# DRAFT STATEMENT OF STRATEGIC PRIORITIES FOR ONLINE SAFETY RESPONSE

1. The Secretary of State's Draft Statement of Strategic Priorities (SSP) for Online Safety was published in November 2024. This response, a version of which was submitted to the Department for Science, Innovation and Technology in December 2024, reflects some of the themes and concerns that have emerged in discussions facilitated by the Network since the publication of the SSP. We do not speak on behalf of the organisations we work with but have provided in the annex a list of 16 of the organisations who have contributed to, or who support, this response.

2. In general, we are pleased to see such a strong statement from the Secretary of State on the importance of the online safety agenda and the need for pace, urgency and greater ambition in Ofcom's implementation of it. The clear statement of the Government's intent here is not just important for Ofcom, but also for the companies preparing to be regulated.

3. We are also broadly pleased with the proposed framework of action set out in the draft SSP, particularly the fact that it covers many of the areas that have been of particular concern to the OSA Network and our civil society partners in the year since Royal Assent. For example, whether Ofcom is focusing enough on "safety by design"; whether the "safe havens" that are created by Ofcom's narrow approach to measures in the codes will mean that platforms' actions to address illegal content and activity will not be robust enough; and whether there would be enough pace and flexibility in Ofcom's iterative approach to the codes to keep up with new evidence of emerging harm. None of these issues have been resolved by Ofcom in the year since they published their illegal harms codes and risk assessment guidance for consultation; in fact they have doubled-down on many of them in the final versions that have recently been published. This SSP is therefore even more important to delivering better online safety outcomes in the months and years ahead. We commend the Secretary of State and his officials on their responsiveness in this regard and the clarity around much of their stated objectives.

4. In our view, the tone and content of the draft statement will be particularly helpful in stretching Ofcom's ambitions and also providing the foundation for a stronger and more robust approach to implementation in the next five years. For example:
    a. The reference to the Secretary of State's expectation that there will be a **"material reduction"** in illegal content and activity online. This addresses some of the concerns expressed by those in our Network that Ofcom's codes are something of a "lowest common denominator", reflecting existing good practice (or "banking what we already

have", as Ofcom have put it in their "Approach to the Codes" recent publication) without pushing companies to go further. Ofcom's approach - combined with the "safe harbour" built into the Act itself - also, potentially, creates conditions where existing safety measures could be removed by regulated services without consequences. We feel the Secretary of State's ambition is particularly important with regard to the work Ofcom is undertaking to improve **protections for women and girls**: while the guidance the regulator will produce in February 2025 is not enforceable, it will be of persuasive effect; the successful enforcement of the illegal content duties in respect of the priority offences that particularly affect women and girls will also be vital in shifting the dial towards a safer overall online experience.

b. In relation to innovation and the adoption of safety tech, the statement that "we expect that this will not only raise the floor of what is expected from online services but will also **raise the ceiling** of what is possible by nurturing a culture of effective innovation in practice" is notable. This also seems to address concerns about the *de minimis* nature of the codes and to incentivise companies to go further. Innovation however does not need to be limited to "bolt on" or *ex post* safety tech solutions but - as we set out further in our section on safety by design below - can be about the inherent design of the services and their business models.
c. The overall focus and emphasis on the criminal end of **misinformation and disinformation** is helpful, particularly in the light of the riots in the UK in the summer and the limitations of the Act in dealing with much of the material that contributed to the offline violence.

***Amending the OSA***

5. The Secretary of State notes that there may be areas where the OSA may need to be amended to easily enable the delivery of these objectives. In light of the above, we would hope Ofcom makes the necessary recommendations in this regard in its response to this consultation and does not wait for more time to elapse before it tells the Secretary of State what it already knows will need to change to meet his objectives. For our part, **we would urge the Secretary of State to make two small amendments to the Act as a matter of urgency to ensure Ofcom acts to close the gap between the scale of risks of harm they evidence in the work and the limited numbers of measures they propose in their codes to mitigate those risks.** There are two specific aspects of the Act that are relevant here: the "safe harbour" provisions in the Act (section 49) which state that a provider is in compliance with their duties if they follow the measures in the codes; and the requirement, set out in schedule 4 of the Act, that "measures in codes of practice must be sufficiently clear, and at a sufficiently detailed level, that providers understand what those measures entail in practice".

