

Implementation of the Safer Social Networking Principles for the EU: Testing of 20 Social Networks in Europe February 2010



Leslie Haddon, London School of Economics and Political Sciences

Introduction

YouTube, owned by Google, is primarily a site for posting videos and viewing other people's videos. The reason why it is included in this test is that it has some SNS elements, mainly user profiles, but also the opportunity for users to communicate e.g. in terms of comments regarding videos posted. The minimum age of users is 13 years old.

At times YouTube provides multiple approaches addressing the same issue, elements that go beyond the minimum stated in the principles and features that exist in practice but are not in the self-declaration (as in the case of material for teachers). Some other elements of the principles have not been addressed in the self-declaration. Sometimes in testing the mechanisms could be better or the information/options could be made more visible (principles 2 and 6) but they are compliant with the claim in the self-declarations that they exist.

The principles

Principle 1: Raise awareness of safety education messages and acceptable use policies to users, parents, teachers and carers in a prominent, clear and age-appropriate manner

The self-declaration includes information on the terms of use to which users should abide, located in the Community Guidelines. Information about safety is covered at various points in the document, mainly in relation to their Safety Tips facility. The self-declaration notes that YouTube encourages and advises upon privacy issues, as well as providing tools (such as the ability to hide personal information) and a complaint mechanism. The declaration indicates the importance specifically of educating children about online safety and points to the various sections in which to find safety information, including additional advice on how to use tools in the Help Centre. It notes these are accessible, by virtue of being at the bottom of every page (there is no comment about it being 'prominent'). The declaration also notes that the tools are written in an easy to understand, user-friendly format and indicates the type of content and conduct that will not be tolerated, as well as the consequences of breaching terms of service. There is information aimed at empowering parents, but in the declaration there are no comments about advice specifically aimed at teachers

All the policy statements (terms of use, safety, privacy, code of conduct) are easy to find through links at the bottom of the page. Safety tips could also be found there and there was also a dedicated Safety Centre, with information for children and resources for parents. Although the declaration itself says little about addressing teachers, apart from mentioning that they too can watch the safety videos, there are, in practice, educational resources for them. The advice was always easy to understand for children of various ages and adults, and certainly sufficient in terms of raising a range of issues. The videos were useful for showing

both how to report problems and illustrating specific situations, and there were some links to other agencies. Clear examples of the types of content and conduct that will not be tolerated are easily accessible, as is an indication of the consequences of breaching the terms of conduct (e.g. account can be terminated). All of the items listed in the test can be found (i.e. hate speech, porn, violence, stranger danger, bullying, divulging personal information, posting sexually provocative photos and images of child abuse – the latter discussed in relation to child exploitation). On the other hand, the list of prohibited items goes beyond those tested to include gory content and content that incites violence as well as videos of reckless and dangerous conduct.

In sum, as regards the self-declaration, there are no comments about teachers, therefore it is partially compliant to the principles. However, in terms of testing the educational material is there, all the policy statements can be found at the bottom of the pages and are clear. The advice is reasonable, videos quite useful and what is not tolerated is in place. Hence this aspect is judged to be compliant.

Principle 2: Work towards ensuring that services are age-appropriate for the intended audience

The self-declaration notes that ‘age-restricted’ content is only viewable by those over 18, that there is a minimum registration age of 13, and that if children enter a birth-date revealing they are below that age they will be denied entry. The provider does not outline the steps taken to delete under-age users, but says that a cookie will be placed on users’ browsers to stop them trying to re-register with a different age. There is no information on any additional means to enforce compliance with minimum age requirements, nor how, in detail, to actually promote the uptake of parental controls, nor if professionally produced content is only shown at particular times of day (although the essence of YouTube is that it is mainly amateur produced content). One key mechanism used for limiting exposure to potentially inappropriate content is the ‘flag’ system, enabling the wider YouTube community to mark video content that is dubious for various reasons. This material can then be excluded from certain listings and areas. In addition, YouTube has implemented automated systems to help classify content.

