ANNEX A: GAPS BETWEEN OFCOM'S ANALYSIS OF CAUSES AND IMPACTS OF ONLINE HARM (VOL 2) AND PROPOSED MITIGATIONS (VOLUME 4, ANNEX 7 & 8)

Ofcom says in Volume 2 - "The Causes and Impacts of Online Harm" that it "*presents our assessment of the causes and impacts of illegal online harms based on the evidence that we have gathered over the past three years. The analysis we set out here forms part of our duty under the Act to assess the factors that can cause a risk of harm to individuals on a service. We expect services to have reference to it when they carry out their own risk assessments*."

Ofcom's assessment focuses on the over 130 priority offences defined in the Act (but not non-priority offences), grouped in volume 2 under 15 broad kinds of illegal harm. Ofcom notes in its introduction to volume 2 that "*The impact of the harms we have looked at can be extremely severe. It is not limited to the online world but can also profoundly affect people's lives offline*". Also that "*the kinds of illegal harm we have looked at occur on services of all types. Services as diverse as social media services, dating services, marketplaces and listings services, search services, adult services, and file-storage and file-sharing services are all used to disseminate some of the types of harmful content we have looked at in this volume.* ***Bad actors use both large and small services to spread illegal content****, although the way in which they use large services sometimes differs from the way in which they use small services*." (our emphasis)

It also notes that "certain 'functionalities' stand out as posing particular risks", picking out in its introduction the following:

- End-to-end encryption
- Pseudonymity and anonymity
- Livestreaming
- Recommender systems

The introduction goes on: "*We expect services to think about these risk factors when doing their risk assessments (see Volume 3). As we explain in Volume 4,* ***we have designed a number of the measures in our Codes of Practice to target high-risk service types and functionalities*** *… The role of the new online safety regulations is not to restrict or prohibit the use of such functionalities, but rather to get services to put in place safeguards which allow users to enjoy the benefits they bring while managing the risks appropriately*." We therefore might expect that volume 4 and the Codes themselves would reflect the level of risk threat that the functionalities identified in volume 2 pose.

Volume 2 is a commendable standalone document within the suite of documents that make up the illegal harms consultation - it brings together a vast amount of evidence as to how the illegal offences covered by the Act are prevalent online and is analytical and thorough in identifying the functionalities that contribute to this prevalence and/or risk of harm to individuals. Many of these functionalities are vectors for multiple harms.

However, this assessment does not flow through to the mitigation measures set out in the Codes of Practice (Annex 7) (for user to user services) and Annex 8 for search, which focus primarily on content takedown and measures to deal, ex-post, with illegal content once it has been identified. The rules-based nature of the Codes (which is NOT required by the definition of "measures"[1]) - specifying specific recommended measures rather than describing desired outcomes - and the fact that these are designed as a "safe harbour" (eg if companies follow the measures they will be judged to have complied with their duties under the Act[2]), means that there is no incentive for companies to implement mitigating measures beyond those described in the codes - even if their risk assessment has flagged that their service poses particular risks from other ex ante functionalities (such as design choices). Furthermore, smaller companies are in many instances exempt from implementing particular mitigating measures due to Ofcom's proportionality analysis.

We set out in our full consultation response more detail on where the choices made by Ofcom in these regards are problematic. In this supporting document we seek to illustrate where the gaps between the analysis of harm and the recommended mitigations of it lie. The following tables provide detailed analysis on the individual functionalities, the number of offences where Ofcom identifies that particular functionality is a contributory factor, and the appearance (or not) of mitigating measures relating to this functionality in the codes of practice for user to user and search services. A summary "at a glance" table is provided for U2U (pages 3-4) and search (p5). Supporting tables for user-to-user services (from p6) and search services (pp19-23) provide more detail and extracts from Ofcom's consultation materials. We have divided the measures in annex 7 and annex 8 into "ex ante" and "ex post", the latter largely applying to measures relating to content moderation and takedown once illegal content has been identified on a service. While we have used the term "ex ante" in relation (generally speaking) to the non-takedown measures, the measures identified are focused on the presence of illegal content on the service (or the search functionality enabling users to find it) so are not what we would term "safety by design" measures, which we would classify as biting at a systemic level separate to the nature of the particular types of content (e.g. business model, or measures that are not directed to a particular type of content for eg rebalancing weighting in recommender tools)

