertisements or accounts being linked to indicators of fraudulent advertising, and therefore more legitimate advertisements being affected by actions taken. This could interfere with the rights of users and advertisers to receive or impart information through advertising. This interference may also occur where a provider elects to conduct its fraud indicator assessment by reference to fraudulent advertising proxy and where it defines the content in relation to which users' access should be restricted more widely than is necessary to comply with the Act. However, as noted in Volume 1 Section 5, 'Approach to Codes', service providers may choose to do this as a matter of their own discretion.

3.77 In addition, we provisionally consider that in most cases using the outputs from the fraud indicator assessment will ensure that actions taken will be targeted more accurately at fraudulent advertisements and accounts posting fraudulent advertisements (as compared to an approach uninformed by the insights from this proposed measure). Further, we note that fraudulent advertisements themselves will not attract protection under Article 10 of the ECHR (see 'Approach to Codes'). As such, we provisionally consider that any potential interference with rights to freedom of expression caused, either directly or indirectly, by this proposed measure is justified and proportionate.

3.78 In relation to the proposal that the fraudulent indicator assessment be conducted every 12 months, we note that were the indicators to become out of date and no longer accurate, this could lead to greater impacts on users' and advertisers' rights from actions taken as a consequence of the outputs from the assessment. However, as noted in 'Reviewing the assessment and tracking new indicators of fraudulent advertising', we consider that setting the frequency of review at every 12 months mitigates against this risk sufficiently, while also acknowledging the costs for service providers of conducting the review.

3.79 To the extent that this proposed measure helps to reduce the frequency with which users encounter fraudulent advertising and with which they are caused harm by fraudulent advertising on the service and thereby makes users feel safer, this could also positively affect their human rights. Where a provider takes action against advertisements or advertising accounts as a result of the outputs from this measure, we consider that our measures on determining advertising appeals[96] and account appeals[97] act as a safeguard for freedom of expression.

3.80 To the extent that this proposed measure involves interference with individuals' and advertisers' rights to freedom of expression (or to the extent applicable such rights of service providers), we consider the interference to be proportionate to the Act's legitimate

[96] See Volume 4, Section 4, 'Advertising complaints'
[97] See Volume 3, Section 6, 'Account appeals'

objective of protecting individuals in the UK from fraudulent advertising (which this proposed measure is intended to help providers of Category 1 and 2A services to secure).

Data protection and privacy

3.81 As explained in Volume 1 Section 5 on our approach to these codes, Article 8 of the ECHR confers the right to respect for an individuals' private and family life. Article 8 is a qualified right, and we must exercise our duties under the Act in a way that does not restrict this right unless we are satisfied that is proportionate to the legitimate aim pursued. As noted in relation to freedom of expression, we start from the position that this proposed measure is prescribed by law, in pursuit of a legitimate aim and corresponds to a pressing social need. Article 8 underpins the data protection laws with which service providers must comply.

3.82 Collating and reviewing evidence for the purposes of this assessment may involve processing personal data where the evidence contains, or is connected to, information about an identified or identifiable individual. This may affect the rights to privacy of users and advertising account holders (where they are identifiable individuals) and their rights under data protection law. However, we consider such individuals' privacy rights will not be disproportionately affected as service providers are required to comply with relevant data protection legislation when processing personal data.

3.83 We consider that, depending on the systems and processes used by service providers, the proposed measure may involve processing personal data, potentially at scale, and also potentially processing of special category data. The UK General Data Protection Regulation (UK GDPR) places specific restrictions on making decisions based solely on automated processing of personal data, where the decision has legal or similarly significant effects. These restrictions are imposed by Articles 22A to D of the UK GDPR.[98] So-called automated decision-making is permitted where service providers have appropriate safeguards in place. Additional restrictions also apply in relation to cases where special category data is used. The Information Commissioner's Office (ICO) has provided guidance on these matters.[99]

3.84 It remains open to service providers to decide how they determine whether a specific characteristic or combination of characteristics indicates a material risk of fraudulent advertising, and what forms of personal data they consider they need to gather to make this assessment (including whether they choose as a matter of their own decision to use machine-learning processes to do so). However, service providers should ensure they, or any third parties that they outsource to, act in accordance with data protection legislation and relevant ICO guidance and consider the data protection principles of fairness, transparency and data minimisation in implementing this proposed measure.[100] Providers will also need to ensure that data protection impacts are limited to what is necessary for the legitimate purpose of complying with the fraudulent advertising duties. We consider that safeguards under data protection law, as explained in the various pieces of ICO guidance, will help ensure that the impact of processing (including automated processing) on data protection and privacy rights is minimised.

3.85 We also note the proposed recommendation that certain types of data should be retained for 12 months. However, where a service provider complies with data protection legislation

---

[98] Articles 22A to D were substituted for Article 22 by section 80(1) of the Data (Use and Access) Act 2025, with effect from 5 February 2026. See The Data (Use and Access) Act 2025 (Commencement No. 6 and Transitional and Saving Provisions) Regulations 2026, regulation 2(j), subject to regulation 5.
[99] See ICO, 2026. Automated decision-making, including profiling. [accessed 8 June 2026].
[100] See ICO, no date. UK GDPR guidance and resources. [accessed 8 June 2026].

and guidance, including in relation to the transparency of processing personal data, we consider that any additional interference by this proposed recommendation will be limited and proportionate to the benefits achieved.[101]

3.86 To the extent that these measures involve interference with individuals' rights to privacy, we consider the interference to be proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (which these proposed measures are intended to help providers of Category 1 and 2A services to secure).

Provisional conclusion

3.87 All Category 1 and 2A service providers must have in place proportionate systems and processes designed to protect users from fraudulent advertising in accordance with their duties under sections 38(1) and 39(1) of the Act. Our proposed approach is to recommend that, to enable this, all such providers assess how fraudulent advertising manifests on their service.

3.88 We provisionally consider that this measure would likely result in fewer users being exposed to fraudulent advertising than would otherwise be the case. Safety measures being tailored to the varieties of fraudulent advertising that they are seeking to mitigate will result in them being more effective. An alternative approach, that did not include a process for providers to understand the characteristics of fraudulent advertising on their service, would be significantly less effective. Providers would need to rely on poorly targeted mitigations, deploy measures that are not applicable to their service, or fail to recognise how the fraud threat is evolving. Providers who have completed this fraud indicator assessment, and track new indicators, will be substantially better placed to detect and take down fraudulent advertising or ensure that it is removed from search results than providers who have not.

3.89 We acknowledge that there could be significant costs in undertaking this proposed measure. However, our provisional view is that these costs would be proportionate given: (a) the scale of the harm fraudulent advertising causes, and (b) the important role this measure could play in working against that harm. Moreover, in most cases, the costs are likely to be lower than what we have set out in this section. This is because, in practice, we do not expect service providers to develop these systems from scratch, as we are aware that providers are already likely to have systems in place to identify and assess risks. We expect providers are more likely to incur some incremental costs from adapting their existing systems and processes to align with our proposed measure.

3.90 Overall, we consider that any impacts on freedom of expression rights and privacy rights are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising.

3.91 Our provisional view is therefore that it is proportionate to recommend that Category 1 and Category 2A service providers assess indicators of fraudulent advertising for advertisements and accounts on their service that present a material risk of fraudulent advertising.

3.92 The full text of the proposed measures can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and they are referred to as FAU B1 and FAS B1 respectively.

[101] See ICO, no date. Transparency. [accessed 5 June 2026].

4. Governance and accountability

What is this section about?

Strong governance arrangements, including clear leadership, effective oversight and accountability structures, help ensure that legal and regulatory obligations are understood, appropriately prioritised and consistently met across an organisation.

In this section, we set out our proposed governance measures, and why we are proposing to recommend them.

Our proposals

Number in our Codes	Proposed measure applicable to providers of Category 1 and 2A services
FAU A1 and FAS A1	The most senior governance body in relation to the service should carry out and record an **annual review** of the measures the provider has taken to comply with the fraudulent advertising duties**.**
FAU A2 and FAS A2	The provider should name an **individual accountable** to the most senior governance body for compliance with the fraudulent advertising duties.
FAU A3 and FAS A3	The provider should have **written statements of responsibilities** for senior managers who make decisions relating to compliance with the fraudulent advertising duties**.**
FAU A4 and FAS A4	The provider should have an **internal monitoring and assurance function** to provide independent assurance that measures taken to comply with the fraudulent advertising duties are effective on an ongoing basis.
FAU A5 and FAS A5	Have a **code of conduct** that sets out the standards and expectations for individuals working for the provider around preventing individuals from encountering fraudulent advertising.
FAU A6 and FAS A6	The provider should secure that that individuals working for the provider who are involved in the paid-for advertising function of the service are **trained in the service's approach to compliance** with the fraudulent advertising duties sufficiently to give effect to them.

Why are we proposing this?

Robust governance arrangements are likely to enable providers to more effectively implement and manage the systems and processes that protect users from fraudulent advertising on their service. Our proposed measures reflect established practice in sectors with mature governance frameworks and clear senior accountability. We expect service providers to embed principles such as accountability, effective oversight, independence, and clear standards of conduct supported by appropriate compliance training for employees. These proposed measures should support effective decision-making and ensure clear responsibility for preventing and responding to fraudulent

advertising, leading to a reduction in the amount of fraudulent advertising that can be encountered by UK users on the service.

Consultation question

- Do you agree with our proposals? Please confirm which proposed measure your views relate to and provide any arguments and supporting evidence.

Introduction

4.1 We consider effective governance and accountability structures provide a strong foundation for service providers to effectively manage their compliance obligations under the Online Safety Act 2023 (the Act), including the duties in sections 38 and 39 (which we refer to as the 'fraudulent advertising duties').[102]

4.2 By embedding principles like accountability, oversight, independence, transparency and clarity of purpose into their operations, we expect providers of Category 1 and Category 2A services to better implement and manage the systems and processes that they rely on to protect users from encountering fraudulent advertising.

4.3 Evidence from examples of high-profile organisational failures highlights the importance of effective internal controls (such as those proposed in this section) in managing and mitigating a range of risks. Root cause analysis of major corporate scandals points to weak controls as a contributing factor to organisational failures.[103] We therefore consider that implementing effective organisational controls relating to governance and accountability is essential for ensuring a service provider meets its fraudulent advertising duties to protect UK users from fraudulent advertising.

4.4 Ofcom research indicates that, while the ability to deliver against campaign objectives is the primary driver of platform choice, advertisers also consider a range of secondary factors. These include seeking to avoid associations with online services perceived to have fraudulent or low-quality advertisements in order to protect their brand from potential reputational damage. Where services are perceived to operate robust systems - such as effective controls over advertisement placements and visible signals of enforcement or oversight - advertisers are likely to regard these as reputable and safer environments in which to place advertising spend.[104]

---

[102] Under section 38(1) of the Act providers of Category 1 services must operate the service using proportionate systems and processes designed to: (a) prevent individuals from encountering content consisting of fraudulent advertisements by means of the service; (b) minimise the length of time for which any such content is present; and (c) where the provider is alerted by a person to the presence of such content, or becomes aware of it in any other way, swiftly take down such content. Under section 39(1) of the Act providers of Category 2A services must operate the service using proportionate systems and processes designed to: (a) prevent individuals from encountering content consisting of fraudulent advertisements in or via search results of the service; (b) minimise the length of time for which any such content is able to be encountered; and (c) where the provider is alerted by a person to the presence of such content, or becomes aware of it in any other way, swiftly ensure that individuals are no longer able to encounter such content in or via search results of the service.

[103] For example, for Siemens – which was subject to regulatory investigations for bribery in 2008 – the failure to embed a programme of compliance and code of conduct for staff has been cited as playing (among other factors) a "decisive role" in the scandal. Source: Primbs, M. and Wang, C., 2016. Notable Governance Failures: Enron, Siemens and Beyond. [accessed 26 April 2026].

[104] Ofcom, 2026. Online advertising pathways: qualitative research report.

4.5 Evidence from the Integrity Institute suggests that strong governance can help reduce fraudulent advertising on a service.[105] Interviewees highlighted that detecting and addressing fraud is easier where a provider clearly recognises, at an organisational level, the harm that fraud causes to the service and the importance of maintaining trust with users and advertisers. Embedding this understanding into governance and day-to-day operations can help ensure that anti-fraud activity is properly prioritised and resourced. Where this clarity is lacking, there is a greater risk that fraud is tolerated unnecessarily or that anti-fraud efforts are under-resourced.

4.6 We consider good practice in governance processes to be multifaceted. This means that there is no single governance and accountability intervention that can ensure an effective response to the scale and impact of fraudulent advertising and the range of ways it might be encountered by UK users on a service. As such, we have proposed a suite of six governance measures. These proposed measures are based on best practice in sectors that have a mature and well-established culture of identifying and managing risk and robust governance, as well as on the existing secondary literature on governance. We set out an overview of these proposed measures in the table at the start of this section and our detailed reasoning for each measure in the rest of the section including how each measure works, the expected costs to service providers of implementing such a measure and the rights implications.

Our proposals

4.7 Our proposed governance measures, taken together, provide a coherent package of governance and organisational arrangements that support effective oversight and accountability for fraudulent advertising. As a package, they are intended to help providers to prevent individuals from encountering fraudulent advertising, minimise the length of time such content is present on their services, and ensure it is removed swiftly when identified, while supporting adaptation to changes in the fraudulent advertising landscape and limiting user exposure.