It was clearly stated on the YouTube site that the minimum age is 13. As the first part of the registration process, one has to acquire Google or YouTube account, which means providing a date of birth, gender and post code. When applying to YouTube in the test the system did indeed reject the application whose birth date meant they were younger than 13 at this early stage in the registration process, before the verification phase described below, with system providing the message that this rejection was ‘based on the information submitted’. If the user then applies as a child over 13, the system moves to the next stage asking for an email address that is to be verified (i.e. YouTube then sends a link to that address that the user needs to click on). If the user uses an address the system recognizes as being used by a previous account, the system asks the user to open that existing account and so the ‘child’ does not get further. However, if the user (‘child’) has set up a different email (e.g. hotmail) address for verification, YouTube lets the user open the new account. So if a cookie is indeed placed on the PC, it does not stop this tactic of setting up a new address for verification. Hence, while the verification plus cookie tactic may stop the fainted hearted, a determined, knowledgeable under-age user can get round it by setting up new accounts. As regards the ‘age-restricted’ content this was, the over 18 adult user could access this material (e.g. when searching for ‘porn’) but the 15 year old user could not. The message said ‘you must be over 18 to view this group – hence the system works. When navigating ‘as a child’ and ‘as an adult’ there were no noticeable messages about this, but this may be an automated process that only operates when searching for videos. There is an extra way of checking age not mentioned in the self-

declaration – we are told on the YouTube website that if a video is flagged by someone, the images on it may give rise to doubts about that user’s age and the account may be closed. In sum, in the self-declaration the provider has addressed sufficient suggestions and volunteered mechanisms to be viewed as being compliant to the principle. In terms of testing, there are controls in place on what under-18s can view, as claimed in the self-declaration. As regard minimum age of access, we need to consider the following (a) the provider’s self-declaration says that a cookie is placed on the browser to stop re-registration of under-age users with a different age b) an under 13 year old re-registering on the same browser simply with a different age is not successful but (c) an under 13 year old re-registering on the same browser and PC with a different name and age (i.e. pretending to be older) can register. In other words, the mechanism could be more effective but since the test shows the provider adhered to what they claimed in the self-declaration, then they have to be judged compliant at this level.

Principle 3: Empower users through tools and technology

Google makes it clear that YouTube’s profile pages are not the same as standard SNS ones, since the aim is to encourage the sharing of user-generated content rather than encouraging social networking per se, and hence searches are searches for videos rather than for profiles (here called ‘channels’). The declaration notes that users can volunteer information about themselves. Although the default is not ‘private’ it is the user name, not the actual name, which will appear. The declaration does not say that users have control over who can access their profile, but it does say that they can choose to only share a video with friends/family, and that they can block comments posted by other, as well as pre-moderate and post-moderate them, which implies the ability to delete unwanted ones. Users can report inappropriate contact such as violation of privacy, harassment and cyberbullying. The introduction to the self-declaration says that parents are given tools to protect children, but does not say how it educates parents in relation to principle 3.

As specified in the self-declaration, it is not possible to search for user profiles, and that means adult profiles let alone child ones. Users have controls to block others (or rather ‘specific others’, rather defining which groups - e.g. by age - can make contact), and they can remove any postings from others on their profile. Some parts of the profile appear to be visible to all viewers by default (e.g. user name and statistics about usage, such as when they joined and number of videos watched). In addition to advising parents to talk to the child as a first step, parents are also given some tools and clear information about how to use them (e.g. ‘hide objectionable words’, ‘hide comments’, ‘restrict search options’). Parents do not have to verify the child’s (initially very limited) profile before it can be used, although they can monitor what the child has viewed (although YouTube acknowledges that the child can get round this). Ultimately the provider can shut down the child’s access, although this involves the parent contacting YouTube.

Google notes that the profile is limited, reducing stranger danger. Of course, if children then volunteer more about themselves they could be identified, but they are given advice about this.

In sum, given that various measures that may be included according to principle 3 are in fact noted in the self-declaration, this must be judged to be compliant with the principle. In terms of testing, the system does what is claimed in the self-declaration and is therefore compliant.

Principle 4: Provide easy-to-use mechanisms to report conduct or content that violates the terms of service

The ‘flagging’ system noted above provides a method for reporting inappropriate content, while the Help and Safety tool provides a way to report contact and conduct, and these

channels are available at all times. The self-declaration does not comment upon whether these reporting procedures are reasonably understandable, age-appropriate or whether reports are acknowledged, but it does say that flagged videos are reviewed promptly (within the hour), although there is no comment specifically about the speed of reacting to other types of report (although there is a note that YouTube reports child exploitation to the police). In the self-declaration there is also no comment about the information to make an effective report (although this is fairly clear when you actually try to do this), nor an indication of how, in terms other than speed, reports are typically handled.