6. As a result of Ofcom's narrow interpretation of those two parts of the Act, a number of recommendations from civil society have been rejected that would have helped address many of the risks the regulator has evidenced. The provisions therefore need to be amended: the "safe harbour" needs to be removed, in favour of an approach more akin to that in the EU Digital Services Act, where following the codes is seen to be desirable but not in itself sufficient for

compliance; and the specificity written into schedule 4 needs to be loosened, such that, where there isn't available evidence to meet the "clear and detailed" threshold for measures, Ofcom can require a "best endeavours" approach to mitigating risks from companies instead. Neither of these amendments would alter the substance or policy intent of the Act and would receive cross-party support in both Houses.

7. While we welcome the tone and impetus behind the Secretary of State's message to Ofcom, we would also remind DSIT that the issue of pace also applies to their own role in ensuring that the OSA framework is implemented - in full - as quickly as possible. For example, the delay in responding to Ofcom's advice on categorisation, which was received in February, only to announce in mid-December that, despite significant civil society concerns, the Secretary of State would accept it has had a material impact on the delivery of many of Ofcom's key activities. Despite the DSIT consultation on **supercomplaints** closing last January, the relevant secondary legislation will not be laid until next Spring. This introduces a significant delay in confirming the designation process for eligible organisations and/or for those organisations to be able to use the supercomplaint system once Ofcom has started using its powers.

8. There is a human cost to this delay: harms increase, new technology creates new risks and individuals' online and offline safety is under threat. Including key milestones and dates - both for Ofcom to be held to account for these priorities and for DSIT, in delivering its own obligations - would be helpful in the final version of the SSP.

***Our views on the government's strategic priorities for online safety***

9. There are risks in setting out what inevitably is a limited set of priorities and the effect that this may have on Ofcom's programme of work and its decisions about financial resourcing. We therefore would like to see a clear statement within the final version of the SSP that **the Secretary of State does not expect Ofcom to *deprioritise* anything that is not included in the SSP**. We welcome assurances we have had from DSIT on this but feel that without a strong statement to that effect, things that are not listed in the SSP might not get the same emphasis and/or Ofcom will argue that it has limited resources to do the "new" things required, without adjusting resources assigned already to existing commitments.

10. Timescales for a number of the deliverables have already slipped. Ofcom deliberately delayed including many issues in the illegal harms codes on the grounds that they would be picked up by its phase 3 work. This work is now further delayed and some of the specific measures to address harms to adults - for example, those arising from unverified and fake accounts - will remain unaddressed for at least another 18 months.

11. Where Ofcom has been given long lead times, e.g. for the report on researcher access to data, it has not demonstrated any desire to go faster than the slowest pace permissible under the Act. The injection of pace and agility that the Secretary of State clearly wants to see needs to apply to the whole programme of Ofcom's work, not just the priorities listed in the SSP: it should be seen

as taking Ofcom's existing activity and commitments as read and providing emphasis and/or stretching objectives on top of that, not instead of it.

Priority 1: Safety by Design

12. Safety by design is a really important principle and one which is front and centre in the Act, in section 1. We are very pleased that the Secretary of State has made this a priority for Ofcom in the next five years, particularly in the context of the pressures from a number of fronts at present to push for bans to deal with harms to children - whether bans for smartphones, or bans for access to social media. Neither phones nor social media are intrinsically bad but the design choices that are made by their developers and manufacturers can make the experience of users - whether adults or children - more or less safe. Focusing on safety by design - particularly while the Secretary of State waits for the evidence from the new review of the evidence of the harms from smartphones and screen time - is a vital way forward in delivering a safer experience for all.

