Online Hate
Prevention Institute

Online Hate Prevention Institute
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au

Online Safety Submission

Dr Andre Oboler
7 March 2014

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Online Hate
Prevention Institute

The Director, Cyber Safety Policy and Programs

Department of Communications

GPO Box 2154

CANBERRA ACT 2601

& via e-mail to: onlinesafety@communications.gov.au

7 March 2014

The Online Hate Prevention Institute (OHPI) is pleased to make this submission offering comments on the Government's Discussion Paper on Enhancing Online Safety for Children (January 2014). OHPI is Australia's only Government registered Harm Prevention Charity dedicated to the online sphere. This consultation is therefore relates to matters that directly form the core of our purpose as an organisation. We hope our insight, as presented in this submission, will be of assistance to the Australian Government. We would be happy to provide further information or to meet with the Department of Communications about this submission, or about our work more generally, should this be of assistance.

Formed in January 2012, the OHPI's vision is to change online culture so hate in all its forms becomes as socially unacceptable online as it is "in real life". This hate ranges from cyberbullying of individuals to attacks on minorities and other segments of society. We have addressed hate based on race, religion, sexuality, gender, military service and targeting police, fire fighters and politicians. At the individual level we have provided assistances in cases of cyberbully, cyberstalking and trolling. Our mission is to be a world leader in combating online hate and a critical partner who works with key stakeholders to improve the prevention, mitigation and responses to such forms of online hate. The Government, its departments and its agencies are of course critical partners we wish to work in the pursuit of this mission. We thank you for this public consultation and the opportunity to share our thoughts on this important topic.

The following provides a general overview of our views in relation to the discussion paper, specific responses to the questions asked follow.

Children's E-Safety Commissioner

We support the establishment of a Children's E-Safety Commissioner

- The Commissioner's remit should involve three distinct functions: the first relates to the removal of content from major social media sites in exceptional circumstances; the second related to setting and enforcing minimum standards the social media platforms themselves should apply to address online safety concerns; and the third relates to real world consequences for people in Australia who engage in online activity that causes harm to others.

- We believe the E-Safety Commissioner's remit should cover all content on major social media platforms that can harm children and is sent to a child or made generally available to all users of a system, or a class of users of a system, where children are not excluded from viewing it.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

- We believe any person in Australia should be able to make a complaint about the failure of a major social media company to take appropriate action in reasonable time in response to a users report regarding content that can put children's safety at risk. We believe anyone, whether or not they are in Australia, should be able to make a complaint to the Commissioner about activity by a person in Australia which is part of a course of conduct which puts children anywhere in the world at risk and which is serious or persistent enough to warrant an offline response.

- We also believe international cooperation is needed and complaints about Australians should be able to be received by the Commissioner from foreign law enforcement agencies or relevant foreign Commissioners or foreign government agencies. The Commissioner should have a mandate for cooperation in this area under the Commonwealth's External Affairs power.

A complaints system backed by legislation

We support the creation of an effective complaints system, backed by legislation, to get harmful material down fast from large social media sites

- While complaints should be addressed to the company concerned, and the cost of administering such a complaints system should be worn by the company concerned, we believe the government has an important monitoring role in ensuring the system is working effectively and efficiently.

- We believe an E-Safety Commissioner should have the power to set minimum standards for these major social media companies in relation to the time it takes for reports of content to be reviewed, and regarding the quality of those reviews.

- We believe it should be possibly, either by the Commissioner or through another agency, for a civil penalty to be applied to companies that fall below the set standard.

- We believe the government needs access to data highlighting how many complaints were lodged with each major Social Media company by people in Australia, and how many of these complaints were resolved by user action, how many by action from the platform, and how many were investigated and dismissed, and how many were dismissed and then later re-evaluated resulting in content being removed. Statistics on penalties imposed on users, such as the number of account suspensions for different period of time should also be made available. This will help inform future policy.

Law Reform on Cyberbullying

We believe there is a need for new national legislation to address cyber-bullying, and a new mid-level criminal offence and a new low level civil penalty

- National legislation is needed to address cyberbullying which may have elements of stalking, harassment, threats and invasion of privacy, yet still falls through the gaps of existing provisions. There is also a strong reluctance by law enforcement to use existing laws as the penalty appears too high for what they (often incorrectly) regard as a minor matter

- National legislation is needed to ensure state borders do not pose an impediment in cases where the abusers and the victim are both in Australia but are in different states

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

- The legislation should contain a list of prohibited activities similar in scope to the provisions regarding acceptable use which are already found in the terms of service of major social media companies

- The legislation should compel the release by major social media companies, in reasonable time, of the identification information of a particular user where the E-Safety Commissioner, Police, the Australian Human Rights Commission, the FairWork Commission or other relevant Commissions or agencies believe on reasonable grounds that the user has engaged in prohibited activities and a real world response is needed

Definition of 'Large Social Media' company

We believe a 'large social media' site should mean one on a list approved by the Communications Minister, published in the government gazette, and reviewed at least once a year.

