Good practice guidance for the moderation of interactive services for children

Updated 2010

UK COUNCIL FOR CHILD INTERNET SAFETY

Contents

Overview

Part one: User interactive services and moderation

Part two: Recommendations for good practice

Appendices

Overview

Executive summary

This guidance was originally produced in 2005 following the reference in the previous good practice guidance – on chat, instant messaging and web based services – to the fact that some interactive services were "moderated" and others were not. The importance of moderated services was seen within the context of public concern about online spaces, and in particular chatrooms, where children were potentially at risk from undesirable contact or behaviour from adults in order to "groom" and sexually abuse them. The guidance addressed the potential risks to children using interactive services including inappropriate content and contact.

The launch of the UK Council for Child Internet Safety (UKCCIS), see www.education.gov.uk/ukccis, has provided the opportunity to reflect upon the good practice guidance in the light of technical developments, legislative changes and current knowledge about the risks to children and young people using interactive services. As a result, this revised version of the ***Good practice guidance for the moderation of interactive services for children*** has been produced.

While these services offer huge opportunities for children to communicate and learn, evidence has shown that these services can be misused and children are potentially at risk. It is, therefore, important to consider child safety issues when providing these types of services. There are a number of tools and processes that can be implemented to address child safety concerns, one of which is moderation.

Moderation is an activity or process following an agreed policy or set of guidelines to encourage safe and responsible use of an interactive service in accordance with the Terms of Service, Acceptable Use Policy or House Rules. Moderation is performed by human moderators or filtering software (or a combination) reviewing content posted by users and removing content or restricting users as necessary, either pre- or post- publication in near real time or following user reports.

The guidance applies to **user interactive services** through which individuals can make contact and exchange content and personal information with other users in a virtual public "space", such as but not limited to:

- Forums/message boards including comments and reviews
- Blogs and micro blogging
- Social networking
- Massively Multi-Player Online Games (MMOG's or MMO's)
- Virtual worlds
- TV chat services
- Video sharing sites

This guidance has been produced to provide a good practice 'risk assessment' framework on the moderation of interactive services aimed at or likely to attract children, to enhance the safety of children using these services.

The purpose is to:

- describe the different types of user interactive services;

- inform organisations of the potential risks to children using interactive services including bullying, sexual exploitation and grooming, self-harm and destructive behaviours;
- inform organisations of the issues they should take into account when considering what safeguards to deploy;
- describe the types of moderation that can be used;
- assist organisations to develop, review or update polices on the recruitment, on the selection, training and supervision of moderators to safeguard against unsuitable individuals gaining contact with children; and reporting of incidents and concerns.

Contributors

Project team

Co-chair: Chris Atkinson – Online safety and content management consultant
Co-chair until Nov 2009: Rachel O'Connell – Bebo (Chief Safety Officer)
Co-chair from Jan 2010: Dawn Shackleton – British Sky Broadcasting Ltd
(Head of Business Operations, Customer Information System)

Trish Church – Everything Everywhere (Mobile and Broadband Services Safety Manager)
Tamara Littleton – eModeration (CEO)
David Lutman – Department of Education
Robert Marcus – Chat Moderators (Director)
Simon Protheroe – Square Enix, Europe (Online Publishing Director)
Graham Ritchie – CEOP Centre (Policy Manager)
Gabrielle Shaw – CEOP Centre (Head of International & Relations)
Dom Sparkes – Tempero: Social Media Management (CEO & Founder)
Kate Tilley – NSPCC (Policy Advisor)
Paul Wakely – BBC (Director, Moderation Services)

Contributors

Charlotte Aynsley – Beatbullying (Director of Practice)
John Carr – Children's Charities' Coalition on Internet Safety (Secretary)
Julian Cole – BBC (Senior Adviser, Editorial Policy)
Will Gardner – Childnet International (CEO)
Carole Hart Fletcher – KidsOKOline (Director)
Tia Fisher – eModeration (Marketing and Communication Manager)
Neil Malone – Jagex, Ltd (Community Safety Manager)
Rebecca Newton – MindCandy.com (Chief Community & Safety Officer)
Dominic O'Brien – Samaritans (Policy Officer)
Nathan Sawatzky – The Walt Disney Company (Director Community Support DOS)
Lucy Woodward – The Walt Disney Company (acting Head of DOS EMEA)

Part one: User interactive services and moderation

1. Introduction

Aim and scope of the guidance

The Internet and communication technologies are transforming the way we live. Children have embraced the new technologies enthusiastically and especially services where they can interact and exchange content with others, such as social networking and interactive gaming. However, children may also be vulnerable to inappropriate or harmful content and contact through these services.

It is, therefore, important that interactive services which are aimed at or likely to attract children consider seriously the safety of children using these services.

This guidance has been produced to provide a good practice 'risk assessment' framework on the moderation of interactive services aimed at or likely to attract children, to enhance the safety of children using these services.

The purpose is to:

- describe the different types of user interactive services.
- inform organisations of the potential risks to children using interactive services.
- inform organisations of the issues they should take into account when considering the necessary safeguards to deploy.
- describe the types of moderation that can be used.
- assist organisations develop, review or update polices on the recruitment, selection, training and supervision of moderators to safeguard against unsuitable individuals gaining contact with children and reporting of incidents and concerns

There are a range of safeguards that can be deployed by an interactive service provider including privacy settings/safety tools, reporting mechanisms to enable users to flag or report concerns and resources such as parental controls and guidance on safe and responsible use.

Moderation is an activity or process following an agreed policy or set of guidelines to encourage safe and responsible use of an interactive service in accordance with the Terms of Service, Acceptable Use Policy or House Rules. Moderation is performed by human moderators or filtering software (or a combination) reviewing content posted by users and removing content or restricting users as necessary either pre- or post- publication in near real time or following user reports.

This guidance is for:

- organisations providing, or intending to provide, an interactive service aimed or likely to attract children,
- moderation companies involved in the moderation of user interactive services aimed at or likely to attract children,
- organisations considering a digital marketing campaign using user interactive services where the campaign is aimed at or likely to attract children and involves direct contact with children, and

- digital marketing agencies and social media companies involved in digital marketing campaigns aimed at or likely to attract children and involves direct contact with children.

Please note: this guidance is directed at providers who are considering what safety measures to deploy or are involved in the use of moderation in the context of services orientated at children and young people, in order to guide them on how to minimise the risk to children.

This guidance is not intended to be prescriptive or legally binding and it is recognised that the interactive service industry is diverse and aimed at different communities so the guidance is not a 'one size fits all' and parts of the guidance may be less relevant to some services, for example major global web-based services, who provide user-generated content and other services on a vast scale. The Home Office Task Force (updated by UKCCIS) document *Good Practice Guidance for the providers of social networking guidance and other user interactive services provides* information and recommendations on safety matters and may be more relevant for these providers.

This document is part of a group of documents originally produced by the Home Office Task Force on Child Protection on the Internet and updated 2010, setting out a series of models of good practice[1] for the provision of different kinds of internet services by a range of companies and organisations active in the online world.

- Good practice guidance for the internet industry on chat services, instant messaging and web based services (2002)[2]
- Good practice guidance for the moderation of interactive services for children (2005)
- Good practice guidance for search providers and advice to the public on how to search safely (2005)
- Promoting internet safety through public awareness campaigns guidance for using real life examples involving children and young people (2005)
- Good practice guidance for the providers of social networking and other user interactive services (2008)

The series of good practice guidance are intended to be used alongside any legal obligations and other relevant codes, for example the code of practice relating to premium rate charged content and services operated by the premium rate regulator PhonePay Plus, the UK code of practice for the self-regulation of new forms of content on mobiles[3] and the *Safer Social Networking Principles for the EU.* [4]

[1] All sets of guidance both in original and updated format available at www.education.gov.uk/UKCCIS/

[2] The guidance has been revised and has the following title " Good practice models and guidance for the internet industry for: Chat Services, Instant Messaging (IM),Internet connectivity, content and hosting providers

[3] UK Code of practice for the self- regulation of new forms of content on mobiles. www.phonepayplus.org.uk

[4] Safer Social Networking Principles for the EU. See http://ec.europa.eu/information_society/activities/social_networking

History/background

The Home Office Task Force on Child Protection on the Internet (HOTF) was established in March 2001 in response to concerns about the possible risks to children after a number of serious cases where children had been 'groomed' via the Internet.

In the face of such concerns, the Task Force was a unique collaboration bringing together, in a positive partnership, representatives from the internet industry, children's charities, the main opposition parties, government departments, the police and others who shared the aim of making the United Kingdom the best and safest place in the world for children to use the Internet. The work of the HOTF was subsumed in the 2008 creation of the UK Council for Child Internet Safety (UKCCIS).

The UK Council for Child Internet Safety (UKCCIS)

UKCCIS[5] brings together over 170 organisations and individuals to help children and young people stay safe on the Internet. It was launched by the Prime Minister on 29 September 2008 and is made up of companies, government departments and agencies (including the devolved governments in Scotland, Wales and Northern Ireland), law enforcement, charities, parenting groups, academic experts and others. The Council was a recommendation in Professor Tanya Byron's report 'Safer Children in a Digital World'. [6]

The launch of UKCCIS has provided the opportunity to reflect upon the good practice guidance produced by the Home Office Task Force on Child Protection on the Internet in the light of technical developments, legislative changes and current knowledge about the risks to children and young people. As a result this revised version of the *Good practice guidance for the moderation of interactive services for children* has been produced.

Contributors to the original good practice guidance for the moderation of interactive services from some of the moderation services have come together alongside other UKCCIS members to update this guidance. The experiences of the moderation services implementing the *Good practice guidance on the moderation of interactive services for children* since 2005 has been an important contribution to this revised version.