13. Despite claims in Ofcom's recent press release, there is no meaningful "safety by design" approach inherent in the illegal harms codes and no changes have been made to codes (which do not once mention "safety by design") since we flagged this issue in our consultation response and, subsequently, in a number of detailed meetings with Ofcom on the same topic. So to ensure that Ofcom understands what the Secretary of State is asking for, we would recommend that the material on safety by design is more clearly structured, starting with **a clear statement as to what safety by design should mean**. Currently, this discussion arrives at 1.4 but the discussion there does not necessarily create a coherent set of expectations for Ofcom, nor is there a clear definition. For example, does it include product testing and what the expectations around that are? The expectation that "proportionate safety by design principles" should be embedded is not clear either: is this relating to principles being embedded in product development processes or within the product itself? (Please see the discussion on this in the recent paper from Professor Woods; we would also remind DSIT, as we have frequently reminded Ofcom, that its predecessor department, DCMS, produced a set of safety by design principles in 2021, which - without explanation - do not seem to be in use any more.)

14. There are some overlaps in the statement with regard to safety by design and other expectations which are relevant but separate; for example, on user empowerment tools (eg paragraph 1.4) or the adoption of safety tech as a means to mitigate harms that are already present, rather than to design them out from the outset. These distinctions and connections should be recognised; for example, it should be clear that user empowerment tools (though a good thing in and of themselves) are not a substitute for safety by design. This then links in to 3.1 (p 20) and the expectation that new tech is safe for users.

15. More specifically, given that the orientation towards services that are safe by design is given no separate implementation in the Online Safety Act, it seems that the risk assessment and safety duties are the means to achieve this goal. It would therefore be helpful to see a stronger **link between the idea of safety by design and the general mitigation duties** (both U2U and search as well as illegal content and harmful to children), especially s 10(2)(b) and (c), s12(2), s 27(2) and (3) and s 29(2). Further, proactive duties are not just about finding examples of content and

taking them down but more general design choices too.

16. We welcome the juxtaposition of the emphasis on safety by design with the reference to **violence against women and girls**, given that the guidance that Ofcom will produce needs to take a cross-cutting, holistic approach if it's to deliver the intent of Parliament that it tackle the disproportionate impact of online harm experienced by women and girls. We look forward to seeing this reflected in the guidance when it is published for consultation in February. This guidance is going to be pivotal to the success of the Government's target to halve VAWG - in this regard, DSIT will need to be held to account as much as Ofcom over the term of this SSP's relevance. Any definition of safety by design should include racialised and gendered aspects from the outset, as well as incorporating ideas, such as "design justice" that are specifically relevant to women and girls and minoritised groups. Consideration of the intersectional impacts of design decisions also needs to be worked in.

17. We would also like to see **a definition of "agile regulation"** so we can understand how this can or should be more clearly linked to safety by design, e.g. with a stronger connection to product testing - this allows services to take responsibility for their products and to minimise an approach which relies on prescriptive rules. It might enable a wider focus from the regulator on products and their overall design, addictiveness, etc than is currently prescribed in the Act - though there were plenty of discussions in Parliament on this aspect, particularly in the Lords, during its passage. Partly as a result of the way in which the Act is drafted, Ofcom has taken an overly content-focused approach to implementation to date. This is another area they have doubled-down on in their final illegal harms codes, with no new measures or revised approaches, despite extensive conversations with civil society organisations in recent months as to what "safety by design" looks like. The default in focusing on content is to consider ex-post solutions (eg bolting on safety tech and a reliance on content moderation, takedown etc) rather than considering the harmfulness of the system and the products and services themselves and leaves services and Ofcom playing "whack a mole" with content cascading through systems designed to encourage virality.

18. There is also a need to be clear that - while the evidence base is important, and Ofcom have been tasked by the Secretary of State, via the draft SSP, with a number of additional reviews and research projects - a key underpinning principle for safety by design has to be **a precautionary approach**: not rolling out or introducing new products or features where there is a material risk that they might cause harm, nor waiting for the evidence to be collected once they are deployed to prove that that is the case. There is a tension throughout the response between the focus on safety by design and the number of **additional evidence reviews** that DSIT is asking Ofcom to undertake.

19. We would hope that the final SSP will be clearer on the fact that these two things are in some ways distinct: safety by design provides the precautionary approach to risk that can prevent harm; collecting the evidence on how harm works and how various functions, techniques or processes work in mitigating that harm, is part of an iterative process that can continuously improve the online safety environment and raise the floor across all services in terms of what is

reasonable to expect all services to do. The former does not have to wait for the latter, and it is important that Ofcom - in its response to the SSP - does not default to timescales that prioritise the latter without driving forward on the former at the same time.

