Protecting people from illegal harms online

Volume 5:
How to judge whether content is illegal or not? (Illegal Content Judgements Guidance)

Consultation

Published: 9 November 2023
Closing date for responses: 23 February 2024

Contents

Section

25. Introduction

25.1 Section 193 of the Act requires Ofcom to produce guidance for services about making illegal content judgements in relation to illegal content, or illegal content of a particular kind.

25.2 In this volume, we set out the approach we have taken to developing our 'illegal content judgements guidance', which explains to services how they should assess whether content is illegal or not.

25.3 We have also separately produced the following draft guidance:

- **Annex 5, 'Illegal Contents Judgement Guidance'**: Ofcom's Illegal Contents Judgement Guidance is intended to provide guidance to all services who may need to make judgements about whether content on their service amounts to an offence in the UK. As appropriate, it can help services judge whether content is illegal or not.

25.4 We are consulting on this draft guidance and invite feedback on our approach to developing it, as well as the draft itself. We have set out specific consultation questions in chapter 26 on issues where we would particularly welcome feedback and any further supporting information to inform our final version of this guidance document. See Annexes 1-4, for more information about how to respond to our consultation.

25.5 Having reviewed responses to this consultation, we will then publish our final decisions in a Statement and our final version of this guidance document.

26.Ofcom's Illegal Content Judgements Guidance

What is this chapter about?

The Act requires us to provide guidance to services about how they can judge whether a piece of content is likely to be illegal. In this chapter, we set out our proposed high-level approach to developing this Illegal Content Judgements Guidance ('ICJG'). We explain key terms relevant to illegal content judgements and key factors we considered when drafting the ICJG. We then set out the more detailed policy and legal considerations we have had to take into account when developing this guidance for specific offences.

What are we proposing?

The Act requires services to take action against content where they have reasonable grounds to infer that it is illegal. Broadly speaking there are two ways services can meet this duty. If they wish to, they can follow the process set out in our ICJG to determine when there are reasonable grounds to infer that a piece of content is illegal. Alternatively, they can draft their own terms and conditions in such a way that at a minimum all content which would be illegal in the UK is prohibited on their service for UK users and make content moderation decisions based on their terms and conditions.[1] In practice we expect that many services will take the second of these approaches, or a hybrid approach.

In the ICJG we are proposing to provide guidance to services to give them greater clarity about how they should assess whether content is illegal or not. The proposed guidance does not look at whether content may facilitate the commission of an offence. In our proposed guidance, we also set out our provisional view on: (a) what a service should consider to determine if it has 'reasonable grounds to infer that content is illegal content', and (b) what may constitute information that is 'reasonably available' to services when making an illegal content judgement.

What input do we want from stakeholders?

- Do you agree with our proposals, including the detail of the drafting? What are the underlying arguments and evidence that inform your view.
- Do you consider the guidance to be sufficiently accessible, particularly for services with limited access to legal expertise?
- What do you think of our assessment of what information is reasonably available and relevant to illegal content judgements?

Introduction

26.1 Section 192 of the Act requires services to take down content where they judge there to be 'reasonable grounds to infer' it is illegal, using 'reasonably available information' to make

[1] Services are free to take down content above and beyond what is illegal under the Act, so long as they make this clear in their terms of service, and that their content moderation practices result in the timely removal of illegal content as set out in the illegal content safety duties.

this judgement. Definitions of 'reasonable grounds to infer' and 'reasonably available information' are set out in paragraphs 26.12-26.15 and 26.21-26.41 respectively in this chapter. Section 193 of the Act creates a duty for Ofcom to produce guidance on how services can make illegal content judgements for the purposes of the takedown duty, the risk assessment duty and the safety duties more generally. In addition, the guidance will assist Category 1 service in relation to their duties to protect news publisher content and duties in relation to fraudulent advertising[2]. Together, these are the '**illegal content duties'**. The duty as set out in section 193 of the Act is fulfilled by Ofcom through the publication of what we refer to as the 'Illegal Content Judgements Guidance'.

26.2 In this chapter, we first set out our proposed high-level approach, explaining key terms relevant to illegal content judgements and key factors we considered when drafting the Illegal Content Judgements Guidance. We then set out the more detailed policy and legal considerations we have had to take into account when developing this guidance for specific offences.

26.3 The draft Illegal Content Judgements Guidance is published in draft form as part of this first online safety consultation and starts with an introduction setting out the background that services will need in order to better understand the guidance. There are then chapters for the priority offences themed by type of harm (e.g. 'terrorism'), setting out how services should consider the priority offences in making illegal content judgments, plus a chapter on relevant non-priority offences. Finally, for reference, there are Annexes in which the constituent parts of the offences are set out in the form of tables. These are drafted in more legalistic language than the main chapters.

26.4 This chapter sets out and consults on the approach we have taken with regard to the Illegal Content Judgements Guidance. We begin by explaining our approach to the new legal threshold of 'reasonable grounds to infer' and then we go on to discuss the nature of the Illegal Content Judgements Guidance and how we anticipate it will be used by services.

Our approach

26.5 The main points we have considered when developing our Illegal Content Judgements Guidance are:

a) the importance of explaining the term 'reasonably available information' and what this means for services in practice;

b) providing guidance to services on how they might infer the state of mind or 'mental element' of a the 'reasonable grounds to infer' test;

c) the need to explain the offences in the guidance, and how we propose to approach jurisdictional considerations; and

d) the importance of explaining our approach to the defence element of the 'reasonable grounds to infer' test.

26.6 In addition to explaining our approach in terms of the points outlined above, we have had to make a number of other decisions of fine detail. These decisions relate to how a specific

[2] These are additional duties in relation to Category 1 services. We will be consulting on the requirements of the Act for Category 1 services at a later date. If we need to amend this guidance we will consult on proposed amendments at the time.

offence can be understood when considering the concepts outlined above. Our approach to these offences can be found in paragraphs 26.82-26.263 of this chapter.

26.7 When considering the processes set out in the guidance, we have sought to balance the following factors, in line with our duties under sections 4A of the Communications Act 2003 and the requirement to carry out our duties in a way that is compatible with the Human Rights Act 1998:

a) user protection and safety;

b) user rights, including the importance of:

i) freedom of expression, and
ii) user privacy;

c) avoiding disproportionate interruption to law-abiding users of services; and

d) proportionality and practicability.

26.8 In particular, illegal content judgments made by services as a consequence of the Act may have a significant impact on the rights of individuals and entities to freedom of expression under Article 10 of the European Convention on Human Rights ('**ECHR**') and to privacy under Article 8 of the ECHR.

a) Any limitation on the right to freedom of expression must be prescribed by law, pursue a legitimate aim and be necessary in a democratic society.

b) Any limitation on the right to privacy must be in accordance with the law, pursue a legitimate aim and be necessary in a democratic society.

26.9 In order to be 'necessary', the restriction must correspond to a pressing social need, and it must be proportionate to the legitimate aim pursued. Both the definition of illegal content and the requirement for Ofcom to prepare this guidance are set out in the Act and pursue the aims of the prevention of crime, the protection of health and morals, and the protection of the rights of others. Ofcom has had careful regard to these rights in producing this guidance.

26.10 The need to balance these factors has been particularly important in our consideration of what information services should have regard to when making illegal content judgements. Further information on what we propose to consider as reasonably available information, and our approach to using that information, can be found in paragraphs 26.21-26.27 of this chapter.

26.11 Once we have issued the Illegal Content Judgements Guidance, we expect to continue to monitor its effectiveness and proportionality at appropriate intervals in order to keep it up to date.

Reasonable grounds to infer

26.12 Section 192 of the Act states that, when making illegal content judgements, a provider should consider whether it has "reasonable grounds to infer that content is illegal content".[3] The Act then states that reasonable grounds to infer content is illegal content might exist where a service:

[3] See section 192(5) of the Act.

a) has reasonable grounds to infer that the conduct or behaviour element (the act which constitutes the offence) is present and satisfied;

b) has reasonable grounds to infer that the mental element (the state of mind when committing the offence) is present and satisfied, and;

c) does *not* have reasonable grounds to infer that a defence to the offence may be successfully relied upon.

26.13 As explained in more detail in paragraphs 26.42-26.58, it is often hard to establish whether all three elements of the test are met. Inferring the state of mind or 'mental' element of a piece of content is a particularly difficult challenge at scale.

26.14 It is important to note that 'reasonable grounds to infer' is a new legal threshold and is different from the 'beyond reasonable doubt' threshold used by the criminal courts. The 'beyond reasonable doubt' threshold is a finding that *only* UK courts can reach. When the 'beyond reasonable doubt' threshold is found in UK courts, the person(s) responsible for the relevant illegal activity will face criminal conviction. However, when services have established 'reasonable ground to infer' that content is illegal according to the Act, this does not mean that the user will necessarily face any criminal liability for the content and nor is it necessary that any user has been prosecuted or convicted of a criminal offence in respect of such content.[4] When services make an illegal content judgement in relation to particular content and have reasonable grounds to infer that the content is illegal, the content *must* however be taken down.[5]

26.15 What amounts to reasonable grounds to infer in any given instance will necessarily depend on the nature and context of the content being judged and, particularly, the offence(s) that may be applicable.

Nature of the Illegal Content Judgements Guidance

26.16 When assessing their compliance with the relevant requirements[6] Ofcom may take into account whether services' judgements follow the approaches set out in the Act and in doing so we are likely to refer to the Illegal Content Judgements Guidance.[7]

26.17 The Illegal Content Judgements Guidance should be used as a reference for services when managing the risk of harm in compliance with their illegal content duties. The Guidance is *not* intended to serve as an alternative to services developing their own clear and accessible terms of service or community guidelines (or in the case of search services, publicly available statements which amount to the same), nor does the guidance amount to a full content moderation system.

[4] However, it should be noted that services have a duty under section 59 of the Act to report illegal CSEA content to the National Crime Agency.

[5] Section 10(3)(B) of the Act states that this duty applies "where the provider is alerted by a person to the presence of any illegal content, or becomes aware of it in any other way." The process of making an illegal content judgement, as set out in the Illegal Content Judgement Guidance, presupposes that the content in question has been brought to the attention of a moderator making such a judgement, and as a result this requirement is fulfilled.

[6] 'Relevant requirements' means (a) duties and requirements under this Act, and (b) requirements of a notice given by Ofcom under this Act (section 192(9) of the Act)

[7] These approaches are set out in section 192 of the Act.

26.18 Services are free to take down content above and beyond what is illegal under the Act, so long as they make this clear in their terms of service, and that their content moderation practices result in the timely removal of illegal content as set out in the illegal content safety duties. For example, if a service's terms of service are written in such a way as to permit less content than the Act allows and such content is removed as content which violates the service's terms of service rather than as illegal content specifically, this would be sufficient to meet the illegal content takedown duty.[8] Therefore, these services would not need to make illegal content judgements. For more information on services' illegal content safety duties, see Chapter 2 of this consultation.

26.19 It is our assumption that most services will take the approach explained above as it allows them to freely moderate content based on their own terms of service (or equivalent), rather than having to make illegal content judgements based on our guidance. Many services will already have terms of service or their equivalent in place that are more expansive than the Act in defining what content may be deemed violative and will already be taking down above and beyond what the law requires in terms of preventing users encountering illegal content. These services may use the Illegal Content Judgements Guidance to check that their terms of service, community guidelines or publicly available statements are capturing everything that UK law requires of them.

26.20 At the other end of the spectrum, there will be services that decide to moderate content in such a way that *only* illegal content as defined by the Act is blocked or removed. This guidance is also aimed at services taking this approach as they will be required to make illegal content judgements.

Reasonably available information

26.21 Section 192 of the Act states that illegal content judgements "are to be made on the basis of all relevant information that is reasonably available to a provider". Below we explain and invite comments on how we have reached our view on the information we might expect services to base their illegal content judgements on.

26.22 The Act states that two factors are particularly relevant in considering the information that is reasonably available to a service provider:

a) "the size and capacity of the provider"; and

b) "whether a judgement is made by human moderators, by means of automated systems or processes or by means of automated systems or processes together with human moderators."

26.23 We used both of these factors when considering the availability and relevance of reasonably available information generally, and in relation to each priority offence.

26.24 Below we have outlined our provisional view on what may constitute information that is 'reasonably available' to services when making an illegal content judgement. However, services will need to consider what is reasonably available on a case-by-case basis as what may be relevant and reasonably available for the illegal content judgement may differ

[8] This duty is set out in section 10(3) of the Act: a duty to "operate the service using proportionate systems and processes designed to… where the provider is alerted by a person to the presence of any illegal content, or becomes aware of it in any other way, swiftly take down such content."

depending on the type of content, the offence it may amount to, the service's Terms of Service or Publicly Available Statement and what other information is available.

26.25 As emphasised below, processing some of the types of information ('data') listed below has potential implications for users' right to privacy.[9] Services also need to ensure they process personal data in line with data protection laws. In particular, the likelihood is high that in considering U2U content a service will come across personal data including special category data and possibly criminal offence data, to which UK data protection laws apply.

26.26 For example, depending on the context, reasonably available information may include:

a) **Content information:** The type of information that is most likely to be reasonably available to services when making an illegal content judgement is the information contained within the content itself e.g., what the text displayed within an image reads.

b) **Complaints information:** Services may also want to consider information they receive which is contained in a complaint from a third party (e.g. law enforcement, a trusted flagger, a user). When using this type of information, the service may also want to consider who that third party is and how robust and reliable the information may be based on this.

c) **User profile information:** This refers to information relating to the user that may be found on their profile. A user might for example give details about their age when they register a user profile or in other user profile features such as in their 'bio'. This information may not be routinely accessible to moderators and services would need to consider privacy and data protection laws.

d) **User profile activity:** This refers to other content posted on a profile, for example content posted immediately before and after the content being considered. This information may not be routinely accessible to moderators and services would need to consider privacy and data protection laws. It could also refer to other kinds of information, for example user behaviour monitoring. However, in this consultation we are not proposing that services should use any user behaviour monitoring technology and so we do not consider that information derived from such technology would be 'reasonably available'.

e) **Published information:** Finally, services may need to consider information that has been published where it is credible and relevant, e.g. whether an entity appears on the list of the Financial Conduct Authority's regulated entities or not.

26.27 However, we recognise that in certain instances services may have access to information, which is relevant to a specific content judgement, but which is not either typically available to all services, which would require significant resources to collect, or the use of which would not be lawful under data protection or privacy laws. When making illegal content judgments, services should continue to have reasonable regard to any other relevant information to which they have access, above and beyond what is set out in the Illegal

[9] We recognise that the majority of specific items of content which services are likely to be considering is likely to be content that has been communicated publicly. However, as explained in our draft guidance on when content should be considered to be communicated publicly or privately, in Annex 9, we consider that there could be some circumstances in which a person could still have a reasonable expectation of privacy (for the purposes of Article 8 ECHR) in relation to content which is nonetheless considered to be communicated publicly for the purposes of the Act.

Content Judgements Guidance but only so long as this information is processed lawfully, including in particular in line with data protection laws.

Reasonably available information and size and capacity of service providers

26.28 As set out above, section 192 states that Ofcom must take into consideration the 'size and capacity of the provider' when determining which information it deems to be 'reasonably available'. As such, when considering which information we would specify as reasonably available information for each offence, we assessed it on its availability to large, small and micro services.