Services that use intermediaries

4.8 We intend these proposed governance measures to ensure that all paid-for advertisements placed on Category 1 and 2A services are subject to governance processes, regardless of the pathway through which the advertisement appears on the service. Generally, this means that arrangements with advertising intermediaries should, where appropriate, be covered by relevant governance procedures.

4.9 As such, the proposed advertising intermediaries measure does not apply for the proposed governance measures, because we expect providers to have full control over their governance processes, even when advertising intermediaries are involved in the placement of paid-for advertisements.

4.10 We explain our rationale for each of our proposed governance measures in turn in this section, including how each measure works, the expected costs to service providers of implementing such a measure and the rights implications.

[105] Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.

Annual review of compliance activities

Explanation of the measure

4.11 Our proposed measure recommends that the service provider's most senior governance body should carry out and record an annual review of the activities it has taken to meet its fraudulent advertising duties.

4.12 The annual review should encompass:

a) the provider's assessment of how fraudulent advertising (or fraudulent advertising proxy[106]) manifests on the service;
b) the effectiveness of actions it has taken to address fraudulent advertising and meet the fraudulent advertising duties (including steps the provider has taken to implement a version of a measure (or the relevant part of it) it has been unable to apply in circumstances where it has insufficient control); and
c) lessons learned and how it intends to address these lessons.

4.13 A provider's assessment of how fraudulent advertising manifests on its service should include findings from its fraud indicator assessment (See Volume 2, Section 3, 'Fraud indicator assessment'), and any other, or alternative, steps the provider has taken to assess how fraudulent advertising or fraudulent advertising proxy is manifesting on its service.

4.14 Details on the effectiveness of actions the service provider has put in place to meet the fraudulent advertising duties could include:

a) how the provider has implemented measures we propose to recommend in the Fraudulent Advertising Codes or any alternative measures the provider has taken to comply with the fraudulent advertising duties; and
b) where a provider does not have sufficient control over relevant components to implement a proposed measure, how the provider has sought to implement a version of the measure(s) not applied (or the relevant part of it) in a way that is as similar as possible to the measure it was unable to apply (see Volume 2, Section 2, 'Advertising intermediaries').

4.15 The service provider should conduct a lessons-learned exercise in relation to the effectiveness of the measures implemented. This should include the steps the provider proposes to take to address any problems, and the expected timeframe for doing so. It should also include an explanation of how the provider has considered, and acted upon, the lessons identified through the previous review.

4.16 The annual review should therefore support the service provider's compliance with the fraudulent advertising duties, including consideration of whether any updates or adjustments are needed to ensure the steps taken continue to be appropriate and effective.

4.17 We do not specify the characteristics of the governance body in this proposed measure, because the description, size, complexity or name given to a governance body of a provider

[106] A fraudulent advertising proxy is an advertisement that a service provider has assessed against its own categories of prohibited advertisements (set out in its terms of service or publicly available statement, advertising contracts (where all of the provider's advertising contracts contain similar prohibitions in relation to fraudulent advertisements), or a combination of these when read together). The provider may do this where it is satisfied that the fraudulent advertisements that it has reason to suspect exist are prohibited by these policies or contracts. For more information, see Volume 4, Section 2, 'Advertising moderation'.

will vary. Service providers may define for themselves what they consider to be their most senior governance body. In our December 2024 Statement on Protecting People from Illegal Harms Online (our December 2024 Statement), we referred to it as the body responsible for the overall governance and strategic direction of a service, and we consider that the same applies here.[107]

4.18 We propose to recommend that the service provider should keep a written record of the annual review so that it can track progress and learn from previous reviews. We consider that such records should be kept for a minimum of three years (consistent with Ofcom's Record-Keeping and Review Guidance), or in accordance with the organisation's record retention policies, if longer.[108]

4.19 We note that other proposed measures include additional recommendations to undertake 'reviews' where appropriate for ensuring their effective implementation. This could include, for example, the proposed account checks measure. While providers have flexibility to combine the annual review with other assessments if they wish, the other reviews and assessments are complementary to, rather than a substitute for, the annual review.

Benefits and effectiveness

4.20 The online advertising space is dynamic and adversarial, with harms evolving rapidly. Fraudulent actors frequently change tactics and methods in response to developments in technology, society and user behaviours, and to evade fraud prevention measures as set out in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'. We consider therefore that a service provider should undertake at least annually a review of the activities it has taken to meet its fraudulent advertising duties to ensure the measures it takes to manage fraudulent advertising remain effective.

4.21 Regular review of regulatory compliance by a governance body is standard practice for many organisations and is required for appropriate oversight over internal controls and practices. Evidence supporting this can be found in good governance practice principles and codes.[109]

4.22 Evidence we reviewed in developing our Illegal Content Codes of Practice and Protection of Children Codes of Practice concludes that where senior governance bodies regularly review risk management activities, safety outcomes will improve.[110] We consider that effective risk

---

[107] Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.69. [accessed 28 April 2026].
[108] Ofcom, 2025. Record-Keeping and Review Guidance. [accessed 24 February 2026].
[109] Under the UK Corporate Code, companies with a premium listing on the London Stock Exchange are already required to follow principles related to board oversight. This includes provision 29, which states that boards "should monitor the company's risk management and internal control framework and, at least annually, carry out a review of its effectiveness". Monitoring and review activities are intended to cover all material controls including financial, operational and compliance controls. Source: Financial Reporting Council, 2024. UK Corporate Governance Code, p.13. [accessed 28 April 2026]; The Principles of Corporate Governance of the Organisation for Economic Co-operation and Development (OECD) similarly suggest that a main function of the boards should be "reviewing and guiding corporate strategy, major plans of action [and] risk management policies and procedures". The Principles suggest that while committees or other sub-bodies may have specific responsibilities of different areas of risk, "the board should retain final responsibility for oversight of the company's risk management system and for ensuring the integrity of the reporting systems". Source: OECD, 2015. G20/OECD Principles of Corporate Governance. [accessed 28 April 2026].
[110] Ofcom, 2024. December 2024 Statement, Volume 1, pp.69 and 70; Ofcom, 2025. April 2025 Statement on Protecting Children from Harms Online, Volume 4: What should services do to mitigate the risks of online harms to children?, p.76.

management supports compliance with the fraudulent advertising duties by helping providers to identify and assess the likelihood of fraudulent advertising arising on their services. By understanding the nature and severity of potential harms, providers are better placed to design proportionate systems and processes to prevent individuals from encountering fraudulent advertising, minimise the length of time such content is present, and ensure its swift removal where identified. Risk management processes also support ongoing review and adjustment of controls as the fraudulent advertising landscape evolves.

4.23 We also consider that an annual review is a reasonable timescale on which to conduct this activity. This would offer service providers the flexibility to tie the fraudulent advertising review into a service provider's overall annual risk management and internal controls frameworks, such as financial reporting. It also aligns with existing online safety governance processes, that is, the illegal content and protection of children annual review measures.[111]

4.24 Keeping a written record of the annual review supports good governance by creating an ongoing account of how the most senior governance body has considered compliance with the fraudulent advertising duties and assessed the effectiveness of measures over time.

4.25 Records of annual reviews can also support organisational learning, continuity and strategic decision-making. By keeping a written record, providers can track whether measures are working and ensure that past findings continue to inform future decisions. This helps avoid reliance on individual memory, supports continuity through organisational or personnel change, and enables the provider to respond more effectively as fraud risks, technologies, best practice and business models evolve. Taken together, this supports a more resilient and well-governed approach to addressing fraudulent advertising as practices and conditions evolve over time.

4.26 An annual review will help service providers to ensure ongoing effectiveness of the actions it has taken to meet their fraudulent advertising duties. We expect service providers to implement any changes required that are identified by the annual review and will consider this when determining whether a provider has properly complied with the fraudulent advertising duties.

Impacts and costs on service providers

4.27 We expect that undertaking and recording an annual review will involve staff time and costs. Firstly, service providers would need to dedicate staff time towards pooling together and reviewing the findings and outcomes of activities the provider has taken to identify and address the likelihood of fraudulent advertising occurring, including the implementation and operation of proposed measures set out elsewhere in the draft Fraudulent Advertising Codes (such as the fraud indicator assessment, testing, proactive technology, advertiser checks, and any relevant governance, training and assurance measures).

4.28 Following this, we expect relevant staff time would be needed to prepare a report detailing this for the senior governance body. The governance body would then need time to read, consider and discuss the report.

4.29 We expect most Category 1 and 2A service providers are already likely to have governance bodies in place for the overall management of their businesses. Therefore, we assume service providers will not incur any additional costs in establishing a governance board for

[111] ICU A1 and ICS A1 in the Illegal content Codes of practice and PCU A1 and PCS A1 in the Protection of children Codes of practice.

this proposed measure. We consider potential additional ongoing costs that may be associated with undertaking the review. We estimate the additional costs of relevant staff preparing and the main board reviewing and scrutinising an annual risk management report could be in the region of £16,000 to £37,000 per year.[112] We note that these costs may be higher for service providers with larger and more highly paid governance bodies.

4.30 For service providers in scope of the equivalent measures in the Illegal Content Codes and Protection of Children Codes,[113] we expect they may already have experience in preparing reports of a similar nature, or that there could even be synergies in the processes involved, which could help to reduce costs.

Rights assessment

4.31 We are proposing a suite of six governance measures, including this measure on annual review, which we have designed to operate together to ensure an effective response to fraudulent advertising. We have therefore assessed the human rights impacts for these measures together and the assessment set out in paragraphs 4.32 to 4.41 applies equally to all proposed governance measures in this section.

Freedom of expression

4.32 As explained in Volume 1, Section 5, 'Approach to codes', Article 10 of the European Convention on Human Rights (ECHR) upholds the right to freedom of expression, which encompasses the right to hold opinions and to receive and impart information and ideas without unnecessary interference by a public authority. Article 10 is a qualified right and we must exercise our duties under the Act in a way that does not restrict this right unless we are satisfied that is proportionate to the legitimate aim pursued. As noted in 'Approach to codes', we start from the position that these proposed measures are prescribed by law, in pursuit of a legitimate aim and correspond to a pressing social need.

4.33 The aim of these proposed measures is to ensure governance and organisational arrangements are put in place that support effective oversight and accountability for fraudulent advertising, thereby helping providers to prevent individuals from encountering fraudulent advertising, minimise the length of time such content is present on their services, and ensure it is removed swiftly when identified. We acknowledge that a service provider may choose to take particular actions that have a more direct impact on the availability of paid-for advertisements or on advertisers, because of the activities associated with these proposed measures. However, these proposed measures do not themselves require any steps to be taken with respect to particular kinds of advertisement or types of advertising

[112] In Annex 8, 'Further detail on economic assumptions and analysis', we assume it takes 10 to 20 days for a professional occupation staff member to prepare the paper for the board and that on average each director on the board spends one to two hours in total to read, consider and discuss the report. We assume on average directors spend 250 hours a year on board-related activities for each company they are a director for. Source: PwC, 2022. PwC's 2022 Annual Corporate Directors Survey. [accessed 27 March 2026]; For total remuneration per board member, we assume $336,352 per year, based on the average for 2025 of S&P 500 independent board directors (based on a report by Spencer Stuart, a leadership consultancy). Source: Spencer Stuart, 2025. 2025 U.S. Spencer Stuart Board Index. [accessed 15 May 2026]; For the number of board members, we assume boards have on average 11 members, based on the average S&P 500 board size. Source: Harvard Law School Forum on Corporate Governance (Spierings, S.), 2022. Diversity, Experience, and Effectiveness in Board Composition. [accessed 27 March 2026].

[113] Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.71. [accessed 28 April 2026].

accounts, nor do they have any impact on providers except in connection with paid-for advertising that is available to UK users.

4.34 We provisionally consider that these proposed measures would not constitute an interference with individuals', providers', or advertisers' freedom of expression rights.

4.35 To the extent that it helps to reduce harm from fraudulent advertising and make users feel safer using a service, it could also positively affect their human rights and act as a safeguard for such rights.

Data protection and privacy

4.36 As explained in Volume 1, Section 5, 'Approach to codes' Article 8 of the ECHR confers the right to respect for an individuals' private and family life. Article 8 is a qualified right, and we must exercise our duties under the Act in a way that does not restrict this right unless we are satisfied that it is proportionate to the legitimate aim pursued. As noted in relation to freedom of expression, we start from the position that these proposed measures are prescribed by law, in pursuit of a legitimate aim and correspond to a pressing social need. Article 8 underpins the data protection laws with which service providers must comply.

4.37 Service providers will have to evaluate information and review trends to implement these proposed measures. Where a provider elects to collect any metrics or is already collecting metrics, or otherwise process information about identified or identifiable individuals, it will need to ensure that it processes any personal data in accordance with the relevant data protection legislation. Providers should refer to relevant guidance from the Information Commissioner's Office (ICO).[114]

4.38 We recognise that individuals whose data is processed may be located in jurisdictions which do not have data protection laws, and to which UK data protection laws do not apply. However, these proposed measures do not recommend that any particular records be held in those jurisdictions nor that any processing more generally take place in those locations. We provisionally consider that these measures can be implemented in accordance with data protection law. We consider that safeguards under data protection law, as explained in ICO guidance, will help ensure that the impact of processing on data protection and privacy rights is minimised.