- We believe the Minister should in turn base his decision primarily on:
  (1) the site being a social media site, that is one primarily driven by content posted by users and allowing them significant interaction with each other
  (2) the popularity of the site in Australia as seen in the rankings provided by Alexa.com or similar providers of statistics.
  These two conditions would exclude mainstream media sites, even where they have significant user comments. They would also exclude Web 2.0 sites that support user publishing, but are not social media sites.
  By focusing not on the number of users but on the total time children spend on the platform (which can be approximated by the level of traffic to the platform) Government can aim to maximize the number of minutes and hours children spend in a regulated (safer) environment compared to the time spend in an unregulated (potentially less safe) environment. This will have the result of increasing the overall level of safety and reducing the overall risk to children.

- Based on Alexa we would recommend the following initial list based on the above: Facebook, YouTube, LinkedIn, and Twitter. Excluded under rule 1 above are: ebay.com.au, Wikipedia, Gum Tree, Wordpress.com, Pinterest, Tumlr, ebay.com, and Instagram.
  We believe the Minister should also consider the risk a particular site is posing to Australian children, and on this basis a site like ask.fm which has been linked to serious cyberbullying and youth suicide may also be listed in the interest of public safety.

Reasonable Response

We believe reasonable time for a company to respond should vary depending on the request:

- In the case of an identity request by an authorised person, such as the E-Safety Commissioner, this should be provided within 24 hours. An automatic fine should apply if this deadline is missed with the fine increasing daily and the amount of daily increase growing weekly.

- In the case of an emergency takedown request by the Minister for Communications or an Acting Minister for Communications, the content should be removed within an hour of notification. We believe such requests should be exceedingly rare and would only be used in the case of a significant and imminent threat to public safety. This needs to be covered, but is not a matter for the E-Safety Commissioner.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

- In the case of a priority takedown request by the Commissioner, within 24 hours. This should also be a rare occurrence used when there is a real and significant threat to a child's life.

- Generally it should be up to the social media company to remove content it has deemed would be prohibited. Once a person within Australia reports content to a company, the company should be required to make an initial assessment within 3 days, and if the content is deemed borderline, they should then be able to take a further 7 days before reaching a decision. A civil penalty should apply where these deadlines are missed, and the Commissioner or their agents should regularly test the process for compliance on efficiency.

- In addition to the above system we believe a triage system should be in effect, allowing an initial assessment on whether the content poses a significant risk of imminent harm to a child. This should be assessed as quickly as possible, preferably within an hour of a relevant report being lodged by the user. To reduce the cost to social media companies, if they provide a mechanism for users to indicate that the report relates to an imminent risk of harm, then only these reports need to be assessed for triage.

- Where the social media company determines content is acceptable, a fine should apply if the decision was manifestly unreasonable. The Commissioner should issue the fine, and the company may elect to pay it at a discount and remove the content within 14 days, or have the Federal court determine if the company's decision was clearly wrong.

- The E-Safety Commissioner should also be able to investigate content and determine that it is unlawful. On notification of such a determination, including the commissioner's reasoning in general terms, the company should remove the content within 14 days or request administrative review. Such determinations generally describing what was deemed unlawful should be made publically available. Past determinations should be taken into account by companies when deciding what content is unlawful, and should be considered by the courts in determining when the company was manifestly unreasonable in regard to later incidents. It would by definition be manifestly unreasonable not to remove content that is an exact duplicate of content previously subject to a determination. Companies can automate the detection of such matches and should be reviewed to do so, even if this is just to notify their reviewing staff.

Additional Support

We believe charities like the Online Hate Prevention Institute provide a significant public service in addressing issues related to the online safety for children and the prevention of harm from major Social Media platforms to the Australian public more generally. This is a growing problem and while education on these issues receives significant government funding, there is currently no funding for the sort of work OHPI does at the coal face: monitoring and addressing online problems and pushing for systemic improvement to reduce the risk of harm.

Many charities operating in the public interest receive significant government funding, but this area being so new is yet to receive such support. We would welcome the Government's consideration on whether such support could form part of the online safety response.