26.29 When developing our policy on reasonably available information for different sizes and capacities of providers, we considered segmenting in the following ways:

a) Having a universal base level of information that is deemed to be reasonably available to *all services*, plus an additional layer of information that is reasonably available only to the *largest services.*

b) Having a universal base level of information that is deemed to be reasonably available to *all services*, plus a carve-out for *the smallest services* who might struggle to meet the standard requirements.

c) Not segmenting reasonably available information on the base of size and capacity at all.

26.30 In the case of approach a) and b), definitions of size would be based on the approach taken as part of our work on the application of Codes to services, which would be based on the approach set out in Volume 4 Section 11 of this consultation.

26.31 We found that an average or larger service may have access to information that a micro-business does not have access to; however, at this stage we did not consider this information relevant to a content judgment about the *illegality* of a single piece of content. In other words, in our view the reasonably available information for an average or larger service, that is relevant to an illegal content judgement, is also reasonably available to the smallest services. This does not mean that this information will always be present at the same scale or frequency or through the same channels, only that it is theoretically available in some circumstances. For example, law enforcement could inform a service of a conviction through a trusted flagger scheme, or through a one-off flag to a service through a standard email address.

26.32 Furthermore, we were unable to identify a specific piece of information that would *always* be available to a larger service (it will be dependent on the type of service and its capacity). We recognise, however, that some services may have access to further information beyond what is specified in the Guidance. Therefore, we advise that where such information is relevant to content judgements as set out in this Guidance, services should consider this information as appropriate so long as this information is processed lawfully, including in particular in line with data protection laws.

Reasonably available information and automated systems or processes

26.33 Section 192 furthermore states that, when determining what information is reasonably available to a provider, Ofcom should also consider "whether a judgement is made by

human moderators, by means of automated systems or processes, or by means of automated systems or processes together with human moderators."

26.34 At this stage, the automated content detection technologies that we are proposing to recommend in our Codes are:

a) hashing technology recognising child sexual abuse material;

b) URL detection technology recognising URLs which have previously been identified as hosting child sexual abuse material (CSAM); and

c) keyword search to detect content containing keywords strongly associated with the sale of stolen credentials (i.e. articles for use in fraud).[10]

26.35 The first two technologies, hashing and URL detection, function by matching a known item stored in a database to another item on the platform. Where the database concerned only contains content which has been judged to be illegal content, an illegal content judgment would have been made before the inclusion of each item of content in the database, and so this technology would not typically require a different approach to illegal content judgements.

26.36 Some databases have been established for different purposes and are based on different governance standards and quality controls – for example, some services' internally maintained databases which contain content found to violate the service's terms of service. These databases may not involve any illegal content judgments being made at all.

26.37 We are not proposing to recommend any other kinds of automated technology that might be sufficiently accurate at picking up illegal content. It nevertheless remains open to services to use such technologies, pursuant to their own terms of service.

26.38 Our draft guidance therefore proposes a 'technology-agnostic approach' to reasonably available information and to illegal content judgements in general. We have set out which information we believe is reasonably available to a service, regardless of technology used to collect it, on an offence-by-offence basis. It is our understanding that, while automated tools could be used to collect *more* of this information or to do so more quickly, there is no additional class of information which automated tools could have access to that human moderators could not. We therefore take the view that information may be collected using any approach the service prefers, so long as when it is factored into an illegal content judgement, this is done in a way which allows a *reasonable* inference to be made.

Reasonably available information and permitted activity

26.39 We have identified that in the case of certain priority offences specified in the Act, consideration of the content alone may not enable services to make an illegal content judgment.

26.40 However, we have given particular thought to two groups of priority offences where the crucial offline information is whether or not a particular user has an appropriate registration or authorisation. This is limited, specific information which a service is unlikely to have, but which it could choose to get. The offences in question are:

[10] Section 7 of the Fraud Act 2006; section 49(3) of the Criminal Justice and Licensing (Scotland) Act 2010

a) offences related to the exposure for sale of firearms which can be legally bought and sold with a proper certificate or authorisation ('firearms offences')[11]; and

b) offences related to financial services as set out in the Act in schedule 7 ('financial services offences')[12]

26.41 For more information on our approach to these offences see sections 26.184-26.190 and 26.210-26.212 of this consultation chapter.

Approach to inferring state of mind or 'mental element' of an offence

26.42 The 'mental element' of the offence refers to the state of mind of the person who is potentially committing an offence. In legal terminology this is known as 'mens rea.' Almost all the priority offences laid out in schedule 7 of the Act have a state of mind requirement, which must be satisfied in order for reasonable grounds to infer to exist. Some of the most common types of mental elements relating to the priority offences involve:

a) acting with intent;

b) acting recklessly;

c) acting dishonestly; or

d) acting with knowledge.

26.43 We acknowledge that inferences about state of mind are particularly difficult in online situations, where contextual clues are often not apparent and, for example, what would be an obvious joke or piece of sarcasm in an offline context might not appear so obvious when online. We also acknowledge that conclusions about state of mind in criminal cases are nuanced, and usually draw upon an extensive suite of evidence which is not reasonably available to a service moderating a single piece of content. However, neither Ofcom nor in-scope services can put aside the state of mind or 'mental element' requirement as this is a part of the 'reasonable grounds to infer' threshold, as established by the Act. Our approach therefore seeks to take sufficient account of state of mind requirements as set out in common law whilst also being realistic about how quickly and accurately these inferences can be made by a moderator assessing a single piece of online content. However, as set out above, so long as their content moderation practices result in the timely removal of illegal content as set out in the illegal content safety duties, services are free to take down content pursuant to their own terms and conditions, community guidelines or publicly available statements, regardless of whether the state of mind requirement is present. Therefore, the state of mind requirement relating to illegal content judgements does not in practice constrain services' ability to take down other harmful, but not illegal, content where they choose to do so.

Inferring conduct, behaviour and state of mind when content has been forwarded, shared or reposted

26.44 The definition of 'illegal content' in the Act means that services must consider someone's conduct and state of mind in order to identify illegal content. Each of these must also be

[11] The offences in question are set out in schedule 7, section 19(a) and (b) of the Act. We use the term 'firearms' to refer to firearms which can be legally possessed, purchased and acquired in the UK with the correct certification.

[12] The offences in question are set out in schedule 7, section 31 of the Act.

considered in context (in particular, there are priority offences which are committed, or not, depending on who is likely to see the content, or the forum in which it is shared).

26.45 It follows that, when a piece of content has been shared, forwarded or reposted by a new user, a service should treat this as a new piece of content for the purpose of an illegal content judgement. Of course, it is possible that both the original uploader, and the user who shares the content onwards, may each have the state of mind needed for the content to be an offence, and so each of their posts could be posts of illegal content.

26.46 On the other hand, the same video may be posted by one user in a way that is supportive of the message conveyed, and posted by another to challenge, criticise or 'call out' the same message. A fundamental part of the content would be the same in this instance, but the inferences that could be made about the state of mind in each case would be very different.

26.47 To give a more specific example, the offence of harassment depends on the relationship between the user posting content (in a particular location at a particular time) and the person or persons accessing the content. Posting a photograph of a kitchen implement used in domestic abuse may amount to harassment when posted by the abuser to the abused person. The same cannot be said if a new user shares the content (with different intentions), or the person viewing the post is different, though the content in both cases has not changed.

Inferring conduct, behaviour and state of mind when content has been posted by a bot

26.48 Bots are an umbrella term that refers to a software application or automated tool that has been programmed by a person to carry out a specific or predefined task without any human intervention. Bots are often employed on services to post content at scale without the need for repeated human intervention. For example, a bot may be used to respond to another user's post on a brand's social media platform, or to automatically 'cross-post' content across multiple services operated by the same user. Bots are also used for malicious purposes such as spreading disinformation, distributing malware or by overloading a system or server with traffic, resulting in a service being unable to operate as normal or at all.

26.49 Section 192 of the Act states that where content has been posted by a bot, inferences about the conduct and the presence of the mental element, and any defences, should be made by considering:

a) the actual person controlling the bot or tool, where this is known to the service; or

b) the person who may be assumed to be controlling the bot, where the actual identity of the person is not known.

26.50 Having considered all offences, we believe that this inference will normally be fairly straightforward to apply, since the analysis will not be very different whether the content is posted by a human directly or by a human controlling a bot. However, it may make a substantive difference to judgements about the foreign interference offence. To make such a judgement, the service will need to establish whether the 'conduct in question, or course of conduct of which it forms part, is carried out for or on behalf of a foreign power'. Knowledge of the offline identity of the person controlling the bot (or assumed to be controlling the bot) may therefore often be essential, compared to most other offences where a bot is likely to be used, where the identity of the user posting the illegal content may not be so material to whether an offence has occurred. For instance, if a bot is being

used to post content glorifying the acts of a known terrorist organisation, this may amount to an offence of inviting support for a proscribed organisation. This will be the case based on the content alone, and the identity of the user posting the content – or, in this case, controlling the bot who is posting the content – would not be relevant when making an illegal content judgement.

26.51 As such, we have provided general principles in relation to bots in our offence-agnostic introductory sections, and specific guidance on making inferences in relation to bots for foreign interference offences *only*. To promote clarity and avoid unnecessarily complicating illegal content judgements, we have not steered services to consider bots in the sections giving guidance on other offences.

Inferring presence or satisfaction of the mental element of 'knowledge'

26.52 Several priority offences, including offences to do with child abuse imagery[13] or possession of extreme pornography[14], include a state of mind requirement (or 'mental element') of 'knowledge'. For an offence to have occurred a defendant must *know* that what they have uploaded or shared etc was the image in question.

26.53 We have taken the view throughout that it is reasonable to infer that users are aware of the nature of the content they upload. While we know this is not necessarily true in all circumstances, and that some users upload content they have only skimmed or reviewed a part of, we do not consider it plausible that most users are unaware of the nature of most content they upload.

26.54 We are aware that there is research to suggest that a significant and perhaps a very significant minority of users do not look at content they forward.[15] However, we have provisionally taken the view that most do (or have read them offline etc), and it is therefore still reasonable to infer that users who forward and onward share content are aware of what it is. We do not consider it would be proportionate or sensible for services to need to investigate this in order to make a judgment on whether, for example, a child abuse or extreme pornography image should be treated as illegal content.

26.55 We have adopted an approach in line with the above reasoning in relation to all offences where 'knowledge' of the nature of the content forms part of the state of mind criteria. However, it is not possible to make generalised statements about other kinds of state of

---

[13] For example, section 1 of the Protection of Children Act 1978; article 3(a) of the Protection of Children (Northern Ireland) Order 1978 -.

[14] Section 63 of the Criminal Justice and Immigration Act 2008; the defendant must have knowledge that they have the image in question in their possession, although not that the image is extreme pornography.

[15] See Gabielkov, M., M., Ramachandran, A., Chaintreau, A. and Legout, A., 2016. Social clicks: What and who gets read on Twitter?, *ACM SIGMETRICS Performance Evaluation Review*, 44(1), pp. 179-192. [accessed 19 September 2023]; Holmström, J., Jonsson, D., Polbratt, F., Nilsson, O., Lundström, L., Ragnarsson, S. and Carlsson, N., 2019. Do we read what we share? Analyzing the click dynamic of news articles shared on Twitter, *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 420-425. [accessed 19 September 2023]; Chan, H. Y., Scholz, C., Baek, E. and Falk, E., 2015. The gap between sharing and reading news on social media: A multi-method investigation, *Social Media + Society*, 1(2) [Accessed 19 September 2023]; and Ward, A. F., Zheng, J. and Broniarczyk, S. M., 2023. I share, therefore I know? Sharing online content-even without reading it-inflates subjective knowledge, *Journal of Consumer Psychology*, 33(3), p. 469-488. [accessed 19 September 2023].

mind criteria, and as such we have addressed these in our guidance on an offence-by-offence basis.

Inferring 'possession'

26.56 Sometimes the conduct part of an offence occurs when content is 'possessed'.

26.57 'Possession' is defined as being met when the images are in the custody or control of the suspect i.e. so that they are capable of accessing, or in a position to retrieve the image(s); and the suspect must have known that they possessed an image or group of images on the relevant device/devices.[16] In addition, the definition of illegal content includes that 'content consisting of certain…images…amounts to a relevant offence if—…the possession…of the content constitutes a relevant offence'.

26.58 Services may therefore reasonably infer that if the content appears, 'possession' is met.

Approach to explaining the offences in the Illegal Content Judgements Guidance

26.59 The priority offences set out in the Act are complex and varied, covering many different aspects of UK law across three different jurisdictions: England and Wales, Scotland and Northern Ireland. Few of the offences were written purely for an online context, many offences also overlap with one another, and one piece of content may have the potential to amount to multiple offences.

26.60 Our approach to offences and their relationship to online content is guided by the law. We have set out our approach to the various challenges posed by the translation of criminal offences into illegal content online in the paragraphs below.

26.61 Recognising the varying degrees to which services are likely to have access to UK legal advice, in the guidance we have endeavoured to keep our summary of offences as clear and simple as possible. However, any accurate representation of the law will necessarily require the use of legal terms. In some areas the law is uncertain and in others it is highly technical.

26.62 In the criminal courts, whether an offence has been committed is often a question of fact which is for a jury to determine, and so there is no body of law providing further guidance on the meaning of terms which may appear general or ambiguous. Ofcom has no power to resolve such uncertainties. Just as the UK criminal law requires juries to use their own judgment in applying the definition of the offence to the facts of the case they are hearing, services must do the same in relation to the content they must consider.

Approach to offences unlikely to result in content which amounts to an offence

26.63 It is our assessment that some kinds of priority offences are unlikely to give rise to illegal content, as defined by the Act and, as a result, services are unlikely to need to make illegal content judgements with reference to these offences.

26.64 Section 11 of the Terrorism Act 2000 provides an example of this issue, as it creates an offence of either 'belonging' or 'professing to belong' to a proscribed organisation. In our

[16] *R v Okoro (No 3)* [2018] EWCA Crim 1929.

view, services are far more likely to encounter content which has the potential to amount to an offence of 'professing to belong' than to an offence of 'belonging' in itself. As such, in order to ensure services are focusing their resource and attention on offences where it is warranted, we have only provided substantive guidance on the 'professing to belong' part of this terrorism offence.

26.65 In other cases, we have focused more on offences related to the priority offence than the priority offence itself. Each of the priority offences can be committed not only in its own right, but also in its 'inchoate' form. Inchoate offences happen when someone is involved in another offence in a way which makes them guilty, without actually committing the offence themself. For example, a person may 'assist' in a robbery if they drive the getaway car. They did not carry out the offence, but they were involved in it. It is our provisional view that the most common ways in which an inchoate offence might be committed online are by encouraging or assisting a priority offence or by conspiring (i.e. making an agreement) to commit a priority offence.

26.66 In the context of the Illegal Content Judgements Guidance, the inchoate offences are particularly important where the conduct elements of the main offence do not take place through content. For example, it is not likely to be possible for an under 18 year old to buy a crossbow on a U2U service or via illegal search content, because both the provision of the item and the provision of money in exchange for the item take place offline or on unregulated services. It is, however, possible to commit a related offence of 'encouraging' or 'assisting' an under 18 year old to buy a crossbow online, and so we have focused on those sorts of related offences in our guidance.[17]

26.67 The offences of encouraging or assisting are subject to a defence where the conduct in question was 'reasonable'.[18] This defence is to protect people who might unintentionally encourage others to break the law (for example, if they falsely believe that circumstances exist, which don't). In our guidance, we have taken the view that the service is only likely to have reasonable grounds to infer this when it (i) no longer has reasonable grounds to infer that the content is illegal in any event; or (ii) a defence which used to exist no longer does. Therefore we mostly do not consider it relevant and have not included it in our guidance.