---

[1]Section 236(1) Online Safety Act

[2] "Services that choose to implement the measures we recommended in our Codes of Practice will be treated as complying with the relevant duty. This means that Ofcom will not take enforcement action against them for breach of that duty if those measures have been implemented. Service providers may seek to comply with a relevant duty in another way, but the Act provides that, in doing so, they must have regard to the importance of protecting users' right to freedom of expression within the law, and to the importance of protecting users from breaches of relevant privacy laws. Where providers do take alternative measures, they must keep a record of what they have done and explain how they think the relevant safety duties have been met. (Volume 4, para 11.7)

COMPARISON OF VOLUME 2 FUNCTIONALITIES WITH USER-CODE OF PRACTICE MITIGATIONS (ANNEX 7): SUMMARY TABLE

Functionality	Illegal harms	Code of practice: ex ante mitigations	Code of practice: ex post mitigations
Content: posting, commenting, hyperlinks, including images and video	15	Limited to user controls measures (eg muting, blocking): A9	Content moderation & takedown: 4A-F
Reposting or forwarding content	5	None	Limited: vague reference to “limiting time”
Livestream & live audio	9	None	None (except as type of “content”)
Use of hashtags to direct to illegal content	5	None	None (except as type of “content”)
Editing visual content	9	None	None
Screen capturing or recording	1	None	None
User tagging	5	None	None
User profiles	10	Limited to user controls: A9	None
User connections	8	Limited to default settings, user controls: A6 & A9	None
User search	2	None	None
User groups	9	None	None
User base profile	3	None	Limited: references in 4E, 5B
Recommender systems	11	None	Limited: A6 (“limited time”), A9 safety metrics
Group messaging	6	None	None
Encrypted messaging	10	None	None
Direct messaging	15	Limited to user controls: A9 And included in A7: Default settings for child users where services are high risk for CSAM/grooming	None

Functionality	Illegal harms	Code of practice: ex ante mitigations	Code of practice: ex post mitigations
Anonymous user profiles	15	A9C has recommendations re user labelling schemes, but this is only limited to services at risk of fraud or the foreign interference offence	None
Fake user profiles	13	As above (A9C)	None
Business model - inc small, fast-growing services; ad revenue	5	None	None
Payment facility	2	None	None
User location	4	Included in A7 default settings measures, but only limited to services at high risk of grooming	None
UGC search facility	3	None	None
Posting goods or services for sale	7	None	None
Building lists or directories	2	None	None

COMPARISON OF VOLUME 2 FUNCTIONALITIES WITH SEARCH CODE OF PRACTICE MITIGATIONS (ANNEX 8): SUMMARY TABLE

NB the analysis in Volume 2 of the search functionalities that cause harm is less detailed and presented in a different way to the evidence in the user-to-user section.

Functionality	**Illegal harms**	**Code of practice: ex ante mitigations**	**Code of practice: ex post mitigations**
Typing in searches for illegal content	8	Limited: provision of warnings for CSAM searches; and provision of suicide prevention information in relation to suicide/self-harm searches	Search moderation & takedown: 4A-F - these measures largely replicate the user-to-user content moderation measures but with 4A applying to deindexing or deranking illegal content. An additional deindexing measure applies to CSAM URLS (4G)
Ranking	-	None	As above
Reverse image search	1	None	None
Search prediction or personalisation	3	None	Limited: requires action when there is a user report that predictive search suggestions are directing users to priority illegal content
Revenue models	2	None	None
Commercial profile/size	-	None	None
Gen AI/chat bots	-	None	None