20. A further aspect would be some consideration, even at a high level, as to what safety by design means for enforcement, complaints and user redress.

21. On some of the specifics in this section, many of the children's charities and campaigners in our network are broadly happy with draft SSP, particularly the focus on **age appropriate-design** and effective **age assurance** which comes through strongly. Our partner organisations will be providing more detail on some of the specifics. One area of feedback relates to the reliance on evidence gathering, which we touched on above: while Ofcom does need to keep doing this, there are lots of examples of age-appropriate design techniques and effective age assurance technologies already in use so it's important to signal to Ofcom that they can expect companies to be mirroring existing best practice and not wait for the compilation of an evidence base, by the regulator, before taking action here.

22. One area of note is that there is no reference to some of the challenges around **private messaging** or any detail on what DSIT expects Ofcom to do with regard to this issue, which we know is a concern to children's charities. In the context of the use of the term "safe haven", and its relevance to the sites where CSAM proliferates, this seems like an oversight.

<u>Priority 2: transparency and accountability</u>

23. Transparency and accountability is important for all users, including children and young people, but they are not mentioned in this section (para 2.3). We would like to see the SSP revised to note the fact that children and young people require clear, child-friendly Terms of Service to be accessible and understood, and therefore to make transparency meaningful for them; and for children and young people to be mentioned in relation to expectations around complaints and reporting processes too.

24. From a VAWG perspective, we would like to see more of a focus on non-carceral/ non-criminalisation approaches including national redress schemes, banning of harmful apps eg deepfake and nudification apps, funding prevention work etc.

25. There is still a gap in the Act relating to consumer redress and alternative dispute mechanisms and there is no reference to the supercomplaints process, which is dependent on DSIT laying secondary legislation (see above).

26. On the issue of small platforms, proportional harm based on platforms' size needs to be addressed and understood. Small platforms that are targeted on causing specific harm (whether targeting abuse at minorities, or encouraging suicide or self-harm) can have a significant and material impact on a large number of users without necessarily acting at the scale that would designate them as "large platforms". There needs to be a recognition that some platforms have a supersized impact when harm is carried out on them. One way forward might be to use a

different measurement of harm for these platforms - for example, where they are named in Coroners' Prevention of Future Deaths reports, in relation to suicide, for example.

Priority 4: inclusivity and resilience

27. We would like DSIT to consider whether some of the issues set out here (eg p24) start pointing to further amendment of media regulation (see also 4.1). There is also no mention of advertising and how the advertising system, and the role of influencers generating monetised content as a subset of that, impact on the spread of misinformation and disinformation. Ofcom's responsibilities for these issues goes beyond its Online Safety Group so it would be helpful to see an acknowledgement in the SSP that its regulatory responsibility for other areas of media regulation needs to be considered when planning how to deliver these objectives.

28. While we welcome the emphasis on action on misinformation and disinformation, there might be a corresponding over-reliance on media literacy which is not borne out by the evidence on effectiveness and where the media literacy strategy, and its implementation, is already weak. Ofcom's proposals for "media literacy by design" are not reflected in their OSA implementation proposals nor are they joined up with "safety by design" - a point we made in our response to Ofcom's consultation on its media literacy strategy, but which was not addressed in the final version, which barely mentions the OSA. This is something that the SSP might address more directly, for example in relation to the amount of friction and what is promoted through a platforms' recommender tools.

29. We welcome the expectation that Ofcom's guidance on protections for women and girls, due next February, will summarise "in one clear place, measures that can be taken to tackle the abuse that women and girls disproportionately face online. This guidance will ensure that it is easy for platforms to implement holistic and effective protections for women and girls across their various duties."

30. On some of the specifics, we note that DSIT expects the new Advisory Committee on Disinformation and Misinformation to play a role in delivering the transparency and accountability priority (section 2.1, p 14) but it is not mentioned in relation to the priorities on inclusivity and resilience, where arguably its role will be more crucial. Notwithstanding the independence of the Committee, we would like to see something clearer on DSIT's expectations here, including who will be appointed to the Committee and how.