Yours sincerely,

Dr Andre Oboler

CEO, Online Hate Prevention Institute

1. Establishment of a Children's e-Safety Commissioner

The Online Hate Prevention Institute supports the establishment of a Children's E-Safety Commissioner. We believe the Commissioner should be a single point of contact for all stakeholders on all online safety issues which can affect children.

We believe the Commissioner needs to interact not only with industry, Australian children and those charged with their welfare, as outlined in the paper, but also with charitable institutes active in the space.

We believe the Commissioner should also serve as a contact point for the wider Australian public when they see content that could be harmful to children. They should additionally serve as a contact point for foreign government agencies in the event that a perpetrator of harm occurring in that country is based in Australia, or in the event that there is a major social media company based in Australia or holding data in Australia. Reciprocity of such arrangements would naturally help improve the safety for Australian children.

The Commissioner should be able to refer a matter to another relevant agency for mediation or for investigation which may result in a briefing for the Commissioner. For example cyber-bullying involving cyber-racism may be referred to the Australian Human Rights Commission for mediation, and if that fails, for investigation resulting in a report the E-Safety Commissioner which may inform the Commissioner regarding further action.

We believe the Commissioner should have three distinct roles:

- Directing the **removal** of content in *exceptional* circumstance
- Setting, monitoring and enforcing minimum **standards** to ensure large scale social media platforms systematically remove harmful content themselves in the vast majority of cases
- Providing a **real world response** to ensure there is a meaningful and effective deterrent to harmful behaviour, this may involve warnings, civil penalties or referral to police for criminal investigation

Q1 What existing programmes and powers should the Commissioner take responsibility for?

We believe there is a gap where online safety concerns can fall between the responsibilities of the Australian Human Rights Commission (AHRC), The Australian Communication and Media Authority (ACMA), the Department of Communications, and Police. As such be believe the three tasks outlined above (removal, standards and real world response) would be new powers and would not involve a transfer from any other agency. We believe the AHRC should retain the ability to request the removal of content through mediation and the ACMA should train the ability to remove content based on classification grounds. The commissioner's removal power would be on a new basis of public safety.

OHPI has a concern about major social media platforms seeking to transfer the cost of enforcement from the company to the State. We believe this is occurring in some other countries including New Zealand where some platform providers will now remove content based on the decision of a government recognised organisation. We believe this is inappropriate as the companies are generating significant profits, including from the Australian market place, and should also bear the bulk of the cost burden when it comes to safety concerns their business creates. On this basis we believe the Commissioner's take down power should be used sparingly, and instead there should be a system of minimum standards and civil penalties for failing to meet the minimum expectations on safety. The Commissioner should play a vital role in administering this system. We believe ASIC may have a role to play in such a system in addition to the Commissioner.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

In the table below we provide comments on the existing programs. In general we believe education should not be the focus of the Commissioner's work as it may detract from other important functions which are not currently being met. We believe existing reporting by other agencies is appropriate, but they should be able to refer things to the Commissioner as a secondary stage of response or for urgent attention. A unified reporting system with different kinds of complaints routed to different agencies would be beneficial, but classification should remain with the ACMA, crime with the police, education with various agencies and partnerships, and research with the Communications Department.

Programmes and resources	Agency	Comments
The ***Cybersafety Help Button***	Department of Communications	This should remain with the Department. The Commissioner should be informed by the department if a platform is not taking appropriate action in appropriate time or if the matter is particularly serious and may need immediate attention.
The ***Easy Guide to Socialising Online***	Department of Communications	The department can continue to provide this education role. The content is not specific to Children.
The ***Australian Government Cybersafety Help Facebook***	Department of Communications	The most popular age of people on this page is 25-44. This is not specific to Children. The page is small. This should stay with the department. The Commissioner and their work should be promoted on this page.
The ***Australian Children's Cybersafey and E-security Project***	Department of Communications	The administration of a research project is more suited to the Department than to the Commissioner.
The ***Online Content Scheme***	Australian Communications and Media Authority	Classification and should remain with the ACMA. Classification may have a role to play in limited circumstances, but it is based on an old media model with high production and distribution costs. It will not effectively address harmful online content which can morph and adapt into thousands of variants in seconds, each of which would need to be separately considered for classification. Responding to cyberbullying content, promotion of self harm or substance abuse, or racism or homophobia aimed at children through a classification model would be highly ineffective. The classification model can however play a role when particular items of problematic content become culturally stable and classification of that content as RC may then be helpful.
The ACMA's ***Cybersmart***	Australian Communications	This is a very successful program. We would recommend it be left in place at the ACMA along with other education roles. This