4.39 These proposed measures set out that certain individuals or senior managers be named in a way that means the service provider holds a record of the information. It can be assumed that a service provider would therefore have to process the personal data of such individuals (such as their name in connection with their role and responsibilities). However, we expect providers would have already collected and processed personal data about these individuals in some form (for example, through employment contracts, role descriptions and performance management). Moreover, these proposed measures do not require the service provider to put the identity of such individuals into the public domain nor disclose it to any specific third parties. In addition, any record relating to the identity of such individuals would relate to their professional role and responsibilities, rather than revealing anything particularly private or sensitive about that person. Therefore, while we recognise that these measures may mean that the service provider records additional information about, or relating to, the identity of the certain individuals whose data they may be otherwise have been recording or processing, we consider that any impact on their privacy would be very

[114] ICO UK GDPR guidance and resources. [accessed 29 April 2026].

limited. We therefore consider that the impact of this measure on those individuals is limited and proportionate to the legitimate aims of the fraudulent advertising duties in the Act.

4.40 We also consider that a well-managed business is, in general, more likely to comply with its obligations under privacy and data protection laws (as well as, for that matter, other laws such as those relating to consumer protection and equality). As such, our proposed measures may help to safeguard these.

4.41 To the extent that these measures involve any interference with individuals' rights to privacy, we consider the interference to be proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (which these proposed measures are intended to help providers of Category 1 and 2A services to secure).

Provisional conclusion

4.42 We consider that it is important for service providers to regularly monitor the systems and processes they have in place to identify and address the impacts of fraudulent advertisements on their service, given the ongoing evolution of the space.

4.43 We also consider that it is important that the most senior governance bodies for Category 1 and 2A services have a full understanding of how fraudulent advertisements may occur on their service and ensure appropriate actions are taken to address these advertisements, in compliance with the fraudulent advertising duties.

4.44 We note that regular review of regulatory compliance by a governance body is standard practice for many organisations and that many or all providers of services within scope of the fraudulent advertising duties should be conducting a similar review to meet their illegal harms and protection of children duties.[115] We therefore provisionally consider that an annual review is a reasonable cadence for the assessment.

4.45 While there are costs associated with implementing this proposed measure, we consider that these are proportionate and should be manageable for providers of Category 1 and 2A services.

4.46 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this proposed measure (if any) are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.47 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services conduct an annual review of how they have assessed how fraudulent advertising can manifest on their services and the steps they have taken to meet their fraudulent advertising duties.

4.48 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU A1 and FAS A1 respectively.

[115] In particular, sections 10(2), (3) and (5) to (9), section 12(2), (3) and (9) to (14), section 20(2), and section 21(2) and (3) of the Act for user-to-user services; and section 27(2), (3) and (5) to (9), section 29(2), (3) and (5) to (9), section 31(2), and section 32(2) and (3) of the Act for search services.

Senior accountability

Explanation of the measure

4.49 The proposed measure recommends that service providers name an accountable individual with responsibility for decision-making on compliance with the fraudulent advertising duties. That individual should be accountable to the most senior governance body.

4.50 Given suspected fraudulent advertisements may be placed on a service through the open-display market, the responsible individual may also need to have knowledge of which intermediaries the provider has arrangements with and how the provider works with them, to ensure that all reasonable endeavours are taken to achieve the outcome(s) of Codes measures not applied due to insufficient control, in accordance with the proposed intermediaries measure.

4.51 The individual should be accountable for the oversight of all activities relating to compliance with the fraudulent advertising duties. The role may also encompass responsibility for the effective application of the governance measures, such as annual review processes, compliance training and codes of conduct.

4.52 The accountable individual could be responsible for overseeing the implementation and operation of the proposed measures in practice. This could include oversight of the allocation of responsibilities through written statements of responsibilities; ensuring that a code of conduct and relevant compliance training programmes are in place and operating effectively; and engagement with the independent internal monitoring and assurance function as part of the annual review, including oversight of how lessons learned are implemented.

4.53 Regarding the individual, as is the case for our illegal harms and protection of children duties, we do not set out any specific qualifications that they should hold. We consider that it is for the service provider to determine who this should be. The provider will not be required to publish the name of the individual or routinely notify Ofcom of the name of the individual. We do not propose to recommend that the individual needs to be based in the UK.

4.54 We do consider that only one person should be named. We view service providers as best placed to decide how this role would work most effectively within their structures. For example, providers may decide to have multiple risk owners and subject matter experts responsible for risk management controls reporting to the accountable individual.

4.55 This proposed measure is not associated with senior manager liability for compliance with information notices issued by Ofcom (unless the service provider wishes to give these roles to the same individual),[116] nor the duty to comply with requirements imposed in confirmation decisions.[117]

[116] See section 103 of the Act. Ofcom has powers to issue information notices under sections 100 to 102 of the Act, with section 103 allowing Ofcom to require a named senior manager to be responsible for compliance with an information notice. Information notices are sent directly to service providers and will not be published by Ofcom. For further information on this, see Ofcom's Online Safety Information Powers Guidance, in particular paragraphs 4.95 to 4.102.

[117] Section 139 of the Act. For further information on this, see Ofcom's Online Safety Enforcement Guidance, in particular paragraphs 6.54 to 6.56.

Benefits and effectiveness

4.56 This proposed measure is intended to support compliance with the fraudulent advertising duties under the Act by ensuring there is a clearly identifiable senior individual accountable for overseeing, and explaining and justifying, how the service provider's systems and processes operate to prevent individuals in the UK from encountering fraudulent advertising and to address such content swiftly where it arises.

4.57 We consider that having accountability at the senior management level for compliance with the fraudulent advertising duties and with the measures that the service provider puts in place means that users will be less likely to encounter fraudulent advertisements.

4.58 Evidence from the Integrity Institute shows that responsibility for reducing fraudulent advertising is distributed across teams, with approaches differing by service. Interviewees agreed "that a clear structure for accountability is important, as situations where accountability is ambiguous can allow problems to slip through the organizational cracks."[118]

4.59 The effectiveness of senior accountability has been proven in other regulatory regimes, such as the Senior Managers & Certification Regime (SM&CR) jointly regulated by the Financial Conduct Authority (FCA) and the Prudential Regulation Authority (PRA).[119] A 2020 PRA review noted positive behavioural change and improvement in risk management practices following implementation of the SM&CR.[120] This is corroborated by findings from analysis of senior leadership failures in financial services in relation to the 2008 financial crisis, which demonstrate how a lack of senior accountability can result in reduced oversight and excessive risk-taking.[121]

4.60 In our December 2024 Statement, we pointed to work commissioned by Ofcom from Milliman which puts individual accountability as the first principle of good governance, drawing on the Institute of Internal Auditors Three Lines Model.[122] Having multiple persons in the accountability role would dilute the effectiveness of this proposed measure. Flexibility would remain for service providers to decide how they delegate and deliver responsibilities, while ensuring there is a clearly identifiable senior individual with oversight of how the proposed governance measures operate in practice.

4.61 We also consider that this proposed measure will ensure that service providers give clarity on roles and responsibilities for managing activities in relation to addressing compliance with

---

[118] Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.

[119] The SM&CR is directly underpinned by legislation and serves different outcomes related to compliance with financial regulation, but we consider the broad lessons and findings from the FCA's implementation of the regime as instructive for other areas of risk management and regulatory compliance. Source: FCA, 2023. Senior Managers and Certification Regime. [accessed 22 June 2026].

[120] This review gathered evidence from 2019-2020 among those who had implemented SM&CR. Source: Bank of England, Prudential Regulation Authority, 2020. Evaluation of the Senior Managers and Certification Regime. [accessed 22 June 2026].

[121] In the aftermath of the 2008 financial crisis, an inquiry into professional standards and culture of the banking sector by the Parliamentary Commission on Banking Standards concluded that many bankers had been allowed to operate with little accountability, and claimed ignorance or hid behind collective decision-making. Source: Parliamentary Commission on Banking Standards, 2013. Changing banking for good. [accessed 25 June 2026].

[122] Milliman, 2021. Report on principles-based best practices for online safety Governance and Risk Management; Institute of Internal Auditors (IIA), 2020. The IIA's Three Lines Model. [accessed 22 June 2026].

the fraudulent advertising duties – leading to a more consistent approach to implementing safety measures.

4.62 Where there is no clear ownership within a service, systems and processes to address fraudulent advertising may not be implemented or operated effectively.[123] The proposed measure is intended to strengthen oversight and scrutiny in support of the duties to prevent individuals from encountering fraudulent advertising, minimise the length of time such content is present, and ensure its swift removal when identified.

4.63 We consider that this would also be the case for the management of activities taken to comply with the fraudulent advertising duties.

Impacts and costs on service providers

4.64 We expect there will be some costs associated with selecting and subsequently training an individual accountable for compliance with the service provider's fraudulent advertising duties. We anticipate that most service providers will choose to add accountability for compliance to the portfolio of an existing senior manager or director who already oversees an online safety compliance or risk function, or who has been appointed to meet equivalent measures under the Illegal Content Codes and Protection of Children Codes. Any providers that have not already appointed a suitable individual will incur greater costs in following this proposed measure as they would need to make changes to their internal structure. However, to comply with the Act, service providers would already need someone in a suitably senior role who understands the service provider's legal duties.

4.65 We estimate the cost of identifying and training the relevant accountable individual could be less than £2,000.[124] We expect the actual cost to vary depending on the complexity of the organisation, and the regulatory requirements the individual will be accountable for.

4.66 Additionally, there will be incremental ongoing costs associated with the accountable individual overseeing compliance with the fraudulent advertising duties. For example, if a senior leader was to spend 10 additional days each year overseeing compliance with the fraudulent advertising duties, we estimate that, while costs will vary, the costs would be in the region of £9,000 per year.[125]

---

[123] See Bank of England, Prudential Regulation Authority, December 2020. Evaluation of the Senior Managers and Certification Regime [accessed 1 June 2026]; These findings were corroborated by our commissioned research on best practice, including a report by Milliman which highlighted individual accountability as the first principle of good governance. Source: Milliman, 2021. Report on principles-based best practices for online safety Governance and Risk Management. [accessed 26 June 2026]; The ICO's guidance about artificial intelligence (AI) risk management regarding data protection states that senior management staff are accountable for addressing the technical complexities of AI, and cannot delegate this responsibility to others. It states that senior management will need to align its internal structures, roles and responsibilities maps, training requirements, policies and incentives to its overall AI governance and risk management strategy. Source: ICO, no date. What are the accountability and governance implications of AI? [accessed 1 June 2026].

[124] This is consistent with our cost assumptions and estimates for the equivalent measure in Illegal Content Codes and Protection of Children Codes. Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.83. [accessed 28 April 2026]. We assume training the accountable person takes two days' time for two staff in professional occupations.

[125] This is consistent with our cost assumptions for the equivalent measure in Illegal Content Codes and Protection of Children Codes. Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.83. [accessed 28 April 2026]. See our labour wage assumptions for a senior leader in Annex 8, 'Further detail on economic assumptions and analysis'.

Provisional conclusion

4.67 We propose to recommend that all providers of Category 1 and 2A services name a single individual who is accountable to the most senior governance body for compliance with the fraudulent advertising duties. Being accountable means that the individual is required to explain and justify to that body the actions and decisions taken in relation to how the provider meets those duties, including oversight of the measures put in place to prevent individuals from encountering fraudulent advertising, to minimise the time such content is present, and to ensure it is addressed swiftly where it arises.

4.68 This accountability is intended to support effective oversight of compliance, while allowing service providers flexibility to determine how responsibilities are delegated and delivered within their organisation.

4.69 We consider that the evidence suggests that clearly defining senior accountability supports effective compliance with the fraudulent advertising duties by providing clear ownership, strengthening oversight and reducing the risk of responsibilities being fragmented across teams. This aligns with evidence from other regulatory regimes and research indicating that clear accountability structures improve governance outcomes and support more effective action to reduce the fraudulent advertising that users encounter.

4.70 Given that the costs of the proposed measure are relatively low and the evidence set out suggests it would deliver significant benefits, we consider that any such costs would be proportionate.

4.71 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this proposed measure are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.72 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services have a person accountable to the most senior governance body for compliance with their fraudulent advertising duties.

4.73 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU A2 and FAS A2 respectively.

Written statements of responsibilities

Explanation of the measure

4.74 We propose to recommend that service providers should have written statements of responsibilities for senior managers who make decisions relating to compliance with the fraudulent advertising duties.

4.75 Our proposed recommendation is that the provider should produce statements that clearly set out the responsibilities of senior managers. These should set out what each senior manager is responsible for in relation to the fraudulent advertising duties, and explain how those responsibilities fit within the service's overall governance and management arrangements. These documents should include each senior manager's main responsibilities for decision-making in relation to compliance with the fraudulent advertising duties.

4.76 Depending on the nature and operation of the service, responsibilities set out in such statements may relate to a range of activities relevant to compliance with the fraudulent

advertising duties. This includes systems and policies for the detection and moderation of paid-for advertising, handling of paid-for advertising reports and complaints, advertiser checks and sanctions, or engagement with advertising intermediaries.

4.77 In practice, service providers may already assign responsibility for these activities to existing teams or managers, with written statements of responsibility serving to clarify how those roles support the prevention and mitigation of fraudulent advertising within the service's systems and processes.

4.78 We do not set out which managers should be included within written statements of responsibility, as we consider it important that this measure retains flexibility to be applied across the diverse range of services in scope.

4.79 We do not propose that service providers publish the written statement of responsibilities (unless they wish to do so), nor do we propose that service providers routinely notify us of their statement.

Benefits and effectiveness

4.80 This proposed measure is intended to ensure that users are prevented from encountering fraudulent advertising by providing clarity on roles and responsibilities of those who work on compliance with the fraudulent advertising duties. We consider a lack of clarity could contribute to inconsistent application of measures implemented to protect users from fraudulent advertising and impede providers' efforts to comply with the fraudulent advertising duties.