Services not aimed at or not very likely to attract children or young people

Although this document is intended for interactive services aimed at or likely to attract children, the good practice points can be applied more generally to ensure that staff or volunteers are familiar with the ways that interactive services can be diverted for improper or illegal use, placing users at risk. For example, any service can be used to exchange illegal images, or make inappropriate use of the service to make contacts for purposes which are not connected with those of the service.

The Guidance may also be useful for those services aimed directly at vulnerable adults. The UK government introduced the Safeguarding Vulnerable Groups Act in 2006 and the provisions cover the use of moderators in relation to both children and vulnerable adults.

[5] www.education.gov.uk/ukccis/

[6] Safer Children in a Digital World: The Report of the Byron Review. Available at www.education.gov.uk/ukccis/

Using the guidance

The guidance provides a good practice 'risk assessment' framework, based on current best practice, rather than an absolute model to be followed rigidly irrespective of the circumstances. The guidance is not intended to be prescriptive or legally binding, but is offered to interactive service providers and others involved in the moderation of services as well as organisations involved in digital marketing campaigns aimed at or likely to attract children, where the campaign involves direct contact with children, with a strong recommendation for its use.

It is recognised that the interactive service industry is diverse, is provided in various and constantly evolving forms and aimed at different communities, so the guidance is not a 'one size fits all'. Providers are responsible for how they deliver their services.

In determining the actions they should take, providers will need to take into account the particular nature of their services so that they can apply the relevant aspects of this guidance. It is for each provider to judge whether and how far to apply any specific point in the guidance.

2. User interactive services and recent developments

This document refers frequently to interactive services.

It is aimed at *user interactive services* through which individuals can make contact and exchange content and personal information with other users in a virtual public "space" such as but not limited to:

- Forums/message boards including comments and reviews
- Blogs and micro blogging
- Social networking
- Massively Multi-Player Online Games (MMOG's or MMO's)
- Virtual worlds
- TV chat services
- Video sharing sites

User interactive services are constantly evolving and some of the types listed above are still a relatively new phenomenon. Some share many of the characteristics of discussions forums and chat services

Social networking services

Since the publication of the original good practice guidance for the moderation of interactive services for children in 2005, social networking services are now hugely popular and have become a compelling activity for many Internet users.

Social networking services allow users to create their own content and share it with a vast network of individuals and potentially with the world. These services are also very popular with children.

The Home Office *Good practice guidance for the providers of interactive services for children* and the *Safer Social Networking Principles for the EU*[7] provide guidance specific to social networking on the protection of children.

[7] Safer Social Networking Principles for the EU. See http://ec.europa.eu/information_society/activities/social_networking

Virtual worlds and online games

Interactive services have also developed around specific communities of interest. These include virtual worlds, online gaming communities and auctioning and trading communities. They share many of the characteristics of social networking and encourage and facilitate social interaction. The following is a description of the various types of virtual and online games:

- ***Three-dimensional (3D) virtual worlds*** such as Second Life , IMVU or PlayStation Home, provide a virtual space where users interact using 3D avatars. Although the best known examples are modelled on real-world environments, virtual worlds can also be composed of entirely imaginary settings. These differ from services, such as Habbo or Club Penguin which provide a 2D (or pseudo-3D) avatar-based environment with more limited interactivity.

- ***Web-based online games*** are usually relatively casual games designed for shorter play sessions. Typically such games will be hosted on a web portal (such as Miniclip or Addicting Games) which offers many other games and may support a membership system, which rewards players for loyalty and participation in the community of the site's members. There is often limited interactivity with other players from within the game itself, with interaction being focussed on the hosting web portal. Mainstream social networking sites such as Facebook also support games of this type.

- ***Online PC games*** use software sometimes called a "client", i.e. a separate program rather than the web browser. These games may be launched via a web browser, downloaded or distributed via physical media such as CD or DVD. Client-based games are usually more elaborate than web-based games and will provide much more in-game interaction with other players. There is an emerging class of hybrid online games (such as The Hunter) which use a PC client to provide a rich graphical experience for the player, but tie the game closely to a website which acts as a focus for the community of players.

- ***Console-based online games*** are played on TV-based consoles such as the Xbox 360, Playstation 3 or Wii, hand-held devices such as the Sony PSP or Nintendo DS or mobile devices such as the iPhone. Online interactivity is usually controlled to a large extent by the console manufacturers, each of which offers an online service connecting users of its consoles. Examples include Xbox Live or PlayStation Network. Interactive services offered typically including matchmaking (i.e. a service finding other players of similar ability, location, language etc. for a multi-player game session), voice-based chat, friends' lists and so on. Increasingly, console manufacturers are opening up their services to allow interaction with other services such as Twitter and Facebook, or even allowing a limited form of web browsing.

- ***Multi-player games***, which can be on any platform, typically connect a few players for the length of a single play session. For example a sports game played online might pit a player against other human players rather than artificial opponents controlled by the PC or games console.

- ***Massively multiplayer online games*** (MMOG) or Massively Multi-Player Online (MMO) such as World of Warcraft usually connect large communities of players in a persistent world which lasts across many play sessions. Most MMOs involve the player creating a

character which becomes the player's alter ego throughout their involvement with the game, which often lasts several months.

Convergence of technical and communication platforms

The convergence of communications platforms has seen the Internet becoming accessible from numerous devices including mobile phones, games consoles, personal digital assistants (PDA's), PC's and MP4 Players. This means that users can interact with each other and post and download content on many different services and devices.

The Byron Review and moderation[8]

In 2007 the Prime Minister commissioned Professor Tanya Byron to carry out an independent review of the risks children face from the Internet and the report of the Byron Review was published in 2008.

The Byron Review recognised the efforts of content hosts/providers in addressing the risk that children might be exposed to inappropriate or harmful content or make inappropriate contacts with others. Those efforts include 'acceptable user policies' against which content hosts/providers moderate themselves by taking down material and warning and banning users who misuse the site. The Byron Review also highlighted the Home Office *Good practice guidance for the moderation of interactive services:*

"One of the key considerations is to make sure that these moderators are properly trained to take difficult decisions about how to handle content or behaviour on the site."[9]

Online marketing campaigns and social media

Advertisers and marketing agencies have recognised the potential appeal of interactive services to engage with people in recent years. Digital marketing campaigns harness the dynamic and interactive nature of the medium to develop direct engagement with users of social networking services and online communities. Direct engagement can involve contact with users who may be children.

Charities, law enforcement and government are also harnessing the opportunities offered by these developments to reach, communicate and engage with supporters and the public including children and young people.

This type of social media engagement may be subject to moderation deployed by an advertiser or marketing agency in partnership with a moderation company.

[8] Safer Children in a Digital World: The Report of the Byron Review.2008. Available at www.education.gov.uk/ukccis/

[9] Safer Children in a Digital World: The Report of the Byron Review.2008. Available at www.education.gov.uk/ukccis/

3. Children's use of the new technologies

There is a gathering body of research about children and young people's use of the new technologies.[10] Views on young people's development are often polarised. In one view, children are seen as vulnerable, undergoing a crucial but fragile process of cognitive and social development to which technology poses a risk by introducing potential harms into the social conditions for development and necessitating, in turn, a protectionist regulatory environment. Another view is that children are competent and creative agents in their own right, whose 'media-savvy' skills tend to be underestimated by the adults around them, with the consequence that society may fail to provide a sufficiently rich environment for them. Finding a position that recognises both characteristics is important.

Indeed, most psychologists now consider development to be a lifelong process, with children of different ages showing different degrees and kinds of understanding of personal and social matters as they grow older and test themselves against, and learn from, more complex experiences. The influence of the peer group grows in importance during adolescence, as the influence of parents declines (although remaining substantial).

See the UKCCIS website for updates including research on children's use of the new technologies, online risks and safety: www.education.gov.uk/ukccis.

Adolescent and social and sexual development and maturity

It is a normal part of adolescence to test boundaries, challenge adult norms, experiment with relationships, play with identities, explore new sexual experiences, keep or break secrets, exclude or be excluded by peers, deceive parents and worry about one's development. All this is to be expected online, as it is offline. But online, such practices may be spread, manipulated or shared in ways that are easier, quicker, and possibly unexpected in their consequences, compared with offline practices.

There is growing consensus that the activities young people have always engaged in offline they will also do online, and that the convenience, ease and reach of the Internet facilitates these activities, making them more commonplace.[11]

Risks to children and young people online

Most children and young people use the Internet positively but sometimes behave in ways that may place them at risk. Some of these actions to them seem harmless but could expose them to potential harm. In addition, some of these risks do not necessarily arise from the technology itself but result from offline behaviours that are extended into the online space. A young person can be a victim of online abuse through exposure to harmful content and cyberbullying. Young people may also engage in behaviour that is risky to themselves, including cyberflirting and cybersex. These situations can quickly escalate to a point where the young person may lose control.

[10] 'Children's online risks and safety: A review of the available evidence' : report by Nfer prepared for UKCCIS, 2010. See www.education.gov.uk/UKCCIS/. Also 'UK Children's media literacy 2009 interim report' . See www.ofcom.org.uk/advice/media_literacy/medlitpub/medlitpubrss?children/

[11] There are problematic gaps in the evidence that mean that some will continue to question this consensus (we lack evidence on how young people tested sexual limits before the Internet, for example).Further, many more will question the assumption that the Internet has introduced, or is solely responsible for changing behaviour (and risks).

