Summary of Reasons – Internet Search Engine Services Code

8 September 2023

eSafety decision

The eSafety Commissioner (**eSafety**) has decided to register the *Internet Search Engine Services Online Safety Code (Class 1A and Class 1B Material)* (the **SES Code**), submitted to eSafety on 14 August 2023. The SES Code meets the statutory requirements set out in section 140 of the *Online Safety Act 2021* (Cth) (the **Act**).

Background

The Act permits eSafety to register an industry code that has been developed and submitted by a body or association that represents a particular section of the online industry. To register an industry code, eSafety must be satisfied that it meets the requirements under section 140 of the Act, including that it provides appropriate community safeguards for any matters of substantial relevance to the community.

On 11 April 2022, eSafety gave a notice to the Communications Alliance and Digital Industry Group Inc (the **Applicants**) under section 141 of the Act requesting that they develop an industry code dealing with certain matters (the **Notice**).

On 18 November 2022, the Applicants submitted a draft SES Code to eSafety pursuant to the Notice. In February 2023, eSafety gave a statement of preliminary views on this draft to the Applicants and invited the Applicants to submit a final version addressing feedback from that statement.

On 31 March 2023, the Applicants submitted a further draft of the SES Code to eSafety for registration (**March SES Code**), with a covering document entitled 'Request for Registration of Online Safety Codes' (the **Request**).

On 31 May 2023, the Applicants were advised that eSafety reserved the decision with respect to the SES Code because the code did not deal with recently announced developments regarding the integration of generative artificial intelligence (**AI**) into search engine services and address associated risks of harm. eSafety invited the Applicants to submit a revised code.

On 6 July 2023, the Applicants submitted a revised version of the SES Code. Following feedback from eSafety, the Applicants submitted a further revised version of the SES Code on 14 August 2023, together with an addendum to the 6 July request for registration and a copy of the previously submitted documentation.

Scope of the SES Code

The SES Code applies to internet search engine services (**SES Providers**), so far as those services are provided to Australian end-users.[1] The definitions of 'search functionality' and 'search results' in the SES Code provide that any search engine service functionality, including features which are enabled by artificial intelligence, would be subject to the SES Code. The SES Code contains measures to address, minimise and prevent harms associated with access and exposure to the most harmful forms of online material accessed through internet search engine services.

Material covered by the SES Code includes:
- **class 1A material**, which is comprised of child sexual exploitation material, pro-terror material, and extreme crime and violence material, and
- **class 1B material**, which is comprised of crime and violence material and drug-related material,

in each case as described in Annexure A to the SES Code Head Terms, which reflects the *Classification (Publications, Films and Computer Games) Act 1995* (Cth) (**Classification Act**) and related instruments.[2]

These types of material are subcategories of class 1 material under the Act, which is material that has been or would be refused classification under the Classification Act. Serious harms are associated with these kinds of material whenever it is produced, distributed or consumed.

A future industry code or industry standard will be developed to address class 2 material under the Act, which includes material that has been or would be classified X 18+, R 18+, Category 1 Restricted or Category 2 Restricted under the Classification Act.

Importantly, the definitions of class 1 (and class 2) material under the Act capture deepfake or synthetically generated material which otherwise meets the requirements of the particular class of material. This means that commitments applying to class 1 content in the SES Code apply to content generated on the service via artificial intelligence.

eSafety assessment of the SES Code

The SES Code sets out a range of minimum compliance measures for SES Providers that the Applicants submit provide appropriate community safeguards in relation to the matters identified in the Request.

eSafety agrees that the matters identified by the Applicants in their Request, which are materially the same as those matters identified by eSafety in the Notice requesting a code, are matters of substantial relevance to the community.

---

[1] 'Australian end-users' is used throughout the industry codes but is defined in clause 2 of the Head Terms as an end-user in Australia to align with the language and scope of the Act. Both terms are used in this document.
[2] Importantly, the nature of the material, including its literary, artistic or educational merit, and whether it serves a medical, legal, social or scientific purpose, is relevant to the assessment of class 1B material – see section 11 of the Classification Act. Material only falls within class 1B if there is no justification for the material.

eSafety found that the SES Code provides appropriate community safeguards in relation to each of those matters by requiring SES to:

1. take reasonable and proactive steps to create and maintain a safe online environment for Australian end-users
2. empower people to manage access and exposure to class 1A and class 1B material, and
3. strengthen transparency of, and accountability for, class 1A and class 1B material.