COMPARISON OF VOLUME 2 FUNCTIONALITIES WITH CODE OF PRACTICE MITIGATIONS (ANNEX 7) - USER TO USER SERVICES - FULL TABLE

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
CONTENT FUNCTIONALITIES			
Posting content	**Terrorism** **Grooming*** **CSAM** **Suicide & self-harm*** **Harassment, stalking, threats and abuse*** **Hate offences*** **Controlling or coercive behaviour*** **Drugs offences** **Unlawful immigration** **Intimate image abuse** **Proceeds of crime offences** **Fraud** **Foreign Interference offence** **False communications offence** **Epilepsy trolling**	**Limited** A9 (enhanced user controls) sets out that large services at high risk of offences marked * in LH column (plus suicide/self-harm & controlling and coercive behaviour) and that hav e user profiles, and at least one of three functionalities (user connection, posting content, communication including DM and commenting on content) allow blocking or muting of users. The Government produced its own "best practice" guide for safety by design for platforms that enabled private or public interaction in 2021: https://www.gov.uk/guidance/private-and-public-channels-improve-the-safety-of-your-online-platform	**Extensive** Content is primarily dealt with in the codes via moderation: • 4A: swift takedown • 4B: internal content policies (only for large and multi-risk services) • 4C: performance targets (ditto) • 4D: prioritisation of review of content (ditto) • 4E: resourcing • 4F: moderator training There are specific, detailed measures re hash-matching for CSAM and detection of CSAM URLs P45: The definition table at the end of the codes says re "content"; "*For the avoidance of doubt, comments, titles and descriptions are considered to be 'content' within this definition, as are livestreaming videos or audio, and hyperlinks.*"
Commenting on content	**Terrorism** **CSAM** **Grooming**	**Limited** A9 also sets out (for services that	**Extensive (as per content above)**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
	Suicide and self harm **Harassment, stalking, threats and abuse** **Hate offences** **FIrearms offences** **Fraud offences** Epilepsy trolling Cyberflashing	meet the same condition as above) that users should be able to disable comments.	
Hyperlinks - eg use to direct users to more extreme content	**Terrorism** **CSAM** Suicide and self-harm **Hate offences** Drugs offences **Extreme pornography** **Foreign interference offence** Epilepsy trolling	**None recommended**	**Extensive (as per content above)**
Reposting or forwarding content	**Suicide and self-harm** **Harassment, stalking, threats and abuse** **Intimate image abuse** Proceeds of crime offences **Foreign Interference offence**	**None recommended.**	**Limited** Section A6 (terms of service) obliquely covers this in referring to specifying "how the provider will minimise the length of time" illegal content is present"
Posting images or videos	**Intimate image abuse**	**None recommended.**	**Extensive (as per content above)** P45: The definition table at the end of the codes says re "content"; "For the avoidance of doubt, comments, titles and descriptions are considered to be 'content' within this definition, as are livestreaming videos or audio, and hyperlinks."

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
Livestream	Terrorism **Grooming** **CSAM** **Suicide and self-harm** **Hate offences** **Sexual exploitation of adults** **Intimate image abuse** Fraud (sextortion) **Cyberflashing**	**None recommended** **NB the government produced its own "best practice" guide to "safety by design" for livestreaming in 2021: https://www.gov.uk/guidance/live-streaming-improve-the-safety-of-your-online-platform**	**Limited (except as type of "content")** P45: The definition table at the end of the codes says re "content"; "For the avoidance of doubt, comments, titles and descriptions are considered to be 'content' within this definition, as are livestreaming videos or audio, and hyperlinks." This does not consider how the functionality of livestreaming is used to facilitate the offences in the first place.
Livestream - Sending messages via livestream	**Grooming**	**None recommended**	**Limited (except as type of "content")** P45: The definition table at the end of the codes says re "content"; "For the avoidance of doubt, comments, titles and descriptions are considered to be 'content' within this definition, as are livestreaming videos or audio, and hyperlinks." This does not consider how the functionality of livestreaming is used to facilitate the offences in the first place.
Live audio	Terrorism	**None recommended**	**Limited (except as type of "content")** P45: The definition table at the end of the codes says re "content"; "For the avoidance of doubt, comments, titles