26.68 Even where the priority offence does not or is unlikely to give rise to priority illegal content, U2U services will still need to consider in their risk assessments whether they may be used to commit or facilitate the commission of an offence. This means there may be a class of content which might be said to 'facilitate' the commission of an offence, without actually being illegal content. Such content would not be subject to the takedown duty, which only applies to content that is itself illegal content, but it may nevertheless be appropriate and proportionate to remove it from services in compliance with the safety duty more generally. However, this raises potentially very significant issues of freedom of expression and proportionality. In any event, the structure of the Act means that any action in regard to this type of content would be outlined in Codes of Practice and not in the Illegal Content Judgements Guidance, which focuses solely on illegal content. It is therefore a question to which we will return in future phases of work rather than in this chapter.

[17] Section 2 of the Crossbows Act 1987.
[18] Section 50 of the Serious Crime Act 2007.

Approach to relevant non-priority offences

26.69 In addition to priority offences, the Act defines a second class of offences as 'other' relevant offences. The full criteria defining this class of offences can be found in section 53(6) of the Act, but broadly it encompasses non-priority offences where the victim (or intended victim) is an individual, with certain exclusions including legislation on intellectual property and trading standards. Relevant non-priority illegal content is subject to the same takedown duty as priority illegal content.

26.70 In recognition of the quantity and complexity of offences which could be included within the scope of the definition of relevant non-priority offences, we propose to take the approach of providing in-depth guidance only on the priority offences set out in schedules 5 to 7 of the Act, with limited exceptions (see below). We do not consider it proportionate to expect services to anticipate all relevant non-priority offences other than the ones we have provided guidance on in the Illegal Content Judgements Guidance. We would, however, expect services to respond to content amounting to relevant non-priority offences where they have been made aware by law enforcement or a court order that it has been implicated in a successful conviction. We propose to set this approach out in a separate chapter of the guidance, in order to give services a clear steer on Ofcom's expectations regarding the majority of relevant non-priority offences.

Inclusion of selected non-priority offences

26.71 Notwithstanding the approach summarised above, we have anticipated a demand from services for additional guidance on certain relevant non-priority offences which service providers are likely to come across relatively often in day-to-day running of their service.

26.72 We have included the following relevant non-priority offences, which are created by the Act, in our guidance:

a) False communications offence; section 179 of the Act

b) 'Epilepsy trolling', i.e. offence of sending or showing flashing images electronically; section 183

c) Assisting or encouraging serious self-harm; section 184

d) 'Cyberflashing', i.e. offence of sending photograph or film of genitals; section 187[19]

26.73 We are consulting on the offence of assisting or encouraging self-harm now on the basis that it, too, may come into force before we issue our first Illegal Content Judgments Guidance. If it is not brought into force by then, we will not include it in our guidance, and will instead update our guidance as appropriate when this happens.

26.74 We are not proposing to give additional guidance on a fourth offence created by the Act, that of threatening communications, because we consider that any content which amounted to this relevant non-priority offence would in any event amount to the priority offence of threatening behaviour under section 38 of the Criminal Justice and Licensing (Scotland) Act 2010.

[19] As a result of section 187 of the Act, this offence was integrated into the Sexual Offences Act 2003 as section 66A.

Jurisdictional considerations

26.75 The priority offences outlined in the Act include offences from each of the three different UK jurisdictions: England and Wales, Scotland, and Northern Ireland.

26.76 The Act states that "[f]or the purposes of determining whether content amounts to an offence, no account is to be taken of whether or not anything done in relation to the content takes place in any part of the United Kingdom."[20] The Explanatory Note to the Act explains that the effect of this is that "content does not need to be generated, uploaded or accessed (or have anything else done in relation to it) in any part of the United Kingdom to amount to an offence under this provision. This is the case regardless of whether the criminal law would require the offence, or any element of it, to take place in the United Kingdom (or a particular part of it)."

26.77 This means that services do not need to consider the jurisdiction in which a user is located when posting content, only that the content in question is accessible to users within the United Kingdom. For example, a piece of content posted in Northern Ireland can be illegal content pursuant to a Scottish offence, regardless of whether an equivalent offence exists in Northern Ireland.

26.78 Due to the significant overlap between laws in the United Kingdom's three legal jurisdictions, the practical impact of jurisdictional differences is limited. There are, however, isolated cases in which a priority offence in one part of the United Kingdom is different from the other jurisdictions. Where this is the case, we have set out an appropriate approach to be taken in our guidance.

26.79 The same logic applies if the content was posted outside the UK. This means that, for example, a non-UK user harassing another non-UK user online will create illegal content for the purposes of the Act if the service is regulated by Ofcom. However, the interpretative rule in the Act applies only to what happens in relation to the *content*. It does not affect, for example, any offline circumstances required for the offence to be committed. We consider, for example, that the word "sale", which is used in several priority offences, should be construed as sale *to persons in the UK* unless the underlying priority offence has extra territorial effect. Similarly, for any inchoate offences to be committed, the offence being encouraged, assisted or conspired to etc would need to be an offence within the territorial jurisdiction of the UK.

Combining of related offences

26.80 Where there is significant overlap between priority offences, we have factored this into our approach in order to simplify the process for services and reduce the number of offences they are being asked to consider. For example: many of the Northern Irish and English/Welsh offences are effectively identical to one another, and we have conflated them for the purposes of the guidance. Some other offences comprise the same or very similar elements. For example, coercive and controlling behaviour is an offence with many elements, but to the extent it manifests in content online it is likely to amount to other priority offences as well. Again, we have combined the offences where possible.

26.81 Finally, where one offence is 'broader' and easier to show than another, we have prioritised the broader offence. By way of example, racially and religiously aggravated versions of

[20] Section 59(11) of the Act.

certain priority offences under the Public Order Act 1986 are also priority offences.[21] In theory, in order to identify a racially aggravated offence, the service would not only need to identify all the elements of the Public Order Act offence, but also all the elements of racial or religious aggravation. But in practice, in order to identify the content as illegal content, the service would only need to show the elements of the underlying Public Order Act priority offence, because that would be all that was needed for the takedown duty to be triggered. The racial aggravation would of course be likely to make the case more serious and urgent, but that would be more a matter of prioritisation of content for review than of identifying illegal content.

Approach to inferring conduct or behaviour where information is not available

26.82 In some cases, the conduct or behaviour element of an offence requires inferences to be made about something which is unlikely to ever be apparent on the face of a piece of online content. In these cases, we have evaluated what inferences may reasonably be made from the contextual information that *is* available to a moderator on a case-by-case basis. Where appropriate, we have indicated ways in which the conduct or behaviour criteria of an offence may be inferred to be present and satisfied based on the *likelihood* that this is the case.

Approach to inferring defences

26.83 If a service has reasonable grounds to infer that a defence may be successfully relied on, the content will not amount to illegal content for the purposes of the Act. The person who needs to be considered here is the same person whose actions and state of mind in relation to the content may involve a criminal offence. This will most often be the person posting, uploading or sharing the content, but this may not always necessarily be the case.

26.84 In law, there are number of general defences which are available in relation to a range of offences rather than those which are available only in relation to a particular offence. They include necessity and duress; and insanity/involuntary conduct. We believe that general defences are unlikely to be relevant to a service's illegal content judgments as it is difficult to imagine circumstances in which services would have reasonable grounds to infer that they arise. As such, we propose not to outline these general defences in the guidance.

26.85 In cases where a relevant defence is that the user has a 'reasonable excuse' to believe something to be the case, we have considered what 'reasonable' might mean in an online context.

Offence-specific considerations

26.86 In addition to the substantial policy proposals outlined above, there are some more offence-specific and detailed policy and legal considerations this chapter proposes guidance on. Please see below for more information on these offence specific considerations.

[21] Section 31 Crime and Disorder Act 1998.

Terrorism offences

26.87 The priority terrorism offences are set out in schedule 5 of the Act and covered in Chapter 2 of our draft guidance.

26.88 Some of the terrorism offences are likely to be much easier to make reasonable inferences about than others. In our draft guidance, we have steered services to begin by considering the offences we think are likely to be least difficult to identify (principally offences with the lowest 'state of mind' requirements), rather than the offences that are most likely to occur.

26.89 Below, we set out the reasoning behind our approach to terrorism offences which have raised particular legal and policy considerations.

Proscribed organisation offences

26.90 In our provisional view, by far the most straightforward offences in schedule 5 of the Act are those relating to 'proscribed organisations' – groups that have been proscribed in the UK because of their terrorist activities. The state of mind requirements for these offences are for the most part low, often involving knowledge and/or recklessness and the list of proscribed organisations is publicly available. We have therefore set these offences out first. For reference, the state of mind requirements for the offence of professing to belong to a proscribed organisation require intent to profess to belong to the proscribed organisation. In our view, as part of its illegal content judgement, once a service has inferred that content is professing to belong to a proscribed organisation, it can infer intent to do so.

26.91 One of the proscribed organisation offences relates to publishing an image of an item of clothing, or any other article, in such a way or in such circumstances as to arouse reasonable suspicion that the person is a member or supporter of a proscribed organisation. We considered whether our guidance should include guidance on what sort of articles may arouse this suspicion. However, we are not aware of any publicly available, reliable list of such articles. Such lists of which we are aware tend to include articles and images associated with particular ideologies, rather than of proscribed organisations as defined in the UK. We therefore do not intend to steer services to consider any such list, although we intend to keep this under review pending the production of any suitable resource. Instead, in our guidance we propose to say that services which are aware of logos, flags or other iconography associated with proscribed organisations should factor these into their content judgements where appropriate. This could be ascertained through in-house specialist teams or through engagement with third party organisations that maintain databases of such information. Services should also have due regard to any evidence about proscribed organisation iconography submitted to them by law enforcement.

Information likely to be of use to a terrorist

26.92 Information likely to be of use to a terrorist is information that is, of its very nature, designed to provide practical assistance to a person committing or preparing an act of terrorism. It is highly likely to be illegal content, because the conduct and state of mind either of the user posting the content, or of other users who are viewing such content, is relevant, and all that is required is for one of the persons posting or viewing the content to be broadly aware of what the content contains, and whether by its very nature, it is designed to provide assistance to commit or prepare an act of terrorism.

26.93 There is a defence of 'reasonable excuse' which may be harder for services to make reasonable inferences about, but they only need to consider it if there are positive grounds

to do so. One kind of reasonable excuse is when the collection or viewing was made for journalistic or academic purposes.

26.94 In such cases, we provisionally consider that reasonable grounds to infer that this defence is available will not arise where content has been communicated to the general public. An audience which is larger and/or more general is more likely to contain users who would *not* access the content for a specific, legitimate reason (that is, for journalistic or academic purposes) and it is therefore less reasonable to say that the user collecting the information had a 'reasonable excuse'. Similarly, each person accessing or viewing the information would need their own 'reasonable excuse'. Furthermore, any content made available outside a limited group has the potential to be shared and spread in a way which the user sharing the information originally cannot control. This means that the user collecting the information online to begin with is unlikely to have a *reasonable* excuse for the collection.

26.95 Generally speaking, our provisional view is that the defence will rarely be relevant on U2U services and search services, and even if it was, a service is unlikely to have any information which would suggest that the defence may be available.

26.96 On that basis, we consider it likely to be fairly straightforward for services to make illegal content judgments in relation to this offence, and have included it next in our draft guidance.

Terrorist training offences

26.97 Two priority offences relate to 'training'.[22] Of these, perhaps the most important is that in section 54 of the Terrorism Act 2000 - 'providing weapons training'. This covers content which, in and of itself, provides instruction or training in the making or use of various weapons. It is triggered whether the training or instruction is being made available generally or to one or more specific persons, and there is no state of mind requirement. This means that the offence is likely to be important outside the context of suspected terrorism.

26.98 Jurisdictional considerations play no part in this analysis. This is not only because of the definition of illegal content - the underlying priority offence applies extra-territorially (see section 17 of the Terrorism Act 2006).

26.99 The majority of the weapons concerned are of a nature which means that questions about the users' and viewers' purpose are unlikely to arise – it is difficult to see why any person would need training found on a U2U service or via a search service for radioactive material or weapons designed or adapted for the discharge of any radioactive material, explosives or chemical, biological or nuclear weapons.

26.100 However, the offence also covers instruction and training in the use of firearms. In particular, as set out below, we consider it likely to be relevant to 3D printing instructions for firearms.

26.101 A defence is available if the services has reasonable grounds to infer that the user's action or involvement was *wholly* for a purpose other than assisting, preparing for or participating in terrorism. Evidence of clear non-terrorist purpose is most likely to arise in relation to firearms. It should be noted that providing weapons training for legal purposes, for example as part of a rifle club, is not illegal. However, services are not required to ask the users posting and users viewing the content about their purposes, before making an illegal

[22] Section 54(1) and 54(3) Terrorism Act 2000; section 6 Terrorism Act 2006.

content judgement. In the case of 3D printing instructions for firearms, we are consulting on our provisional view that it is unlikely that a service would have reasonable grounds to infer that the purpose was wholly non-terrorist.

Dissemination of terrorist publications and encouragement of terrorism

26.102 In order for content to amount to the offence of 'dissemination of terrorist publications'[23], a service must have reasonable grounds to infer that the publication in question was posted in a location where it could be seen by at least one person who could possibly (as opposed to will probably or certainly) be encouraged by it to commit an act of terrorism, and that the user who posted it either intended or was reckless that this would happen.

26.103 In considering what would amount to reasonable grounds to infer this, we thought about the likelihood of people posting content of this nature *without* recognising the risk that a person might be encouraged by it to commit a terrorist offence. We took the view that if a terrorist publication has been uploaded to a location that can be accessed by anyone (for example a website or social media profile accessible by other users), it is reasonable to infer that it may be seen by somebody who could be encouraged to commit, prepare or instigate terrorism, and that most users posting such content would recognise this. We are therefore proposing to steer services to remove such content whenever it has been posted in a location that is easily accessible by other users as stated above, absent relevant defences and dependent upon the satisfaction of the other elements of the offence.

26.104 Beyond that, we are proposing that – in most cases – it is reasonable to infer that content is likely to be seen by someone who could be encouraged to commit terrorism where the content is posted to a location accessible by at least one other user. This is due to the highly interconnected nature of the internet, and the ability of content such as terrorist publications to spread quickly despite a relatively small original audience. We note, in particular, evidence that terrorist propaganda sites use a network of private chats to which joining links are posted openly but secretively in order to disseminate their publications to groups of people who onwards post them.[24]

26.105 We recognise that this is a judgement which may not be true in every case - for example, if content were shared with a group composed only of academic researchers into terrorism, or journalists writing about terrorism, all of whom were motivated by a dislike of terrorism, those users would be unlikely to be encouraged to terrorism and the original poster may not be reckless in assuming that they would be unlikely to disseminate the content further. However, from the point of view of the service provider, the majority of cases they see are very unlikely to involve groups of this nature. In the time available, it is likely to be particularly difficult for a service to recognise such a group if it did exist, given that it is possible for persons who hold themselves out to be or actually are journalists and academics to be sympathetic to terrorism.

26.106 On balance, we therefore provisionally consider it reasonable for services to draw the inference that such dissemination amounts to the offence.

---

[23] Section 2 Terrorism Act 2006. The definition of "terrorist publication" is set out in Annex 10 of our consultation.

[24] Hall KC, J. 2023. Online Safety Bill: Distinguishing between public and private communication. [accessed 19 September 2023].