The Byron Review set out the risks to children posed by the Internet and illustrated by the following grid.[12]

	Commercial	Aggressive	Sexual	Values
Content (child as recipient)	Adverts Spam Sponsorship Personal Info	Violent/hateful content	Pornographic or unwelcome sexual content	Bias Racist Misleading info
Contact (child as participant)	Tracking Harvesting personal info	Being bullied, harassed or stalked	Meeting strangers Being groomed	Self-harm Unwelcome persuasions
Conduct (child as actor)	Illegal downloading Hacking Gambling Financial scams Terrorism	Bullying or harassing another	Creating and uploading inappropriate material	Providing misleading info/advice

Potential risks to children using social networking and other user interactive services can include, but are not limited to:

- bullying by peers and people they consider 'friends';
- exposure to inappropriate and/or harmful content;
- involvement in illegal or inappropriate content;
- posting personal information that can identify and locate a child offline;
- theft of personal information;
- sexual grooming, luring, exploitation and abuse through contact with strangers;
- exposure to information and interaction with others who encourage self-harm;
- exposure to racist or hate material;
- encouragement of violent behaviour, such as 'happy slapping';[13]
- glorifying activities such as drug taking or excessive drinking;
- physical harm to young people in making video content, such as enacting and imitating stunts and risk-taking activities;
- leaving and running away from home as a result of contacts made online; and
- addiction/overuse.

Illegal content within a child safety context refers to images of child abuse and inappropriate content usually refers to pornography or sexual content, violence or other content with adult themes which may be inappropriate for children. However, assessments of what constitutes inappropriate or harmful content for children is a matter for parents to decide and can vary between interactive services. Recently there has been a growth in websites that promote self-image related issues such as anorexia, self-harm and suicide as well as extremist sites linked to crime and violence.

It is also important to remember that content posted online can impact on a young person's reputation, both positively and negatively, now or in the future.

[12] EUKidsOnline project : Hasenbrink, Livingstone, Haddon, Kirwil and Ponte, see www2.lse.ac.uk

[13] 'Happy slapping' is a term which typically describes the filming on mobile phones of violent attacks. Happy slapping has been called a youth craze which began in school playgrounds in which teenagers slap or attack unsuspecting children or passers-by while capturing the attacks on camera or videophones.

While interactive services offer great opportunities for children to be creative and express themselves online, they are often unaware that their words or images, although intended for a small audience, can quickly attract a far larger one and may have a lasting impact on other people's perception of them. Some individuals have become notorious as a result of their online postings, which have had both negative and positive impacts on their lives.

Recent research from the EU Kids Online project provides further insights into risks and safety issues from the perspective of European children and their parents.[14]

Bullying and harassment

Bullying via communications technology and victimisation has the potential to be witnessed by a wide audience if it is recorded and shared on the Internet. This may extend the humiliation and embarrassment of the victim. It is difficult to stop abusive content spreading and reappearing, as it can be easily and widely distributed on the Internet. Some victims may therefore find it difficult to manage or recover from the abuse or get closure as the content could reappear at any time, particularly if they do not know who the aggressor is.

Figures with respect to the prevalence of cyberbullying vary widely, from one-in-ten children (Smith et al 2008)[15] to around a third of all children in the UK. Recent research from Beatbullying shows that one in three 11-16 year olds have experienced bullying through the use of mobile phones and the Internet. For a quarter of these this experience was on-going, meaning that one-in-thirteen children were persistently cyberbullied.[16] Cyberbullying can have serious consequences and is known to have led to suicide.

Individual and group disputes are more often than not an extension of arguments and tensions that originate in the offline world. It is therefore no surprise that cyberbullying and harassment are concerns for the children and young people who use social networking and user interactive services. It can manifest itself in the following ways:

- **personal intimidation** – posting personally abusive and threatening comments on the victim's or other people's website, blog or profile;
- **impersonation** – setting up fake webpages that are attributed to the victim of bullying, and may involve the publishing of manipulated pictures and comments;
- **exclusion** – blocking an individual from a popular group or community, deleting them from friendship lists, and/or using 'ignore functions';
- **posting images of bullying incidents** – users sharing and posting images or videos of victims being abused or humiliated offline;
- **stealing a password to take over a user's website , blog, profile or account** – to post comments and images which are attributed to the original user; and
- **making false reports to the service provider**– reporting other users for a range of behaviours with a view to having the user's account or website deleted.

Bullying in any form is distressing. With the proliferation and use of technology by children and young people, victims may feel they cannot escape and perpetrators may believe, falsely, that they are anonymous. Since the bully cannot see the effect they have on their victim, incidents

[14] Report: Risks and safety on the internet: The Perspectives of European Children : Initial findings from the EU Kids Online survey of 9-16 year olds and their parents. www2.lse.ac.uk

[15] Smith,PK, Mahdavi,J, Carvalho,M, Fisher,S, Russel,S and Tippet,N (2008) Cyberbullying: Its nature and impact in secondary school pupils. Journal of Child Psychology and Psychiatry 49,367-385

[16] Riachrdson,B, Cross EJ, Douglas,T, Von Kaenel-Flatt,J (2009) Virtual Violence : Protecting Children from Cyberbullying. Beatbullying. London

can become more serious than in the offline world. The potential for bullying to gain a very wide audience very rapidly is particularly destabilising and highlights the potential impact of this form of bullying.

As well as young people bullying their peers, some adults (particularly teachers) have also found themselves targets of online abuse and harassment. This has caused some concern within schools, not only about the individuals depicted in postings but also the reputation of the school. In some instances, these situations have resulted in investigations being initiated by law enforcement and education authorities.

In the UK, the Department for Education is responsible for policy relating to bullying and cyberbullying which involves children[17]. Industry has also produced an online resource with advice about safe and responsible use of new technology to help the school community including teachers deal with concerns. [18]

Self-harm and destructive behaviours

There has been longstanding concern about children and young people accessing websites, chat rooms and information forums that promote and/or incite risk-taking or dangerous behaviours, self-harm, suicide and eating disorders.

On social networking and user interactive sites, young people seek opportunities to inform one another and express themselves, and therefore may choose to upload content relating to these behaviours:

- eating disorders,
- dieting and body image,
- depression,
- drug and alcohol misuse,
- isolation and loneliness,
- bullying, and
- self-harm and suicide.

To the extent that it allows them to express their feelings and seek support, this can be a positive experience for young people dealing with life's challenges in this period of social development. They can seek out and create networks of likeminded young people who wish to explore these issues and access information. However, there can be negative or worrying aspects of this exploration and engagement which can manifest themselves in the apparent promotion or encouragement of self-harm, for example, filming and publishing these activities.

It is important that providers of interactive services promote opportunities for support and guidance for users related to the issues listed above by having links to helpful information and support organisations. Where a service is moderated, moderators can play an important role by managing what content is posted and identifying a situation where a user maybe in distress or has requested assistance. In these circumstances a moderator can provide links to specific information and support organisations. See Appendix B for information on sources of further advice and information.

[17] See www.education.gov.uk

[18] See www.teachtoday.eu

Sexual exploitation of children and young people online

There is also concern that the capabilities of social networking services and other user interactive services, combined with children's own high-risk behaviour, may increase the potential for sexual exploitation of children and young people by adults, or sometimes by other young people. Some children and young people may not be aware that their behaviour is high risk.

This exploitation can include:

- exposure to harmful content, including adult pornography and illegal child sexual abuse images,
- engaging in sexually explicit communications and conversations that may reduce children and young people's inhibitions. When children send sexual messages and images to each other this is commonly referred to as sexting,
- manipulation and exploitation, which can include being encouraged or paid to pose in sexually provocative ways and pose naked and/or perform sexual acts via webcams, and
- grooming and luring of children to meet offline to sexually exploit them.

The 'grooming' process[19]

Grooming is a process by which someone makes contact with a child with the motive of preparing them for abuse either online or offline. Abusers can use online interactive spaces to find and meet children and young people. Indeed, children and young people can be exploited online without actual physical contact taking place in the real world, for example by sending and exchanging sexual images, and/or by persuading children and young people to send explicit images of themselves. Abusers may also record young people performing sexual acts through webcams.

There have been a number of cases where adults have used social networking services and other user interactive service as a means of contacting and grooming children and young people for sexual exploitation.

Once contact is established in the public space, the adult will often try to move the child to a more private means of communication, such as instant messaging or by texting. In some cases, this has resulted in actual contact abuse.[20] Abusers use a range of techniques to make contact and establish relationships with children and young people, including:

- gathering personal details, such as age, name, address, mobile number, name of school and photographs,
- offering opportunities for modelling, particularly to young girls,
- promising meetings with pop idols or celebrities, or offers of merchandise,
- offering cheap tickets to sporting or music events,
- offering material gifts, including electronic games, music or software,
- offering virtual gifts, such as rewards, passwords and gaming cheats,
- suggesting quick and easy ways to make money,
- paying young people to appear naked and perform sexual acts via webcams,

[19] See NSPCC information and advice to parents: www.nspcc.org.uk/helpandadvice/publications/leaflets/protecting_children_pdf_wdf36296.pdf Also see Sexual Offences Act 2003 for 'meeting a child following sexual "grooming" offence www.legislation.gov.uk

[20] For further information on grooming statistics see www.ceop.gov.uk

- gaining a child's confidence by offering positive attention and encouraging the child to share or talk about any difficulties or problems at home, and providing a sympathetic and supportive response,
- bullying and intimidating behaviour, such as threatening to expose the child by contacting their parents to inform them of their child's communications or postings on a social networking site, and/or saying they know where the child lives or goes to school,
- using webcams to spy and take photographs and movies of victims,
- asking sexually themed questions, such as 'Do you have a boyfriend?' or 'Are you a virgin?',
- asking children and young people to meet offline,
- sending sexually themed images to a child, depicting adult content or the abuse of other children,
- masquerading as a minor or assuming a false identity to deceive a child, and
- using school or hobby sites to gather information about a child's interests, likes and dislikes.

Having made contact with a child or young person, abusers may also use that young person as a means to contact and get to know their friends by using the links to their 'friends' in user profiles.

