Annex 1

Written evidence submitted by Clean up the internet (OSB0238)

Verification: What Works?

What aspects of someone's identity should be verified?

We would suggest that the two most obvious aspects of a social media user's identity for which a verification option should be offered are name and nationality (or residency, in certain circumstances e.g. for asylum seekers). The reasons for focusing on these two aspects of a user's identity are:

- Both name and nationality are readily capable of being proved in a number of different ways as they are used in such a wide range of scenarios, indicated on a number of documents, and already held by a range of organisations.

- A number of digital, and scalable, approaches to verifying name and nationality have already been developed and are discussed below.

- Studies suggest that these identifiers have a significant bearing on subjective feelings of anonymity and therefore on disinhibition, so are associated with large reduction in negative online behaviours associated with the online disinhibition effect.

- The inclusion of a verification requirement for claims of nationality would have an immediate bite on foreign influence operations by making it much harder to create large numbers of fake accounts purporting to be of a different nationality to spread disinformation or distort political debate.

We would expect that the visibility of a verified user's true name and nationality would act to reduce other false claims which they could still in theory make about their identity, even if these additional claims are not being formally verified. This is firstly because of a reduction in the online disinhibition effect which would make them less inclined to behave badly online, and secondly simply because they'd be more likely to be "caught" by other users. A user is much less likely to dishonestly claim that they, say, live in a particular town, or have a medical qualification, or are a lifelong member of a particular political party, whilst also using their real name and nationality - and if they do, another user is much more likely to detect the falsehood.

We would also expect a reduction in the potential for false claims by unverified users to cause harm, simply because they would appear less credible. If someone is making very specific claims about their identity such as where they live, or what they do for a

living, or their experience or background, but are choosing not to verify some very basic aspects of their identity, other users would have reason to question the veracity of such claims.

However, different platforms may take different approaches, and this could include verifying additional or alternative aspects of their users' identities. Other aspects of a social media identity where verification could potentially be considered could include:

- Profile picture - It would be possible to verify that profile pictures were a true likeness of a user, with approaches similar to those currently used on dating sites or as part of the verification processes used by some banks. Users could be restricted from using a human image as their profile photo unless it was verified as a true likeness. This would serve to restrict the use of stock or stolen images for profile photos, as is currently common in fake accounts forming part of disinformation operations. Studies suggest that the use of a true likeness photo, particularly when combined with real name, would further contribute to a reduction in the online disinhibition effect.

- Location - There may be circumstances where it would be desirable to verify a user's location to a greater degree of granularity than nationality or country of residence would permit. The proposals in the SNP report mentioned below for a "verified Scottish" status on social media would fall into this category. This proposal arose from a concern that the social media debate regarding Scottish independence was subject to foreign interference. There may be other instances where more specific claims of location are pertinent to a disinformation or electoral interference operation. One existing method of verifying a precise location claim is to verify a postal address by posting a confirmation code to that address - as Facebook currently does as part of its verification process for political advertising.

- Claims of specific experience or expertise - There may be scenarios where it is judged important to verify certain specific claims of experience or expertise. For example, during a time when the fight against coronavirus is threatened by false information about the coronavirus and the safety and efficacy of vaccines, there may be a case for restricting a user's ability to include unsubstantiated claims of medical expertise in their profile.
  Verification of such claims could build on existing documentation of qualifications and national licensing and registration processes - for example in the UK the General Medical Council already provides an online facility enabling checks to verify that someone is registered with a licence to practise as a medical doctor. Other examples of the kind of claims where verification could be considered – because inauthentic claims of these kinds have been commonly linked to scams or disinformation operations - would be for membership of other regulated professions, such as lawyer or accountant, or of having served in the armed forces.

To be clear, we are not proposing that all social media users be subject to these additional forms of verification. We are merely suggesting that they are the kind of options which a regulator could expect a platform to evaluate when developing its approach to reducing the harm caused by anonymity and identity deception. Where such additional verification was applied, it need not mean every social media user was required to use a verified photo, or

proof of address/qualification – only those users' who chose to use a profile picture, or to include such claims in their profile.

Options for verification mechanisms and processes

i. Verification processes already in use by large tech platforms

A variety of verification models already exist and are currently deployed at scale bylarge tech companies:

- Facebook operates three different, and differently rigorous, verification processes- two levels of verification for "pages", a "blue tick" and a lower level "grey tick", and a separate and more robust verification process for political advertising on its platform, which includes verification of nationality and involves a confirmation passcode being sent by post.

- Twitter offers a restricted availability account verification process (the "blue tick").For the tiny proportion of users for whom it is available, the "blue tick" informs other users an account is authentic. Verified users are given additional filtering options, such as being able to see feeds containing only other verified users. Twitter announced the suspension of the program in 2017 stating a concern that verification was being confused with endorsement (a concern which would be eliminated if verification was universally available rather than a highly restricted perk). They announced in May 2021 that it is being relaunched with a more transparent set of qualification criteria and more details on how verification is carried out.
  Availability is restricted to accounts which fit notability criteria and looks set to remainextremely limited.