26.107 The separate offence of 'encouraging terrorism'[25] involves some similar ideas – many terrorist publications also encourage terrorism. However, this offence can only be committed where the content concerned has been 'published' to members of the 'public'. There is no comprehensive definition of these terms, but it is clear from section 20(3) that 'public' can include a group access to which is conditional. It is clear from section 20(4) that publication can include using a U2U service to enable or to facilitate access by the public to the statement.

26.108 In our draft guidance, we take the view that content posted to a site or forum which is accessible to anyone is, by definition, published to members of the public. We also propose to say that a members-only group which may be joined or accessed by any user without prior approval from an administrator or similar should still be considered accessible to the public.

26.109 We recognise that terrorist publications are often disseminated in closed groups.[26] We also recognise that the law in this area is likely to develop over time.[27] We provisionally consider that, without detailed investigation and substantial interference with the privacy rights of the members of the group, together with case specific legal advice, services are unlikely to be in a position to make nuanced judgments about whether publication to the 'public' has taken place when the content is being shared via a 'closed', invitation- or prior-approval-only group, or a private social media account where follow requests must be approved. We also note that where law enforcement has carried out the appropriate investigation and has concerns, any constable can take action against the content directly by issuing a notice to the service provider under section 3 of the Terrorism Act 2006. We are therefore consulting on the view that where content has been posted to such a group, a service will not usually have reasonable grounds to infer that content has been published to the public.

Preparation of terrorist acts

26.110 Section 5 of the Terrorism Act 2006 outlines an offence of 'engaging in any conduct in preparation for giving effect to an intention of committing acts of terrorism or assisting others to commit such acts is an offence'.

26.111 We are consulting on our view that the high state of mind requirement of this offence means that it is difficult to conceive of content which would amount to it without also amounting to one of the offences above.

26.112 However, the section 5 offence is particularly relevant for U2U services and search services when considering an account or a website which appears to be run for and on behalf of a proscribed organisation. This is because the definition of terrorism means that *any* action taken for the benefit of a proscribed organisation should also be considered to be an action taken for the purposes of terrorism.

26.113 The offence may also be relevant to services when considering content relating to proscribed organisations which does not obviously fall within one of the specific proscribed organisation offences. Furthermore, it may be relevant to U2U services when considering

[25] Section 1 of the Terrorism Act 2006.

[26] Closed groups, also known as private groups, are forums, group chats or other isolated communication spaces where access and/or membership is limited and controlled by a user administrator or moderator (as opposed to the service itself). Also see: Hall KC, J. 2023. Online Safety Bill: Distinguishing between public and private communication. [accessed 19 September 2023].

[27] See paragraph 7.50 of Hall KC, J. 2023. The Terrorism Acts in 2021. [accessed 19 September 2023].

whether to take down content which is associated with an account run for and on behalf of a proscribed organisation. *Any* content posted to such an account would be likely to be posted for the benefit of the organisation concerned.

26.114 However, it may not be straightforward to identify such accounts in practice. As set out in our draft Codes of Practice, we provisionally consider that relevant user profile information from which to draw inferences about an account would include: the account name; user profile images such as profile, account or background images; user profile information such as 'bio' text, descriptive text on the account; and other user profile information. We also provisionally consider that reasonable grounds to infer that an account is operated by or on behalf of a proscribed group may also arise where a significant proportion of a reasonably sized sample of the content recently posted by the user amounts to a proscribed group offence.

Publishing information about members of the armed forces etc.

26.115 The offence of publishing information about members of the UK's armed forces, UK intelligence services or a constable (a UK police officer)[28] is one which may not be obvious to services. It is rarely prosecuted, so there is not much information available on how to interpret it. Many soldiers and police officers have social media accounts.

26.116 In our guidance, we are proposing to say that, for example, information on the specific location or activity of military units during a specific current or future time period may be information of a type likely to be useful to a person committing or preparing an act of terrorism.

26.117 There is a defence of 'reasonable excuse'. We are consulting on our provisional view that such a defence may be reasonably inferred where the true purpose of the publication is academic or journalistic. For example, reasonable excuse may exist where a journalist or academic shares information on military exercises or movements in a way that presents them as matters of historical or journalistic record and which could not be reasonably said to risk the safety of the personnel involved.

Terrorist threats and directing a terrorist organisation

26.118 Although our proposed guidance covers the offences of making terrorist threats and directing a terrorist organisation, we propose to do so only briefly because we provisionally consider that in practice content which amounts to these offences will also amount to other less specific priority offences. Terrorist threats can be considered along with other kinds of threats (we consider other kinds of threats below). The offence of directing a terrorist organisation is likely to be very difficult for services to identify. If the content were sufficiently clear to make an illegal content judgment, it would likely also amount to the offence of preparation of terrorist acts, above.

Terrorist financing offences

26.119 The Act incudes in schedule 5 a series of offences to do with financing of terrorism.[29] Of these, we provisionally consider that only the offence of inviting someone to provide money or other property for terrorism may be committed online through the posting of content.

---

[28] Section 58A of the Terrorism Act 2000.

[29] Sections 15, 16, 17 and 18 of the Terrorism Act 2000.

26.120 The remaining offences are use of money or property for terrorist purposes, possession of money or property for terrorist purposes, involvement in terrorist funding arrangements or laundering of terrorist property. Although the offences of encouraging, assisting or conspiracy to commit these are relevant in theory, the state of mind requirements are high (intent) and this content is, of its very nature, unlikely to be obvious on its face.

Threats, abuse and harassment offences (including hate)

Broad approach to the chapter

26.121 The priority offences which relate to threats, abuse and harassment overlap with one another to a very significant degree. It is likely to be repetitive and inefficient for services to consider each offence in turn.

26.122 We therefore propose to approach this chapter in a thematic manner, grouping offences by type, rather than going through offence by offence as we have with the majority of other chapters in the Illegal Content Judgements Guidance. We believe this will allow services to work through several complicated and interlinked offences in a manageable and efficient way. Therefore, we consider all the offences to do with threats first, then those which involve insults/abuse, before moving on to the offences which are more specific.

26.123 A consequence of this approach is that some offences which can be committed by threats or abuse are split across more than one section of the draft chapter (for example, the offence of fear and provocation of violence is handled partly in the 'threats' section and partly in the 'provocation' section). However, those preferring to see what the individual offences are, can consult the Annexes to our draft guidance.

26.124 Our Threats, Abuse and Harassment chapter sets out our approach to the following priority offences relating to race, religion, and sexual orientation:

a) offences relating to the stirring up of hatred on the basis of race, religion and sexual orientation (Public Order Act 1986[30]); and

b) other priority offences which concern:

i) racially-aggravated harassment[31]; and
ii) the commission of offences under the Public Order Act 1986 and the Protection from Harassment Act 1987[32] which are racially or religiously aggravated.[33]

26.125 In our guidance on 'Threats, abuse and harassment (including hate)', we propose to only give substantive guidance on the Public Order Act 1986 offences of stirring up hatred on the basis of race, religion and sexual orientation. We do *not* provide separate guidance on the

[30] Specifically: section 18 (use of words or behaviour or display of written material); section 19 (publishing or distributing written material); section 21 (distributing, showing or playing a recording); section 29B (use of words or behaviour or display of written material); section 29C (publishing or distributing written material); and section 29E (distributing, showing or playing a recording).
[31] Section 50A(1)(a) and (b) Criminal Law (Consolidation) (Scotland) Act 1995
[32] Sections 31 and 32 of the Crime and Disorder Act 1998.
[33] We are aware that for sentencing purposes, any offence is to be treated as aggravated if it demonstrated or was motivated by racial hostility, religious hostility, hostility related to disability, hostility on the basis of sexual orientation, or hostility related to transgender identity, as set out in section 66 of the Sentencing Act 2020. However, the presence or absence of an aggravating factor for sentencing purposes is not material to the identification of illegal content under the Act.

racially or religiously aggravated priority offences. This is because once a provider has established that the elements of the non-aggravated offence are present, it is not necessary to go on to consider whether the offence is racially or religiously aggravated. The service should take down the content regardless. By way of example, making an illegal content judgement that content amounts to an illegal threat, for example, is easier than showing it amounts to an illegal threat which is racially or religiously aggravated. Therefore, if a service has already identified illegal content because, for example, it amounts to an illegal threat causing fear or alarm, there is no need to separately consider whether it is also illegal content because the offence is racially or religiously aggravated. It is noted that sometimes the characteristics or identity of the victim are relevant to how reasonable it is for them to feel fear, alarm, harassment or distress.

Importance of freedom of expression

26.126 While, as set out above, the right to freedom of expression is engaged by all the guidance we are giving, we consider it particularly strongly engaged by the offences relating to threats, abuse and harassment (including hate).

26.127 The right to freedom of expression has been held not to be engaged by content which is 'gratuitously offensive'.[34]

26.128 However, robust debate in a healthy democracy often involves the expression of highly emotive and sometimes offensive opinions which touch upon issues of, for instance, politics, religion or race. Similarly, humour often involves an aspect of controversial speech which some people might find offensive and consider to be hateful or abusive.

26.129 We have sought to balance this in our draft guidance.

Threatening and abusive behaviour

26.130 A number of different priority offences may be committed by threats, and a slightly smaller number by abuse. In our guidance, we have dealt with the following offences under the following headings:

a) Threats: section 38 of the Criminal Justice and Licensing (Scotland) Act 2010 (asp.10); section 4 Public Order Act 1986 (fear or provocation of violence); section 4A Public Order Act 1986 (intentional harassment); section 5 Public Order Act 1986 (harassment, alarm or distress); section 16 Offences against the Person Act 1861 (threats to kill); section 31 of the Crime and Disorder Act 1998; sections 18, 19 and 21 of the Public Order Act 1986 (incitement to racial hatred); section 29(b)(c) and (e) of the Public Order Act 1986 (incitement to religious hatred and incitement to hatred on grounds of sexual orientation); section 50A of the Criminal Law (Consolidation)(Scotland) Act 1995 (racially aggravated harassment).

b) Abuse: section 38 of the Criminal Justice and Licensing (Scotland) Act 2010 (asp 13)); section 4 of the Public Order Act 1986 (fear or provocation of violence); section 4A of the Public Order Act 1986 (Intentional harassment, alarm or distress); section 5 of the Public Order Act 1986 (Harassment, alarm or distress); section 31 of the Crime and Disorder Act 1998; sections 18, 19 and 21 of the Public Order Act 1986 (incitement to

[34] *Otto-Preminger-Institute v Austria* (1995) 19 E.H.R.R. 34; *Wingrove v United Kingdom* (1997) 24 E.H.R.R. 1; *Gündüz v Turkey* (2003) 41 E.H.R.R. 5; *Giniewski v France* (2007) 45 E.H.R.R. 23.

racial hatred); section 50A of the Criminal Law (Consolidation)(Scotland) Act 1995 (racially aggravated harassment).

26.131 Of these, we consider the broadest, and therefore most important, is the offence in section 38 of the Criminal Justice and Licensing (Scotland) Act 2010 (asp.10) (the 'section 38 offence'). This offence is committed if a person behaves in a threatening or abusive manner, and the behaviour would be likely to cause a reasonable person to suffer fear or alarm. The state of mind requirement is that the person intends by the behaviour to cause fear or alarm or is reckless as to whether the behaviour would cause fear or alarm. There is a defence if the behaviour was, in the circumstances, reasonable.

26.132 One reason why this offence is broader than others is that 'recklessness' here is the same as it is in Scottish criminal law generally. A person is reckless as to whether the behaviour would cause fear or alarm if they failed to think about or were indifferent as to whether the behaviour would have that result. By contrast, most of the priority offences under the Public Order Act 1986 generally require some sort of positive awareness on the part of the user, that their conduct would have the effect concerned.[35] It is easier to infer that a person failed to think about whether their conduct would have the effect concerned than it is to infer that a person was positively aware that it would.

26.133 Another reason why this offence is broader than others is that the fear or alarm caused need not be of 'immediate' violence. The offences under section 16 Offences Against the Person Act 1861 (threats to kill) and section 4 Public Order Act 1986 (fear or provocation of violence) both require the threat to be of immediate violence, which is likely to be particularly difficult for online services to infer. By contrast, the Scottish courts have interpreted the section 38 offence fairly broadly. In particular, the test of whether a reasonable person would suffer fear and alarm is objective – it is not rebutted by evidence that the particular person concerned, did not.[36] In the appellate case of Jamie Love v PF Stirling, the defendant had posted sectarian abuse on his Facebook page with no suggestion of immediate violence and no evidence that anyone in particular had suffered fear or alarm. His conviction for the section 38 offence was upheld.[37]

26.134 This offence overlaps greatly with most of the other 'threat' and 'abuse' offences and is easier to show than most of them. It therefore makes sense for services to consider it first. However, it is not possible *only* to consider this one offence, because other offences contain some elements which do not overlap. By way of illustration, our reasoning for the threat offences was as follows:

a) The two important non-overlaps with section 38 are the offences in section 5 of the Public Order Act 1986 (threatening and abusive conduct) and in sections 18, 19 and 21

---

[35] Section 4 Public Order Act 1986 (fear or provocation of violence) – requires at least awareness that the conduct may be threatening, abusive or insulting (see section 6(3) Public Order Act 1986); section 4A Public Order Act 1986 (intentional harassment) – requires intent; section 5 Public Order Act 1986 (harassment, alarm or distress) – requires at least awareness that the conduct may be threatening or abusive (see section 6(4) Public Order Act 1986); section 16 Offences against the Person Act 1861 (threats to kill) – requires intent, section 29(b)(c) and (e) of the Public Order Act 1986 (incitement to religious hatred and incitement to hatred on grounds of sexual orientation) – requires intent.

[36] *Paterson v Harvie* [2014] HCJAC 87. In this case, the individual who was prosecuted for the offence had shouted and made threats while confined in the back of a police van. Notwithstanding that neither of the police officers concerned in fact felt fearful or alarmed, the court held that a reasonable person would have done, knowing they would have to remove the individual from the van at the end of the journey.

[37] *Paterson v Harvie* [2014] HCJAC 87.

of the same Act (stirring up racial hatred). Where section 5 talks about threatening or abusive conduct which is likely to cause alarm, it overlaps with the section 38 offence. However, it can also be committed when the threatening or abusive conduct is likely to cause no alarm, but only harassment or distress. Harassment in particular is a fairly low threshold. However, content likely to cause harassment or distress will only be illegal content if there are reasonable grounds to infer that the person posting it was at least aware that what they were doing may be threatening or abusive, and that a person likely to be caused harassment or distress was nearby. This tends to make the offence less likely to be identifiable in practice.

b) Threatening and abusive conduct likely to stir up racial hatred is next. In practice much content which is likely to stir up racial hatred is also likely to amount to the section 38 offence, which is easier to show and should therefore be considered first. It is also possible that as set out above it may amount to the section 5 offence. However, in theory it is possible that content could exist which, even though it was likely to stir up hatred, was neither likely to cause a reasonable person to suffer fear or alarm *nor* was used within sight or hearing of a person likely to suffer harassment or distress. In that case, services would need to go on to consider whether it was likely to stir up racial hatred.

c) Similar reasoning applies to conduct likely to stir up religious hatred or hatred on grounds of sexual orientation, but for these offences there must be intent to stir up hatred.

d) Finally, to the extent that the section 4 Public Order Act offence relates to *fear* of violence, it overlaps with the section 38 offence, and because it only relates to immediate violence it is unlikely to take place online in any event. But very rarely, content online may *provoke* immediate violence – for example in the context of ongoing serious public disorder or a genocide. In that case, services would need to consider the section 4 Public Order Act offence.