Functionality	Related Offences *offences in bold are where the functionality is highlighted in the introductory summary	Code of practice: systemic or ex-ante mitigation?	Code of Practice: recommended ex-post mitigation?
			and descriptions are considered to be 'content' within this definition, as are livestreaming videos or audio, and hyperlinks."
Content tagging - Eg hashtags	**Suicide and self harm** **Hate offences** **Drugs offences** **Intimate image abuse** **Epilepsy trolling**	**None recommended.**	**None recommended.**
Screen capturing or recording	**Terrorism** **Grooming** **CSAM** Intimate image abuse Cyberflashing	**None recommended**	**None recommended**
USER FUNCTIONALITIES			
User tagging	**Harassment, stalking, threats and abuse** Controlling or coercive behaviour **Firearms offences** Foreign interference offence **Epilepsy trolling**	**None recommended.**	**None recommended.**
User profiles	**Grooming*** **Harassment, stalking, threats and abuse*** **Hate offences*** **Drugs offences** **Unlawful immigration** **Sexual exploitation of adults**	**Limited** A9 (enhanced user controls) sets out that large services at high risk of offences marked * in LH column (plus suicide/self-harm & controlling and coercive behaviour) and that	**None recommended**

Functionality	Related Offences *offences in bold are where the functionality is highlighted in the introductory summary	Code of practice: systemic or ex-ante mitigation?	Code of Practice: recommended ex-post mitigation?
	Proceeds of crime offences **Fraud** **Epilepsy trolling** **Cyberflashing**	have user profiles, and at least one of three functionalities (user connection, posting content, communication including DM and commenting on content) allow blocking or muting of users NB the Government produced its own "best practice" guide for "safety by design" for user profile functionality in 2021: https://www.gov.uk/guidance/users-account-details-and-activity-visible-to-others-improve-the-safety-of-your-online-platform	
User connections	**Terrorism** **Grooming*** **Harassment, stalking, threats and abuse*** **Controlling or coercive behaviour*** **Drugs offences** **Fraud** **Foreign Interference offence** **Epilepsy trolling**	**Limited** Section A7 includes recommendation (only for services at high-risk of grooming, or a large service at medium-risk of grooming) that default settings do not include children in network expansion prompts and connection lists A9 (enhanced user controls) sets out that large services at high risk of offences marked * in LH column (plus suicide/self-harm) and that have user profiles, and at least one of three functionalities (user connection, posting content, communication including DM and commenting on content) allow	**None recommended**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
		blocking or muting of users	
User search	Grooming Cyberflashing	**None recommended**	**None recommended**
User groups	Grooming CSAM **Suicide and self-harm** Controlling or coercive behaviour **Drugs offences** **Unlawful immigration** **Extreme pornography** **Fraud** Foreign interference offence	**None recommended**	**None recommended**
User base profile	Terrorism (demography) **Grooming, CSAM (children)** **Harassment etc (women)**	**None recommended**	**Limited** Recommendation 4E re content moderation says the services needs to take into account "the particular needs of its United Kingdom user base as identified in its risk assessment, in relation to languages." Recommendation 5B re complaints says "In designing its complaints processes for relevant complaints, including its reporting tool or function, the provider should have regard to the particular needs of its United Kingdom user base as identified in its risk assessment. This should include the particular needs of: a) children (for services likely to be accessed by children and considering