- Several digital challenger banks such as Monzo operate identity verification processes robust enough to comply with EU Money Laundering Regulations. In the case of Monzo, verificationprocesses have been applied successfully to over2 million users. These processes usually involve users submitting a photo of an identity document such as a passport or driving licence, accompanied by a video of themselves confirming some of the information in the documents.

- A number of mainstream dating sites such as Tinder and Bumble have developed profile photo verification to confirm that photos are an authentic likeness of the user. These processes tend to involve requiring the user to submit additional photos copying specificposes, which are then compared by human or automated means to their profile photos.

ii. Verification via a third party provider

Another option which could be explored would be for tech platforms to partner with verification systems provided by a third party verification system. This could be provided by a government or government-backed body, or by a commercial

entity.

- The UK government's "GOV.UK Verify" scheme allows registered users to access government services by proving their identity via tools provided by commercial partners. Once a Verify account has been created it allows continued access to a range of services, without further verification being required.

- A number of independent entities sell identity verification services to companies, which can be integrated into that company's own website or app. For example Yoti offers a variety of identity and age verification services to businesses, including embedded document scanning portals, and verification via a Yoti app which customers can install on their phone.

- An investigation into inauthentic activity in Scottish Twitter, commissioned by a SNP MEP,

proposed that a "Scotland Verified" system could be piloted by the Scottish government.

- Think tank Demos has proposed that the UK government could create a "British Identity Corporation (BIDC)", which would be an "independent body to verify identities online. The body would allow people more control over how they're identified as they travel across the internet, whilst also providing a method to help law enforcement tackle persistent online harms more effectively, in a way open to public scrutiny." Several countries already offer some form of state backed digital identity which would have the potential to underpin verification on social media, such as Estonia.

Safeguarding accessibility, diversity and inclusion

Any verification system would inevitably introduce additional steps which a user would need to take in order to achieve verification, and may require them to have access to specific means of proving their identity such as official documents. This would have the potential to raise some accessibility and inclusion issues.

Care would need to be taken to avoid over-reliance on a narrow range of national identity documents, such as passports and driving licenses, which sizeable minorities of the population do not possess. Options for "vouching" as a form of verification for those without documents would need to be developed. The specific needs of different minority groups would need to be considered, for example to ensure people with no fixed abode were not excluded through not having a permanent address, or that there was a straightforward way for trans people to transition their accounts to their new name. The best way to ensure diverse needs are adequately considered would be for the independent regulator, Ofcom, to consult and involve a diverse range of users in the development of a Code of Practice for verification processes, and to monitor and review the inclusivity of those processes on a regular basis. There is extensive research and best practice in this area which could be drawn on, and a range of charities serving communities with differing needs who could offer advice. The regulator should ensure that these perspectives are heard as part of its "user advocacy" function.

Social media companies should be required to demonstrate that all their terms and conditions, including those relating to verification, are compliant with relevant equalities legislation, and that they have been developed with due regard to diversity and inclusion.

As a final safeguard, our recommendation that verification be made optional would

significantly mitigate any remaining issues. Retaining the ability to access social media without verification -

albeit with certain limitations, but with limitations designed to restrict malign rather than legitimate use - no user would be at risk of losing access to a platform through not taking part in verification, whether by choice or because they had some difficulty with the process.

Safeguarding privacy

In general, social media companies have deservedly poor reputations for respecting their users' privacy. Their business models rely heavily on the collection and processing of users' behavioural data for the purposes of targeted advertising. The ways this data is processed and the uses to which it is put are often opaque, and users' consent often not meaningfully obtained.

It's therefore reasonable to have concerns that improving options for identity verification could lead to further privacy violations. It will be crucial to address such concerns, both to protect individual users' rights and to ensure that a critical mass of users are willing to undergo some form of verification.

A regulator should insist that data gathered for the purposes of identity verification should be used only for that purpose, and only retained as long as is necessary for purposes of verification. For example, should a document or image be uploaded as part of a verification process, it should be destroyed once verification status is confirmed.

Verification processes and outcomes should be communicated in plain english to the user. That should include making it clear to users that they have a genuine choice as to whether or not they verify - i.e. it should be made clear that it is possible to continue as a user of the platform (albeit with some limitations) without verification. Ideally users should also be offered more than one option for how they verify e.g. via a platform's in house system or via a service provided by a third party.

These principles are consistent with the principles of the GDPR. Strict enforcement of existing privacy rules as they relate to social media identity verification will therefore be a key pillar of
protecting users' rights and ensuring trust. Ofcom, as the online harms regulator, will need to work closely with the ICO, as the data and privacy regulator, to ensure proper enforcement of privacy rules with regard to verification systems, and to identify (and fill) any regulatory gaps.

2 December 2021