26.135 The order in which we have taken the offences, and the amount we have written about them, is not a sign of their seriousness.

26.136 It is particularly important for Ofcom to have regard to the right to freedom of expression in considering the safety duty in relation to the offences relating to insults and abuse causing harassment or distress, because of the risk that an over cautious approach to these would lead to disproportionate takedown, including (for example) of political and religious discussion.

Approach to other harassment and coercive and controlling behaviour offences

26.137 Once services have considered all the offences which are *necessarily* carried out by threats or abuse, a set of offences remain which can be carried out by threats or abuse, but which need not be – they are the offences to do with harassment, stalking and coercive and controlling behaviour.

26.138 These offences include conduct which is very serious indeed and which disproportionately affects women and girls. Victims of coercive and controlling behaviour are deprived of their independence, exploited and subject to having their day-to-day life regulated by the perpetrator. We recognise the severity and impact of such behaviour, and its seriousness as an offence both online and in wider society. Our approach to controlling or coercive

behaviours goes far beyond the Illegal Content Judgements Guidance and will be addressed through multiple channels and approaches.

26.139 However, for the purposes of the Illegal Content Judgements Guidance, the sensible way to approach these offences is not necessarily to consider the most serious offence first.

26.140 In practice, it is not likely to be straightforward for a service to identify specific instances of coercive and controlling behaviour (at least not consistently with the privacy rights of their users) because the service would need to know whether the possible victim and possible perpetrator are in an intimate personal relationship, or are living together either as members of the same family or because they have previously been in an intimate personal relationship. However, the coercive and controlling behaviour offence also requires the perpetrator to repeatedly or continuously engage in behaviour towards another person that is controlling or coercive, in a way that has a serious effect on them. A serious effect is where the victim fears at least twice that violence will be used against them, or is caused 'serious alarm or distress' which has a substantial adverse effect on the victim's usual day-to-day activities. The perpetrator is only guilty of the offence if they know or ought to know that the behaviour will have a serious effect on the victim.

26.141 Any case like this involving threats or abuse causing fear of violence, or alarm or distress, will be caught by the threats and abuse priority offences set out above. A case of fear of violence, or alarm or distress which is not caught by those will be caught by the harassment offence in section 2 of the Harassment Act 1997 and/or Article 4 of the Protection from Harassment (Northern Ireland) Order 1997 (S.I. 1997/1180 (N.I. 9), which applies when a person engages in a course of conduct (a minimum of two instances, but these can include offline as well as online instances), which amounts to harassment of another, and which a reasonable person in possession of the same information would know or ought to know amounts to harassment. In other words, before a service had sufficient information to make a reasonable inference of coercive and controlling behaviour, it would have already identified harassment and the takedown duty would already have been triggered. The same reasoning applies in relation to stalking and the racially or religiously aggravated harassment offences, since all involve harassment.

26.142 In our draft Illegal Content Judgements Guidance, we therefore focus on harassment. The other, more serious offences need not be considered in order to make an illegal content judgment, though they may well be relevant in considering the seriousness of the content and how it should be prioritised.

Epilepsy trolling

26.143 We are proposing to include one relevant non-priority offence in this section of our Illegal Content Judgements Guidance – that is, the newly created offence of epilepsy trolling. This occurs when a person sends flashing images to a person known to have epilepsy, with the intention of causing them harm. We include it in the section on threats, abuse and harassment because the type of conduct concerned is likely also to potentially amount to harassment, but epilepsy trolling may be easier to show since there is no need to show that there has been a course of conduct.

Child sexual exploitation and abuse (CSEA): offences relating to child sexual abuse material (CSAM)

26.144 Child sexual abuse material, or 'CSAM', refers to indecent or prohibited images of children (including still and animated images, and videos, and including photographs, pseudo-photographs and non-photographic images such as drawings). CSAM also includes other material which contains advice about grooming or abusing a child sexually or which is an obscene article encouraging the commission of other child sexual exploitation and abuse offences. Furthermore, it includes content which links or otherwise directs users to such material, or which advertises the distribution or showing of CSAM.

26.145 The priority CSAM offences are set out in schedule 6 of the Act. Although they are very serious offences, they are amongst the least complex priority offences analytically. The state of mind requirements are very low (knowledge, which, as set out above, we consider is met by the content being present on the service). It is unlikely that a service will have reasonable grounds to infer that a defence is available. However, challenges can arise where it is not clear what the content represents, and we discuss our approach to these below.

26.146 There are multiple image-related offences that are to do with indecent images of various kinds. They include making, taking, distributing, showing or possessing this kind of material.[38] The prohibited image offence is committed by possession only. For the purposes of services making illegal content judgments, there is no need to consider the verbs used in the offences in detail. If an indecent picture is available on the internet, it has been 'made'. If a prohibited image is available on the internet, it is 'possessed' by at least the user who uploaded it. Our guidance therefore focuses on what kinds of pictures are indecent or prohibited pictures. For the purposes of the guidance, we considered the comparative offences across the nations. It should be noted that the Scottish version of the 'making' offence includes additional defences relating to what was reasonably believed by the person 'making' the image in respect of the child's age. Given the England, Wales and Northern Ireland offences do not include this, and are applicable to illegal content regardless of which part of the UK is concerned, we have directed services to consider the England, Wales and Northern Ireland offences.

Inferring the age of a subject in an image

26.147 When it comes to offences relating to CSAM, the key question will often be whether or not the person depicted in an image is a child.

Inferring age from the content, captions and comments alone

26.148 We propose that generally speaking, the age of a subject in an image should be inferred based on the general appearance of the subject(s) in the content itself and any contextual information that is available. Such contextual information may include captions to the image or comments.

26.149 Where there is no hard evidence of the subject's age, but a reasonable person would assume from the appearance of the subject that they are under the age of 18, the age criteria should be assumed to be met and a service should proceed on the basis that the content is an image of a child.

[38] For example, section 1 Protection of Children Act 1978; article 3 Protection of Children (Northern Ireland) Order 1978 (S.I. 1978/1047 (N.I. 17)); and section 52 and 52A of the Civic Government (Scotland) Act 1982.

Inferring age from account information

26.150 However, we are concerned that some children may look older than they are. We are therefore proposing to recommend that reasonable grounds to infer that the subject of the image is under 18 may exist where:

a) age estimation or age verification measures indicate that the subject in the image is aged under 18;

b) the subject in the image itself states in a report or complaint that they are aged under 18 or were aged under 18 at the time when the potentially illegal content was posted; or

c) account information indicates that the subject in the image is aged under 18, except where the subject concerned has been using the service for more than 18 years.

26.151 We recognise that services may hold quite a lot of other information which may help them determine a possible victim's age – for example, they may be using the service to exchange messages with friends which include references to school. However, our provisional view is that asking services to consider information of this type would require them to engage in a very significant interference with all users' rights to privacy. At this stage, we therefore do not propose that this type of information should be considered 'reasonably available' to services.

26.152 We also recognise that sex workers exist, whose business may depend on them looking younger than they are. If a service had good evidence that a person who looked underage was in fact over 18, our guidance would not require it to take the content concerned down. However, in the absence of good evidence we consider it reasonable for services to infer that a person is underage if they look underage.

26.153 In order to protect children from online harms, the Online Safety Act requires services that are likely to be accessed by children to use age estimation or age verification measures. We are continuing to build our evidence base in relation to available age estimation and verification measures and expect to return to this matter in our work focusing on the protection of children online.[39] Once this work is complete, many services will be expected to use age estimation or verification measures that are highly effective at determining whether or not a particular user is a child or not. This may make it easier for services to identify potential victims whom it is reasonable to infer are children, particularly in the case of self-generated content on U2U services.

Sharing of URLs

26.154 We consider it particularly important to be clear that if the user of a U2U service posts a URL which leads to an indecent or prohibited image of a child, an obscene article or a paedophile manual, that is illegal content.

a) Under section 7(4) of the Protection of Children Act 1978, reference to a photograph includes data stored by electronic means which is capable of conversion into a photograph.

b) It is reasonable in any event for a service to infer that a person sharing a URL of that nature knows what it leads to and *intends* the person with whom they share it to click on

[39] For more information on our current position with regard to age assurance measures and our work focusing on protection of children online, please see Volume 4, Chapter 21 of our Phase One consultation on 'User Access'. paragraphs 21.106 – 21.110.

the link. As such, it is reasonable to infer that they are intentionally encouraging or assisting the commission of an offence of making an indecent image of a child or of possessing a prohibited image of a child or a paedophile manual.[40] This is the case even if the person concerned is only doing it to express outrage about the content.

c) Dissemination of URLs is likely to amount to distribution or showing of indecent images as the case may be.

d) The definition of 'publish' for the purposes of section 2 of the Obscene Publications Act 1959, in relation to obscene publications, includes 'distribute' (see section 1 of that Act).

Child sexual exploitation and abuse (CSEA): Grooming and exploitation of children

26.155 The remaining CSEA priority offences are in the chapter we have called Grooming and Exploitation of Children. These offences are more complex to identify in practice.

26.156 We have structured this draft section by looking at how much the service needs to know in order to make an illegal content judgement and discuss the offences with the lowest requirements first. Therefore, we have begun with the offences in relation to which the service only needs to be able to draw inferences about the age of the potential victim. We then turn to offences in relation to which the services need to consider the age of both the potential victim and the potential perpetrator. Then we turn finally to the more complex offences involving offline conduct.

Meeting a child offences

26.157 The priority offences relating to grooming and exploitation of children include offences related to meeting a child following sexual grooming or preliminary contact. Meeting in relation to these offences means a physical, face-to-face encounter in the real world rather than online (unlike the terrorism offences). For this reason, we do not deal with the 'meeting' offences in our guidance. However, the preceding communications leading up to the offence may amount to illegal content by virtue of one or more of the other priority offences[41], and any online 'meeting' which was unlawful would be likely to amount to one or more other priority offences too.[42]

---

[40] In *R v Jayson [*2002*]*; *Regina v Smith* [2002] EWCA Crim 683, it was held that the mere act of downloading a photograph or pseudo-photograph from the internet to a computer screen could be said to constitute the "making" of a photograph or pseudo-photograph. It was not necessary to prove that the individual did any act with a view to saving the image on their computer.

[41] For example, sexual communication with a child (section 15A of the Sexual Offences Act 2003; Article 22A of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2)); or communicating indecently with a child (sections 24 and 34 of the Sexual Offences (Scotland) Act 2009).

[42] For example, causing or inciting a child to engage in sexual activity (sections 8 and 10 Sexual Offences Act 2003; Articles 15 and 17 of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2)), causing a child to watch a sexual act (section 12 of the Sexual Offences Act 2003); Article 19 of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2), arranging or facilitating commission of a child sex offence (section 14 of the Sexual Offences Act 2003; Article 21 of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2)); sexual communication with a child (section 15A of the Sexual Offences Act 2003; Article 22A of the Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2)); causing a child to participate in a sexual activity (sections 21 and 31 of the Sexual Offences (Scotland) Act 2009); causing a child to look at a sexual image (sections 23 and 33 of the Sexual Offences (Scotland) Act 2009); communicating indecently with a child (sections 24 and 34 of the Sexual Offences (Scotland) Act 2009).

Sexual activity offences where the child is under 16

26.158 Several of the offences which deal with sexual activity with a child break the offences down depending on whether the child is under 13, or whether they are between 13 and 15 years old. The main difference between the offences is the severity of the potential penalty, which is not relevant to the question of whether the content is illegal content. In our guidance, we propose to deal with both groups of offences as content relating to potential victims under the age of 16.

26.159 However, for some offences, such as the offence of causing or inciting a child to engage in sexual activity, there is an additional element to be considered where the child is aged 13, 14 or 15, which is not required where the child is under 13. That is, for content to be considered illegal content, there must be reasonable grounds to infer that the potential perpetrator did not *reasonably* believe that the child in question was 16 or over.

26.160 When deciding our proposed approach to this discrepancy, we took into account that our guidance will be in place at a time when some services may not yet have robust age verification or age assurance measures in place enabling them to determine whether a child is under 13, or is 13, 14 or 15 years old. As a result, we propose to say that, where services are able to reasonably infer that a potential victim is under 16, this provides reasonable grounds to infer that the potential victim is not *generally* seeking to represent themselves to others as being over the age of 16. In these cases, services can infer that the potential perpetrator did not *reasonably* believe the child in question was 16 or over, and we propose that the content should be treated as illegal and taken down, *except* where the victim has a made a positive statement that they have represented themselves to the other user as being aged 16 or over.

Inferring the age of a potential victim of grooming

26.161 In the chapter on grooming and exploitation of children, there may not be an image, or at least not a current one, which the service can necessarily use as the basis for drawing inferences about age.

26.162 Generally speaking, in our view, self-declaration is not a good way to infer age. This is partly because children may declare themselves to be over 18 in order to access age-restricted content, and partly because would-be abusers may declare themselves to be children in order to gain access to children.

26.163 However, for the specific purposes of making illegal content judgments about grooming, we are consulting on our view that a *potential victim* of grooming, who declares themselves to be a child, should usually be believed. This is because:

a) many children do give their age truthfully;

b) abusive adults who claim to be children are unlikely to make complaints about grooming; and

c) although there is some risk of malicious reporting, the content itself would need to meet the definition of the offence, which would be relatively difficult for malicious reporters to achieve.

26.164 We therefore propose that services should use information where a potential victim states their age (for instance in the relevant content itself or in other places associated with the potential victim's account) as a way to infer their age.

26.165 We do not consider that the same can be said of potential perpetrators. We are consulting on the view that reasonable grounds to infer that a perpetrator is 18 or over may arise in any of the following ways:

a) The potential perpetrator states they are aged 18 or over;

b) The potential perpetrator has been using the service for 18 years or more;

c) The potential victim provides evidence that the potential perpetrator is aged 18 or over, and the service is not aware of any strong evidence to suggest the contrary.

26.166 As set out above in relation to CSAM, our provisional view is that asking services to consider all the activity on a given account to determine the age of the account holder would require them to engage in a very significant interference with all users' rights to privacy. At this stage, we therefore do not propose that this type of information should be considered 'reasonably available' to services.

Sexual exploitation of a child

26.167 The offences relating to sexual exploitation of children are designed to penalise those involved in child sexual exploitation at many levels. For example, the offence of controlling a child aged 17 or younger in relation to sexual exploitation, would capture the activities of a person at a higher level of a criminal gang involved in the exploitation, as well as the gang member directly controlling a child day-to-day.

26.168 However, the more remote from the child victim the individual is, the greater the evidential difficulties of proving that the content amounts to the offence are likely to be. We consider that the child exploitation offences that services are most likely to encounter online will be when, in the content being considered, a child is being incited or coerced into providing indecent images of themselves online.

26.169 The child exploitation offences have a fairly high state of mind requirements.

26.170 First, where the child is over 13, the service must have reasonable grounds to infer that the potential perpetrator did not reasonably believe that the potential victim was 18 or over. We recognise that a service is not likely to have a direct statement from the potential perpetrator of their beliefs, reasonable or otherwise. More importantly, it would be a very great interference with users' rights if services were to go looking in their account activity for evidence of potential perpetrators' beliefs, and perhaps the activity of other users to see whether they had said or done anything to make a belief reasonable.

26.171 However, we are consulting on the view that if the service itself is in a position to infer that the potential victim is under 18, it is sufficiently obvious that a potential perpetrator's belief is unlikely to be reasonable.