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
			the likely age of the children using that service); and b) disabled people" Neither of these address the way in which the service design might ensure that users identified in the risk assessment might be protected in the first instance from harm.
RECOMMENDER SYSTEMS			
Recommender systems	Terrorism* **Grooming/CSAM*** **Suicide and self harm*** Harassment, stalking, threats and abuse* **Hate offences*** Controlling or coercive behaviour **Drugs offences*** **Extreme pornography*** Intimate image abuse* **Foreign Interference offence*** **Epilepsy trolling**	**None recommended**	**Limited** Section A6 (terms of service) obliquely covers this in referring to specifying "how the provider will minimise the length of time" illegal content is present" Section A8 (recommender system testing) requires (but only for services that conduct test and are at a high risk of two types of the harms marked * in the LH column) that it analyse the safety metrics from its tests to understand if changes to the recommender system would increase the risk of users encountering illegal content There is no upstream requirement in the code to ensure that services to consider the design of their recommender systems in the first place.

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
MESSAGING FUNCTIONALITIES			
Group messaging	Terrorism **CSAM** **Suicide and self-harm** **Intimate image abuse** Intimate image abuse Fraud	**None recommended**	**None recommended**
Encrypted messaging	**Terrorism** **Grooming** **CSAM** **Drugs offences** **Sexual exploitation of adults** Intimate image abuse **Proceeds of crime offences** **Fraud** **Foreign Interference offence** **False communications offence**	**None recommended**	**None recommended**
Direct messaging	**Terrorism** **Grooming*** **CSAM** **Harassment, stalking, threats and abuse*** **Hate offences*** **Controlling or coercive behaviour*** **Drugs offences** **Firearms offences** **Sexual exploitation of adults** Intimate image abuse **Proceeds of crime offences** **Fraud**	**Limited** A9 (enhanced user controls) sets out that large services at high risk of offences marked * in LH column (plus suicide/self-harm) and that have user profiles, and at least one of three functionalities (user connection, posting content, communication including DM and commenting on content) allow blocking or muting of users A7 includes recommendation (only	**None recommended**