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

programme	and Media Authority	will allow the Commissioner to better focus on enforcement, standard and response.
Tagged, ***Connect.ed***, ***Zippep's Astro Circus, The Cloud: Dream On***.	Australian Communications and Media Authority	These programs and Cybersmart should be with the same agency. As stated aboive we believe they should remain with the ACMA. They could also move to the Communications Department.
The ACMA's ***research programme***	Australian Communications and Media Authority	We believe this should remain with the ACMA or move to the Department. We do not believe the Commissioner needs to manage research, but we do believe they should contribute to our understand of what it is we need to know. The Commissioner could achieve this through an ex-officio position on a joint committee into internet research.
The Australian Federal Police's ***ThinkUKnow*** programme	Australian Federal Police	The reporting element should remain with law enforcement, but an efficient system of referrals should be created. The AFP should focus more on issues such as serious cyberbullying, impersonation, intimidation, stalking, etc. The Commissioner should refer such matters to the AFP if the Commissioner believes a particular instance requires a criminal rather than civil penalty response. This may related to the severity of an incident, or relate to a perpetrator who has not been deterred by civil penalties.
The Line campaign	Department of Social Services	This appears to be with the right department.
The ***Safe Schools Hub***	Department of Education	This appears to be with the right department.
Bullying No Way!	Commonwealth, state and territory education authorities	This is a collaborative effort and should remain as such
The ***BackMeUp*** campaign	Australian Human Rights Commission	The AHRC has an important role to play in combating cyberbullying and online hate, particularly when the harmful content is based on racism or other forms of unlawful vilification. Educational campaigns by the AHRC will support this work. The AHRC should also have the ability to refer matters to the E-Safety Commissioner when reconciliation fails or a party refuses to engage with the AHRC.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Q2 Considering the intended leadership role and functions of the Commissioner, which option would best serve to establish the Commissioner?

The Online Hate Prevention Institute believes the E-Safety Commissioner would primarily be scrutinising large social media companies. As such it is independence from these companies that is paramount and not independence from Government. In this instance a closer relationship with Government would provide the greatest level of relevant independence.

We strongly recommend Option 2 with administrative support provided by the Department of Communications. This will give the Commissioner the weight needed to ensure compliance by large social media companies. It will also ensure the focus can be on setting minimum standards that major social media companies must comply with, and then monitoring compliance.

We believe that major social media companies are making significant profits from the Australia market and evaluating what content is harmful and removing it should be part of their operating costs. The role of government in this sphere should be regulating to ensure the companies are doing this effectively and efficiently. It should be no different to the approach taken with environmental pollution, government's job is not to come and clean up a company's pollution, it is to set standards for permissible levels of emissions (or the permissible time harmful content can remain online in this case) and to impose penalties if companies seek to cut corners and put the public at risk.

We believe the New Zealand model shifts the role of compliance from the company making the profits, and causing the problem, to the charitable sector. This indirectly transfers a large part of the cost to the public purse (directly or through tax deductions). It also puts the monitoring in a resource poor situation where only the tip of the iceberg can be addressed. This leaves the problem substantially in place.

Q3 Are these definitions of 'social networking sites' suitable for defining 'social media sites' for the purposes of this scheme?

We belief the definition used should be tied to the problem being addressed, rather than adopting a generic definition of what a social media or social networking site is for some other purpose.

We would define social media as "*a site primarily driven by content posted by users and allowing them significant interaction with each other*". This definition is based on what we see as the two defining risks social media poses: The first is the ability to upload content which the network can then spread; the second is the ability to interact not just with an individual but also with their network. The first risk allows a broad based attack on an individual or group, while the second risk allows a targeted, highly personal, and highly damaging attack on a person support network.

Q4 Should the proposed scheme apply to online games with chat functions?

We do not believe this is appropriate. A user can block another user in most games. They can also walk away from the game and stop playing. Alternatively they can create a new account and often transfer their in game assets to it. In short they can escape the bully in a way that is not possible in a site like Facebook or LinkedIn where one's digital identity is tied to a real world identity and walking away from the platform can have a serious real world impact including social exclusion and an impact on job prospects (for older children).

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Q5 What is the best criterion for defining a 'large social media site', and what available sources of data or information might be readily available to make this assessment?