Whatever its guise, the grooming process can result in many young victims feeling guilty and responsible for inappropriate interactions, exploitation and actual abuse. They can find it extremely difficult to seek help or disclose their abuse because of their sense of personal responsibility, feelings of guilt or shame, and fear that they may not be believed or may be 'blamed' and lose access to the Internet. In some cases they may not identify the experience itself as abuse.

Often the child's feelings may be manipulated, so they genuinely believe they are 'in love' with the abuser.[21]

Risks to children and online gaming

In addition to the above, children may experience risks in relation to online gaming, including cyberbullying and grooming.

Many games allow users to create modified or entirely new game content and to share this with other players. Neither the game developer/publisher nor the console manufacturer, in the case of console games, may be able to control the precise content created and it is this level of creative control and freedom that appeals to players. However, it is therefore possible for users to create content which might be offensive to other players and to make this content publicly available.

Many games, particularly on the PC, provide "level editors" that allow users to create and share game environments. In addition to the general risks of offensive content mentioned in the preceding paragraph, it is also possible for these tools to be used to create environments which, while not offensive in themselves, may nevertheless prove disturbing for children. For example, there have been cases where someone has used a level editor to recreate a specific school or similar environment as a setting for a violent or horror-based game, which can obviously be an unsettling association for a child used to seeing their school as a safe and secure environment.

[21] Ybarra,ML, Mitchell,KJ, Finkelhor,D and Wolak,J (2007), Online Victimization of Youth : Five Years Later. Journal of Adolescent Health,40,116-26 (CV135).Available at www.unh.edu/ccrc/pdf/CV138.pdf

Cheating, whether for financial gain or not, spoils the enjoyment of a game for other players and can, under some circumstances, also be a form of cyber-bullying. In MMOs particularly, much of the game play involves developing a character over a period of weeks or months, with the character gaining in abilities and powers as the player gains experience. Players will often invest a very large period of time in building their characters and can feel a strong affinity with the character. If their game character is victimised, beyond the normal scope of playing the game itself, that can be perceived as an attack on the player.

A recent review of research carried out for UKCCIS on children's online risks and safety concluded:

" A lot of the research focuses on adults' and young people's perceptions of online risks , rather than on their experiences of engaging in risky behaviour or accessing inappropriate content. More evidence is needed to quantify the extent to which children encounter different types of online risks, in particular in relation to engaging in underage or coercive sexual contact or behaviour and other risks such as identity theft or gambling".[22]

4. What is moderation?

Moderation is an activity or process following an agreed policy or set of guidelines to encourage safe and responsible use of an interactive service in accordance with the Terms of Service, Acceptable Use Policy or House Rules. Moderation is performed by human moderators or filtering software (or a combination) reviewing content posted by users and removing content or restricting users as necessary, either pre- or post- publication in near real time or following user reports.

Automated or software-assisted moderation tools attempt to filter words and phrases that they have been programmed to identify, such as personal identifying information, profanities and explicit language that may cause offence. Technical interactive solutions that can also limit a participant's communication to a choice of pre-scripted words and phrases have proven to be effective in significantly minimising risk to children.

The latest generation of software assisted moderation tools can analyse users' behaviour over time and combined with human moderation can help to identify and track user patterns and activity history more effectively. Such tools can help identify potential grooming and bullying behaviour.

Different approaches to moderation

There are different approaches to moderation as set out in the table below

Different approaches to moderation adapted from the table contained in the Byron review[23]

	Description	Benefits	Drawbacks

[22] Children's online risks and safety: A review of the available evidence, report by Nfer prepared by UKCCIS,2010,p25 available at www.education.gov.uk/ukccis/

[23] Table taken from the Byron report with additions. Byron Report available at www.education.gov.uk/UKCCIS

Professional/ Human moderators	Staff employed or contracted by the service provider to provide : *Pre-moderation*: in a pre-moderated service all material/content supplied by users reviewed by the moderator for suitability **before** it becomes visible to other users**;** *Post-moderation*: in a post-moderated service, all material/content supplied by users reviewed **after** it becomes visible to other users and action taken to remove inappropriate content and warn/ban users who break the rules; *Sample moderation*: a moderator may 'patrol' a number of spaces or otherwise examine a sample of content but not all content is reviewed after publication, and *Reactive moderation*: in a service of this type moderation will take place only after a report is made.	Can take a very subtle approach, understanding the nuances of what makes a particular kind of content or behaviour inappropriate and how to respond effectively. Pre-moderation can, in theory, prevent disallowed content appearing at all.	Staff are expensive and require proper training in order to moderate effectively. Often impractical for staff to view everything that is uploaded, especially where large volumes of content are concerned.
Automatic scanning	A computer programme scans for words, phrases, properties of pictures and videos and/or patterns of behaviour by users to identify inappropriate content of behaviour	Can 'look' at huge volumes of content at once, saving time and money	May lack nuance of a human moderator. For example may flag up harmless pictures of a swimmer because of the presence of bare flesh. Alternatively, may miss inoffensive comments where a usually inoffensive word is used in an offensive context
Community moderation[24]	User's report or 'flag' content and behaviour which they believe is inappropriate and contravenes the site's terms of use. Reports or 'Flags' are then reviewed by moderators for contravention of the site's terms of use.	Potentially, every user on the site can play a role in enforcing the rules of their online community.	Users may not report inappropriate content before others users including children have had a chance to see it. Some users may not wish to abide by the rules. Different users may interpret the rules inconsistently.
Reputation-based systems	This is a version of community moderation where the 'reputation' a user has built up on a site (e.g. based on their level of activity, or rating they have been given by other users) gives particular weight to reports they make.	Genuinely empowers users, including children themselves, to become responsible, respected members of an online community with a role in keeping themselves and others safe.	May need to be supported by professional moderators to make sure that users with a high reputation do not adopt a vigilante approach which could lead to some users being bullied.

The Byron report concluded that most social networking services use a combination of different approaches to moderation,

"Techniques for moderating online spaces are evolving and improving all the time, and often the most effective approaches find new ways of combining different techniques – for example, with a site's moderation team supported by technical tools which draw on user reports, weighted by the user's reputation."

Human moderation

[24] Community moderation can also be referred to as reactive moderation.

In general, there are three main approaches to the use of human moderators, each of which has different implications for risk and employment practice:

- **sub-contractors**
- Moderators employed by a company which is contracted to provide moderation services to another company;
- **volunteers**
- Users of the community service who have applied to the provider to become moderators of the service and who might not be paid for their time; and
- **in-house employees**
- Members of staff of the service provider who are specifically required to moderate the service.

The role of a 'moderator'

There are a range of terms and terminology to describe the different roles and responsibilities that take place within interactive environments but there is no fixed definition of a moderator. The role a moderator undertakes will depend on the kind of service offered and can be in either pre- or post- publication in near real time or following reports from users.

For the purposes of this document, we have considered the following to be separate roles:

- **moderator** – this term is used to describe an individual who has a clear and defined role to monitor and filter user-generated content, and who will intervene where interactions break the 'House Rules' or cause concern. Moderators in some services also take action against users who break the *"House Rules"* or *"code of conduct"*, ranging from sending them a warning through to denying the offending user access to the service. Moderators in some services may also take action when a user posts a message which causes concern, for example a message containing an intention to commit suicide, and signpost the user to links to support organisations or contact the emergency services. They may therefore have a position of trust and authority over a child user, and may also have access to data about users;

- **host** – this is a common term used to describe an individual in an interactive environment who hosts a particular chat room, forum or message board. Sometimes their role is simply to meet and greet new members and offer information about the interactive service and respond to any questions by the new user. Sometimes they may also try to facilitate discussion in the interactive service, which may have a particular theme or not. They may or may not have authority over a child user or access to data about users;

- **abuse team** – respond to reports from users and may also take action against users who break the *"house rules"* or *"code of conduct"*, ranging from sending them a warning through to denying the offending user access to the service.

There are other terms, which may be used to describe people with these or similar roles, for example "*guide", "monitor", "animator"* or *"text-jockey"*. A single individual may sometimes undertake all or different aspects of these roles.

It is important to recognise that moderation on its own is not a panacea, but can play an important part in keeping children safer online, alongside a range of other tools including privacy tools/settings, filters and parental controls as well as education on safety and responsible use.

5. Recruitment and selection of human moderators

Background

Following a series of public enquiries in the 1990s into the abuse of children in local authority homes in the UK, there has been growing recognition of the potential of sex abusers to gain employment with children in order to abuse and exploit them. The need for clear, transparent and rigorous recruitment and management procedures within organisations working with children was a key recommendation throughout the enquiries.

In recent years, cases of sexual abuse of children in other sectors such as the sport, leisure and entertainment sectors and youth and faith-based organisations have extended the need for child safety beyond child welfare organisations. Similarly, interactive services are raising child safety issues both in terms of the posting of inappropriate content and potential contact with children from sex abusers.

In response to concerns to make environments safer for children, the UK government asked the Children's Workforce Development Council (CWDC) to produce guidance and online training on safer recruitment. *"Recruiting safely: Safer recruitment guidance helping to keep children and young people safe"* was published in November 2009,and applies to everyone who has a role (paid or volunteer) in an organisation working with children, including those who may not have direct contact with children, and are likely to be seen by children as a safe and trustworthy adult.

The guidance provides advice on a range of topics including recruitment and selection, vetting and checking candidates, safer practice and managing allegations against staff. The guidance also covers the use of volunteers and recommends that where an organisation 'is actively seeking volunteers about whom it knows little, it should follow the same safe recruitment measures as it would for paid staff.'