26.172 Secondly, the possible perpetrator must have intent – for example, for the offence of obtaining the sexual services of a child, the potential perpetrator must intend to obtain sexual services. Again, we recognise that in these types of instances, the potential perpetrator is unlikely to have stated their intent explicitly. However, we are consulting on the view that where content is identifiable by a service as meeting the 'conduct' part of the offence (for example, if the content comprises a direction to the child to provide sexual services, coupled with an offer of payment), it is reasonable to infer that the state of mind requirements are also met. It is difficult to conceive of any reason why a person would send such a request, absent that intent.

Fraud and financial services offences

26.173 The financial services and fraud offences are amongst the most technically difficult offences in the Act but, as set out in our Register of Risks, this type of content causes widespread and serious harm.

26.174 The guidance will also apply in relation to fraudulent advertisements when those duties come into force. When the time comes, we will consider whether any revisions are needed to it and will consult if we think there are.

False claims to be authorised or exempt

26.175 We propose that as a starting point for making illegal content judgements on offences to do with fraud and finance, services should first consider whether the firm offering those services is claiming to be authorised. That is because it should be relatively straight forward for services to identify content containing a false claim to be authorised or exempt. Determining whether a claim to be authorised is true is a fairly straightforward matter of checking the content, including address and other contact details, against a register the FCA publishes on its website (the Financial Services Register ('FS Register')). We consider that this is a check services can be expected to make where alerted to a possible false claim to be authorised. A service will have reasonable grounds to infer that a claim to be authorised is false and the content is illegal content if the firm is not included as an authorised firm on the FS Register or the details referred to in the online content do not match the details of the authorised firm on the FS Register. Similarly, the FCA gives firms a unique Firm Reference Number (FRN) when the firm becomes authorised. Using an FRN which does not appear on the FS Register, or providing different contact details than those included on the FS Register would, in our provisional view, provide reasonable grounds to infer that the content contains a false claim to be authorised.[43]

26.176 However, that is only one of the priority financial services and markets offences. We consider that the other offences in this category are some of the most technically complex to interpret. We are proposing to structure this chapter in a way which is intended to enable services to capture the relevant content most easily. For this reason, we do not propose to deal with the more complex financial services and markets acts offences until later in the chapter.

26.177 Following on from false claims to be authorised or exempt, we therefore move on to the offence of fraud by false representation, as we consider that many types of fraud are likely to amount to this offence, in one way or another. We also consider that many of the most egregious examples of fraud will fall under this offence.

Fraud by false representation

26.178 The offence of fraud by false representation is undeniably complex. In order for content to amount to this offence, a service would need reasonable grounds to infer that it contains a statement which is false, that it is dishonest, and that the user intends to make a gain or cause a loss. All these things are matters which involve drawing inferences about circumstances offline. For a statement to be false, there must be a 'truth' which exists outside the content. Dishonesty and intent are both parts of the user's state of mind.

[43] Financial Services Authority, 2023. The Financial Services Register. [accessed 25 September 2023]

26.179 Recognising the difficulty of these judgements, however, we are provisionally of the view that it is possible to draw reasonable inferences in some circumstances, based on the content and the context in which it appears. This is likely to be the case for the most egregious examples of this type of content.

26.180 We are proposing the use of a 'filter system' to identify content which may reasonably be inferred to amount to fraud by false representation. This is because whilst certain features of online content might raise concerns about fraud by false representation, it would be unusual for a single representation to provide on the face of it reasonable grounds to infer that it is false; that it is dishonest; and that the user intends to make a gain or cause a loss.

26.181 For example, if a user posts a false location, this may amount to a false representation (even though there may be a legitimate purpose for making it), but even if made dishonestly, it does not amount to a fraud by false representation unless the intention to make a gain or cause a loss can be reasonably inferred. A false representation can be made honestly and can be made without any intention to make a gain or cause a loss. On its own, it is not necessarily even grounds for concern. Users who purport to be located somewhere different from their true location may be acting as agents for companies, charities or other groups located in their purported location. They may afraid that their true location would put other users off from interacting with them. Or they may be afraid to reveal their true location because of possible retribution for what they say and do online (for example, if they are in a country where LGBTQ+ sexualities are persecuted, or if they wish to criticise a government which does not tolerate dissent).

26.182 However, the *combination* of a false location with other factors may (in our view) give rise to reasonable grounds to infer that content amounts to fraud by false representation. For example, a false location coupled with a false statement that the entity is regulated by the UK's financial services regulator and an invitation to invest all together would be, in our view, reasonable grounds to infer that the content amounted to fraud by false representation.

26.183 The filter system we are proposing contains a (non-exhaustive) list of suggested 'red flag indicators', split into three categories. In each category there are examples of types of content which might be indicative of elements of the offence.

26.184 The first category 'Disguised account information or activity', such as a user masking their location (discussed above) and the final category, 'Account and content characteristics commonly associated with fraudulent behaviour' are focused on identifying features of that content which might point to dishonest intention, and might in context amount to reasonable grounds to infer that the representation being made is false (if not apparent on the face of the content).

26.185 The second category filters by content which contains a relevant "representation" (such as 'requests, invitations, or inducements to invest, send money, send identification documents, or send financial information'). Without a representation, which is made with the intention to make a gain *or* to cause another person loss (or expose them to the risk of loss), there can be no offence of fraud by false representation.

26.186 We emphasize in the chapter that no single example in any of the categories is capable of being reasonably inferred to amount to fraud by false representation. It is only in cases where there is content of the type suggested in each category where may be reasonable grounds to infer fraud by false representation *except* where services have evidence to suggest the contrary. Whether or not there are reasonable grounds to infer fraud by false

representation in relation to any piece of content, will ultimately rest with services and will be a case by case decision.

26.187 We recognise that by introducing a system where we in writing describe examples of content which may indicate fraudulent behaviour, there is a risk that content which is not fraudulent, but looks like it is fraudulent, is wrongfully judged to be illegal content and removed from a service. This has freedom of speech implications as users that are not posting illegal content may still have their content removed. However, it is our duty to provide guidance on what illegal content may look like on a service and due to the complexity of the offences relating to this type of content, we believe this filter system will help services take down illegal content.

26.188 We also recognise that there is a risk that our guidance is too narrow and sets the bar too high to capture all instances of fraud by false representation. We consider this to be an inevitable consequence of the need to avoid over-takedown, but welcome views in response to consultation on whether we have struck the balance appropriately.

26.189 We recognise that some services are likely to have far more sophisticated approaches to identifying frauds taking place on their services, possibly involving use of proactive technologies. However, the Illegal Content Judgements Guidance is just that: guidance. Nothing in our guidance or the Act prevents services from continuing to use systems and processes which they consider more effective at identifying illegal content.

26.190 Finally, we acknowledge that there is a risk of gaming, in that bad actors may use our illegal content judgements guidance to design frauds which evade our proposed system. We consider this risk to be tempered by the fact that we do not expect most services necessarily to adopt the UK's definition of illegal content for all their users worldwide.

26.191 There are a couple of other priority offences relating to frauds: fraud by abuse of position and participating in fraudulent business carried on by a sole trader etc. We are proposing to deal with these only very briefly in our guidance because we are not aware of any circumstances in which they could be identified in online content, where the content would not also amount to a fraud by deception.

Approach to the Financial Services and Markets Act 2000 (FSMA)

26.192 A number of the priority offences in the Act are from the Financial Services and Markets Act 2000 (FSMA).

26.193 This Act regulates the provision of financial services in the UK. It creates a number of activities which are subject to regulation and may only be carried out by authorised persons, including a regime specifically for the regulation of financial promotions[44] (this is applicable to persons anywhere in the world promoting investments to UK users). The definitions of regulated activities and investments are lengthy and technical, and so is the application of the various relevant exemptions.

26.194 We carefully considered whether it would be correct to say that the Act requires U2U and search services to become sufficiently expert in UK financial services regulation to apply the FSMA offences correctly to the content they see. Our provisional view at this stage is that this is not likely to be proportionate, even for the larger services, due to the significant

[44] A financial promotion is content which seeks to persuade or incite the recipient to engage in 'investment activity' or engage in 'claims management activity' – both defined terms.

expertise and time we consider would be required, which go well beyond the typical knowledge base of a content moderator. We consider this is evident simply from the definitions themselves.[45]

26.195 We believe it would also not be appropriate to attempt to simplify the FSMA offences in a way which reduces their technical complexity, as this would almost certainly result in Ofcom misleading both services and the general public on the meaning of the offences.

26.196 We considered whether it may be appropriate, the in light of the technical complexity, to ask services which choose not to equip themselves with appropriate expertise and time to make these judgments to take down *all* content which appears to promote investments.

26.197 However, this would be likely to affect a lot of legal content and as such would need to be a measure taken under the safety duty, which would need to be proportionate. The impact of such an approach on businesses in the UK, including small and microbusiness, appears to us potentially to be very significant indeed. We are not currently in a position to take the view that it would be proportionate.

26.198 Our proposed approach at this stage therefore steers services to make illegal content judgements about content that may amount to the FSMA offences by relying on reasoned reports or flags from expert bodies such as the FCA or the courts. We recognise that this means services are likely to rely on those bodies' judgment heavily, with possible unfairness to users and risks to their commercial interests and to their rights to freedom of expression. However, they are public bodies bound by their own duties of fairness, and they are bodies with significant technical expertise and experience. Overall, we consider that this is the best way to balance the competing interests of users and services in a way which secure that the most damaging content is likely to come down.

Approach to articles for use in frauds

26.199 It is an offence to make, adapt, supply or offer to supply any article, knowing that it is designed or adapted for use in the course of or in connection with frauds. It is also an offence to make, adapt, supply or offer to supply any article, intending that it be used to commit, or assist in the commission of, fraud. An 'article' includes data or software.

26.200 As set out in Chapter 6O ('Fraud and financial offences') of Ofcom's draft Register of Risks, we are aware that both search and U2U services are used to offer to supply, and sometimes to supply, data and/or software for use in frauds – for example, lists of stolen passwords.

26.201 While the state of mind requirement for this offence is fairly high (intent), we are consulting on our view that in practice, it is difficult to conceive of any reason why a person would be disseminating or offering to disseminate certain information online, other than for use in a fraud.

Buying and selling offences

26.202 Schedule 7 of the Act includes priority offences relating to the marketing, buying and selling or supply of drugs/psychoactive substances and of weapons. We refer to these collectively

[45] See, for example, the definitions of 'regulated activity', 'controlled activity', 'controlled investment' and 'claims management activity' in: Financial Conduct Authority, 2023. 'FCA Handbook: Glossary Terms [accessed 19 September 2023]. See also: the definition and list of 'exempted persons' in the Financial Services and Markets Act 2000 (Exemption) Order 2001.

as the 'buying and selling offences.' They raise particular interpretative challenges in relation to jurisdiction.

Jurisdiction

26.203 The general purpose of the Act is to make the use of regulated internet services safer for individuals in the United Kingdom.[46] The safety duty extends only to the design, operation and use of the service in the United Kingdom, and in the case of a duty that is expressed to apply in relation to users of a service, the design, operation and use of the service as it affects United Kingdom users of the service.[47]

26.204 However, as set out above, the definition of illegal content is not limited to conduct that takes place in the UK or that affects UK users. The Act states that "[f]or the purposes of determining whether content amounts to an offence, *no account is to be taken* of whether or not anything done in relation to the content takes place in any part of the United Kingdom." The Explanatory Note to the Act explains that the effect of this is that "content does not need to be generated, uploaded or accessed (or have anything else done in relation to it) in any part of the United Kingdom to amount to an offence under this provision. *This is the case regardless of whether the criminal law would require the offence, or any element of it, to take place in the United Kingdom (or a particular part of it)*" (Ofcom's emphasis).

26.205 Not every country or jurisdiction in the world prohibits the buying and selling of the items covered by UK priority offences. In particular, some countries take a more liberal approach than the UK to the sale of drugs like cannabis. Similarly, many countries take a less restrictive approach than the UK to the selling of knives and guns. We recognise the tension between protecting UK users from illegal content and the commercial interests of services in hosting content (for U2U services) or indexing search content (for search services) for jurisdictions in which it is lawful.

26.206 As we have explained above, the interpretative rule in the Act applies only to what happens in relation to the *content*. It does not affect, for example, any offline circumstances required for the offence to be committed. In the case of the buying and selling offences considered below, and having regard to the intention of Parliament, we consider that the words 'sale' and 'supply', and the linked phrase 'expose for sale' are best construed as relating to sale etc *to persons in the UK*.

26.207 This still creates challenges. In particular, online content which may amount to a buying or selling priority offence may not be at all clear about whether or not UK users are included in the offer. If our guidance suggested that all content should be considered legal unless it is expressly targeted at UK users, it would be likely to mean that services judged very little content associated with the buying and selling of drugs and weapons to be illegal content and may create a loophole that could be exploited by bad actors. This appears unlikely to be the intention of Parliament in including these offences as priority offences in the Online Safety Act.

26.208 On the other hand, we do not consider it practical to suggest that all over the world, overseas users and URL providers should expressly state that UK users are not allowed to buy.

---

[46] Section 1 of the Act.
[47] Section 8(3) and 25(1) of the Act.

26.209 In our view, there is no simple proxy which services can use to decide whether an exposure for sale etc. has potentially been made to UK users. Language is relevant, but the English language is widely spoken worldwide and is not the only language used day to day by UK users. The location of the seller is relevant, but not necessarily determinative since goods may be sold for export.

26.210 We therefore consider that services will need to make sensible, nuanced judgments on this point, having regard to the content itself, its context and – in particular – any evidence from users (via complaints) or from law enforcement that goods are being marketed unlawfully to users in the UK. If a piece of content explicitly or implicitly excludes UK consumers from its customer base, it follows that it cannot be said to amount to illegal content. If a piece of content makes it clear that the item in question may only be purchased in person in a location within the jurisdiction where it is legal, or if it makes clear that delivery to a buyer is restricted to those within the same jurisdiction, then in our view the buying and selling priority offences have not been engaged.

Drugs and psychoactive substances offences

Drugs

26.211 In drafting our guidance on illegal content relating to offers to supply drugs and psychoactive substances, we considered whether it was appropriate to identify drugs only by their legal (chemical) names and to make it the responsibility of services to keep their moderators up to date on the drugs' 'street names'. In so considering, we recognised that street names used by dealers and drug users change often and so any list compiled by Ofcom would risk being incomplete and quickly outdated. We would not want Ofcom's guidance to be an excuse for services to fail to take appropriate steps to keep their knowledge of drugs slang properly up to date.

26.212 However, our provisional view is that the Illegal Content Judgments Guidance is for all services – including smaller services based overseas – and that a potentially incomplete list of drugs' street names is therefore better than no list. We have drafted on that basis.

Offering to supply

26.213 The priority drugs and psychoactive substances offences relate to the unlawful supply, or offer to supply, of controlled drug or psychoactive substances respectively. 'Offer' here takes its natural meaning in English rather than its technical meaning in the law of contract. We considered whether we could provide more guidance than that in our Illegal Content Judgement Guidance, but considered that – absent judicial authority – we would risk misdirecting services by doing so.

26.214 By its nature, an offer to supply must be made intentionally. Therefore, if the content amounts to an offer, the service will have reasonable grounds to infer that the state of mind requirements are met. We therefore do not propose to discuss state of mind separately in our guidance.