We believe a 'large social media' site should mean one on a list approved by the Communications Minister, published in the government gazette, and periodically reviewed. We support the proposal of using objective criteria, but believe internet traffic may be a more useful measure than user accounts, and is also easier to obtain. An account of user accounts will over expose sites that have many dormant accounts, as well as sites that see very occasionally use. We believe it is more important to address safety in those places where people spend most of their time. We do agree that the metrics used should be based on Australia.

We would recommend using a site such as Alexa to rank the traffic to the various sites and then focusing on the social media sites at the top of the list.

Based on Alexa and our definition we would recommend the following initial list: Facebook, YouTube, LinkedIn, and Twitter. Excluded on the basis they are not social media sites in the required way are ebay, Wikipedia, Gum Tree, Wordpress, Pinterest, Tumlr, and Instagram. These sites all allow content to be uploaded, but only the first list directly allows the high degree of interaction our definition requires.

We believe the Minister should also consider the risk a particular site is posing to the Australian public, on this basis a site like ask.fm which has been linked to serious cyberbullying and youth suicide may be listed in the interest of public safety even if its popularity is not high enough to see it included on that basis. This could be converted into a metric if "actual harm" was factored in, in other words, a count of incidents leading to serious consequences which are linked to that platform.

Q6 Is the coverage of social media sites proposed by the Government appropriate and workable?

While a voluntary form of compliance would be welcome, we believe it should be different to that applied to large social media companies. Otherwise there will be a push to reduce the requirements on the large companies in order to encourage more smaller companies to join. While this will increase the number of companies covered, it will reduce the amount of online time that is effectively protected. It will reduce the protection for Australian children.

Instead we would suggest other companies could sign up to a voluntary scheme:

- committing to respond to user reports on the same time scale as required for large companies
- agreeing to prohibit the same categories of content large companies are required to prohibit
- agreeing to follow the rulings of the commissioner in relation to specific items of content (which may be shared across both large and small social media sites)
- agreeing to follow the guidelines and past ruling of the Commissioner in determining whether certain types of content or acceptable or not

The Commissioner could provide certification to companies that participate in a voluntary scheme by conducting an audit. Such a review could have two parts, the first examining how long the company takes to respond to user complaints, and the second inspecting the last 100 complaints the company reviewed and assessing them for accuracy. This would ensure harmful content was removed, and on the other hand would ensure freedom of speech was protected (and that companies weren't over blocking content).

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Q7 Should the scheme allow children who are unsupported by adults to be active participants (either as complainants or notice recipients)? Having regard to the vulnerability of children, what procedural safeguards should be in place?

We believe "specific target" needs to be more clearly defined. It would clearly include a child who is being targets with material that includes their picture or name. It is unclear if it would include a child who was receiving messages against their race, for example, if an Indigenous child was being send Aboriginal Memes as a form of harassment or if a Jewish child was being sent antisemitic material.

If both of these situations are specific targeting, there is still the situation where there are only one or two people of a specific ethnicity in a sports team and racist material against that ethnicity is for example posted to the teams Facebook page.

We believe the specific targeting may overly restrict the eligibility of complainants. At the same time, we don't want to see children being used as complainants on general issues of racism (for example) simply because the E-Safety Commissioner is a more effective response to online hate and only a complaint on behalf of child will activate the appropriate mechanism.

We believe the solution is to allow anyone to make a complaint about content that may harm a child. This means platforms, or parts of platforms, which have in place age based verification would be exempt. In all other cases if the content is likely to cause harm to a child, and can be viewed by a child, then it should be removed.

We are also concerned that the proposal is focused only on the victim, we believe the proposal should also apply if the perpetrator is in Australia (even if the victim is not). This may protect other Australian children from harm and provides a basis for international cooperation which could in turn see enforcement locally in other countries when a perpetrator there attacks a child in Australia.

We believe a child should be able to make a complaint. Any barrier may leave a child at risk. We believe a child receiving a notice should be in the presence of a parent, guardian, teacher, or another appropriately trained and responsible adult who is able to support them and look after their interest.

Q8 What type of information would it be necessary to collect from complainants in order to assess their eligibility under the proposed scheme (including age verification), and also to adequately process complaints with minimal investigation required?

We believe complainants should be able to lodge a complaint with the Commissioner immediately after lodging it with the social media site. The complaint could then remain in the system without action until a fixed period had passed and the complainant could then be sent an automated e-mail asking them to confirm if the complaint had been resolved or not. Only unresolved complaints would require investigation by the Commissioner.