The test shows that users can report inappropriate contact and conduct (e.g. with the Health and Safety tool), with slightly different mechanisms operating for different types of reporting, e.g. content of videos vs. hate speech. These reporting tools are easy to find and understand – the user is offered various options. For example, in the case of unwanted contacts the system asks the user name of the person and how they are harassing you – the user does not write a message, but picks from choices, (which means that the actual wording of the test could not be used). While there is no message notifying the user with the words ‘the report has been received’, in the test the automated system made one first check of the claim of harassment and immediately displayed results that no-one could be harassing in the test because there were no messages from outsiders to the new account.

In sum, the self-declaration does not comment on some of the provisions of the principles (reporting procedures are reasonably understandable, age-appropriate, whether reports are acknowledged, the speed of reacting to reports other than those relating to content) and so should be judged partially compliant. In the test the reporting mechanism is clear enough and the automated system detected no harassment and therefore has to be judged compliant with the self-declaration claim that a reporting mechanism is in place.

Principle 5: Respond to notifications of Illegal content or conduct

As indicated above, the self-declaration provides information about the content reviewing process, noting that offending material will be ‘dealt with appropriately’. There is additional information about technologies to prevent the re-upload of removed material – indicating that some such material might be removed. The self-declaration says that there are arrangements to report criminal content and conduct to the relevant law enforcement agencies, and there are links with a variety of other relevant organisations. If these measures are in place then the provider has to be judged compliant.

Principle 6: Enable and encourage users to employ a safe approach to personal information and privacy

A range of privacy setting options is implied by the declared ability to hide information or share videos selectively, and the self-declaration says that users are provided with tips to make informed decisions about the information they post, tips that are accessible at all times (the declaration does not comment on the availability of privacy options, although in practice they are always there). While there are no comments on the implications of automatically uploaded registration information for profiles, notification to users that this information is used in profiles and the ability of users to edit this information, to put this into perspective, the initial profiles on YouTube are very limited (user name, when joined, when last signed in, number of videos viewed, country). Later one can choose to add more information. There is no comment on the ability to view privacy settings ‘at all times’, but self-declaration says they are ‘on the site’ and so this if they are always on the site it is implied that they are always available.

In the test, a user can change the privacy setting in Google Accounts at any time e.g. so that only friends can send messages and share videos. The privacy policy online indicates the type of information that YouTube collects about users and that email addresses may be passed on to third parties. On registering for YouTube, users provide information about age, their email and gender, but, unless they have read the privacy policy, the user is not warned at this point about what information might be used in the initial profile that is generated. The location of the profile details could be clearer – the option appears only when moving the cursor over the user name, so if you just look for it on the screen it is not visible.

In sum, the provider has to be judged compliant with the principles, given that the initial profiles generated are very limited. In terms of testing, the information provided could be clearer, but in general the site does what is claimed in the declaration and so is compliant.

Principle 7: Assess the means for reviewing illegal or prohibited content/conduct

The flagging system noted earlier provides the human form of moderation (to flag potentially illegal or prohibited content), which is complimented by automated systems to classify content. The declaration notes that it responds to this flagging of content, but says nothing in detail about its response to conduct reports. There is no comment on the steps taken to vet any human moderators (although nothing has been said, either, about moderators contacting children in the first place, so it is unclear whether this is an issue).

In the self declaration that there are multiple systems in place, and so this must be judged compliant with the principle

Summary of results and conclusion

At times YouTube provides multiple approaches addressing the same issue, elements that go beyond the minimum stated in the principles and features that exist in practice but are not in the self-declaration (as in the case of material for teachers). Some other elements of the principles have not been addressed in the self-declaration. Sometimes in testing the mechanisms could be better or the information/options could be made more visible (principles 2 and 6) but they are compliant with the claim in the self-declarations that they exist.

Assessment of the Principles vs. the Self-declaration

Principle	Compliant	Partially Compliant	Not Compliant	Not Applicable	Comments/ Clarification
1		Yes			There are no comments in the self-declaration about teachers Does not comment on some of the provisions of the principles
2	Yes				
3	Yes				
4		Yes			
5	Yes				
6	Yes				
7	Yes				

Assessment of the Self-declaration vs. the measures implemented on the SNS

Principle	Compliant	Partially Compliant	Not Compliant	Not Applicable	Comments/ Clarification
1	Yes				Since the test shows the provider adhered to what they claimed in the self-declaration, then they have to be judged compliant. But the information provided could be clearer
2	Yes				
3	Yes				
4	Yes				
5	Not Tested				
6	Yes				
7	Not Tested				

The copyright of this report belongs to the European Commission. Opinions expressed in the report are those of the authors and do not necessarily reflect the views of the EC.