Exemptions

26.215 In the guidance, we make reference to the Misuse of Drugs Regulations 2001 (SI 2001/3998) ('Misuse of Drugs Regulations 2001'). The regulations provide certain exemptions from the provisions of the Misuse of Drugs Act 1971 . In some cases, these regulations are relevant to offering to supply controlled drugs and drugs article. It is our provisional view that providers

of U2U services will not encounter examples of exempted content on their services. However, we recognise that it may be more challenging for search services to distinguish between illegal content and content which is legal due to the circumstances of its posting being exempted under the Misuse of Drugs Regulations 2001. We propose that, where providers of search services encounter content which could possibly to be exempted under the Misuse of Drugs Regulation 2001, they should take a pragmatic view, considering the context available and consider whether the controlled drugs appear to be sold in the UK. We would welcome further evidence and comments on this approach.

Weapons offences

26.216 The weapons offences in schedule 7 split broadly into four. There are: offences relating to firearms (broadly construed); offences relating to banned knives and 'offensive' weapons; an offence relating to the marketing of knives; and a series of offences relating to the buying and selling of various weapons to a person who cannot lawfully buy or sell them. We consider these below in turn.

Firearms

What is a firearm?

26.217 For the purposes of the guidance, we have focused on the priority firearms offences from the Firearms Act 1968 (the "Firearms Act"). This is because they are the most comprehensive set of priority offences, differ only in minor technical detail from the equivalent Northern Irish legislation (Firearms (Northern Ireland) Order 2004 (S.I. 2004/702 N.I.3)) and advisors are likely to be more familiar with the Firearms Act because it applies to a greater territory. It is difficult to find a term that is clear about the weapons this Act covers. It includes 'firearms' as the Act defines them, but it also includes other weapons that it defines as not being firearms – for example air weapons. It also includes component parts and ammunition. Within the definition of firearms there are a number of types of weapon that a layperson may not intuitively consider to be 'firearms'; for example, pepper sprays, stun guns (often known by a brand name, tasers), and rocket launchers. In what follows, we use 'firearms' broadly, to cover all the types of weapon, parts and ammunition that are subject to the Firearms Act.

26.218 In our draft guidance, due partly to this technical complexity and partly to the likelihood that content moderators will lack a detailed specialist understanding of types of guns, we have attempted to draft in a way that avoids services having to grapple with the detail of what type of firearm they are considering unless it is absolutely necessary.

Sale or exposure for sale and structure of our guidance

26.219 Most of the priority firearms offences in schedule 7 of the Act relate to the actual sale or purchase of the firearm concerned. However, such a transaction almost certainly takes place offline (for example, with the exchange of money) and cannot take place through the posting of user-generated content on a U2U service or in search content. What takes place online, either on a U2U service or in search content, is almost always only the lead-up to a sale or a purchase rather than the purchase itself. It is the marketing and advertising or 'exposure for sale' which encourages a potential buyer to contact a potential seller.

26.220 One priority offence for firearms, in section 3 of the Firearms Act, relates to the activity of 'exposing for sale'. Our draft guidance therefore focuses on this. The offence in question

takes place when an *unauthorised* person exposes a relevant firearm for sale *by way of a trade or business*. This offence applies to most types of firearm, but there are exceptions.

26.221 It is therefore necessary to consider each part of the offence in turn. In order to help services navigate through the detail, we are therefore proposing to draft our guidance in the form of a series of questions. We particularly welcome comments on whether our proposed approach is sufficiently clear for services to understand and use in practice.

Approach to 'by way of a business or trade'

26.222 Although the section 3 Firearms Act offence covers almost all types of firearms, the phrase 'by way of a trade or business' means that we believe that – in practice – it is appropriate for our guidance to distinguish between certain types of firearms. This is because the Firearms Act creates a class of weapons, 'prohibited weapons', which it is unlawful even to possess in the UK without specific authority from the Secretary of State in England and Wales and Scottish Ministers in Scotland. Such authority is normally only granted to those with a legitimate *commercial* need to possess prohibited weapons, rather than for private use or speculative business interest. It follows that a person dealing in such weapons lawfully will, by definition, be trading a business asset.

26.223 In our provisional view, the limits on *lawful* possession and trade of 'prohibited weapons' are likely to make it difficult for any person to acquire such weapons for *unlawful* onward sale. These are not the sort of weapons which it is likely that a casual seller might find in an attic and decide to place for sale online. Usually, it would take effort and knowhow which may be associated with fairly significant expense. The offence is also serious - possession of such items for sale is subject to a statutory minimum term of imprisonment of 5 years.[48] Altogether, for these reasons, we consider it unlikely that a person in the UK dealing in such weapons *unlawfully* would be in a position to do so other than by way of a generally unlawful trade or business of some kind. The likelihood is therefore that 'prohibited weapons' are being dealt by way of an (unlawful) trade or business, and it is reasonable for services to draw this inference.

26.224 The same is not true of less heavily restricted firearms that are not 'prohibited weapons', such as shotguns, air weapons, and 'lethal barrelled weapons'. For these types of weapons, in our provisional view, positive evidence would be needed to make a reasonable inference that trading in the UK was taking place by way or business. We are proposing that it would be reasonable to infer that trading was taking place by way of a business or trade only if:

a) the person's account or website appears to be a marketplace containing multiple items for sale;

b) the person is holding themselves out as acting by way of a trade or business, for example by describing themselves as a professional, a gun trader or as doing business, or is using a company or business name; and

c) law enforcement provides evidence that the person is acting by way of a trade or business.

26.225 We welcome comments on this approach in response to consultation.

---

[48] Section 311 of the Sentencing Act 2020 provides that defendants must receive a minimum custodial sentence of 5 years for possession of a prohibited weapon committed when they were aged 18 or over (or a minimum of 3 years for the same offences committed when they were aged under 18).

Authorisation

26.226 We understand that there is no central, public, easily consulted register of which persons are authorised to deal in firearms in the UK. However, we understand that authorised dealers behave in ways which are likely to make unlawful sales identifiable to services. In particular, a website purporting to sell directly and remotely to UK users would not be authorised.[49]

3D printing of firearms

26.227 The Firearms Act does not cover 3D printing instructions for guns. However, we consider that in practice this type of content would be caught by one of the priority offences in schedule 5 of the Act. The offence in section 54 of the Terrorism Act 2000 relates to 'providing weapons training'. As set out above, this covers providing instruction or training in the making of firearms, making it available either generally or to one or more specific persons, and there is no state of mind requirement. Jurisdictional considerations play no part in this analysis.[50]

26.228 A defence is available if the user concerned can prove that their action or involvement was *wholly* for a purpose other than assisting, preparing for or participating in terrorism. We are consulting on our provisional view that this is likely to be difficult to show in relation to content circulated on the internet in places readily accessible to the general public.

Knives and 'offensive' weapons

26.229 A disparate set of weapons are caught by the legislation relating to knives and offensive weapons: section 1 of the Restriction of Offensive Weapons Act 1959; Article 53 Criminal Justice (Northern Ireland) Order 1996 (S.I. 1996/3160 (N.I. 24) (flick knives and gravity knives); and section 141(1) of the Criminal Justice Act 1988 (offensive weapons). The Government has recently announced that it intends to add to the list of offensive weapons.[51] If this takes place before we issue our final guidance, we would update our draft to take account of the change.

26.230 These are fairly straightforward offences which apply to any exposure for sale and do not have any state of mind requirements. Acknowledging the jurisdictional issues discussed above, the main challenge arises in correctly identifying the weapon itself. We propose to list the weapons themselves but also to provide a description of them, which is taken from UK government guidance.[52]

26.231 However, these offences are also subject to a series of defences which may be important for the creative, historical and religious sectors. At present, we have little evidence of how these defences are applied in practice or what effects are likely to follow from the way in which the Act defines illegal content. Nor do we have evidence of the risk of gaming by bad actors as a result of the content of our guidance. We welcome further evidence on this. We note however that we do not have discretion to change the definition of illegal content which is set by the Act. All we can do is set out the basis upon which we consider it reasonable for a

---

[49] [**CONFIDENTIAL✂**].
[50] This is not only because of the definition of illegal content - the underlying offence applies extra-territorially (see section 17 of the Terrorism Act 2006).
[51] Home Office and Philp MP, Chris, 2023. Government bans machetes and zombie knives [press release]. [accessed 19 September 2023].
[52] His Majesty's Government. Selling, buying and carrying knives and weapons. [accessed 25 September 2023]

service to infer that a defence exists. The applicable defences are set out below relating to offensive weapons only.

a) A defence arises if the weapon in question was being sold for the purpose of being used in a theatrical performance (including rehearsals for such performance) or in the production of a film or television show. We are proposing to give guidance that it would be reasonable to infer this if, for example, the weapon is blunt and is marketed as being for this use.

b) A defence arises if the weapon in question is being sold for the purpose of being used in historical re-enactment or in a sporting activity for which public liability insurance covering third parties has been obtained. It appears likely to be difficult for services to draw any inferences about whether or not events or activities in the future will be properly insured, and we do not currently hold any evidence ourselves which would warrant drawing inferences about the likelihood that such activities are insured. If there are no grounds to infer that activities are insured, the defence would not be available and content marketing weapons for use in such activities would be illegal content. We are consulting on saying that it would be reasonable to infer that this defence is available if, for example, appropriate wording is included in the content (e.g. 'buyers must have insurance').

c) A defence arises if the weapon in question as being sold for the purpose of being used for a religious ceremony or for religious reasons. Again, it appears to us likely to be fairly difficult for services to draw inferences about this and we would welcome evidence from representatives of any religions which use weapons in ceremony or for religious reasons as to when it may be reasonable for internet services to draw this inference. Meanwhile, we are proposing to say that it would be reasonable to infer this if, for example, appropriate wording is included in the content. There is a similar, but separate defence for curved swords presented by a Sikhs to another person in a religious or ceremonial event. Here, we propose that it would be reasonable to infer this defence may be successfully relied upon if, for example, appropriate wording and imagery is included in the content and if the language used in the advertisement is one commonly in use in the Sikh community (for example, Punjabi or English).

d) A defence arises if the weapon in question is an antique, that is made more than 100 years before the content was posted, or in the case of a curved blade of over 50cm in length, before 1954. We consider that specialist knowledge is usually likely to be required to date weapons advertised and that such knowledge is unlikely to be available to services. We therefore propose to say that it would be reasonable to infer that this defence may be successfully relied upon if the content states that the conditions for antiquity are met, so long as this is not obviously inconsistent with any other description or depiction of the item concerned.

e) Finally, a defence arises if the weapon in question was made at any time according to the traditional methods of making swords by hand. We are consulting on the view that it would be reasonable to infer this defence may be successfully relied upon, for example, the content says this is the case, as long as this is not obviously inconsistent with any other description or depiction of the item concerned, or with the volumes of such items sold by the same seller.

26.232 There are also defences when the conduct is only for the purposes of making the weapon available to a relevant museum or gallery and (for most offensive weapons) where it is

carried out only for the purposes of functions carried out on behalf of the Crown or of a visiting force.[53] While we are including this in the annexes to our guidance, we are not proposing to say anything about these in the main body of our guidance, because they do not think they would be relevant to content on a U2U service or a search service.

'Marketing' offence

26.233 A separate offence exists in section 1 and 2 of the Knives Act 1997 for the marketing of otherwise lawful knives in a way which indicates, or suggests, that the knife is suitable for combat; or is otherwise likely to stimulate or encourage violent behaviour involving the use of the knife as a weapon.

26.234 This offence is defined in the legislation in substantial detail, and we are proposing to refer services to guidance published by the UK's Crown Prosecution Service for examples of how it may manifest in practice.

'Buyer' offences

26.235 A large number of the priority offences relating to weapons are not absolute prohibitions, but partial prohibitions. It is lawful to trade in weapons, but (where relevant) the buyer must be appropriately authorised, the right age, and not a criminal. As set out above, the actual sale takes place offline, and so in considering these offences, the offences of encouraging, assisting and conspiracy are more likely to be relevant.

26.236 However, a person cannot encourage, assist or conspire with themself. The user responsible for the content is not the same person as the person committing the main offence. The jurisdictional issues considered above mean, in addition, that the content is unlikely to be illegal content unless there are reasonable grounds to infer that the user responsible for the content was aware that the purchase or sale itself would take place in the UK.

26.237 In our proposed guidance, we have grouped all these offences together based around the nature of the offence and the nature of the buyer.

26.238 Notwithstanding that it is likely to be difficult for services to identify individual items of illegal content, they will still need to consider the risk of such illegal content being present, and U2U services will also need to consider the risk that they will be used to facilitate the commission of these offences.

Sexual exploitation of adults

26.239 The sexual exploitation of adults offences comprise causing or inciting prostitution for gain[54] and controlling a prostitute for gain[55].

Causing or inciting prostitution for gain

26.240 The offence of causing or inciting prostitution for gain has been an issue in an online context principally in relation to so-called 'sex for rent' advertisements on user-to-user 'classified

[53] Section 1(3) Restriction of the Offensive Weapons Act 1959; section 141(8) of the Criminal Justice Act 1988.
[54] Section 52 of the Sexual Offences Act 2003; article 62 Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2).
[55] Section 53 of the Sexual Offences Act 2003;Article 63 Sexual Offences (Northern Ireland) Order 2008 (S.I. 2008/1769 (N.I. 2).

ads' services, where the victims are incited into prostitution in exchange for accommodation.[56]

26.241 The offence requires the potential victim to be at risk of *becoming* a prostitute as a result of the action.[57] It is therefore implicit within the definition of the offence that the potential victim was not already involved in sex work prior to accessing the content in question, and that the content would cause or incite them to *become* a prostitute by engaging in sex work.[58]

26.242 We recognise that most services are unlikely to be in a position to know whether or not their users are already sex workers. However, we are consulting on our view that this is not always a barrier to them drawing a reasonable inference that content incites prostitution. We consider it uncontroversial that *most* users of *most* U2U and search services are not working as sex workers. Save as set out below, we therefore consider it reasonable to say in our guidance that, absent evidence to the contrary, it is usually reasonable for services to infer that a user viewing content causing or inciting prostitution is not a sex worker. This applies where content has been posted in a public forum other than a service primarily or solely used for the selling of sexual services.

26.243 The clear exception to this argument is services (or accounts within services) which are specifically dedicated to sex work, where it is conceivable or even likely that the majority of users viewing the advertisements in question are already engaged in sex work. In these cases, we propose that it is *not* reasonable to assume that users viewing such content are not already engaging in sex work, and therefore the criteria for illegality *cannot* be reasonably be inferred to have been met.

Controlling a prostitute for gain

26.244 For content to amount to the offence of controlling a prostitute for gain, a service must have reasonable grounds to infer: a. that the user uploading the content is, *through the content*, controlling the activities of a separate person or persons related to their prostitution in any part of the world; *and* b. that the person uploading the content does this for, or in expectation of, gain for themself or a third person.

26.245 We consider it unlikely to be clear from online content alone whether content relating to this offence is posted by a person (or persons) acting on their own behalf or whether it is posted by someone that is controlling a prostitute or prostitutes for gain.

26.246 We have considered whether there are any 'warning signs' that may indicate that a sex worker is being controlled for gain by another person or persons. However, we have provisionally concluded that it is not possible to identify any of the factors we considered as signs of control, as they are also likely to be present in arrangements which are unlikely to amount to control. For example, a piece of content advertising sexual services on behalf of two or more sex workers may indicate that it has been posted by someone who is

---

[56] See, for example 'Sex for rent arrangements and advertisements' in: Crown Prosecution Service, 2019. Prostitution and Exploitation of Prostitution. [accessed 28 August, 2023].

[57] *R v Ubolcharoen* [2009] EWCA Crim 3263.