We believe the address of problematic content is required, ideally in the form of a link. The date it was reported to the social media company would be required. The complainants name, age, e-mail address and phone number would be required. Both the e-mail and the phone should be verified, for example by sending a link that needs to be clicked in an e-mail to the nominated address, and by sending an SMS with a code to the phone. This would help ensure the identity of the complainant and prevent fraudulent complaints being lodged in the name of another person.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Q9 How would an eligible complainant demonstrate that the complainant has reported the content to the participating social media site?

The social media companies should provide a receipt for reports and the complainant should enter the receipt number as part of their complaint lodgement. The commissioner should be able to enter a receipt number and view whether the complaint is (a) awaiting review, (b) reviewed and removed, (c) reviewed and complaint dismissed.

An alternative approach is for complainants to lodge their report with a third party system and for the social media companies to have access to these reports. This would shift the burden of checking the list from the user to the social media company.

Q10 What should the timeframe be for social media sites to respond to reports from complainants? Is 48 hours a reasonable timeframe, or is it too short or too long?

We believe there should be two stages to the response. The first stage should occur within 3 days and remove content that is clearly harmful and prohibited by the scheme. The second stage may take up to 7 days more if the platform believes greater consideration of the particular case is needed. The second stage should be available where a clear position on the same or very similar content already exists and has been published by the Commissioner.

There is a risk that too short a period leads to poor quality decision making, which in turn leads to many valid complaints being rejected.

Q11 What level of discretion should the Children's e-Safety Commissioner have in how he/she deals with complaints?

Complaints which in the Commissioner's view are frivolous, vexatious or not made in good faith should be rejected, but this should be open to administrative review.

Q12 What is an appropriate timeframe for a response from the social media site to the initial referral of the complaint?

We believe this approach is problematic. There should be no initial notification by Government of the intention to investigate. If content has not been reviewed quickly enough, or has been reviewed and not removed, there should be a, perhaps very small, penalty for the company upon notification of a breach. This should grow slowly and then at an increased rate the longer the material remains online.

As the model is worded, the cost to the social media platform for not removing harmful content is no more than notification. There is no incentive for the Social Media company to remove the content before receiving a notification. This in turn means there is no incentive to properly police their system to avoid notification. This would be to the companies advantage as it would allow increased revenues (advertising in relation to the content, advertising to those posting the content) and reduced costs (and less effective monitoring system). This shifts the burden of assessment from the company, who is making a profit off the harmful content, to the public purse.

We believe a better approach would be for the social media company to police their own platform at their own expense. The Commissioner should set minimum standards and should see these standards are met. The approach might be similar to regulators like the Environmental Protection Authority. Under such a model we believe there should be no notice of an intention to investigate, as providing this would allow the companies to only deal with the content the Commissioner notifies

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

them about. Instead we believe notice should only be provided after investigation, and this should come with a fine which is automatically reduced if it is paid without going to court. If the company believes the content is not prohibited, they take the matter to court and it would be assessed by the court, potentially resulting in the full fine having to be paid or in the dine being cancelled.

Q13 Are the nominated factors, the appropriate factors to be taken into account when determining whether the statutory test has been met? Should other factors be considered in this test?

We believe the proposed statutory test sets an impractically high barrier. It captures not an effort 'likely to cause harm or distress' as intended, but rather an effort, 'reasonably likely to result in death'. The factors assessed need to be adjust back to covering harm or serious distress to a child.

In particularly we believe the "risk of triggering suicide or life-threatening mental health issues for the child" is too strict. At a minimum we believe: *the risk of triggering self harm, substance abuse, or other serious mental health issues such as depression, anxiety and eating disorders should also be included*.

Q14 Is the test of 'material targeted at and likely to cause harm to an Australian child' appropriate?

We also believe the test for assessment must be objective. The risk of triggering harm to the child may be a factor in a prosecution of a poster, but it would be very difficult (and expensive) to factor this in when deciding if the content should be removed. We believe the harm done by removing negative content about a child in the case where the child is particular resilient and not negatively impacted by the content is minimal. This is very different to a situation where there may be penalties on another person for posting the content.

We believe the test should favour removal if: *a reasonable child would, if in the place of the targeted victim, be at an increased risk of harm*. Safety problems are not fixed when the fault is likely to cause harm, they are fixed then the problem increases the risk of harm. The same should apply here. Only fixing it when the problem is likely to cause harm means allowing a significant increase in risk, which statistically means a certain number of child suicides are deemed acceptable. If there is only a 10% change content will lead to harm, that should be enough to see the content removed, otherwise, statistically, one in ten children will suffer that harm.