***Our views on where Ofcom's role in contributing to the strategic priorities could be clearer***

31. The language in the document might be clearer in places to give the Secretary of State greater assurance that the outcomes he is seeking to be delivered are understood and actionable. While we respect the fact that the Secretary of State cannot "direct" Ofcom, and that may be a factor in the looseness of some of the language, we feel that there are some specific areas that might be tightened up.

32. We would like to have seen more detail on how Ofcom will be held to account on its progress against the priorities in this statement and what it will need to provide to DSIT to illustrate this. While an annual report is required, it is not clear what the Secretary of State will be expecting to see here. We would hope that in Ofcom's first formal response to the final report to see their own detailed set of measures and milestones. We also hope that they use that first response to set out - clearly - where, in their assessment of what the Secretary of State is asking them to focus on, they feel "unable to use regulatory options to contribute to achieving the strategic priorities set out in this SSP due to the existing statutory framework". This clarity is needed early on if the Secretary of State is to have the information he needs in considering whether to "bring forward legislation to allow [Ofcom] to do so"; any delays or prevarication from Ofcom here will have a knock-on impact on bringing forward that legislation. (p7)

33. We think there is an oversight in the paragraph that sets out the "foundational protections of the Act", which "includes pursuing a reduction in illegal activity online. The Act requires online platforms to proactively identify and remove illegal content, including content related to terrorism, foreign interference, fraud, illegal abuse and threats, and stirring up hatred offences". There is no mention of child sexual abuse material here and it should be included in the final section. Similarly, the draft SSP refers to the fact that "The government is also clear that terrorism content must be tackled in our online environments. To achieve this, we expect Ofcom to use powers at its disposal to oversee a reduction in such content." But there is no equivalent reference to CSAM, despite both these offences being rolled into the provisions under section 121 of the Act.

34. While the SSP recognises the significance of violence against women and girls especially when seen in its impact on the individual, there is no discussion addressing indirect harm in relation to violence against women and girls and a statement on indirect and societal harms should be included. There are references to the societal impact of misinformation and disinformation - in the context of the riots in the summer in the UK - but there is a similar societal impact from misogynistic content online, which goes beyond the impact on individuals. See for example the recent report from the Government's former counter-extremism adviser on the impact of misogynistic influencers and content-creators such as Andrew Tate.

35. We also note that there is no reference to the Part 5 duties on pornography services. With reference to our concern above that things that are not listed in this SSP might not get appropriate attention from Ofcom going forward, we think this is also an oversight.

Section 2: transparency and accountability

36. We welcome the fact that the Government is putting such emphasis on greater understanding of harms and how they are best tackled. We are not sure that the focus on p16 that the transparency reports are can be meaningfully used by the public is necessarily helpful for Ofcom, who also need to ensure that the information they receive from platforms via this route also works for civil society organisations to be able to hold tech platforms to account as well as for Ofcom itself in their enforcement. We would not want to see transparency reports shortened or watered down for public use at the expense of others being able to use them in a more rigorous

and detailed way. We do not think that one report will meet all three purposes and, if publicly accessible reports are DSIT's priority, they therefore need to be a separate product.

Section 3: Agile regulation

37. In addition to a definition as to what DSIT means by this, we feel Ofcom would benefit from having this more clearly **linked to the general duty that underpins the legislative framework that regulated services need to be responsible for identifying and mitigating the risks on their services** and to be accountable for the outcomes as a result. Ofcom's approach to the codes of practice has been to be prescriptive about the measures, based on the evidence that it has for their effectiveness. The Act does not require Ofcom to take on that responsibility for itself. An approach that is more about outcomes, requiring regulated services to address the risks that they have identified in ways that are appropriate to their services - and to be more responsive to harms as they identify them, through whatever routes are open to them - will deliver more innovation as well as more responsiveness. It will also help Ofcom build the evidence they need to provide future recommendations on measures for adoption. We do think there is a missed opportunity here for the Secretary of State not to call for services to address all the risks they have identified in their risk assessment. We do not accept Ofcom's rationale, in their recent illegal harms publication, that they cannot do this because the Act's safety duties only require "providers to take proportionate steps and we can only make recommendations we are satisfied are proportionate, having impact assessed them."