4.81 Evidence from the Integrity Institute notes that "research, policy development, measurement, and monitoring are all steps of any significant effort to curb violating impressions universally, and so all of these can be brought to bear on fraud."[126] This highlights that tackling fraud is not confined to a single function within a service, but instead relies on a range of distinct organisational activities, each with a role to play in prevention, detection and mitigation. The evidence therefore supports an expectation that responsibility for addressing fraud will be distributed across different teams and functions within a service (which may include but is not limited to policy, operational delivery, monitoring and evaluation), and that these responsibilities should be clearly articulated and linked back to the service's approach to managing fraud risks.

4.82 Further evidence suggests that providing clarity in how responsibilities are owned within an organisation leads to safer design of services, and more effective safety mitigations, delivering material benefits. For example, findings from a 2020 review of the FCA's SM&CR reported that many firms surveyed said the requirements of the regime had resulted in clearer articulation of authority and had improved focus on accountability and responsibility. In a 2014 cost–benefit analysis of the SM&CR, the large banks that were surveyed anticipated that statements of responsibility would positively affect behaviour around decision-making and risk.[127]

---

[126] Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.

[127] "Large banks and investment firms did consider it likely that the policies would result in behavioural changes as senior managers sought to ensure they would be protected in the event that misconduct or a regulatory breach was discovered, driven by the statement of responsibilities and the presumption of senior responsibility. Such behaviour includes increased due diligence, monitoring and sign-off processes, as well as

4.83 Based on this research, we consider that clearly defining senior managers' responsibilities will support effective compliance with the fraudulent advertising duties by ensuring that relevant systems and processes are applied consistently and overseen effectively. We also expect this proposed measure to help embed organisational awareness of the fraudulent advertising duties and reduce the chance that unclear or overlapping responsibilities undermine efforts to prevent users from encountering fraudulent advertising.

4.84 Written statements of responsibility can also support service providers in ensuring that responsibility for these activities is clearly allocated to named managers or teams, while allowing service providers to determine how best to organise and deliver these functions within their own structures.

Impacts and costs on service providers

4.85 Service providers would incur one-off and ongoing costs in developing these statements of responsibilities. This would consist of staff time and costs associated with agreeing and discussing responsibilities within teams, as well as producing and keeping up to date the written statements of responsibilities.

4.86 We expect that the costs of agreeing and discussing responsibilities will vary by services according to the number and complexity of fraudulent advertising activities they need to manage. We estimate that producing written statements of responsibilities for about 10 senior managers would result in an upper-bound initial one-off cost in the region of £18,000-£19,000.[128] We consider our assumption on the number of senior managers to be broadly representative of the resourcing that service providers may need. However, we are aware that a very large service with complex management could require more senior managers and therefore incur higher costs.

4.87 We also expect that service providers will incur ongoing costs in maintaining and keeping these statements up to date. In line with our standard assumptions for ongoing costs,[129] we estimate that this would represent 25% of the initial costs and so could cost providers approximately £5,000 per year.

4.88 Overall, we note that our proposed measure is not specific on the level of detail that should be included in these statements. This added flexibility is designed to help manage the likely cost of the proposed measure.

Provisional conclusion

4.89 We consider that users will be better protected from encountering fraudulent advertisements if service providers have adequate governance practices in place to manage

---

more formalised and considered decision-making. These actions are all likely to contribute to an increased likelihood that potential and actual regulatory breaches are identified and prevented." Source: Europe Economics, 2014. Cost Benefit Analysis of the New Regime for Individual Accountability and Renumeration. [accessed 4 February 2026].

[128] This is consistent with our cost assumptions and estimates for the equivalent measure in our Illegal Content Codes and Protection of Children Codes. Source: December 2024 Statement, Volume 1, paragraph 5.124. [accessed 15 May 2026]; We assume that on average it would take three days to develop and agree each statement, and that the time would mostly be of senior managers. We assumed an annual salary of £103,000 for senior managers of a large service and used the non-wage uplift assumption. See our labour wage assumptions for a senior manager in Annex 8, 'Further detail on economic assumptions and analysis'.

[129] See Annex A8. 'Further detail on economic assumptions' for more details.

and monitor the activities that they are undertaking to prevent fraudulent advertising on their services.

4.90 Our analysis suggests that there are considerable benefits from improved outcomes from having statements of responsibilities for members of staff who perform functions relevant to addressing fraudulent advertising.

4.91 Given that the costs of the proposed measure are likely to be low and the evidence referenced in paragraphs 4.80 to 4.84 suggests that a written statement of responsibilities would deliver significant benefits, we consider that any such costs would be proportionate.

4.92 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this measure (if any) are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.93 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services have written statements of responsibility for senior managers in relation to fraudulent advertising.

4.94 The full text of the proposed measures can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and they are referred to as FAU A3 and FAS A3 respectively.

Internal assurance and compliance

Explanation of the measure

4.95 We propose to recommend that service providers have an independent function to provide assurance that the measures in place to comply with the fraudulent advertising duties are operating effectively.

4.96 The function should be independent of the activities it assesses. This can be achieved through direct accountability to the overall governance body and access to the people, resources and data necessary to carry out its work. It should operate free from bias and interference and be able to report objective findings on the effectiveness of controls.[130]

4.97 The function should report to the overall governance body or an audit committee, which should be responsible for considering its findings and providing appropriate oversight. Providers should ensure that their assurance arrangements enable issues identified through this activity to be escalated and considered, and that decisions can be taken on whether further action is required.

4.98 Independent assurance may be delivered through an existing internal audit function or equivalent arrangements. Where a dedicated function is not feasible, providers may structure responsibilities to ensure that oversight and challenge are carried out by individuals who are not directly involved in the activities being assessed.

4.99 We do not consider that independence would require service providers to engage an independent third party (such as an external auditor) to confirm effectiveness of mitigations, although providers may choose to do so.

[130] IIA, 2020. The IIA's Three Lines Model. [accessed 28 April 2026].

4.100 The independent assurance function is intended to support the annual review by providing the overall governance body with objective assurance on the effectiveness of the measures in place to meet the fraudulent advertising duties.

Benefits and effectiveness

4.101 The overall objective for independence of the monitoring and compliance function is to ensure that service providers find a way to achieve as much independent oversight and challenge as possible for each main task.

4.102 We consider that having an internal assurance function is important to support effective compliance with the fraudulent advertising duties. Such a function can help service providers evaluate whether the systems and processes they have put in place are operating effectively to protect users in accordance with these duties.

4.103 Where providers do not have processes to independently review and test the effectiveness of these measures, any failures in compliance activities may go undetected, increasing the likelihood that fraudulent advertising persists on the service. Putting in place structured assurance processes enables providers to scrutinise how controls operate in practice and to identify where improvements are needed.

4.104 We also consider that the effectiveness of an assurance function is strengthened where it operates independently from the activities it assesses. Independence supports objective, authoritative and credible challenge, and helps ensure that judgments about compliance with the fraudulent advertising duties are robust and not unduly influenced by operational or commercial considerations.

4.105 In our December 2024 Statement on the Illegal Content Codes, we noted the benefits of independent oversight of internal controls in supporting effective governance.[131] We consider that the same principles apply in the fraudulent advertising context, where independent assurance can support reliable assessment of whether governance arrangements and controls are sufficient to meet the fraudulent advertising duties.

4.106 Evidence from wider corporate governance practice, including consultations by the (former) Department for Business, Energy and Industrial Strategy and the European Commission, indicates that strengthening internal controls and independent oversight improves the management of compliance risk and the quality of organisational decision-making.[132]

4.107 Analysis of serious corporate governance failures similarly highlights the role of independent assurance and clear reporting lines to senior governance bodies in identifying and addressing weaknesses in internal controls. These findings highlight the role of independent assurance

[131] Ofcom, 2024. December 2024 Statement, Volume 1, pp.96 and 97.

[132] The European Commission's consultation on corporate reporting found overall support from respondents in favour of ensuring effective internal controls to improve the effectiveness and efficiency of corporate governance mechanisms. Notable responses to this consultation included comments from professional services organisations, which pointed to evidence that establishing and embedding a system for monitoring and reporting of internal controls improves the quality of financial reporting (PwC) and reduces the risk of corporate failure and fraud (Deloitte). Source: European Commission, 2022. Corporate reporting – improving its quality and enforcement. [accessed 5 February 2026]; We also found support for stronger internal control frameworks reflected in response to a 2022 Department for Business, Energy and Industrial Strategy (BEIS) consultation, which cited improved reporting and audit and better corporate governance as important outcomes. Source: Department for BEIS, 2022. Restoring trust in audit and corporate governance. [accessed 5 February 2026].

in supporting sustained compliance with the fraudulent advertising duties and helping to prevent systemic failures from undermining their intended outcomes.[133] Findings from independent assurance activity can inform the governance body's assessment of how fraudulent advertising may arise on the service, the effectiveness of actions taken to address it, and any lessons learned. This helps ensure that the annual review is informed by appropriate challenge and evidence, and that it supports effective oversight and continuous improvement over time.

Impacts and costs on service providers

4.108 Establishing and operating an independent function to provide assurance that the measures a provider has taken to comply with the fraudulent advertising duties are effective could involve substantial costs.

4.109 We do not propose to specify the frequency or timing with which a service provider undertakes its independent assurance activities, as this should form part of the provider's own strategy for managing fraudulent advertising. However, for the purposes of estimating the costs associated with this proposed measure, we calculate these on an annual basis.

4.110 To estimate the respective costs for setting up an independent assurance function from scratch, we used the number of staff employed in internal audit functions as a reference to inform our views on how many staff members service providers might need for their internal monitoring and assurance function. These functions evaluate an organisation's internal controls, especially its corporate governance and accounting processes. They may also oversee the organisation's risk management processes and may focus on a specific area of a business, such as cybersecurity. This approach is consistent with our proposal that this independent assurance can be provided by an existing internal audit function.

4.111 The size of the internal monitoring and assurance function needed will vary by service. If a provider of a larger, riskier, more complex service requires ten additional people to fulfil this function, we estimate the costs to be approximately between £550,000 to £1,100,000 per year. This is because costs are likely to be higher for providers of larger services that face more risks, and we would expect the costs to increase with the potential benefits to some extent.

4.112 A provider that has a limited range of fraudulent advertising risks and measures may only require a single person to fulfil its monitoring and assurance function, in which case the estimated annual costs would be around £55,000 to £110,000. Because costs are likely to be higher for providers of larger services that face more risks, we would expect the costs to increase with the potential benefits to some extent. We expect that the costs of this proposed measure for providers of smaller services would tend to represent a higher proportion of their annual revenue.

4.113 We are aware that some service providers have internal assurance processes already in place to deal with fraudulent advertising, or providers within scope of our Illegal Content Codes and Protection of Children Codes may have similar processes in place to meet these proposed recommendations. This means that there may be cost synergies and as a result,

[133] Krahnen, P.K., Langenbucher, K., Leuz, C. and Pelizzon, L., 2020. Wirecard Scandal: When All Lines of Defense Against Corporate Fraud Fail, Oxford Business Law Blog, 23 November. [accessed 25 February 2026]; European Parliament ECON committee (Langenbucher, K., Leuz, C., Krahnen, P.K. and Pelizzon, L.), 2020. What are the wider supervisory implications of the Wirecard case. [accessed 25 February 2026].

the cost estimates set out in paragraphs 4.111 to 4.112 are likely to significantly overstate the additional costs that many such providers would incur in practice.

Provisional conclusion

4.114 Our analysis has shown that there are benefits in ensuring that organisations have independent oversight over internal controls to ensure that governance is effective.

4.115 We have identified ongoing costs associated with these proposed measures. However, we consider that this proposed measure is sufficiently integral to good risk management and governance that any such costs would be proportionate to the benefits for users.

4.116 By establishing an independent function that reports to the overall governance body, senior decision-makers will receive independent assessment of how effectively measures are operating to mitigate fraudulent advertising on their service. This will provide clearer visibility of issues and any gaps in existing mitigations and controls, enabling more informed decision-making on where improvements are needed. In turn, this will support providers in taking appropriate steps to strengthen their approach and meet the fraudulent advertising duties.

4.117 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this proposed measure (if any) are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.118 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services have an independent function to provide assurance that the measures the provider has taken to comply with the fraudulent advertising duties are effective.

4.119 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU A4 and FAS A4 respectively.

Code of conduct

Explanation of the measure

4.120 We propose to recommend that service providers should have a code of conduct that sets out the standards and expectations for individuals working for the provider around protecting users from fraudulent advertising in accordance with the fraudulent advertising duties.

4.121 The proposed measure is intended to ensure that individuals with a role that involves the service's advertising function understand the provider's approach to dealing with fraudulent advertising, embedding the approach into the operation of the organisation.

4.122 This proposed measure is of narrower application than the equivalent measures in the Illegal Content Codes and Protection of Children Codes.[134] Those measures state that a code of conduct should apply to all individuals working for the provider. In the context of fraudulent advertising, we consider it appropriate for the code of conduct to apply only to individuals whose role involves the online advertising function. This reflects the fact that, while most

[134] Illegal Content Codes measure ICU A6 and ICS A6; Protection of Children Codes measure PCU A6 and PCS A6.

individuals working for a provider of a regulated service are likely to have roles connected to the operation of the service and would therefore fall within scope of the Illegal Content Codes and Protection of Children Codes, not all individuals will have responsibilities relating to online advertising.

4.123 We propose that the code of conduct should be specific to the service. It could include recognition of the scale of impact of fraudulent advertising on UK users, a clear organisational statement around protecting users and expectations and guidelines for reporting instances of concern relating to fraudulent advertising on the service.