Further information on safer recruitment including the guidance and online training can be found at the CWDC website at www.cwdcouncil.org.uk

Legislative developments in the UK

There have been considerable legislative developments in recent years to improve the protection of children from abuse by those in positions of trust and authority. These include the introduction of the abuse of trust offences in the *Sexual Offences Act 2003, and the extension of Criminal Records Bureau (CRB) disclosures to include persons whose suitability is being assessed for* "employment which is concerned with the monitoring, for the purpose of child safety, of communications by means of the Internet". The Rehabilitation of Offenders (Exceptions) Order (Northern Ireland) 1979 applies in Northern Ireland. At this stage, employment in regulated activity excepted, which creates an entitlement to both standard and enhanced disclosure certificates through AccessNI. The most recent *developments are* the Safeguarding Vulnerable Groups Act 2006 and the establishment of a new vetting and barring scheme in England and Wales, the Safeguarding Vulnerable Groups (Northern Ireland) Order 2007 which does the same in Northern Ireland and the Protection of Vulnerable Groups (Scotland) Act 2007 which will shortly do the same in Scotland.

The Vetting and Barring scheme

The Vetting and Barring Scheme under the Safeguarding Vulnerable Groups Act 2006, the Safeguarding Vulnerable Groups (Northern Ireland) Order (SVGO) 2007 and the Protection of Vulnerable Groups (Scotland) 2007 (PVG) Act reduce the risk of harm to children and vulnerable adults by barring individuals from gaining access to such people by doing certain work ("regulated activity") with them in a paid or unpaid capacity, where the individuals are guilty of a serious offence or were referred following serious misconduct in a post. The schemes each maintain a barred list, and each scheme recognises the other schemes' lists' so, an individual included in the children's barred list in one part of the UK is barred from regulated activity relating tochildren across the whole of the UK.

At the time of release of this guidance in late 2010 the Government is in the process of reviewing the scope of the Vetting and Barring scheme in England and Wales. The description that follows relates to the Vetting and Barring scheme as it currently stands, but the Scheme is likely to change considerably following the review. Further announcements are expected in January 2011. For further information, please see: *http://www.homeoffice.gov.uk/media-centre/press-releases/vetting-and-barring.*

Regulated activity and moderators

Regulated activity (or, in Scotland, 'regulated work') covers the range of activities where people are working or volunteering with children or vulnerable adults, which a barred person must not do, and where referral duties apply. This includes the moderation of a public electronic interactive communication service which is likely to be used wholly or mainly by either children or vulnerable adults.

Safeguarding Vulnerable Groups Act 2006 Schedule 4 – Regulated activity relating to children
Safeguarding Vulnerable Groups (NI) Order 2007 Schedule 2

Activities
2 (1) The activities referred to in paragraph 1(1) are—
… (e) moderating a public electronic interactive communication service which is likely to be used wholly or mainly by children;…
(4) For the purposes of sub-paragraph (1)(e) a person moderates a public electronic interactive communication service if, for the purpose of protecting children, he has any function relating to-
(a) monitoring the content of matter which forms any part of the service,
(b) removing matter from, or preventing the addition of matter to, the service, or
(c) controlling access to, or use of, the service.
(5) But a person does not moderate a public electronic interactive communications service as mentioned in sub-paragraph (4)(b) or (c) unless he has—
(a) access to the content of the matter;
(b) contact with users of the service.

The same provision is repeated in relation to regulated work relating to vulnerable adults.

To note: the Northern Irish legislation uses the same provision in relation to both children and vulnerable adults; the Scottish legislation uses the same provision in relation to children but 'regulated work relating to protected adults' does not include moderation of interactive services

Those who moderate, either paid or not, and employers of moderators of a public electronic interactive communication service which is likely to be used wholly by or mainly by children or vulnerable adults are affected by the scheme.

Key points – the new Vetting and Barring scheme (VBS)

The Vetting and Barring Scheme (which was due to start on 26 July 2010) has been halted to allow the government to carry out a review. However, the regulations introduced in October 2009 will still apply.

From 12 October 2009 the barring and referrals aspects of VBS started:

- *You must not knowingly use in regulated activity* (paid or unpaid), a barred person. *That is, a wider range of activity than covered by previous bars.*
- *If you use people (paid or unpaid) in regulated activity or controlled activity, and subsequently dismiss or cease using them because you think they have harmed or pose a risk of harm to children or vulnerable adults, you must refer the case to the Independent Safeguarding Authority (ISA).*
- *If you yourself are barred from regulated activity with either children or vulnerable adults you must not work, or seek to work, in regulated activity with that group.*
- *The previous lists of people barred from certain work with children or vulnerable adults in England, Wales and Northern Ireland have been phased out and replaced by two lists: the ISA Adults' Barred List and the ISA Children's Barred List.*

While the barring and referral aspects of the VBS provide significant safeguards, it is important to remember that they are part of a wider framework of safe recruitment practices. In particular they do not replace Criminal Records Bureau (CRB) or Access NI disclosures, which are still required by law or recommended by Government guidance for some positions; nor does it remove the need for employers to develop and apply robust recruitment procedures, including checking identity, qualifications and references and enquiring into career history.

For further information on the ISA and the Vetting and Barring Scheme see the ISA website: www.isa-gov.org.uk; for further information on the Protecting Vulnerable Groups scheme in Scotland please see www.infoscotland.com/pvgscheme.

For further information about disclosures, contact the Criminal Records Bureau on 0870 90 90 811 or AccessNI on 028 90 25 91 68.

Recruitment and selection for moderators outside the UK

It is a matter of fact that communication technologies operate on a global basis and a number of companies run their operations across a number of different territories. Some services may fall outside the scope of the UK legislation but may be covered by laws in the country where they operate. It should be noted that a new EU Directive on combating the sexual abuse, sexual exploitation of children and child pornography is currently being negotiated. This Directive would, if passed in its current form, require all Member States to consider disqualifying anyone convicted of sex offences involving children from (at least) professional activity involving regular contacts with children, and for information involving such offences or disqualifications to be available to employers.

Neither the Criminal Records Bureau nor AccessNI is able to conduct criminal records checks overseas. Some countries, including most in the EU, have arrangements in place to provide information to prospective employers upon request. The level of information varies from country to country. Many countries use criminal records as a starting point. However, what constitutes a

criminal offence will depend upon the legal framework in the relevant country. It is also worth noting that, in some countries, only judgements given by criminal courts are recorded. In other cases decisions by administrative authorities are included in the check. Some countries also operate a “disqualification from working with children” system.

If a UK provider employs moderators from outside England, Wales and Northern Ireland, equivalent checks should be made with national agencies (if they exist) in other countries. The CRB may be able to assist by providing details of what is available in a range of countries.

Part two: recommendations for good practice

The following recommendations provide good practice guidance for the moderation of interactive services aimed at or likely to attract children. The guidance also contains recommendations for organisations including marketing agencies and social media companies considering a digital marketing campaign aimed at or likely to attract children where the campaign includes direct contact with users who may be children.

1. General principles

- Each of the recommendations below should be included as part of a larger focus on user protection by responsible interactive service providers, moderation service providers, marketing agencies and their clients considering digital marketing campaigns to engage with children or likely to attract children. None of them should be viewed as a panacea.

- These recommendations apply to all platforms, fixed and mobile, while recognizing that the different characteristics of each platform (for example, the different screen sizes and methods of navigations) may require modified or alternative approaches to safety.

- Language and terminology should be accessible, clear and relevant for all users, including children, young people, parents and carers, especially in relation to the service’s terms and conditions, privacy policy, safety information and reporting mechanisms.

- When developing new services or digital campaigns, providers, marketing agencies and their clients considering digital marketing campaigns to engage with children or likely to attract children should consider the existing good practice guidance produced by the Home Office on Child Protection on the Internet.[25]

2. Undertaking a risk assessment

It is important for interactive service providers to undertake a risk assessment of their own service and the potential for harm to children, in order to decide what safeguards to deploy, including the use of moderation.

The following points are key areas for consideration:

- whether the service is specifically aimed at children;

[25] All sets of Guidance available at www.dcsf.gov.uk/ukccis

- whether the service is likely to attract children and younger users due to the theme of the service such as football or celebrities;
- whether the service enables users to have contact and interaction with strangers;
- whether the service enables users to manage their contact with others including friends, friends of friends and strangers;
- whether the risks to children and young people associated with interactive communication services may arise;
- whether the service enables users to post personal information such as contact details;
- whether the service enables users to upload and share content with others;
- the ease with which users may be able to move from a public moderated area to a private un-moderated area within the same service; and
- whether users of the service are anonymous and identity is not verified and stored.

Part 1 of this guidance provides some useful information on risks to children to help inform a risk assessment.

The target age group for the service may also be a useful indicator when considering appropriate safeguards – a service for younger children may require a more restrictive approach to interaction with others than an interactive service for older children.

Having undertaken the risk assessment, it is necessary to decide what safeguards are necessary and this could include a combination of moderation, privacy settings/safety tools, reporting mechanisms to enable users to flag or report concerns, and resources such as parental controls and guidance on safe and responsible use.

For further information and good practice recommendations on privacy settings/safety tools and controls, safety information and reporting mechanisms see the Home Office *Good practice guidance for the providers of social networking and other interactive services*. [26]

Where moderation is employed it will be necessary to decide which form of moderation or combination of forms is appropriate.

3. Online marketing campaigns and social media

The misuse of new technologies and the potential risks to children and young people is an important consideration for a wide range of organisations using interactive services to engage with audiences online for marketing purposes as well as providers of such services. Direct engagement can involve contact with users who may be children.

The recommendations in this guidance should be considered as part of an assessment and management of risk to this audience and may also be useful to address as part of a brand reputation strategy.

Following a risk assessment your organisation may wish to consider the use of moderation as part of your digital marketing campaign to meet your specific safety requirements. Understanding the safeguards provided by the service provider is an important consideration. Your organisation may also wish to partner with the service provider to establish joint

[26] All sets of Guidance available at www.education.gov.uk/ukccis/

agreements and/or procedures to complement or support the site moderation provided by the interactive service provider.