We also believe the requirement that the content 'relate directly' to the child is overly restrictive. We believe it should relate to the child, but the idea of directness might imply the child must be named or depicted in the content, or that the child should have been personally selected to receive the content. We believe harm such as cyber-racism or online homophobia should also be removed under these provisions. This type of content might be uploaded to communities where children who are at risk of harm from this content participate. The content may be generic in nature, but can still result in serious harm to the child. As platforms already prohibit hate speech, and such content is likely to already be unlawful, we believe the commissioner should be able to take action on it when it occurs in an environment open to children.

Q15 What is an appropriate timeframe for material to be removed?

For an individual three days after receipt of notification. For a company 24 hours. We believe the notification should include a small fine (e.g. $10), with a daily increase if the content is not removed within the designated period. This should apply to both users and social media platform providers, though possibly with differing fine levels. The fine is necessary as a deterrent, but needs to be small to be effective. Larger fine should apply to repeat offenders.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Q16 What would be the best way of encouraging regulatory compliance by participating social media sites that lack an Australian presence?

The G20 when it met in Sydney highlighted the need to address global loopholes used by internet companies. We believe it is fundamentally wrong that an Australian company can pay a social media company for advertising which is then displayed to an Australian audience, yet no Australian tax is paid on this transaction. We believe social media platforms that are gathering data from Australians or displaying advertising to Australians are doing business in Australia and should be paying tax on these transactions in Australia. This should involve processing in Australia through an Australian subsidiary. This would ensure such companies do have a presence in Australia.

In the absence of a presence in Australia, we favour the approach used by the High Court of Paris (France) in 2000 in the LICRA V Yahoo! case. In this case the court issued a fine which increased daily so long as non compliance continued. Yahoo! went to a use court to have the decision declared void for violating First Amendment rights, and the US court held that the French Court has the right to impose the fine and Yahoo! could not rely on the First Amendment unless the plaintiff brought an action for recovery in the US. With the fine continuing to grow the plaintiff was in no hurry to bring such an action. A similar approach could be applied here, and eventually enforcement action could be taken in another jurisdiction, for example in Europe, if needed. Companies that are unwilling to comply with decisions outside of their home country will end up seriously restricted in the international market place.

Another approach would be to apply a tax of 20% payable by Australian Companies who advertise on platforms who refuse to pay the fines levied against them. This would result in a market shift away from advertising on such platforms and would encourage compliance, or at least acceptance of the penalties for non-compliance.

The Government should also, as a administrative action, prohibit government advertising on non-complaint social media platforms. Ideally the States should agree to follow any Commonwealth boycott of a platform over non-compliance. This allow will be worth a significant sum of money and may help ensure compliance.

Q17 Should the proposed scheme offer safe harbour provisions to social media sites which have a complying scheme, and if so, what should they be?

Social media companies who have an effective reporting system should not be liable for content that has not been reported to them. A safe harbour provision would be a reasonable incentive, but the requirements of an “effective reporting scheme” need to be outlined carefully.

The Online Hate Prevention Institute has repeated highlighted to one major social media company that their reporting system has significant flaws. One of these flaws puts those trying to report content at risk of making themselves a target for the person they are trying to report. Another flaw involves the workflow and user interface design of the reporting system which is either exceedingly badly design, or is intentionally design to reduce the number of complaints that are actually lodged. Such flaws should result in a notice form the commissioner that the reporting system is not compliant and the safe harbour provision will not apply until the problem is corrected.

A complaint scheme should also require the social media provider to make certain statistics available to the Commissioner, for example the number of reports made by Australians, the number deemed resolved as a result of the poster removing the content, the number deemed resolved as a result of the platform removing content, and the number of user complaints rejected.

A complaint scheme should also require that reports are responded to in time and at a reasonable level of quality. A scheme which allows users to report content, then automatically rejects all reports,

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

is clearly not compliant. There is a point where the quality of the handling of reports is so poor that the reporting system is counter-productive and reinforces children's sense vulnerability and adults sense of powerlessness to protect them.

Q18 Is merits review by the Administrative Appeals Tribunal the most appropriate review mechanism and if so, which parties and in relation to which types of decision is it appropriate? What are the alternatives?

We support appeal to the AAT on the basis outlined in the paper.

In the case of the Commissioner refusing to remove content, we believe any party suffering harm should have a right to appeal as well as any party at risk of increased harm. We believe schools should have the risk of appeal on behalf of their students. We believe peek bodies representing sections of the community should have standing on behalf of children from their community if the content relates to the community. Such peek bodies may include ethnic and religious community bodies as well as, for example, peek bodies representing people with specific disabilities or illnesses.