4.124 The service provider should ensure that the code of conduct is kept up to date and that the provider puts in place procedures to make individuals aware of the code of conduct, for example, as part of the onboarding process.

4.125 We also consider that it should be simple, concise and easy to understand and consistent with other policies and communications. We consider that a good code of conduct implemented in accordance with this proposed measure would also be reviewed by multi-disciplinary teams.

Benefits and effectiveness

4.126 A code of conduct can be an effective mechanism for service providers to set clear expectations regarding behaviour and responsibilities for individuals working for them, supporting the compliance objectives of the regulatory requirements.[135] Drawing on our supervisory and enforcement engagement with providers, we consider that these principles also apply in the fraudulent advertising context.

4.127 Where a service provider has a clear code of conduct in place, it can support a more consistent understanding among individuals working for the provider of how fraudulent advertising should be prevented and addressed. This, in turn, can support more consistent application of the provider's approach in day-to-day decision-making, including the identification, reporting and appropriate escalation of concerns relating to fraudulent advertising.

4.128 Where individuals working for the service provider are not subject to a clear and enforceable code of conduct relating to fraudulent advertising, there is a risk that the potential harm to users is not consistently considered in day-to-day decision-making and operations. In the absence of appropriate oversight or assurance that such expectations are understood and applied in practice, the effectiveness of the proposed measure in driving consistent behaviours across the organisation may be limited.

4.129 Where individuals working for a service provider do not understand, or are not trained in, the provider's approach to fraudulent advertising, mitigations may be applied inconsistently or ineffectively. This may result in missed opportunities to prevent individuals from

[135] The FCA requires regulated firms to have codes of conduct for staff under its SM&CR scheme (cited in footnote 119), which are in line with firms' duties to comply with financial regulations. The US Department of Justice in guidance to prosecutors on evaluating corporate compliance programmes, notes that "any well-designed compliance program utilizes policies and procedures to give both content and effect to ethical norms and to mitigate risks identified by the company as part of its risk assessment process" and advises prosecutors to consider whether a code of conduct is accessible and applicable to all employees, alongside the extent to which compliance policies and procedures are embedded in operations. Source: US Department of Justice, 2023. Evaluation of Corporate Compliance Programs. [accessed 6 April 2026].

encountering fraudulent advertisements, to minimise the length of time such content is present, or to ensure its swift removal, in accordance with the fraudulent advertising duties.

Impacts and costs on service providers

4.130 We estimate that developing a code of conduct would involve one-off costs of up to around £10,000 for service providers that do not already have an existing code of conduct in place. Our estimates assume that developing a code of conduct would require input from a specialised professional occupation staff member (for example, senior manager or lawyer) on a full-time basis for a period of less than 20 days.[136]

4.131 We acknowledge these costs could be higher for service providers that identify more indicators of fraudulent advertisements and accounts that have posted fraudulent advertisements (as part of their fraud indicator assessment). Another factor that could influence the costs of developing a code of conduct is the complexity of internal due diligence processes (for example, legal review) and of integrating this proposed measure with existing staff policies. As is the case with other governance measures, there could be some synergies realised between the proposed Fraudulent Advertising Codes governance measures and the governance measures recommended under the Illegal Content Codes and Protection of Children Codes.

4.132 We also expect there would be some additional costs associated with disseminating the code of conduct to all individuals working for the service provider, as well as individuals spending time to review and understand it. However, we assume that the document would be short and would not take a significant amount of time to read and understand.

4.133 There would also be some ongoing costs associated with reviewing and keeping the code of conduct up to date, especially as the nature of fraudulent advertising evolves and hence service providers' approach to dealing with fraudulent advertising would require periodic updating. We do not expect these costs to be substantial as we assume that the document would be short and would not take a significant amount of time for individuals to read and understand.

Provisional conclusion

4.134 Our analysis shows evidence that staff policies and processes, including having codes of conduct, are effective in ensuring that service providers communicate compliance requirements. This, in turn, is important for supporting a culture of awareness of fraudulent advertising among staff.

4.135 Having a code of conduct around protecting users makes it more likely that service providers will identify, consider and adopt opportunities to protect users from encountering fraudulent advertising. Therefore, the proposed measure would deliver significant benefits.

4.136 Set against this, we consider that costs related to the proposed measure regarding codes of conduct are likely to be small for all Category 1 and 2A services.

---

[136] This is consistent with our cost assumptions and estimates for the equivalent measures in our Illegal Content Codes and Protection of Children Codes. Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.111. [accessed 28 April 2026]. See our labour wage assumptions in Annex 8, 'Further detail on economic assumptions and analysis'.

4.137 There is a possibility that without these efforts for service-wide understanding, people working in different areas of a service will not understand how the provider is approaching regulatory compliance, or how it manages fraudulent advertising on its service.

4.138 Together with our other proposed governance measures, this measure will contribute to the creation of a service-wide culture to support compliance with the fraudulent advertising duties.

4.139 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this measure (if any) are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.140 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services have a code of conduct that sets out the standards and expectations for individuals working for the provider around preventing individuals from encountering fraudulent advertising.

4.141 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU A5 and FAS A5 respectively.

Compliance training

Explanation of the measure

4.142 We propose that service providers should secure that individuals working for the provider with a role that involves the service's paid-for advertising function are trained in the provider's approach to compliance with the fraudulent advertising duties sufficiently to give effect to them.[137]

4.143 We do not propose to stipulate how the training should be carried out nor by whom. We consider it important that there is flexibility so that this proposed measure can be implemented by the range of services in scope of the fraudulent advertising duties. Service providers are best placed to inform how they can achieve the outcomes set out by this proposed measure. This is consistent with the approach we have taken in our Illegal Content Codes and Protection of Children Codes. It also aligns with our approach on training of individuals working in advertising moderation (see Volume 4, Section 2, 'Advertising moderation').

Benefits and effectiveness

4.144 We consider that compliance training is important to achieve good safety outcomes for users by establishing a common understanding and expectation around actions to address fraudulent advertising on a service. Training is a key mechanism through which service providers communicate compliance requirements across an organisation and, by shaping individual decision-making, supports the consistent identification, consideration and day-to-day addressing of fraudulent advertising.

[137] Our definition of 'working for the provider' does not include volunteers. Volunteers are those who, in relation to the activities in question, are not employed by the provider or anyone else, remunerated, or acting by way of a business.

4.145 Evidence suggests that the absence of compliance training can have negative impacts, including by undermining company-wide understanding of regulatory requirements and of how risks – such as those related to fraudulent advertising – are managed and mitigated. Evidence from organisations that have experienced significant governance failings further highlights the importance of such training. For example, in our November 2023 Illegal Harms Consultation we referenced a Siemens case study which focused on strengthening compliance programmes through staff training and provided evidence that these measures were associated with improved perceptions of how risks were managed.[138]

4.146 We consider that this proposed measure will ensure that service providers clearly communicate the importance of compliance requirements to staff involved in advertising-related functions, including those responsible for the design, operation and oversight of advertising systems and processes. This will support the embedding of awareness, management and mitigation of fraudulent advertising risks across relevant teams, reducing the likelihood of inconsistent or ineffective implementation.

4.147 Taken together, this proposed measure reinforces the role of compliance training in supporting effective and sustained compliance by establishing clear expectations around how fraudulent advertising risks are identified, mitigated and kept under review across a provider's organisation.

Impacts and costs on service providers

4.148 We consider that a service provider would require input from one professional occupation staff member on a full-time basis for a period of two to four weeks to develop the relevant training materials. We estimate this would cost around £2,000 to £10,000.[139]

4.149 Service providers would also incur costs associated with delivering the compliance training to their employees who are involved in the advertising function. We have estimated the training cost per person per day to be in the range of £170 to £300 for general staff involved in the design and operation of the service.[140] We have also estimated the respective costs for senior staff to be between £620 for the lower bound and £900 for the upper bound for one day.[141]

4.150 We also expect costs to be higher for services hosting larger volumes of advertisements as they will need to train more people, as well as for services with a high volume of fraudulent advertising as they may need to provide more detailed training.

4.151 There may be some ongoing costs in keeping the relevant training materials up to date, especially as the nature of fraudulent advertising evolves and a service provider's approach

---

[138] Following a 2008 bribery scandal, Siemens attempted to redress governance failings identified by strengthening its compliance programmes. This included ensuring that employees "in different levels have been provided with trainings specific to their roles and responsibilities". Source: OECD, 2010. Compliance Program@Siemens, cited in Ofcom, 2023. Consultation: Protecting people from harms online Volume 3, Section 8, p.85.

[139] This is consistent with our cost assumptions and estimates for the equivalent measure in Illegal Content Codes and Protection of Children Codes. Ofcom, 2024. December 2024 Statement, Volume 1: Governance and Risk Management, p.112. [accessed 28 April 2026]. See our labour wage assumptions in Annex 8, 'Further detail on economic assumptions and analysis'.

[140] More specifically we have calculated the cost for training as the daily rate of the full-time equivalent of a general professional (for example, software engineer, content moderator, and so on).

[141] More specifically we have calculated the cost for training as the daily rate of the full time equivalent of a senior professional (for example, senior manager, director, and so on).

to compliance may change. In line with our standard assumptions for ongoing costs, we estimate that this would represent 25% of the initial costs of creating the relevant training materials and so could cost providers approximately £1,000 to £2,500.

Provisional conclusion

4.152 For similar reasons to the proposed measure on having a code of conduct, we consider that our proposed measure on staff compliance training will deliver significant benefits to users. It will establish a shared understanding of compliance expectations and will support consistent action across relevant staff, driving the effective implementation of processes and controls to meet the fraudulent advertising duties.

4.153 The cost of the proposed measure is likely to be relatively modest, and we consider it to be justified given the scale of the benefits and the foundational importance of robust governance in ensuring users are prevented from encountering fraudulent advertising.

4.154 Overall, we consider that the impacts on freedom of expression rights and privacy rights from this measure (if any) are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (see paragraphs 4.31 to 4.41).

4.155 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services secure that individuals working for the provider who are involved in the advertising function of the service are trained in the service's approach to compliance with the fraudulent advertising duties sufficiently to give effect to them.

4.156 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services, and it is referred to as FAU A6 and FAS A6 respectively.

5. Testing advertisement generation tools

What is this section about?

Service providers increasingly offer advertising account holders artificial intelligence (AI) tools (advertisement generation tools) that can create text, image, audio and video advertisements. Where advertisement generation tools have vulnerabilities or lack effective safeguards, they can be exploited by bad actors to create fraudulent advertisements.

In this section, we set out our proposed measure for testing advertisement generation tools (testing measure) and why we are proposing to recommend it.

Our proposal

Number in our Codes	Proposed measure applicable to providers of Category 1 and 2A services where a provider makes an advertisement generation tool available to advertising account holders (regardless of whether the tool is made available on or off the service).
FAU F1 and FAS F1	The provider should ensure that an **advertisement generation tool is tested** to identify whether and, if so, how the tool could be used to create fraudulent advertisements or fraudulent advertising proxy.

Why are we proposing this?

Our proposed measure should ensure service providers adopt effective testing of their advertisement generation tools which will help them identify key risks of these tools being used to produce fraudulent advertisements. This would enable them to better mitigate these risks, preventing users from encountering this harmful content.

Consultation question

- Do you agree with our proposal? Please provide any arguments and supporting evidence.

Introduction

5.1 Many service providers offer advertising account holders artificial intelligence (AI)[142] tools that can create text, images, audio or video advertisements.[143] We refer to such tools as 'advertisement generation tools'.

[142] AI, in this case, refers to the use of technologies – such as machine learning, natural language processing and computer vision – to perform tasks, such as categorising images and analysing data. Generative AI (GenAI) is a new form of AI which can create new content (text, images, video and audio) in response to a user prompt. Source: Ofcom, 2024. Ofcom's strategic approach to AI 2024/25.

[143] Ofcom research found that advertisers reported using Meta, Google Ads and Amazon's advertisement generation tools alongside external third-party AI tools. Source: Ofcom, 2026. Online advertising pathways: qualitative research report; Service providers have told us that they make advertisement generation tools available. Sources: [✂] response to our formal information request issued 30 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal

5.2 AI tools that create or modify content can lower the cost and expertise required to create convincing fraudulent advertisements.[144] This is likely to contribute to an increase in the overall volume, sophistication and variety of AI-generated fraudulent advertisements.[145] Additionally, AI-generated fraudulent advertisements can be difficult for users to distinguish from legitimate advertising.[146] Some evidence suggests AI-generated advertisements may be more persuasive than human-created advertisements.[147]

5.3 Advertisement generation tools present these same risks to UK users. These tools automate the production of advertisements and can enable advertising account holders to rapidly create and publish large numbers of polished advertisements at scale and at very low marginal cost, which may be viewed by millions of users.[148] While many AI-generated advertisements are legitimate, the capabilities described in paragraph 5.2 can – and indeed appear to have been – exploited by perpetrators of fraud.[149]

5.4 Where advertisement generation tools have vulnerabilities or lack effective safeguards, they may be misused to create fraudulent advertisements at scale, increasing the likelihood that users encounter fraudulent advertisements on a service. The risk of creating fraudulent advertisements may be amplified where tools are directly embedded into platform advertising workflows or selected as the default option for creating advertisements.[150]

5.5 In this section, we set out an explanation of the proposed measure, including:

- what an advertisement generation tool is;
- how providers should carry out testing of their tools (including examples of testing methods and frequency);

---

information request issued 26 June 2025; [✂] response to our formal information request issued 24 November 2025.