The obligations outlined below include requirements on SES providers to take steps to address the risk associated with class 1 material being discoverable and surfaced in search results, whether generated by AI or surfaced from the traditional indexing of material sourced from the World Wide Web.

Creating and maintaining a safe online environment for Australian end-users

Specific obligations relating to the first of these objectives include requirements on SES Providers to:
- take appropriate steps to improve the effectiveness of machine learning algorithms and/or models with the aim of reducing the accessibility or discoverability by Australian end-users of class 1A materials in search results
- implement appropriate policies, processes, systems and technologies designed to:
  - reduce the accessibility or discoverability of class 1A material by Australian end-users, and
  - limit Australian end users' exposure to class 1B materials in search results
- ensure there are designated personnel to oversee the safety of the service including compliance with the Act and SES Code
- conduct a review of the risk that class 1A and class 1B materials are accessible by Australian end-users in search results prior to implementing any new feature or functionality of the service that significantly increases risk, and take reasonable steps to mitigate any additional risk
- for SES Providers with more than 500,000 monthly active Australian end-users, implement appropriate procedures for collaborating with eSafety, law enforcement, non-governmental or cross-industry organisations to facilitate the safe, secure and lawful sharing of information that prevents child sexual exploitation material from appearing in search results
- refer complaints from the public concerning non-compliance with the SES Code to eSafety where the SES Provider is unable to resolve the complaint within a reasonable timeframe
- update eSafety on any significant changes to the functionality of the service that are likely to have a material effect on the risk of access or exposure to, or distribution of, class 1A or class 1B materials
- take appropriate steps to improve systems, processes and/or technologies that aim to reduce the safety risks to end-users concerning synthetic materials generated by artificial intelligence that may be accessible via the internet search engine service
- take appropriate steps to ensure that any features integrated in an internet search engine service that are enabled by artificial intelligence such as longer form answers, summaries

or materials do not return search results that contain child sexual abuse material
- take appropriate steps to ensure that autocomplete or predictive entries that appear on the internet search engine service do not include, without justification, terms that have known associations to child sexual exploitation material based on keyword searches and input from independent organisations that have expertise in combatting child sexual exploitation material, and
- take appropriate steps to ensure that search results specifically seeking images of known child sexual abuse material are accompanied by deterrent messaging that outlines the potential risk and criminality of accessing images of child sexual abuse material.

Empowering people to manage access and exposure to class 1A and class 1B material

Specific obligations relating to the second of these objectives require SES Providers to:
- where relevant, make clear when a user is interacting with any features using artificial intelligence
- provide age-appropriate safety settings and tools to Australian end-users, such as 'safe search' functionality, which enable users to limit exposure to explicit and/or graphic materials
- make available clear and accessible information to end-users about online harms and the measures that users of the service can take to improve the safety of themselves and children in their care
- respond to report and delist requests from end-users, and
- promptly notify law enforcement or appropriate non-government organisations about child sexual exploitation material on their services if the provider forms a good faith belief that the material provides evidence of an immediate threat to a child/adult in Australia (this is intended to supplement existing laws).

Strengthening transparency of and accountability for class 1A and class 1B material

Specific obligations relating to the third of these objectives require SES Providers to:
- publish easily accessible and plain language information on their policies and approaches to class 1A and class 1B material, and
- submit an annual report to eSafety on request, setting out steps taken to comply with the SES Code, including information on the volume of child sexual exploitation material or pro-terror material flagged and responded to by the service, and the number of complaints about code compliance.

eSafety considers that these requirements will together create effective and meaningful obligations that reflect the role of SES Providers in the online safety ecosystem in addressing the risks of class 1A and class 1B content. Specifically, eSafety considers that the obligations will help address the risks of generative AI features integrated into search engine service being misused to create synthetic child sexual abuse material.

Next steps

The SES Code will shortly be published on the eSafety Commissioner's Register of industry codes and industry standards for online safety. The SES Code will come into effect six months after registration.