38. We would note, however, that while the "future-proof" nature of the Act is mentioned a couple of times in the draft SSP, the OSA is only future-proofed insofar as it applies to user-to-user services and search; there are many harms and risks arising from various different technologies and products that are not in scope of the OSA and many areas of intersection between services that may be part "in" and part "out" of the regime. Setting out an expectation as to how Ofcom might address this in terms of its iterative approach and evidence-gathering, particularly with regard to the likely need for future legislation, would be helpful.

39. On a specific point in this section, regarding the **threats from AI-generated content and activity** (section 3.2), while we understand why DSIT might link some of the dialogue with Ofcom on this to the annual publication of its updated Strategic Approach to AI, we have concerns about this, The Strategic Approach to AI is not OSA-specific and annual reporting in relation to AI-generated harms on online services is not frequent enough given the acceleration of technology and the related acceleration of harm arising from it. Given that the Government is also considering how AI regulation might evolve in future, we are concerned that this commitment may be out of date long before the timeframe covered by the draft SSP ends. There is a much broader consideration of AI - for example, how it will underpin some of the mitigation measures, such as content moderation and recommender tools - that is not addressed here either.

40. Without rehearsing the details again (see the recent letter to the Prime Minister, which we co-signed), we remained concerned - along with a number of our partners - about the consequences arising from Ofcom's advice to DSIT, that the Secretary of State has now accepted, that **small but risky services** will not be included in category 1. As Baroness Morgan said recently

in Parliament when news of the Government's decision had reached her, this was in "direct contravention of the amendment passed in this House": an amendment which Morgan had won, defeating the then Government in a vote with cross-party support.

41. This part of the draft SSP is a restatement of DSIT's position, taking as read Ofcom's reassurances about their approach to illegal harm capturing much of the activity on smaller platforms, without acknowledging the motivation behind the amendment to the OSA to bring those small platforms into the regime precisely because many of them are set up to be harmful, while frequently operating at a level just below the threshold for criminality. (This was a point that Morgan made in the recent debate: "if the Government seriously wants to tackle violence against women and girls they need to be consistent across all legislation and treat the platforms carrying this content as seriously as they should be treated".) It also does not take into account the fact that many users are likely to gravitate towards smaller platforms when more stringent duties apply to the larger, more mainstream platforms precisely because enforcement is more difficult.

42. We would also flag to DSIT that they have effectively said (in section 3.4) that it is "ok" for adults to be exposed to content that is harmful and - in the case of suicide and self-harm sites - potentially lethal: "Where relevant, such sites must also keep children safe from content which does not meet the criminal threshold but is nonetheless harmful to them. This includes preventing children from seeing harmful legal self-harm and suicide content." This underlines why there is such a deep concern amongst civil society organisations that Ofcom has decided that these small platforms don't need to be covered by the most robust duties: they are set up to be harmful, whether for children or adults, and they should be placed in category 1.

43. Having accepted DSIT's advice and laid the regulations, the Secretary of State now needs to set out a clear statement as to how and when Ofcom will be required to keep this under "continual review" and his what steps he will take to amend the Online Safety Act, urgently, to ensure that Ofcom's next advice on this matter fully delivers on Parliament's intent when the OSA was revised to allow for small and risky services to be included in category 1.

44. We are grateful to the organisations and individuals below who contributed to this response and hope that the concerns above will be reflected in the final SSP in due course.

**Online Safety Act Network**

**December 2024**

**CONTRIBUTING ORGANISATIONS**

Glitch
Thomas William Parfett Foundation
Antisemitism Policy Trust
Parent Zone
Center for Countering Digital Hate (CCDH)
NSPCC
Barnardo's
Institute for Strategic Dialogue (ISD)
Suzy Lamplugh Trust
End Violence Against Women Coalition (EVAW)
Samaritans
Internet Watch Foundation (IWF)
Marie Collins Foundation
5Rights Foundation
Clean up the Internet
Mental Health Foundation