[144] Ladish, J., Lermen, S., Rogers-Smith, C. and Gade, P., 2024. BadLlama: cheaply removing safety fine-tuning from Llama 2-Chat 13B. [accessed 10 March 2026]. PwC and Stop Scams UK, 2023. Impact of artificial intelligence on fraud and scams. [accessed 10 March 2026].

[145] We discuss this further in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising', sub-section 'Tools and techniques used in fraudulent advertising'.

[146] Deepfake detection accuracy among humans averages around 55% across a meta-analysis of 56 studies involving over 86,000 participants, indicating that people struggle to reliably distinguish real from AI-generated media. Source: Diel, A., Lalgi, T., Schröter, I., MacDorman, K., Teufel, M. and Bäuerle, A., 2024. Human performance in detecting deepfakes: A systematic review and meta-analysis of 56 papers. [accessed 24 June 2026].

[147] Lee, H., Todri, V., Adamopoulos, P. and Ghose, A., 2025. The impact of visual generative AI on advertising effectiveness. [accessed 10 March 2026].

[148] Advertisers explained that when using AI for their campaigns, "you just set the budget, set the tracking up, and then just click publish. Within about 5 minutes I can have an ad with 100 different variations". A few advertisers made connection between AI tools and the ability of fraudulent actors to generate professional looking ads. Although none of the advertisers interviewed experienced this in practice. Source: Ofcom, 2026. Online advertising pathways: qualitative research report.

[149] [✂] response to our formal information request issued 26 June 2025 explained that some advertising campaigns that were created in part using their advertisement generation tool were restricted for a reason under their 'fraud-related proxy' umbrella.

[150] Many advertisers in our qualitative study exploring how advertisement creation interfaces are structured said that they felt that service providers nudged advertisers to use an on-platform AI advertising tool, because providers auto-select native tools in the campaign settings. Some advertisers said that they "gravitated" towards platforms that offered AI features or tools that enabled efficiency in the process of producing advertisements, for example, those that designed advertisement creatives. Source: Ofcom, 2026. Online advertising pathways: qualitative research report.

- the types of metrics that providers should produce and analyse to assess the risk that their tool could produce fraudulent advertisements; and
- decisions to take following the results of testing.

5.6 We then set out the expected benefits, costs and rights implications, before explaining why we consider the proposed measure to be proportionate.

Testing advertisement generation tools

Explanation of the measure

5.7 We propose that providers that make advertisement generation tools available to advertising account holders[151] to create paid-for advertisements for placement on the categorised service should test the tool for vulnerabilities that could enable the creation of fraudulent advertisements. We also propose that service providers should take appropriate steps to fix or mitigate any vulnerabilities identified with the tool.

5.8 In the context of this proposed measure, an advertisement generation tool refers to a tool that:

a) is made available by a provider of a Category 1 or 2A service to the holder of an advertising account for the purpose of creating a paid-for advertisement that can be encountered by UK users when posted on the categorised service; and
b) works by generating new content or substantially modifying existing content in response to an account holder's prompt, which can be an input of text, image, audio or video.[152] We consider this would include, for example, tools that create video scripts, animations, voiceovers, new images or backgrounds, advertising copy or descriptions, or content in new formats or languages. Such content could be based on existing content that an advertising account holder inputs into the tool or wholly new content that the tool creates.

Test the advertisement generation tool

5.9 Firstly, we propose that service providers should test the advertisement generation tool (either themselves or by appointing an appropriate third party) before launch, as reasonably practicable after deployment, and at regular intervals for as long as the tool remains in use by advertisers to generate advertisements available to UK users.[153] Providers should choose testing methods that enable them to identify whether, and how, the tool could be misused to create fraudulent advertisements. For example, this could be through testing methods such as red teaming or A/B testing.[154] We also propose that providers should determine a schedule or points at which they (or their appointed third party) would undertake testing. When seeking to understand how a tool could be misused, providers should consider the results of their latest fraud indicator assessment (FIA), or other relevant

[151] This is regardless of whether the tool is made available on or off the service.
[152] Section 236(1) of the Act defines content as "anything communicated by means of an internet service, whether publicly or privately, including written materials or messages, oral communications, photographs, videos, visual images, music and data of any description".
[153] For the purpose of this proposed measure, 'deployment' refers to an operational technology being put into use on a categorised service.
[154] Red teaming involves 'attacking' the tool to find vulnerabilities associated with fraudulent advertising, while A/B testing can compare how different versions of the tool perform against fraudulent advertising metrics.

insights they hold from an alternative measure.[155] Providers should choose how frequently to test, provided that testing does indeed take place at these stages. For example, this could be at a given cadence or point, such as following a significant update to the tool or underlying technology.

5.10 Secondly, we propose that service providers – or appointed third parties – should produce and analyse metrics which indicate the risk that the advertisement generation tool could produce fraudulent advertisements. Providers should seek to identify metrics that are relevant to fraudulent advertising. For example, this may include the number and type of fraudulent advertisements created during the test (of the total share of advertisements created), or the techniques used to bypass safeguards to create fraudulent advertisements. Providers should also consider the results from their latest FIA (or equivalent) to determine the metrics that might helpfully be compiled to assess the risk that their tool could produce fraudulent advertisements. Providers (or appointed third parties) could use metrics relevant to fraudulent advertisements or those relevant to fraudulent advertisement proxies (see Volume 4, Section 2, 'Advertising moderation').[156]

Address vulnerabilities identified

5.11 Thirdly, we propose that service providers should review the test results and take appropriate steps to fix or mitigate the identified vulnerabilities. For example, where results indicate that the tool creates content depicting well-known public figures or references known scams, a provider may adopt prompt and output filters or introduce additional safeguards to reduce misuse.

5.12 Fourthly, we propose that service providers should maintain a log comprising the test, findings, and resulting decisions related to the design or operation of the tool. We consider that such logs should be kept for a minimum of three years (consistent with our Record-Keeping and Review Guidance), or in accordance with the organisation's record retention policies, if longer.[157] A provider should establish a log and include information in their log regarding (but not limited to), who was involved in the testing of the tool, the testing method used, the metric(s) they collected (including whether a proxy has been used), the results obtained from the test (including an analysis of the metrics collected), and any design and operational decisions a provider (including a third-party supplier) has made, in response to the results of the testing (if any). Examples of the types of decisions that a provider should log include:

- a decision to alter the design of the tool that introduces friction in the process of creating fraudulent advertisements;
- a decision to adopt a new safety mitigation to tackle risks identified;
- a decision to undertake further testing to understand more about the risk identified; or

[155] See Volume 2, Section 3, 'Fraud indicator assessment' for more information.
[156] A fraudulent advertising proxy is an advertisement that a service provider has assessed against its own categories of prohibited advertisements (set out in its terms of service or publicly available statement, advertising contracts (where all of the provider's advertising contracts contain similar prohibitions in relation to fraudulent advertisements), or a combination of these when read together). The provider may do this where it is satisfied that the fraudulent advertisements that it has reason to suspect exist are prohibited by these policies or contracts. For more information, see Volume 4, Section 2, 'Advertising moderation'.
[157] Ofcom, 2025. Record-Keeping and Review Guidance.

- a decision to pause the deployment of the tool where a significant risk is identified.

5.13 Finally, we propose that service providers should ensure that the log is made available and referred to by individuals who work for the provider, directly or indirectly, in the development, testing or operation of the tool. This includes individuals involved in carrying out their FIA or equivalent measures to understand how fraudulent advertising manifests on the service. Providers should also ensure that the log is available to and referred to by individuals involved in risk and governance associated with the tool as part of our proposed accountable individual and annual review measures, as these individuals are responsible for ensuring the provider complies with its fraudulent advertising duties.[158]

5.14 We have designed the proposed measure in a way that allows service providers some flexibility in exactly how it is implemented. As we set out in paragraphs 5.9 to 5.13 providers can, for example, determine a testing method, a schedule for undertaking testing and metrics relevant to fraudulent advertising. We consider this flexibility appropriate because it enables providers to make context-appropriate decisions about testing that correspond to, for example, the type of tool they make available, their existing testing architecture and their own internal expertise.

Scope of the proposed measure

5.15 We acknowledge that advertisement generation tools, or underlying technologies that power the advertisement generation tools, may be provided by a third party. We consider that any service provider who makes an advertisement generation tool available on their categorised service should achieve the outcome of the proposed measure regardless of how (or by whom) the advertisement generation tool is developed. This is because the proposed measure is sufficiently flexible to enable a provider, or a third party, to undertake testing of the tool.

5.16 Where an advertisement generation tool is made available by a provider of a Category 1 or 2A service to the holder of an advertising account for the purpose of creating a paid-for advertisement capable of being encountered by UK users when posted on the service, it is in scope of this measure. We consider that the tool is made available where this is done either on the service itself, or in circumstances where the provider has effective control of the tool and it is similarly available to advertising account holders as if it was on the service. This includes circumstances in which, by virtue of a contractual or other arrangement, the tool is not technically provided on the categorised service but is made available through different means within the provider's control, where the tool nonetheless enables advertising account holders to create paid-for advertisements for publication on the categorised service. This proposed measure intends to capture cases where the tool performs a substantially similar function, and serves the same purpose, as a tool made available on the categorised service. The proposed measure therefore recommends that any service provider who makes an advertisement generation tool available in these ways should achieve the outcome of the measure regardless of where the advertisement generation tool is deployed.[159]

---

[158] See Volume 2, Section 4, 'Governance and accountability' for more information.

[159] We consider that our proposed intermediaries measure does not apply in this case (see Volume 2, Section 2, 'Advertising intermediaries'). This is because the testing measure applies depending on whether the provider has effective control of the advertisement generation tool. As such, even where some functions of advertisement placement have been outsourced to advertising intermediaries, where the provider retains effective control of the advertisement generation tool, they will still be in scope of the testing measure.

5.17 We acknowledge that advertising account holders use many AI tools for the purposes of creating advertisements.[160] However, this proposed measure is not applicable to general-purpose AI tools that providers make available to all users to create other types of content that are not for the purposes of paid-for advertising.[161] Additionally, we do not consider a tool that only applies minor or superficial edits to content (without the means to create or substantially modify content) would be an advertisement generation tool for the purposes of this proposed measure. This includes, for example, a tool which can only alter the colour saturation of an advertisement.[162] Neither would we consider an advertising tool that creates an advertising campaign schedule, or which creates insights on a campaign's performance to be an advertisement generation tool. This is because we do not consider these types of tools to pose a material risk of creating fraudulent advertisements.

5.18 This proposed measure is likely to interact with other measures proposed in the draft Fraudulent Advertising Codes of Practice, including the FIA measure, annual review measure, and advertising moderation measures. We propose that providers who adopt this proposed measure should do the following:

- Consider the results of product testing (paragraphs 5.11 and 5.12) when determining whether any characteristics of accounts or advertisements are associated with fraudulent advertising, as part of their FIA.
- Include, when undertaking an annual review of their compliance with the fraudulent advertising duties, any actions they have taken in response to the results of product testing (paragraphs 5.11 and 5.12).
- Consider insights from their moderation of fraudulent advertisements where relevant to the advertisement generation tool when determining how often they should test the tool (paragraph 5.9).

Benefits and effectiveness

5.19 Advertisement generation tools can make it easier for bad actors to create fraudulent advertisements at scale (see paragraphs 5.2 to 5.4 and Volume 1, Section 4, 'Causes and impacts of fraudulent advertising'.[163]

---

[160] Advertisers reportedly used a mixture of third-party tools, like ChatGPT, Canva and Nano Banana, as well as on-platform tools. Source: Ofcom, 2026. Online advertising pathways: qualitative research report.

[161] Sections 38(3), 38(3)(a), 39(3) and 39(3)(a) of the Act. Section 236 of the Act defines a paid-for advertisement as "[A]n advertisement is a paid-for advertisement in relation to an internet service if (a) the provider of the service receives any consideration (monetary or non-monetary) for the advertisement (whether directly from the advertiser or indirectly from another person), and (b) the placement of the advertisement is determined by systems or processes that are agreed between the parties entering into the contract relating to the advertisement".

[162] We note that an advertisement could be produced across several components, including a service's advertisement generation tool and one or more out-of-scope tools. We do not consider how an individual advertisement is produced to be relevant in determining whether a tool is or is not in scope, only whether or not the provider's advertisement generation tool can create or substantially modify or alter a paid-for advertisement.

[163] In response to our formal information notice, some providers told us that, in the context of their specific service and existing controls, they did not consider that their AI capabilities or advertisement generation tools were a risk factor for fraudulent advertising. We provisionally disagree with their view, based on the evidence we set out in Volume 1, Section 4, 'Causes and impacts of fraudulent advertising', sub-section 'Tools and techniques used in fraudulent advertising', and in paragraphs 5.2 to 5.4. Additionally, we consider that, as advertisement generation tools become more sophisticated and are increasingly deployed, and as providers

5.20 We provisionally consider that the steps outlined in this proposed measure will help service providers to identify vulnerabilities in advertisement generation tools before they are exploited and to respond as new risks emerge. It will also enable providers to take targeted steps to address vulnerabilities, therefore reducing opportunities for bad actors to misuse the tool, leading to fewer fraudulent advertisements on the service.