Functionality	Related Offences *offences in bold are where the functionality is highlighted in the introductory summary	Code of practice: systemic or ex-ante mitigation?	Code of Practice: recommended ex-post mitigation?
	False communications offence **Cyberflashing**	for services at high-risk of grooming, or a large service at medium-risk of grooming) that as a default, child users should not receive messages from a non-connected user; and if the service does not have user connections, child users can actively confirm if they want to receive a direct message from someone they don't know	
Direct messaging - Sending images via messaging	Grooming	**Limited (see above)**	**None recommended**
ANONYMOUS/FAKE ACCOUNTS			
Anonymous user profiles	Terrorism Grooming **CSAM** **Suicide and Self-Harm** **Harassment, stalking, threats and abuse** **Hate offences** Drugs offences **Firearms offences** **Extreme pornography** **Intimate image abuse** **Fraud** **Foreign Interference offence** **False communications offence** **Epilepsy trolling** Cyberflashing	**Limited** A9C: user verification/labelling schemes sets out that large services at high risk of either or both of fraud and the foreign interference offence; and has user profiles under a relevant scheme (notable users or monetised scheme) should have consistently applied policies to reduce the risk of harm to users associated with that scheme. These policies should include "how the provider will treat relevant users and the content they post including recommender systems, content	**None recommended**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
		curation, user reporting and complaints, quality assurance, fact checking, content moderation, account security" There are no recommended measures to address the role of anonymous or fake user profiles in the list of offences in the LH column NB the Government produced its own "best practice" guide to "afety by design" for anonymous or multiple account creation in 2021; https://www.gov.uk/guidance/anonymous-or-multiple-account-creation-improve-the-safety-of-your-online-platform	
Fake Profiles	Grooming **CSAM** Suicide and self-harm **Harassment, stalking, threats and abuse** **Controlling or coercive behaviour** **Unlawful immigration** **Sexual exploitation of adults** **Intimate image abuse** **Proceeds of crime offences** **Fraud** **Foreign Interference offence** **False communications offence** **Epilepsy trolling**	**Limited** A9C: user verification/labelling schemes sets out that large services at high risk of either or both of fraud and the foreign interference offence; and has user profiles under a relevant scheme (notable users or monetised scheme) should have consistently applied policies to reduce the risk of harm to users associated with that scheme. These policies should include "how the provider will treat relevant users	**None recommended**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
		and the content they post including recommender systems, content curation, user reporting and complaints, quality assurance, fact checking, content moderation, account security" There are no recommended measures to address the role of anonymous or fake user profiles in the list of offences in the LH column.	
MISCELLANEOUS			
Business model: • Low capacity and early-stage services	Terrorism	**None recommended**	**None recommended/** This is an issue re the small vs large differentiation, covered elsewhere in our analysis.
Business model: • Ad revenue	**Foreign Interference offence** **Hate Offences** **Sexual Exploitation of Adults** **Extreme Pornography**	**None recommended**	**None recommended**
Payments/transactions capability	Terrorism **CSAM**	**None recommended**	**None recommended**
User location	Grooming **Harassment, stalking, threats and abuse** **Controlling or coercive behaviour** **Sexual exploitation of adults**	**Limited** Section A7 includes recommendation (only for services at high-risk of grooming, or a large service at medium-risk of grooming)	**None recommended**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
		that as a default automated location information displays are turned off child users. This does not address the role of user location functionality wrt to VAWG.	
Editing visual media	Grooming CSAM Hate offences Controlling or coercive behaviour **Extreme pornography** Intimate image abuse Foreign interference offence **False communications offence** Epilepsy trolling	**None recommended**	**None recommended**
Downloading content	CSAM **Extreme pornography** Intimate image abuse	**None recommended**	**None recommended**
UGC content searching or filtering	Suicide and self harm Drugs offences **Firearms offences** **Extreme pornograrphy** **Proceeds of crime offences** **Fraud**	**None recommended**	**None recommended**
Posting goods or services for sale	**Drugs offences** **Firearms offences** **Unlawful immigration** **Sexual exploitation of adults** **Extreme pornography**	**None recommended**	**None recommended**

Functionality	**Related Offences** *offences in bold are where the functionality is highlighted in the introductory summary	**Code of practice: systemic or ex-ante mitigation?**	**Code of Practice: recommended ex-post mitigation?**
	Proceeds of crime offences **Fraud**		
Building lists or directories	CSAM Extreme pornography	**None recommended**	**None recommended**

COMPARISON OF VOLUME 2 FUNCTIONALITIES WITH CODE OF PRACTICE MITIGATIONS (ANNEX 8) - SEARCH SERVICES - FULL TABLE

The analysis on the functionalities related to user access to illegal content via search services is presented in a different way by Ofcom in volume 2: a high-level summary narrative that talks about functionality in relation to particular offences, rather than an offence-by-offence analysis. The table below includes some of the core narrative for each functionality in volume 2, along with a similar assessment of ex-ante or ex-post measures as per user-to-user services. NB the Government produced its own "best practice" guide for "safety by design" for search functionality in 2021: https://www.gov.uk/guidance/search-functionality-improve-the-safety-of-your-online-platform (It is not referenced by Ofcom.)