Based on your risk assessment some or all of the following recommendations may be relevant:

- where a digital campaign seeks to engage with, or is likely to attract children it should be clear who has responsibility to report behaviours considered a breach of the providers' Terms of Service, Terms of Service or House Rules and also potentially illegal activities. This includes, but is not limited to :

    - posting images depicting child abuse or exploitation,
    - suspicious behaviour towards children and young people, including behaviour indicative of grooming,
    - signs that may warrant intervention (for example, invitations to meet off-line or requests for personal details from another user),
    - bullying and harassment,
    - posting of inappropriate content, such as information promoting or encouraging self-harm, suicide or eating disorders, and
    - other potentially illegal or criminal behaviour.

- advertisers should also follow relevant local guidelines or codes for advertising to minors. In the case of the UK, this is the British Code of Advertising, Sales Promotion and Direct Marketing (the CAP code).[27]

The Advertising Standards Authority announced an extension of its digital remit.[28] From 1st March 2011 the rules in the CAP code will apply in full to marketing communications online. This includes the rules relating to misleading advertising, social responsibility and the protection of children. The new remit will cover:

- advertisers' own marketing communications on their websites; and
- marketing communications in other non-paid for space under their control, such as social networking sites.

4. Safety information, awareness and education

The provision of safety advice for users by interactive service providers including information on the specific safeguards deployed on a service to help users manage their online experience and protect themselves from harm is crucial.

The Home Office *Good practice guidance for the providers of social networking and other interactive services* as well as the other Home Office Good practice guidance documents, contain recommendations on safety information, awareness and education for services providers.

Recommendations

Based on your risk assessment, some or all of the following recommendations may be relevant:

[27] The CAP code can be found at http://bcap.org.uk/The-Codes/CAP-Code.aspx. Following a recent review, a new set of advertising codes will come into force in September 2010.

[28] www.asa.org.uk

Providers of public interactive communication services should provide clear and prominent information:

- about the kind of interactive service offered. For example, ‘this service is for children aged 11-14 years of age.’;

- about the specific safeguards provided on the service for users, for example safety tools, privacy settings, block/ignore tools and reporting mechanisms;

- about whether the service is moderated or un-moderated . If moderation is used, what type of moderation is used and how it works from a user perspective, for example pre-moderated or post moderated, in near real time or following user reports; and

- if human moderators are used, about :

    - the type and level of criminal record check carried out for human moderators;
    - what the moderator is expected to do and if there is a facility to alert and/or make a report to the moderator how this works; and
    - provide a means of reporting a moderator’s failure to meet expectations.

Parents need to be advised by interactive communication providers of the importance of communicating with their children about their safe and responsible behaviour online on a regular basis and to be aware that moderation and other safeguards are not infallible.

5. Personal information and data security

The misuse of personal data can present a risk to users including children. Children may willingly provide their and others’ personal information without being aware of the consequences.

The posting of personal details such as full name, address, mobile number, email address or school name on their social networking profile, for example, can leave a child vulnerable to cyberbullying, identity theft and being contacted online and located in the offline environment.

Many interactive services collect personal data, as a means of authenticating user identification and obtaining agreement to the terms and conditions, by asking users to register and provide a certain amount of personal information.

Interactive and moderation service providers need be aware of the potential misuse of personal data internally by people who have legitimate access to data and may wish to use this data to initiate contact with children, either to make inappropriate contact themselves, or to pass to third parties outside the organisation. It is recognised that in some situations a moderator may not have access to a users’ registration details or other online data other than their online identity for example, a user name.

Recommendations for personal information and data security

Based on your risk assessment some or all of the following recommendations may be relevant:

Service providers, including moderation companies, where relevant should:

- minimise the amount of personal data collected from children;

- have a clear data protection policy in place to protect end user privacy and help employees comply with the law. This should include situations where moderators work from home. In the UK, organisations are required to comply with the Data Protection Act 1998 (DPA 1998); and

- consider the necessary safeguards to minimise the risk of personal information being misused including the use of moderation. For further information and recommendations on registration, user profile and controls see the *Good practice guidance for the providers of social networking and other user interactive services.*

For further information on personal information and the data security see the *Personal information online code of practice* published by Information Commission Office: www.ico.gov.uk

6. Reporting incidents and concerns

The internet industry and law enforcement agencies have achieved a great deal of success in co-operating effectively to combat illegal activities online using well-established protocols and procedures, in line with applicable law. These arrangements for providers to report potentially illegal incidents and suspicious behaviour where interactive services have been misused are also subject to the applicable laws and local jurisdiction.

Depending upon the role of a moderator within a service, it may involve identifying a range of concerns. These concerns could range from offensive communications or other behaviours which breach the provider's terms and conditions, to potentially illegal activities, including but not limited to :

- posting images depicting child sexual abuse or exploitation,
- suspicious behaviour towards children and young people, including behaviour indicative of grooming,
- bullying and harassment,
- posting of inappropriate content, such as information or encouraging self-harm, suicide or eating disorders,
- incorrectly tagged adult or age-inappropriate content, and
- other potentially illegal or criminal behaviour.

It is for each service provider to make an assessment of how their services are used, which behaviours are likely to occur and how concerns can be addressed.

The Home Office *Good practice guidance for the providers of social networking and other interactive services* also provides recommendations on reporting concerns and should be cross referenced with this guidance.

Recommendations

Based on your risk assessment some or all of the following may be relevant:

Service providers should:

- deploy and give due prominence to a system for the user to report incidents or inappropriate content to moderators and/or abuse teams and enable the user to alert the moderator/abuse teams. The reporting mechanism should be clear and accessible to the user,

- ensure that moderators are familiar with links to sources of help and advice and/or be able to signpost users to links where they can choose an appropriate agency or organisation to contact if they have a concern or want to make a report. Links to sources of help and advice may be available on a general page available on the service[29] ,

- ensure that moderators are familiar with the reporting policy and should know what to do, who to escalate reports to, when and how. The policy should also cover situations where the moderator observes behaviour that gives cause for concern on the service and involves escalation to a moderator's supervisor, manager or abuse management team,

- have in place procedures for the role of the moderator when potential illegal images of children are identified . This should include the handling, storage and reporting of potential illegal child abuse images to the appropriate authority. In the UK this is the IWF or the police, in accordance with the Sexual Offences Act 2003 and the accompanying Memorandum of Understanding, and

- have in place clear procedures, whether moderation is done internally (e.g. by a service provider) or externally (e.g. by a moderation company) for reporting an incident, including emergency incidents where there is an immediate threat to life or where a child is at immediate risk of harm. It may be necessary to arrange out-of-hours contact if the service is available outside office hours.

Disclosing communication data to law enforcement

Based on your risk assessment some or all of the following recommendations may be relevant:

Service providers should, where relevant and consistent with applicable laws:

- have in place clear procedures for the disclosure of data and other non-public information to law enforcement which are compliant with relevant data protection and privacy laws. The service provider is usually responsible for disclosure of data rather than a moderation company,

- where a moderation company has explicit responsibility for the disclosure of communication data to a public authority this should be made clear in the relevant service and contractual agreements with the service provider, and

- where the Regulation of Investigatory Powers Act (RIPA) 2000 requirements apply, have in place clear procedures to allow for the disclosure of communication data and authenticating communication data requests from public authorities in accordance with relevant legislation (RIPA 2000). These procedures may include:

[29] See Appendix A for sources of further advice and information.

- what communication data can be disclosed and to which public authorities;
- authenticating a request from a public authority to disclose data;
- a designated person or contact point for the purpose of liaising with public authorities, including law enforcement, and
- a record of disclosures made to public authorities.

For further information about RIPA 2000 see www.legislation.gov.uk and www.homeoffice.gov.uk

7. Recruitment and selection of human moderators

There is always a risk that any role that allows access to children will be attractive to child abusers. In the case of interactive services, access to children could be obtained by becoming a moderator. This is because the role may provide:

- opportunity for direct contact with children,
- a perceived position of trust and authority, and
- access to personal information about children.

Those who are responsible for moderators and hosts will need to assess the risk to children based on:

- what opportunity for contact a moderator has with children through the service;
- the extent to which they are in a position of trust and authority in relation to children;
- what access if any they have to personal information about children; and
- how closely the process of moderation is supervised and managed.

Based on your risk assessment some or all of the following recommendations may be relevant:

Service providers, including moderation companies, should:

- make efforts to adopt safer recruitment and selection procedures based on relevant good practice guidance. See CWDC " Recruiting safely: Safer recruitment guidance helping to keep children and young people safe" at www.cwdcouncil.org.uk ,

- where relevant, carry out the appropriate CRB or AccessNI (*Enhanced Disclosure*) and ISA registration check prior to an appointment to a position, paid or unpaid, where the duties involve a regulated activity involving moderation as set out in the Safeguarding Vulnerable Groups Act 2006 and the Safeguarding Vulnerable Groups (NI) Order 2007 – subject to the outcome of the ongoing review - and

- interview face-to-face all prospective volunteers and employees for moderation positions involving contact with children.

8. Training of moderators

The training of moderators needs to cover a number of key areas so they have an awareness of relevant issues and policies and can operate effectively.

All moderators should have a reasonable level of awareness of child protection issues relevant to the service being provided. The aim should be to provide moderators with the necessary skills to identify child protection issues and know when to hand over to trained professionals. The depth to which this is necessary will vary with the level of service being provided. For example, a person moderating an interactive service aimed at children between 5 - 11yrs will need a different mix of knowledge from a person moderating a teens service. Where a service is aimed at a particular subject area, for example, bullying or drugs, moderators may also need specific training.

It is not critical whether training is provided in-house or by use of outside expertise. What is important is that the overall result is to prepare the moderator to apply their knowledge effectively. The training should reflect the realities of what is expected of the moderator in the particular environment.