[58] We use the term sex work here to refer to the specific acts of prostitution that are covered by this section of the Sexual Offences Act 2003. However, we acknowledge that sex work also has a broader meaning which encompasses activity which does not amount to prostitution under the Sexual Offences Act, and that many sex workers would not recognise themselves as prostitutes. We use the term where necessary in order to properly reflect legislation.

controlling several sex workers for gain. However, it could equally be the case that two sex workers have chosen to advertise together in an effort to stay safe or because of language barriers faced by one or more of the parties. We recognise the importance of ensuring that sex workers are able to conduct their business on online services in a way which protects and promotes their safety. We are concerned that setting out prescriptive guidance on when the offence of controlling a prostitute for gain may be inferred, based on indicators such as the ones mentioned, we may undermine this effort, and potentially drive sex workers towards less safe environments in conducting their business, such as offline settings.

26.247 As a result of this conclusion, we believe services are unlikely to be able to reach the reasonable grounds to infer threshold for this offence unless they receive information from a credible third party (like for instance law enforcement) that the content has been implicated in a successful conviction or otherwise amounts to an offence.

Adult image-based sexual offences

Extreme pornography

26.248 Extreme pornography is a 'possession' offence and we set out our reasoning on inferring possession above.

26.249 Knowledge of the content of extreme pornography images is not required – the statutory defences deal with that. Knowledge that the person has uploaded an image is required, however we think it is reasonable for services to infer that.

Intimate image abuse offences

26.250 The Act includes as priority offences the intimate image abuse offences of both England/Wales, and Scotland. [59] These are similar to one another, but not identical.

26.251 In addition, the Act will, eventually, replace the existing English/Welsh offence of intimate image abuse with a new, wider one.[60] This has not yet been brought into force. Once it has, the old offence will be revoked. The new one would become a priority offence if the Secretary of State decides to make regulations under section 222 of the Act, adding it to schedule 7.

26.252 For the purposes of this consultation, we have assumed that the offence will be brought into force and will be a priority offence before we issue our final Illegal Content Judgments Guidance.

26.253 As set out above, for the purpose of identifying illegal content, it does not matter what country a user is posting the content from if the service it is being posted to is being regulated by Ofcom. In effect, content is illegal content if it amounts to *either* the English/Welsh offence or the Scottish offence.

26.254 However, considering each offence separately in turn is likely to be onerous for services and may be confusing to content moderation teams as well. After careful thought about the similarities and differences between the offences, and for the reasons set out below, we are

[59] The Scottish offence is that in section 2 Abusive Behaviour and Sexual Harm (Scotland) Act 2016.
[60] The new English/Welsh offence will be multi-limbed. In our guidance we have focussed on sub-sections 66B(1) and 66B(2). We have not dealt with sub-section (66B(3) as we believe that in practice, most if not all content which would be identifiable as amounting to this offence, would also amount to an offence under sub section 66B(1).

consulting on a version of the guidance which collapses the two offences together, led mostly by the English/Welsh version of the offence which on balance we consider likely to be identifiable first.

26.255 The key differences between the offences are:

a) **Consent:** the principal reason why the English/Welsh offence is easier to consider than the Scottish one is that to show the Scottish offence, the service would need 'reasonable grounds' on which to infer a negative - that the photograph or film concerned has not previously been disclosed to the public at large, or any section of the public, by the individual or with the individual's consent. While it would be possible to build a content reporting form which asked this question specifically, we are not aware that services generally do, so they may have no information on previous disclosure. By contrast, the English/Welsh offence only requires positive evidence about consent in relation to the content itself. In many cases reasonable grounds to believe that the disclosure was non-consensual are likely to be provided by the fact of there being a complaint from the person depicted, or by contextual information around the content. The English/Welsh offence is therefore likely to be easier to show.

b) **What content is caught:** the English/Welsh definition of the offence is both more detailed and broader than the Scottish one. It captures a photograph or film if it shows or appears to show the person participating or engaging in an act which a reasonable person would consider to be a sexual act; the person doing a thing which a reasonable person would consider to be sexual; all or part of the person's exposed genitals, buttocks or breasts; the person in an act of urination or defecation, or the person carrying out an act of personal care associated with the person's urination, defecation or genital or anal discharge. The reference to all or part of a person's 'exposed' genitals, buttocks or breasts includes a reference to all or part of them being visible through wet or otherwise transparent clothing, them being exposed 'but for the fact that they are covered only with underwear', and them being exposed 'but for the fact that they are obscured, provided that the area obscured is similar to or smaller than an area that would typically be covered by underwear'. This is broader than the Scottish offence in that it definitely captures deepfakes, in that it captures urination/defecation and associated personal care which may not be sexual, and in that it captures exposure through wet clothing or obscuring.

c) **State of mind:** the English/Welsh offence occurs when the user uploading the content does not 'reasonably believe' that the person depicted consents. The Scottish offence applies the Scottish definition of recklessness. A person is reckless as to whether the disclosure would cause fear, alarm, or distress if they 'failed to think about or were indifferent as to' whether the disclosure would have that result. However, we provisionally consider that for the purposes of the Illegal Content Judgements Guidance in practice, on the information likely to be available to services, this is likely to be a distinction without a difference. A person who failed to think about or was indifferent as to causing fear, alarm or distress would not have reasonable grounds to believe in consent, and the basis for services to draw either inference is likely to be the same.

Threats to disclose intimate image abuse content

26.256 In addition to the illegal act of disclosing intimate image abuse content, it is also an offence to *threaten* to disclose a photograph or film which shows (or appears to show) a person in

an intimate state, where there the person making the threat has the appropriate state of mind.

26.257 Here, we have drafted taking the Scottish offence as the main offence. Although as set out above, the Scottish definition of content caught by the offence is narrower than the English/Welsh definition, we consider that in practice a service is very unlikely indeed to have enough information about the nature of the image concerned to know for sure that it falls within the English/Welsh but not the Scottish definition.

26.258 On the other hand, for the threat offences there are material differences in the state of mind requirements. Both England/Wales and Scotland provide that the offence may be committed if the user is 'reckless'. However, the Scottish definition of 'recklessness' is wider than the English/Welsh one. We have therefore drafted our guidance with reference to the Scottish definition.

'Cyberflashing' and inferring a purpose of sexual gratification

26.259 The Act creates a new offence, colloquially known as 'cyberflashing'. This is not a priority offence, but it is a relevant non-priority offence. We have considered it here because it takes place online.

26.260 Cyberflashing refers to the unsolicited sending of a photograph or film of someone's genitals to someone through digital communication channels. Whilst a person of any gender may be victim of cyberflashing, evidence shows that this behaviour disproportionately affects women and girls, and that a majority of the perpetrators are men. Cyberflashing can cause victims severe distress, and often leaves victims feeling unsafe, vulnerable and upset. We are committed to reducing harm from cyberflashing as part of our wider effort to make the online space safer for women and girls.

26.261 While the law now criminalises acts of cyberflashing in some cases, the act of sending an unsolicited photograph or film of someone's genitals is not *in itself* illegal, and this means that not all online content resulting from such behaviour can be said to amount to illegal content either. The state of mind requirement for this offence is intent to cause distress, alarm, or humiliation or recklessness as to whether this would be caused, combined with a purpose of sexual gratification in sending the photograph or film.

26.262 We recognise that cyberflashing can be very harmful to its victims. However, it will often be difficult to infer whether the state of mind element of this offence is present. Provisionally, we consider that the existence of an image of genitalia taken on its own with no further context is unlikely to provide a sufficient indication of the user's intent for a service to have reasonable grounds to infer that an offence has been committed. That said, we recognise that many services may choose to take down such images pursuant to their terms of service, without necessarily making an illegal content judgement. We welcome views on this point, and in particular any further evidence as to the state of mind of those engaging in cyberflashing, from stakeholders.

26.263 We note that in many cases cyberflashing images are sent via direct messages. In this context it appears to us that the important thing is not so much that services remove the content (the recipient can, after all, delete it), but that victims have the opportunity to prevent further such messages being sent to them. As set out in volume 4, we are proposing measures in our Codes recommending that services with a high risk of harms including harassment, offer users the ability to block the sender. We consider this tool would go some way towards enabling users to protect themselves from unwanted contact of all kinds.

Immigration and human trafficking

26.264 The Act contains five priority offences to do with immigration and human trafficking. They are offences relating to illegal entry into the UK[61]; facilitating unlawful immigration[62]; and the human trafficking offences[63].

26.265 Of these, the offences relating to illegal entry into the UK cannot be committed online as it is not possible for a person to 'enter the UK' except physically. The related offences of 'encouraging' or 'assisting' may be relevant. However, the state of mind requirements are high (including 'intent') and the analysis is complicated by the fact that, for example, it is not necessarily unlawful to cross the channel or to invite others to take trips by boat.

26.266 We are therefore consulting on our provisional view that reasonable grounds to infer that content is illegal may exist in cases where the following has been made available to services via law enforcement:

a) information justifying an inference of intent; and

b) information that the entry being encouraged or assisted is illegal;

or

a) information justifying an inference that the person uploading the content believes an offence will (not may) be committed and their act will (not may) encourage the commission of that offence; and

b) information that the entry being encouraged or assisted is illegal.

26.267 The offence of facilitating unlawful immigration is only committed where the content posted amount to an 'act' that facilitates the breach or attempted breach of immigration laws in a range of different countries, and the person posting it has knowledge or reasonable cause for believing that the individual whose breach is facilitated is not a national of the United Kingdom. The range of possible acts which might facilitate the commission of such a breach is very broad, but it is difficult to see how any of them could be committed online. Applying the offence in practice would also require services to have a detailed knowledge of the immigration laws in many countries.

26.268 Our provisional view is that it is not proportionate or practical to expect services to be able to do this. We are therefore consulting on our view that in cases where information justifying this inference has been made available to services by law enforcement or a court order, reasonable grounds to infer may exist.

26.269 Finally, it is our provisional view that the Scottish version of the human trafficking offence is broader than the English/Welsh and Northern Irish ones, we have therefore focused on that in our guidance[64]. The Scottish version of the offence takes place when a person (Person A) takes a 'relevant action' with a view to another person (Person B) being exploited.

26.270 In the context of online content, 'relevant actions' are most likely to be the recruitment of another person, or the arrangement or facilitation of acts of transport or transfer, or of

[61] Section 24(A1), (B1), (C1), (D1) of the Immigration Act 1971.
[62] Section 25 of the Immigration Act 1971.
[63] Section 1 of the Human Trafficking and Exploitation (Scotland) Act 2015 (asp 12) (human trafficking); section 2 of the Modern Slavery Act 2015; and section 2 of the Human Trafficking and Exploitation (Criminal Justice and Support for Victims) Act (Northern Ireland) 2015 (c. 2 (N.I.)) (human trafficking).
[64] Human trafficking is not a legal term, it is just a way to refer to the offence.

harbouring, or of receiving of another person – so long as all of these actions are done with a view to exploiting the person involved. 'Exploiting' is a defined term which we propose to set out in detail in our guidance.

26.271 The offence contains a number of different provisions relating to jurisdiction/connection to the UK. However, amongst them is the provision that the offence is committed if 'any part of the relevant action takes place in the UK'. We consider that this will *always* be met, in the case of online content and therefore do not propose to discuss it in our guidance. This is because content is only illegal content if it 'amounts' to the offence, so there can be no conceptual distinction between the content and the relevant action. If the content is accessible to users in the UK, the relevant action will take place in the UK. And, in any event, the Act provides that for the purposes of determining whether content is illegal content, it does not matter whether anything done in relation to it takes place in any part of the UK. The Explanatory Note to the Act confirms that this means content will 'amount to an offence' regardless of whether the criminal law would require the offence, or any element of it, to take place in the United Kingdom (or a particular part of it).

26.272 We are proposing to consult on the basis that reasonable grounds to infer that content amounts to an offence are likely to exist where content makes explicit reference to the exploitation of another person. However, most perpetrators of human trafficking will not be honest about their intentions to exploit people through their actions. For example, they may use false job advertisements which appear legitimate. We propose to say that services should have regard to any evidence provided by UK law enforcement agencies that in their view there are reasonable grounds to infer that content is posted for the purposes of exploitation.

Suicide and self-harm

26.273 The offence of encouraging or assisting suicide is a priority offence under the Act. The offence of encouraging or assisting serious self-harm is a relevant non-priority offence created by the Act, which, as set out above, we propose to include in our Guidance if it is brought into force in time.

26.274 In preparing guidance on these offences, we have been very mindful that political discussion of the law on assisting suicide is likely to be found on both U2U and search services, and that it is not unlawful to discuss the fact that assisting suicide is lawful in some countries. Nor is it unlawful to portray suicide or self-harm in content (for example, to make a movie in which a person ends their own life), or report on suicides or self-harm. Finally, a person who feels suicidal or may want to harm themselves may well express themselves on a U2U service, and this may be helpful to them. The right to freedom of expression protects all these kinds of content.

26.275 In our draft guidance, we have therefore described types of content which we do not consider to be illegal content, as well as types of content which may be. We have also discussed in some detail the basis on which a reasonable finding may be made on the state of mind of intent. We consider the context to be particularly important here. We are consulting on our view that where specific, practical or instructive information on how to end one's life is posted to a forum or within a chat in which suicidal ideation is discussed, it may be reasonable to infer that intent to assist (attempted) suicide exists by virtue of information having been posted. Where an encouragement to end one's life is posted in

response to what appears to be a credible threat by another user that are about to take their own life, it may also be reasonable to infer intent.

Foreign interference offence and false communication

26.276 A new offence, which is a priority offence, was created by the National Security Act 2023 and will be known as the 'foreign interference offence'. In addition, the Online Safety Act creates a new relevant non-priority offence of false communication. Neither is yet in force, though the false communications offence will come into force two months from the date of the Online Safety Act.

26.277 As both these offences are new, they lack a body of case law or academic discussion on which Ofcom can draw for their interpretation. They are both also likely to be particularly difficult to identify in practice, because they depend heavily on context and on circumstances offline.

26.278 For the time being, our proposed approach is that our guidance should describe the offences and the questions a service should ask itself. Bots play an important role in generating and spreading content which is likely to amount to a foreign interference offence and we propose also to draw attention to this in our guidance.[65]

Animal welfare

26.279 At a fairly late stage in the progression through Parliament of the Bill which became the Online Safety Act, the offence in section 4(1) of the Animal Welfare Act 2006 (unnecessary suffering of an animal) was added to it. We will consult in due course on how we propose to include that offence in our guidance.

Non-priority offences

26.280 We have discussed some specific non-priority offences above. These are offences we consider likely to come to services' attention. With those exceptions, we do not propose to attempt to give guidance on all possible relevant non-priority offences in our Illegal Content Judgments Guidance. We consider that the risk that our guidance may be incorrect or out of date would be very significantly greater if we were to attempt to do so. And we consider the volume of information we would need to provide may be so great as to overwhelm services, in particular smaller services.

26.281 We are proposing instead that services should ensure that they respond appropriately to information given to them by law enforcement or through a court order regarding content that has been implicated in a successful conviction of a non-priority offence, or (taking legal advice as appropriate) where a reasoned case is put to them by a law enforcement body.

[65] Bots are an umbrella term that refers to a software application or automated tool that has been programmed by a person to carry out a specific or predefined task without any human intervention. Bots are often employed on services to post content at scale without the need for repeated human intervention. In many cases bots are used for benign purposes, however, bots may also be used to spread spam and malicious content, including misinformation and phishing attempts.