In the case of the commissioner removing content, we believe the poster of the content and the social media platform provider should both have standing. We believe there may be groups for the Human Rights Commissioner of the Australian Human Rights Commission to have standing if they believe a decision has been wrongly made and overly impinges on freedom of expression.

Q19 What do industry representatives consider are the estimated financial and administrative impacts of compliance with the proposed scheme? How are these estimated impacts derived?

We believe Social Media Companies are substantially unregulated and are profiting from activities which create a cost to the public health system and the general well being of Australian society. Quite apart from the harm that is done to individuals, we believe this public impact means there is a need for the companies who creating this environment as a by-product of profit making activities to contribute to solving the problems they create. This will involve increased costs for them, including compliance costs. This is not only unavailable, it is entirely reasonable. We believe it is important that their turn over from Australia, which is likely to be substantial, is considered along with the costs.

Q20 In light of the Government's proposed initiatives targeting cyber-bullying set out in Chapters 1 and 2; do the current criminal laws relating to cyber-bullying require amendment?

Yes the current laws do requirement amendment. Training is also urgently needed for law enforcement so they appreciate that these laws do exist, understand their scope, and realise they can be used to secure convictions. The Online Hate Prevention Institute regularly deals with serious incidents which breach the existing laws and are within Commonwealth jurisdiction, yet we regularly encounter a reluctance by police and prosecution services to use the existing laws, even in the case of repeat offences. Documentation by our organisation has resulted in referral by State Government Ministers to Federal Ministers, ultimately resulting is communications outlining that the above laws exist. The existence of the law is not the issue, it is the willingness to apply them and first investigate and then proceed to prosecution. This needs to improve, and *a lower level offence would help to secure more convictions and prior to this, more willingness to investigate*.

We would also like to highlight the need to focus on bullying of children by adults. In the context of protecting children this situation also needs to be considered.

Online Hate
Prevention Institute

OHPI
306 Hawthorn Road,
Caulfield S. Vic. 3162

T (61-3) 9272 5594
E info@ohpi.org.au
www.ohpi.org.au

Regarding ACORN, the Online Hate Prevention Institute will be releasing our own advanced reporting system in April 2014 with additional functionality being added throughout the year. It is intended that our system will be able to provide reports which can be used by Police, the E-Safety Commissioner (should one be appointed), and the Australian Human Rights Commission, and other relevant Government departments and agencies at Federal and State level. Those who report content will be able to remain anonymous if they wish, or may authorize the release of their contact details to specific partners in our system. We would be happy to discuss this further with the Government and with relevant agencies.

Q21 Is the penalty set out in section 474.17 of the Criminal Code appropriate for addressing cyber-bullying offences?

In serious cases of cyberbullying where there is an intent to cause serious harm or serious disregard to whether serious harm may occur as a result of the offence, the penalty is appropriate.

A lower penalty, preferably a fine, be more appropriate in other circumstances, for example where there in an intent to engage in the conduct but not an intent to cause serious harm.

Q22 Is there merit in establishing a new mid-range cyber-bullying offence applying to minors?

Why would a lesser offense only exist when the victim is a minor? This sends the message that bullying of children (even by adults) is more acceptable. We believe a midrange offence should exist for bullying regardless of the age of the victim or the age of the offender.

Q23 Is there merit in establishing a civil enforcement regime (including an infringement notice scheme) to deal with cyber-bullying?

Yes, however, this system should only be used for additional offline penalties when online penalties (such as having content removed or accounts temporarily suspended) are not sufficient to cause a change in behaviour, or where a social media platform fails to take appropriate action.

Q24 What penalties or remedies would be most appropriate for Options 2 and 3?

Option 2 is for a cyberbullying offence, we believe a penalty of imprisonment for up to 3 months, or a fine not exceeding $5,000, would be appropriate. This would allow the courts scope to issue a good behaviour bond on a first offence, a fine on a subsequent offence, suspended imprisonment if a fine is not adequate deterrent, and ultimately imprisonment if the offences continue.

Option 3 is for a civil enforcement regime, we believe this approach is highly desirable. We believe the Commissioner should have the power to impose a range of fines starting at $200 for a first penalty, automatically increasing to $500 for a subsequent penalty, and at the Commissioner's discretion a warning should be able to accompany a fine such that a further breach within 12 months will result in an elevated fine of $2000.