Recommendations for areas moderators should be trained in

Based on your risk assessment, some or all of the following recommendations may be relevant:

- understanding when and how moderators are expected to intervene, and the activities that are prohibited to moderators, for example, unauthorised communication or meeting with service users, together with the reasons for such prohibitions,

- the use of a reporting and escalation procedure (including how, when and why moderators should refer particular types of incident and to whom a report should be made). Procedures should include urgent and serious incidents and how to contact the appropriate child protection and law enforcement agencies. Where there is an immediate risk to life or where a child is at immediate risk of harm, users should be advised to contact the emergency services by phoning 999 (UK) 911(USA) 000 (Australia) or 112 (Europe),

- the full range of behaviour that may constitute child abuse, and address the ability of moderators to identify such behaviours manifested on the interactive service. Experience has demonstrated that there are some behaviour patterns which, while not immediately obvious as signs of abuse, may merit further investigation,

- the serious risk of harm to children posed by child abusers who misuse interactive services in order to gain contact with children in order to abuse them. This can sometimes involve child abusers pretending they themselves are children,

- child development issues, and associated behaviours that can be expected from different age groups, for example early teens will be forming and testing their sexual identities and may engage in using explicit and sexual language and flirting,

- other associated risk-taking and dangerous behaviours linked with drug or alcohol misuse, isolation and loneliness,

- behaviours identified or considered a breach of the providers' Terms of Service, Terms of Service or House Rules and also potentially illegal activities. This includes, but is not limited to:

- posting images depicting child abuse or exploitation,
- suspicious behaviour towards children and young people, including behaviour indicative of grooming,
- signs that may warrant intervention (e.g. invitations to meet off-line or requests for personal details from another user),
- bullying and harassment,
- posting of inappropriate content, such as information promoting or encouraging self-harm, suicide or eating disorders, and
- other potentially illegal or criminal behaviour.

- the relevant law applicable to their work. This should include awareness of material that is potentially illegal – such as images of child abuse – and also procedures to report potentially illegal content to the appropriate authorities in accordance with legal requirements and arrangements with law enforcement. See the IWF website[30] for further information on the law and reporting potentially illegal images of child sexual abuse,

- recognition of and how to respond appropriately to users of their service who are vulnerable, or are at risk. Where, for example, they appear to be in need of counselling or support, there should be a clear escalation procedure. This is important with any user, but is particularly important when the moderator believes the user may be a child. Where there is an immediate risk to life or where a child is at immediate risk of harm, users should be advised to contact the emergency services by phoning 999 (UK) 911(USA) 000 (Australia) or 112 (Europe),

- an understanding of, and ability to identify, material that is inappropriate and/or harmful to children. Inappropriate and/ or harmful content typically refers to pornography or sexual content, violence or other content with adult themes which may be inappropriate for children. Assessments of what constitutes inappropriate or harmful content for children can vary between interactive services. The Terms of Service, Acceptable Use Policy or House Rules usually contain what is inappropriate and/or harmful content for that particular service, and

- the most recent advice on staying safe online, so moderators are able to signpost users to the relevant online safety resources and help organisations when a user contacts them with a concern, or if they need to intervene to deal with a situation. This may include advice from the various multi-stake holder or cross-agency collaboration initiatives.

9. Management, supervision and accountability of moderators

Experience in a variety of settings, including children's homes, nurseries, youth work and faith and educational contexts, has shown the importance of good and informed management systems for the protection of children.

As far as possible, professional childcare standards should be incorporated where interactive services are aimed at or likely to attract children to enhance the safety and protection of children. Because of the crucial role managers play, they need to be fully informed of the child protection issues related to the operation of interactive services.

[30] www.iwf.org.uk

Recommendations for the management and supervision of moderators

Based on your risk assessment, some or all of the following may be relevant:

Service providers should:

- have in place effective policies and management systems for moderators;

- have in place procedures to ensure that any concerns that a moderator has harmed or poses a risk of harm to a child are referred to the ISA for consideration. Specifically, a 'Regulated Activity Provider' must refer the case to the ISA if they think the individual:

    - has committed an offence that would lead them to be automatically included on a barred list under the automatic barred provisions,
    - engaged in relevant conduct, or
    - the harm test is satisfied [31].

Further information about making a referral and a downloadable referral form is available from the ISA. See www.isa.homeoffice.gov.uk;

- have in place procedures to cover situations where a response to training gives a manager cause for concern about a particular trainee moderator to limit the services that person will moderate to those which exclude children, impose particular supervision, or reconsider their employment, and where appropriate make a referral to the ISA;

- have in place procedures to ensure moderation practice fosters awareness of child safety and protection and managers and moderators are aware of their responsibilities in respect of child protection;

- keep a record of which moderator is responsible for any service at any particular time in order to facilitate investigations of any complaints after the fact;

- carefully consider management and supervision in situations where moderators are working at home to take account of the added difficulty of supervision at a distance. Service providers will need to consider a range of measures, which might include instant messenging, video conferencing and telephone contact, to ensure:

    - the moderator is who they should be;
    - the work-station is set up in a way to keep users' data secure; and
    - the organisation can monitor what the moderator does; and

- ensure that managers supervise the work of moderators so that:

- they are able to monitor the impact on moderators, particularly for stress, burnout or behaviours that may give rise to concern for the staff member or for safety and security of the service and have procedures in place to deal with these concerns;

[31] A referral to the ISA is also required irrespective of whether an individual has been removed from the moderation duties which fall under the regulated activity requirement as set out in the SVG Act 2006, dismissed from employment, no longer used by the employer, or if the individual has left while under investigation for allegedly causing harm or posing a risk of harm. An individual may also be referred by an employer if they are concerned about their conduct and think the ISA ought to be aware of it. Referrals should not be made on the basis of allegations suspected to be unfounded or malicious. See *The Vetting and Barring Scheme Guidance,Oct,2009 at www.isa.homeoffice.gov.uk* .

- they can raise any specific concerns relating to users or patterns of behaviour observed in the course of their work, and these can be escalated to senior management or law enforcement in accordance with escalation procedures; and

- if there are concerns that a moderator or former moderator employed to moderate a service wholly or mainly for children or vulnerable adults has caused harm or poses a risk of harm to a child or vulnerable adult ,a referral is made to the Vetting and Barring Scheme for consideration.

Appendix A: The criminal law

It is important to note the general principle that an action that is illegal if committed offline is also illegal if it is committed through an interactive service.
This applies both to issues such as distributing illegal material and also harmful behaviour if it amounts to a course of harassment, or grooming. Inciting someone to commit an offence is also no less an offence simply because it is carried out through a computer or mobile device. Other criminal activity may include fraud and identity theft.

Each case will be different, and it is impossible to set out in a document of this sort a definitive explanation of the law. Nevertheless, it is hoped this brief and general guide to a few relevant offences, particularly those involving children, will be helpful. Please note that this is not a definitive list of all of the relevant legislation and that the laws summarised below can be subject to amendment.

No-one using an interactive service should be under the illusion that the criminal law does not bear on what they do. Some of the legislation below applies only to England and Wales, although Scotland, Northern Ireland and other jurisdictions will have equivalent legislation.

The Acts below appear in alphabetical order. All Acts can be found in full at www.legislation.gov.uk

Communications Act 2003 (also extends to Northern Ireland)

Section 127 (1) provides that it is an offence if any person sends a message or other matter by means of a public electronic communications network which is grossly offensive, indecent, obscene or menacing, or if a person causes any such message or matter to be sent.

Section 127 (2) provides that a person is guilty of an offence if, for the purpose of causing annoyance, inconvenience or needless anxiety to another, he sends or causes to be sent by means of a public electronic communications network a message he knows to be false, causes such a message to be sent, or persistently makes use of a public electronic communications network.

The offences carry a penalty of a maximum of six months' imprisonment and/or a level five fine (£5000).

Computer Misuse Act 1990 (also has implications in Northern Ireland)

This may come into play when online bullying – cyberbullying - takes the form of hacking into someone else's account.

Defamation Acts 1952, 1996

Defamation is a civil "common law" in respect of which the Defamation Acts of 1952 and 1996 provide certain defences. It applies to any published material that damages the reputation of an individual or an organisation, and it includes material published on the internet. A civil action for defamation can be brought by an individual or a company, but not by a public authority. It is up to the claimant to prove that the material is defamatory. However, the claimant does not have to

prove that the material is false – the burden of proof on that point lies with the author/publisher, who has to prove that what they have written is true.

Where defamatory material is posted on a website the person affected can inform the host of its contents and ask the host to remove it. Once the host knows that the material is there and that it may be defamatory, it can no longer rely on the defence of innocent dissemination in the Defamation Act 1996. This means that the person affected could (if the material has been published in the jurisdiction, i.e. in England and Wales) obtain a court order (an injunction) to require removal of the material, and could sue either the host or the person who posted the material for defamation.

The following provisions of the 1996 Act extend to Northern Ireland: S1, 2 to 4 [except 3(9)], 6, 7, 8 to 11, 12(3), 13, 14 &15 and Sch. 1, S16 and Sch. 2, S17(1), 18, 19 and 20

Malicious Communications Act 1988

Section 1 of this Act makes it an offence to send an indecent, grossly offensive or threatening letter, electronic communication or other articles to another person with the intention that it should cause them distress or anxiety.

Section 2 refers to corresponding legislation made under the Northern Ireland Act 1974.

Obscene Publications Act 1959

It is an offence under this Act to publish an obscene article. Publishing includes circulating, showing, playing or projecting the article or transmitting that data. An obscene article is one whose effect is such as to tend to deprave and corrupt persons who are likely to read, see or hear the matter contained or embodied in it. Depravity and corruption are not confined to sexual depravity and corruption. This Act does not extend to Northern Ireland.