Functionality	Related Offences	Code of practice: systemic or ex-ante measures?	Code of practice: ex-post measures?
Typing in searches for illegal content 6T.33 Functionalities related to general search "include the underlying potential for illegal content on webpages indexed by search services to appear in, or via, search results; the features visible to users to optimise search results (such as recommended searches, autocomplete suggestions); and those which determine results behind the scenes (such as ranking algorithms) … These service characteristics are designed largely to optimise the accuracy and usefulness of search results to users. Where a user is intentionally seeking out illegal content – which is considered the most likely situation in which a	Terrorism Hate Extreme pornography CSAM Firearms offences Drugs offences Fraud Suicide and self harm	Limited 7B: provision of CSAM content warnings - applies to large general search services "The provider should employ means to detect and provide warnings in response to search requests of which the wording clearly suggests that the user may be seeking to encounter CSAM and uses terms or combinations of letters and symbols that explicitly relate to CSAM. Warnings should not be provided in response to search requests using terms which, on their face, do not relate to CSAM." 7C: provision of suicide crisis prevention information - this is to be provided in response to a) "general queries regarding suicide; and b) queries seeking specific, practical or instructive information	Extensive Content is primarily dealt with in the codes via the search moderation duties Eg: 4A: The provider should have systems or processes designed to deindex or downrank illegal content of which it is aware (a 'search moderation function') - applies to all services. 4B: internal content policies (large and multi-risk) 4C: performance targets (ditto) 4D: prioritization for review (ditto) 4E: resourcing (ditto) 4F: training (ditto) Plus 4G: deindexing CSAM URLS (all services)

Functionality	Related Offences	Code of practice: systemic or ex-ante measures?	Code of practice: ex-post measures?
user would encounter content that amounts to an offence – these same optimising characteristics have the unintended consequence of helping that user encounter illegal content.		regarding suicide methods.	
Ranking 6T.28: “General search services use proprietary algorithms (‘ranking’) to perform this prioritisation function. The ranking process uses factors such as how closely the search query is matched and the website’s functionality and authority (the perceived value of the site’s content and how often it is linked to by other sites). As with all functionalities, the ranking process is designed to provide accurate and reliable content, but it can be manipulated by users to increase the likelihood of illegal content being displayed to users. For example, the tactic of keyword stuffing (filling a web page with keywords or numbers in an attempt to manipulate rankings in		None recommended	Extensive (see above) 4A: The provider should have systems or processes designed to deindex or downrank illegal content of which it is aware (a ‘search moderation function’)

Functionality	Related Offences	Code of practice: systemic or ex-ante measures?	Code of practice: ex-post measures?
search results) has been identified in research looking at how easily illegal content relating to fraud can be accessed via search services."			
Reverse image search Vol 2 notes that evidence of how this is used in relation to searches to purchase drugs and that, while the evidence is limited on other offences, "it is possible that the reverse image search functionality also presents opportunities to access content relating to other prohibited items" (para 6T.36)	Drugs offences	None recommended	None recommended
Search prediction or presonalisation 6T.37 "It is reasonable to assume that these functionalities can increase the risk of accessing illegal content amounting to a range of offences, unless effective mitigations are in place to prevent this, or indexed content is blocked."	Suicide or self harm Hate Fraud	None recommended	Limited 7A: removal of predictive search suggestions (large general search services that use predictive search functionality) NB This measure only requires those services to provide a "means to easily report predictive search suggestions which they consider to direct users towards priority illegal content" NOT ex-ante measures to prevent such predictive search suggestions arising in

Functionality	Related Offences	Code of practice: systemic or ex-ante measures?	Code of practice: ex-post measures?
			the first place.
Revenue models - ad-based models Evidence suggests that advertisements on search services may be misused for illegal activity.	Coercive control Foreign interference offences	None recommended	None recommended
Commercial profile/size “Despite the limited evidence, we consider that *search services that are low-capacity or at an early stage in their lifecycle may face an increased risk of harm on their services*” (6T.46)		None recommended	None recommended
Gen AI/chat bots Volume 2 says “Research indicates that search services integrated with GenAI chatbots could be used to facilitate fraud whereby a perpetrator could covertly collect personal information including the user's name, email, and credit card information. There is also evidence illustrating how such services could be used to	Fraud	None recommended	None recommended

Functionality	Related Offences	Code of practice: systemic or ex-ante measures?	Code of practice: ex-post measures?
share malicious links and steer search results towards manipulated content.” (para 6T.18)			