Benefits of testing an AI tool

5.21 Testing a tool before it is launched will help service providers to spot vulnerabilities that could allow fraudulent advertisements to be created. For example, one evaluation of an AI model found that it could create "unauthorized voice generation" which could be used to impersonate well-known individuals in fraudulent advertisements.[164] Meanwhile, another evaluation found that another model whose safeguards had been circumvented was able to create "a scam email requesting 10,000 dollars".[165] This step enables providers to take targeted steps to fix specific problems before bad actors have an opportunity to exploit them.

5.22 It is critical to test a tool as soon as reasonably practical after its launch. Testing after launch enables service providers to identify vulnerabilities that would be difficult to identify pre-deployment.[166] We note that advertising campaigns (that use AI tools) can reach large numbers of users in a very short space of time. So, if there is testing as soon as reasonably practical after launch, it will help mitigate the risk that vulnerabilities remain unaddressed for a long period of time, and reduce the resultant harm to users.

5.23 Regularly testing a tool after it is deployed helps service providers to identify new vulnerabilities in the tool and ensures that any previously implemented fixes are working correctly. As bad actors develop new tactics to exploit vulnerabilities in advertisement generation tools, testing should give providers better visibility of emerging risks. Indeed, evidence from the Integrity Institute emphasises that the technical systems deployed to support detection and mitigation of fraudulent advertisements should be continually developed and evaluated.[167] This is because it allows providers to identify vulnerabilities as they emerge, reducing the risk that those vulnerabilities are undetected and exploited by bad actors to create fraudulent advertisements.[168]

5.24 We also consider that collecting and analysing metrics relevant to fraudulent advertising during testing should help service providers to understand the scale of the potential harm from fraudulent advertisements. The collection and analysis of metrics can give providers

---

meet the outcomes of this proposed measure (namely, testing), we expect to see further evidence that such tools have created fraudulent advertisements. Source: [✂] response to our formal information request issued 30 June 2025.

[164] OpenAI, 2024. GPT-4o System Card. [accessed 10 March 2026].

[165] While this evaluation was undertaken on an AI model who safety features had been circumvented, we have included the evidence because we consider that it illustrates how testing identifies issues relevant to fraudulent advertising. Source: Ladish, J., Lermen, S., Rogers-Smith, C. and Gade, P., 2024. BadLlama: cheaply removing safety fine-tuning from Llama 2-Chat 13B. [accessed 10 March 2026].

[166] A survey of over 200 model safety evaluations found that most of the tests were model-centric and did not consider deployment context, which left entire risk categories unaccounted for. Source: Wang, Z., Knight, C., Kritz, J., Primack, W. and Michael, J., 2025. A Red Teaming Roadmap Towards System-Level Safety. [accessed 26 June 2026].

[167] Ofcom and Integrity Institute, 2026. Fraudulent advertising and account integrity: Expert insights on best practice.

[168] Google, 2026. Google Workspace's continuous approach to mitigating indirect prompt injections. [accessed 24 June 2026].

insights into how the tool can be abused by bad actors. For example, it can enable them to understand the types of fraudulent advertisements (or fraudulent advertising proxy) the tool could create, how often the tool creates fraudulent advertisements (or fraudulent advertising proxy) and the methods bad actors can use to get around safeguards. We consider that where providers do not collect and analyse fraudulent advertisements metrics, they may not fully understand the tool's vulnerabilities and therefore not be able to take informed decisions about appropriate action to take in response to the results to prevent the tool being exploited (see paragraph 5.26).

Benefits of responding to the vulnerabilities identified in the test

5.25 However, testing alone will not necessarily make it harder for bad actors to create fraudulent advertisements. To be effective, service providers need to act on the results of testing by taking targeted steps to fix or mitigate specific vulnerabilities identified.

5.26 Evidence indicates that where providers review the test results and adopt appropriate fixes, the amount of harmful content created by a tool is likely to reduce. In one internal evaluation of an AI model, iterative rounds of red teaming and targeted fixes produced an average 75% reduction in the amount of harmful content the model generated.[169] Taking targeted steps to address vulnerabilities identified through testing can therefore result in a tool that is much harder for bad actors to exploit. This is because it adds friction to fraudsters' efforts to create fraudulent advertisements, by increasing the time and skill required to create such content. The flexibility offered by our proposed measure allows service providers to tailor their responses to the tests results, best ensuring a reduction in users encountering fraudulent advertisements on the service.[170]

5.27 We consider that keeping a log can help make sure that everyone involved in the design, operation and governance of the tool is well informed about the tool's risk in relation to fraudulent advertising. Documenting information regarding the type of test undertaken, vulnerabilities identified, how the advertisement generation tool performed, and any design or operational decision that has come as a consequence of the test results can improve the service provider's oversight of the advertisement generation tool, and reduce future risk of misuse. Additionally, recording the details of individuals involved in testing also ensures that decisions are traceable. This aligns with the international AI management standard ISO/IEC 42001, which mandates that organisations that create, develop or use AI must have auditable documentation containing logs of what was tested, who was involved in the process, any vulnerabilities found, and plans to fix them.[171]

5.28 Keeping a log of subsequent tests enables service providers to assess the effectiveness of the tool over time, for example, by identifying whether particular updates or specific mitigations have increased or reduced the fraudulent advertisements created by the tool. It can also support the work of individuals involved in risk and governance associated with the tool as part of our proposed accountable individual and annual review measures by creating an ongoing record of how the provider is complying with its fraudulent advertising duties. Taken together, we consider that maintaining a test log is likely to support appropriate accountability and oversight of the tool, including by informing a provider's FIA.

---

[169] Microsoft, 2025. Responsible AI transparency report. [accessed 17 March 2026].
[170] Ofcom, 2024. Red teaming for GenAI harms – Revealing the risks and rewards for online safety.
[171] International Organization for Standardization, 2023. ISO standard 42001:2023. [accessed 6 April 2026].

5.29 We consider that making the log available to individuals who work for the provider, directly or indirectly, in the development, testing or operation of the tool should increase the likelihood that decisions relating to the design, operation and governance of the tool are based on accurate and complete information. Experts in AI management standards compliance have noted that AI tool management systems commonly fail where operational decisions are not traceable.[172] Without having a log to refer to, service providers risk developing gaps or blind spots in their understanding of a tool's vulnerabilities, which in turn is likely to make it harder for them to meet their fraudulent advertising duties.

5.30 Overall, these steps will make it more difficult and time-consuming for bad actors to bypass safeguards to create fraudulent advertisements by means of an advertisement generation tool. As a result, users will ultimately see fewer of these types of fraudulent advertisements on the service. We consider that without this proposed measure, it would be more difficult for providers with advertisement generation tools to meet their fraudulent advertising duties under the Online Safety Act 2023 (the Act) including to prevent users from encountering fraudulent advertisements.

Impacts and costs on service providers

Direct costs

5.31 We have sought to estimate the costs associated with setting up and operating an effective testing process for advertisement generation tools from scratch. We consider that in practice, these costs would depend on many factors, including the service provider's size and internal expertise, the provider's existing testing infrastructure, the scale of advertising on a service, the number and type of tools being tested, and the testing approach (that is, methodology for testing, degree of automation, whether the testing is done in-house or through a third party). The relationship between the costs for implementing this proposed measure and the factors listed at the start of this paragraph is not linear.

5.32 Our baseline cost estimate quantifies the cost of testing one multimodal text-and-image advertisement generation tool using a hybrid red-teaming approach, with a seed library of around 500 adversarial prompts crafted by subject matter experts.[173] Building such a testing pipeline from scratch could involve technical and specialist resources, as a service provider would need to establish all core components, including the testing methodology and prompt library, the automated evaluator, the testing infrastructure and the audit-log pipeline, and to run a baseline test against the resulting setup.

5.33 We consider that service providers would be in a position to build and run an initial testing pipeline in stages over a period of two to three months, with some activities running in parallel, using a mix of leadership, specialist and technical staff to design and implement the testing pipeline, oversee delivery, manage milestones and support sign-off.

---

[172] ISMS, 2025. Documentation Required Under ISO 42001. [accessed 1 April 2026].

[173] We decided to base our cost estimates on the assumption that providers would use 500 seed or behaviour prompts in their testing because it aligns with established industry benchmarks and would be sufficient to account for the main fraud categories. Sources: Zou, A., Wang, Z., Carlini, N., Nasr, M., Zico Kolter., J. and Fredrikson, M., 2023. Universal and Transferable Adversarial Attacks on Aligned Language Models. [accessed 4 March 2026]; Mazeika, M., Phan, L., Yin, X., Zou, A., Wang, Z., Mu, N., Sakhaee, E., Li, N., Basart, S., Li, B., Forsyth., D. and Hendrycks, D., 2024. HarmBench: A Standardized Evaluation Framework for Automated Red Teaming and Robust Refusal. [accessed 4 March 2026].

5.34 More specifically, we consider that services would involve input from a team of software engineers[174] to (a) lead with building the testing infrastructure (two to three software engineers working on a full-time basis for four to six weeks), and (b) perform technical remediations following the baseline test results (one to two software engineers working on a full-time basis for two to three weeks).

5.35 In addition, services would involve input from a core team of fraud subject matter experts tasked with (a) designing the taxonomy and test framework (two to four experts working on a full-time basis for two to four weeks), and (b) analysing the baseline test results and log and synthesise findings in a report (one to two experts working on a full-time basis for two to three weeks).[175] Services would also involve input from three to five red teamers on a full-time basis for one to two weeks for the test execution.[176] We also account for input from one to two legal professionals over a period of one to two weeks, who would be responsible for reviewing and approving the testing approach and managing sign-off and internal governance.

5.36 Using our standard wage assumptions,[177] we estimate build costs, including the first baseline test, would range between £27,100 and £160,000 depending on the scale and complexity of the testing programme.[178]

5.37 We acknowledge that service providers would incur additional costs where they test multiple advertisement generation tools. We consider that in most cases there will be synergies for setting up a testing pipeline for subsequent tools, and the resulting costs are likely to be lower than the initial build costs. Respective costs for each subsequent tool will therefore vary depending on whether providers choose to set up new testing pipelines or adapt an existing one. This is primarily a business decision, though we acknowledge that to some extent, it will depend on the features of the advertisement generation tools and how straightforward it would be to adapt an existing testing pipeline.

5.38 Service providers will incur non-staff costs including compute, software and available data storage to set up the testing pipeline.[179] We do not quantify these precisely since they are likely to vary significantly based on a provider's existing capabilities and needs.

5.39 In addition to the costs associated with the setup of the testing pipeline, service providers would also incur costs when running the tests at regular intervals. This is because the proposed measure recommends that providers conduct testing before and after deploying their advertisement generation tool, and to repeat the test at regular intervals, as explained in paragraphs 5.9 and 5.10.

---

[174] In practice, services would involve a mix of software engineers, data engineers and data scientists. As explained in Annex 8, 'Further detail on economic assumptions and analysis', for the purposes of estimating labour costs, we use SOC (Standard Occupational Classification) 2134 (Programmers and software development professionals) to also capture these occupations.

[175] For the purposes of estimating labour costs, we consider that the most appropriate occupational code for subject matter experts would be SOC 24 (Business, media and public service professionals).

[176] For the purposes of estimating labour costs, we consider that the most appropriate occupational code for red teamers would be SOC 24 (Business, media and public service professionals).

[177] As set out in Annex 8, 'Further detail on economic assumptions and analysis'.

[178] We further detail the steps involved and any additional assumptions underpinning these estimates in Annex 8, 'Further detail on economic assumptions and analysis'.

[179] We use the term 'compute' to refer to processing power, memory, networking, storage, and other resources required for the computational processing of software.

5.40 We expect that subsequent tests will likely be less expensive than the first test, ranging from £3,400 to £26,100, as service providers build experience and the testing process becomes more efficient. This is because providers will have developed a library of testing prompts, improved their automated tools (for example, which may classify test outputs) and established their testing pipeline.[180] We therefore assume that providers would be in a position to run subsequent tests in two to three weeks with some activities running in parallel using input from:

a) One to two software engineers on a full-time basis for three to five days tasked with evaluating whether the model performance remains fit for purpose and implementing any changes, and over a period of five to ten days to analyse and perform technical remediations;
b) One to two fraud subject matter experts on a full-time basis for three to five days tasked with reviewing flagged fraud cases; and
c) One to two legal professionals on a full-time basis for one to three days to manage sign-off and review the audit log.

5.41 Several factors could affect the recurring testing costs incurred by service providers. This includes the type of advertisement generation tool that is tested, with a multimodal tool that includes video generation likely being towards the upper bound of our estimated range (or above), and a simpler text generation tool towards the lower bound of our estimated range (or below). Similarly, testing a greater number of prompts (our baseline estimate included a seed library of around 500 adversarial prompts) will incur higher costs. The frequency of repeating the test for each tool would also affect the costs providers would incur. As explained in paragraph 5.40, for each subsequent test, a provider would be burdened with costs ranging between £3,400 and £26,100. Therefore, ongoing costs associated with recurring testing would be higher the more regularly the test is repeated.

5.42 We further expect service providers to incur annual maintenance costs of between £6,800 and £40,000.[181] We note that these costs only capture regular upkeep of the tool's systems and processes. Providers would incur additional costs when implementing more significant updates on their testing pipeline, for example, to include additional testing metrics or to respond to new tactics employed by perpetrators of fraudulent advertising. Though the proposed measure does not explicitly recommend providers do that, we expect that it is in the interest of providers to periodically implement such updates on their testing infrastructure to ensure that it remains effective.

5.43 A service provider may incur costs to fix or mitigate vulnerabilities identified in the test, as explained in paragraph 5.11. We have not assigned a cost for adopting mitigations, as these will vary depending on the action a provider decides to take. Providers would also need to

[180] We note that "costs may decrease when a service repeats a red teaming exercise or as it gains experience from conducting increasing numbers of different red teaming exercises." This should apply to other types of testing as well. Source: Ofcom, 2024. Red teaming for GenAI harms – Revealing the risks and rewards for online safety.

[181] Where implementing a measure involves an initial build cost, we typically assume there is also an annual maintenance cost of 25% of the build cost. We note however that several factors could affect the cost of maintenance, potentially departing from the 25% assumption. For example, if there are frequent updates from application programming interface (API) providers (which could mean maintenance have to spend extra time fixing the connection and retest the system) the provider may have to check that the large language model remains accurate. Moreover, the first year likely includes extra time for debugging and security patches as the testing algorithm stabilises.

account for costs associated with maintaining a log as set out in paragraph 5.12, though we consider these costs to be negligible.

5.44 While our baseline reflects the costs to set up testing infrastructure from scratch, we know that some service providers [✂] already test their advertisement generation tools to determine whether the tools create content that violates their fraudulent advertising policies or can be misused.[182] We therefore consider that some providers are likely to have relevant testing architecture in place for their advertisement generation tool, and may not incur each of the costs we set out in paragraphs 5.34 and 5.35.

5.45 Even where service providers do not already test their advertisement generation tools, many providers will have established testing architecture for other AI systems, tools or functionalities on their service. While we acknowledge that any adaptation of an existing testing process to align with this proposed measure will not necessarily be straightforward, we consider that the costs incurred by providers in this scenario are also likely to be lower than those that we have set out in our baseline.[183]

5.46 As we explain in paragraph 5.14, service providers have some flexibility on exactly how to adopt each aspect of this proposed measure. This enables them to adopt the measure in a way that can be scaled up or down to complement their internal expertise and leverage their existing infrastructure and testing architecture in a cost-effective way.[184]

5.47 Some service providers may opt to outsource a product testing exercise to a third party to undertake the testing on their behalf. How the costs of third-party testing compare to in-house testing is not straightforward as it depends on a number of factors, including the approach to testing, the scale of the testing and the cadence of testing. Indicatively, one compliance organisation sets a price of between US$16,000 and US$60,000 (£11,620 and £43,575) for a typical two-week red team assessment.[185] Third-party input can take different forms and may include governments, safety technology providers and civil society groups, particularly where a provider needs outside expertise specific to the harm area.[186]

Indirect costs

5.48 We have not identified substantive evidence that the proposed measure would have an adverse impact on a service provider's revenue or their ability to compete and innovate with providers of non-categorised services. We have also not identified substantive evidence that providers would be likely to withdraw advertisement generation tools rather than test them. We know some providers [✂] already appear to undertake relevant

[182] [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025.

[183] In our November 2023 Illegal Harms Consultation, we assumed that the costs incurred for recommender system testing can be "limited considerably" where a service already runs on-platform tests, because "such services would already have an established testing environment in place, as well as the specialist staff needed to execute on-platform tests and implement the recommendations put forward..." While this consideration refers to testing a different product, we expect that the same factors remain relevant for testing advertisement generation tools. See Ofcom, 2023. November 2023 Illegal Harms Consultation, Volume 4: How to mitigate the risk of illegal harms – the illegal content Codes of Practice, paragraph 19.36.

[184] Ofcom, 2024. Red teaming for GenAI harms – Revealing the risks and rewards for online safety.

[185] Schellman, 2026. Penetration testing: AI red teaming and Penetration testing: Red team assessment. [accessed 1 April 2026]; Costs calculated at a conversion rate of 1.00 USD = 0.73 GBP.

[186] Ofcom, 2024. Red teaming for GenAI harms – Revealing the risks and rewards for online safety.

testing,[187] while many others are likely to have existing testing infrastructure for AI tools more generally that could be adapted for this purpose. We therefore consider that there is minimal risk that the proposed measure would either directly or indirectly negatively impact the development of AI-powered advertisement generation tools or have any negative implications for the development of AI more broadly.[188] The proposed measure also offers some flexibility in how it may be implemented, which should enable providers to use existing systems and internal expertise in a cost-effective way.

5.49 In fact, where a service provider adopts this proposed measure, they may reduce costs in other areas, for example, in content moderation. Industry evidence suggests that rectifying an error can cost four to five times more if it is rectified after the product is released, compared with if the error had been addressed during the design phase. This evidence also indicates that the cost to rectify an error at the maintenance phase could be up to 100 times more than addressing it during the design stage.[189] A global survey suggested that organisations who regularly audit or assess the performance and compliance of their AI systems and tools are "over three times more likely to achieve high GenAI value than organizations who do not".[190]

Rights assessment

Freedom of expression

5.50 As explained in Volume 1, Section 5, 'Approach to Codes', sub-section 'Approach to human rights assessments', Article 10 of the European Convention on Human Rights (ECHR) upholds the right to freedom of expression, which encompasses the right to hold opinions and to receive and impart information and ideas without unnecessary interference by a public authority. Article 10 is a qualified right and we must exercise our duties under the Act in a way that does not restrict this right unless satisfied that it is proportionate to the legitimate aim pursued. As noted in Volume 1, Section 5, 'Approach to Codes', we start from the position that these proposed measures are prescribed by law, in pursuit of a legitimate aim and correspond to a pressing social need.

5.51 This proposed measure does not recommend that service providers block or remove any particular content or advertisements. Instead, it proposes to recommend that providers identify and take steps to respond to vulnerabilities in any advertisement generation tools they operate, to reduce the likelihood that fraudulent advertisements can be created

[187] [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025; [✂] response to our formal information request issued 26 June 2025.

[188] Department for Business and Trade, 2024. Growth duty: Statutory Guidance. [accessed 5 June 2026]; See Volume 1, Section 6, 'Combined impact assessment', sub-section 'Other relevant considerations'.

[189] This article focuses on software testing (for mobile applications), with the broad concept referred to in software engineering as "Shift left testing". Models and red teaming thus likely fit a similar pattern, as vulnerabilities identified during design or development can typically be mitigated at far lower cost than those discovered after deployment or exploitation. Source: Global App Testing, 2025. How Much Does Software Testing Cost in 2025? [accessed 17 March 2026].

[190] Survey was conducted in May – June 2025 among 360 respondents from organisations with at least 250 full time-employees. The reported relationship between AI system assessment and compliance practices and GenAI value reflects an observed association and should not be interpreted as causal. Gartner, 2025. Gartner survey finds regular AI system assessments triple the likelihood of high GenAI value. [accessed 17 March 2026].

through those tools. We therefore consider any interference with service providers' rights is likely to be limited.

5.52 Interference with advertising account holders' and individuals' rights to freedom of expression may arise if fixes or mitigations implemented following testing (in accordance with paragraph 5.11) restrict advertising account holders' ability to create or modify paid-for advertisements that would have been posted on the service. In such cases, individuals may no longer be able to encounter them. We also note that this proposed measure operates at the level of the advertisement generation tool itself, rather than in relation to specific advertisements, and so could theoretically affect a high number of potential paid-for advertisements which could otherwise have been created or modified.

5.53 However, the proposed measure recommends only that service providers take actions to fix vulnerabilities associated with fraudulent advertising, meaning that the scope of such fixes should be narrowly targeted to address the production of fraudulent advertisements, rather than legitimate content more broadly. Additionally, the proposed measure recommends that testing be conducted at regular intervals which we consider would lead to providers refining the steps they take to more accurately address fraudulent advertisements. Providers also have incentives to meet users' and advertising account holders' expectations in relation to what advertising can be encountered, and they may also be incentivised to minimise the number of legitimate advertisements wrongly affected by any fixes they make in order to maximise their advertising revenue.

5.54 We also note that to the extent that any affected paid-for advertisements are fraudulent advertisements, such advertisements would not attract protection under Article 10 of the ECHR.

5.55 To the extent that this proposed measure helps to reduce harm from fraudulent advertising and make users feel safer using a service, it could also positively affect their human rights. Further, our proposed FIA measure acts as a safeguard for freedom of expression to the extent that it could help a provider to direct its testing at content more likely to consist of fraudulent advertising, rather than affecting legitimate advertisements.

5.56 Overall, we do not consider that there is a less intrusive means to achieve the same aims and we consider any interference to be proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (which this proposed measure is intended to help providers of Category 1 and 2A services to secure).

Data protection and privacy

5.57 As explained in Volume 1, Section 5, 'Approach to Codes', Article 8 of the ECHR confers the right to respect for an individual's private and family life. Article 8 is a qualified right, and we must exercise our duties under the Act in a way that does not restrict this right unless we are satisfied that it is proportionate to the legitimate aim pursued. As noted in relation to freedom of expression, we start from the position that these proposed measures are prescribed by law, in pursuit of a legitimate aim and correspond to a pressing social need. Article 8 underpins the data protection laws with which providers must comply.

5.58 As noted in Volume 1, Section 5, 'Approach to Codes', we consider that in the majority of cases, there will be minimal expectations of privacy in relation to the content of paid-for advertisements given their public nature. However, this proposed measure does not operate at the level of individual advertisements. The testing of an advertisement generation tool could in theory involve information not currently in the public domain, for

example, when testing the tool prior to its launch. While there may therefore be some limited expectations of privacy in relation to the use of information in such a setting, we provisionally consider that this proposed measure is unlikely to significantly interfere with individuals' rights to respect for private and family life.

5.59 The proposed measure does not specifically recommend the processing of any data relating to identified or identifiable individuals. However, depending on the input data on which an advertisement generation tool is trained, the tool may be able to create or modify advertisements which contain personal data. To the extent that this is the case, a provider may therefore process such personal data, as part of the testing process. A service provider may also choose to specifically test the tool by reference to personal data, for example, where they seek to test the ability of the tool to create 'deepfake' images of well-known and identifiable individuals.

5.60 We consider that, depending on the systems and processes used by service providers, the proposed measure may involve processing personal data, potentially at scale, and also potentially processing of special category data. The UK General Data Protection Regulation (GDPR) places specific restrictions on making decisions based solely on automated processing of personal data, where the decision has legal or similarly significant effects. These restrictions are imposed by Articles 22A to D of the UK GDPR.[191] So-called automated decision-making is permitted where service providers have appropriate safeguards in place. Additional restrictions also apply in relation to cases where special category data is used. The Information Commissioner's Office (ICO) has provided guidance on these matters.[192]

5.61 Service providers should ensure they, or any third parties that they outsource to, act in accordance with data protection legislation and relevant ICO guidance and consider the data protection principles of fairness, transparency and data minimisation in implementing this proposed measure.[193] Providers will also need to ensure that data protection impacts are limited to what is necessary for the legitimate purpose of complying with the fraudulent advertising duties. We consider that safeguards under data protection law, as explained in the various pieces of ICO guidance, will help ensure that the impact of processing (including automated processing) on data protection and privacy rights is minimised. As such, we do not consider that there is a less intrusive means to achieve the same aims.

5.62 Where a service provider collects any metrics because of this proposed measure, it will also need to ensure that any personal data is processed in accordance with the relevant data protection legislation.

5.63 To the extent that this proposed measure involves interference with individuals' rights to privacy, we consider the interference to be proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising (which this proposed measure is intended to help providers of Category 1 and 2A services to secure).

---

[191] Articles 22A to D were substituted for Article 22 by section 80(1) of the Data (Use and Access) Act 2025, with effect from 5 February 2026: see the Data (Use and Access) Act 2025 (Commencement No. 6 and Transitional and Saving Provisions) Regulations 2026, regulation 2(j), subject to regulation 5.
[192] ICO, 2026. Automated decision-making, including profiling. [accessed 30 April 2026].
[193] ICO, no date. UK GDPR guidance and resources. [accessed 30 April 2026].

Provisional conclusion

5.64 We provisionally consider that this proposed measure will deliver significant benefits. It will enable providers to identify, understand and take proportionate steps to address vulnerabilities that could allow fraudulent advertisements to be created, before and following deployment of an advertisement generation tool. This is likely to make it much harder for bad actors to create fraudulent advertisements via these tools, reducing the volume of fraudulent advertisements that users may encounter on the service. This is important as advertisement generation tools become more advanced and capable of producing increasingly convincing and persuasive content.

5.65 We acknowledge that there could be significant costs in developing testing infrastructure from scratch. However, in practice, we do not expect service providers to develop this infrastructure from scratch, as we are aware that they are already likely to have testing infrastructure in place for advertisement generation tools specifically, or AI tools more generally. Instead, we expect providers may incur some incremental costs in adapting their existing processes to align with our proposed measure. This means that in practice the costs of the proposed measure would likely tend to be materially lower than the upper bound of the estimates we have set out in paragraph 5.36.

5.66 In view of the prevalence and impact of fraudulent advertising, the growing risks posed by AI-generated fraudulent advertisements and the important role that testing can play in helping mitigate these risks, we provisionally consider that these incremental costs are proportionate.

5.67 Overall, we consider that any impacts on freedom of expression rights, privacy rights and rights in relation to data protection are proportionate to the Act's legitimate objective of protecting individuals in the UK from fraudulent advertising.

5.68 Our provisional view is therefore that it is proportionate to recommend that providers of Category 1 and 2A services that make advertisement generation tools available to advertising account holders to create advertisements that are placed on the service should undertake testing to identify and respond to vulnerabilities in the tool such that the tool is less likely to create fraudulent advertisements.

5.69 The full text of the proposed measure can be found in the draft Fraudulent Advertising Code of Practice for user-to-user services and in the draft Fraudulent Advertising Code of Practice for search services and is referred to as FAU F1 and FAS F1.