Protection of Children Act 1978

The 1978 Act essentially prohibits creation or distribution of indecent photographs of children, in whatever form. Proscribed activities are taking, making, permitting to be taken or made, distribution or showing, possessing with intent to possess or show, or publishing an advertisement for such photographs. The maximum penalty is 10 years imprisonment. Simple possession of such a photograph is an offence under s 160 of the Criminal Justice Act 1988, and carries five year maximum penalty. Although there are defences specified in the Acts, it is unlikely in the extreme that any of these could apply to images that might be sent over a public interactive service, so anything discovered in the course of moderation which appears to be an indecent photograph of a child needs to be reported and properly investigated.

A Memorandum of Understanding (MOU) concerning the defence to “making” an indecent photograph of a child is provided by s 46 of the Sexual Offences Act 2003. The MOU addresses the handling of illegal images by a range of professionals included those involved in IT. The key points covered by the Memorandum are:

- It is an identified role
- The speed at which the illegal image is reported and any delay was reasonable
- The handling and storage was appropriate and secure

Further information on the MOU is available on the CPS website, www.cps.gov.uk and the IWF website www.iwf.org.uk.

Section 8 of the Act refers to corresponding legislation made under the Northern Ireland Act 1974.

Protection from Harassment Act 1997

The Protection from Harassment Act 1997 extends to any form of persistent conduct which causes another alarm or distress.

Section 4 of the Act makes it a criminal offence for a person to pursue a course of conduct which he knows, or ought to know, will cause another to fear violence. This offence is intended to catch the most serious cases where behaviour is so threatening that victims fear for their safety. It carries a penalty of a maximum of five years' imprisonment and/or an unlimited fine.

Section 2 of the Act provides for a further offence in cases of a course of conduct which the perpetrator knows, or ought to know, will cause harassment. This offence is intended to catch the sort of persistent conduct which, although it may not make the victim fear that violence will be used, nonetheless can have devastating effects. It carries a penalty of a maximum of 6 months' imprisonment and/or a level five fine.

A court sentencing someone convicted of an offence under either of these sections may also impose a restraining order prohibiting specified forms of behaviour. Breach of a restraining order is a criminal offence punishable by up to five years' imprisonment.

In addition to these criminal offences, section 3 of the Act provides a civil remedy which enables a victim to seek an injunction against a person who is harassing them or may be likely to do so.

Section 13 refers to corresponding legislation made under the Northern Ireland Act 1974.

Public Order Act 1986

Section 5 makes it an offence to, with the intent to cause harassment,
alarm and distress, use threatening, abusive or insulting words, behaviour, writing, signs or other visual representation within the sight or hearing of a person likely to be caused harassment, alarm or distress. This offence may apply where a mobile phone is used as a camera or video rather than where speech writing or images are transmitted.

Sections 38, 41 and 43 extend to Northern Ireland.

Sexual Offences Act 2003
Sexual Offences (Northern Ireland) Order 2008

The Sexual Offences Act 2003 contains a number of offences which could capture online activity aimed at exploiting children.

Section 8: Causing or inciting a child under 13 to engage in sexual activity
Sexual Offences (NI) Order equivalent: Article 15

Section 8 makes it an offence for a person intentionally to cause or incite a child (B) under the age of 13 to engage in sexual activity. Similar to Section 10 (below) this offence captures a wide range of sexual activity and including the situation where a person incites (encourages) the child to take part in the sexual activity, even if the activity itself does not take place.

The offence has a maximum penalty of life imprisonment.

Section 10: Causing or inciting a child (under 16) to engage in sexual activity
Sexual Offences (NI)

Section 10 makes it an offence for a person to cause or incite a child to engage in sexual activity. This encapsulates all sorts of sexual behaviour, including when a person is seeking to get a child to perform a sex act on itself. For example, if A asks B (a child) to touch herself or to pose in her underwear before a webcam it is quite possible that a jury may consider this to be a sexual act.

The offence has a maximum penalty of 14 years' imprisonment
Section 12: Causing a child to watch a sexual act
Sexual Offences(NI) Order equivalent: Article 19
Section 12 makes it an offence for a person aged 18 or over to intentionally cause a child aged under 16, for the purposes of his own sexual gratification, to watch a third person engaging in sexual activity, or to look at an image of a person engaging in a sexual act. The act can be live or recorded, and there is no need for the child to be in close physical proximity to the sexual act. Examples of this offence would be where a person, for the purposes of his own sexual gratification, enables a child to watch two people have sex, either in the physical presence of the activity or remotely, for instance via a webcam; or where someone invites a child to watch a pornographic film.

The offence has a maximum penalty of 10 years' imprisonment.
Section 15: Meeting a child following sexual "grooming"
Sexual Offences(NI) Order equivalent: Article 22

Section 15 makes it an offence for a person aged 18 or over to meet intentionally, or to travel with the intention of meeting, a child under the age of 16 in any part of the world, if he has met or communicated with that child on at least two prior occasions, and intends to commit a "relevant offence" against that child either at the time of the meeting or on a subsequent occasion.

This offence is intended to protect children from adults who communicate (not restricted to on-line communications) with them and then arrange to meet them with the intention of committing a sexual offence against them, either at that meeting or subsequently.

An offence is not committed if the adult reasonably believes the child to be 16 or over. In cases where the defendant claims to have reasonably believed that the child was 16 or over, it is for the prosecution to prove that he held no such belief or that his belief was not reasonably held.

The offence has a maximum penalty of ten years imprisonment.

Section 14 Arranging and facilitating a child (under 16) sex offence
Sexual Offences(NI) Order equivalent: Article 20

The purpose of this offence is to prevent people from making it possible for a child under 16 to be sexually abused. A person must intentionally arrange or facilitate for himself or another something that he intends or believes would happen that would result in a commission of a child sex offence in any part of the world.

The offence carries a maximum sentence of 14 years on indictment

Risk of sexual harm orders (RSHOs)

Sections 123 to 129 of the Sexual Offences Act 2003, which extend to Northern Ireland, provide for a civil preventative order, the risk of sexual harm order (RSHO). This is a civil order that can be applied for by the police against any person thought to pose a sexual risk to children aged under 16. The orders originally arose out of the work of the Home Office Task Force on Child Protection on the Internet which identified a gap in the law concerning the "grooming" of children by paedophiles.

The RSHO should not be used as a substitute for prosecution. The requirement that an order is necessary to prevent serious harm means that those with a genuine and benevolent interest in children (such as those providing advice on sexual health matters) should not be caught by the legislation.

A person subject to a RSHO will not be subject to the notification requirements in Part 2 of the Sexual Offences Act but breach of a RSHO will be a criminal offence and will entail compliance with the notification requirements.

Suicide Act 1961 (which extends to Northern Ireland)

Encouraging or assisting suicide (England and Wales) - under section 2(1) of the Suicide Act 1961 (as amended by section 59 of the Coroners and Justice Act 2009) it is an offence to carry out an act capable of encouraging or assisting the suicide or attempted suicide of another person with the intention to so encourage or assist. The law applies to online actions in exactly the same way as it does offline. It applies whether or not the defendant knows or has identified the person encouraged or assisted and whether or not a suicide takes place.

The maximum penalty for an offence under section 2(1) is 14 years' imprisonment.

Other Acts which may be relevant include: the Crime and Disorder Act 1998 (currently under review, with S S36(1), 2(a) and (d), 6(b) and 118 extending to Northern Ireland), the Criminal Justice Act 1988; the Fraud Act 2006 (of which S1 to 9, 11 to 13 extend to NI); the Offences against the Person Act 1861 (which extends to England, Wales and Ireland); and the Theft Act 1968.

Appendix B: Sources of further advice and information

Organisations

Advertising Standards Authority – See www.asa.org.uk

Children's Workforce Development Council (CDWC) – for guidance and online training on safer recruitment. See: www.cwdcouncil.org.uk

Crown Prosecution Service – for further information on the MOU Sec 46 Sexual Offences Act 2003. See: www.cps.gov.uk

European Commission – Information Society – *Safer Social Networking Principles for the EU.* 2008. See ec.europa.eu/information

Home Office – for information on RIPA 2000. See: www.homeoffice.gov.uk
See also www.legislation.gov.uk

Independent Safeguarding Authority – for information on the current status of the Vetting and Barring scheme. See: www.direct.gov.uk/vetting

Information Commissioner's Office – for information on data protection legal requirements for organisations and also advice for children and young people on keeping personal information personal on social networking services. See: www.ico.gov.uk

UKCCIS – for information on the work of UKCCIS and the Home Office Task Force on Child Protection on the Internet set of guidance. See: www.dcsf.gov.uk/UKCCIS/

Support organisations

Beat
Beat is the leading UK charity for people with eating disorders and their families - www.b-eat.co.uk

Beat bullying
Beat bullying works with children and young people across the UK. www.beatbullying.org. Beat bullying also runs the 'cybermentors' programme. www.cybermentors.org.uk

Child Exploitation and Online Protection Centre (CEOP)
CEOP has the legal remit and authority for tackling child sexual exploitation within the UK including the online environment, as well as dealing with its offline consequences. www.ceop.police.uk.

ChildLine (0800 111)

ChildLine is the UK's free, 24-hour helpline for children in distress or danger. Trained volunteer counsellors comfort, advise and protect children and young people who may feel they have nowhere else to turn. www.childline.org.uk

Internet Watch Foundation (IWF)
The Internet Watch Foundation is the UK 'Hotline' for the public to report incidences of illegal content online. Its remit covers child sexual abuse images hosted anywhere in the world; criminally obscene adult content hosted in the UK; and incitement to racial hatred content hosted in the UK. For more information or to report content see www.iwf.org.uk

NSPCC Child Protection Helpline (0808 800 5000)
The National Society for the Prevention of Cruelty to Children's (NSPCC's) purpose is to end cruelty to children. The NSPCC has 177 community-based projects and runs the Child Protection Helpline and ChildLine in the UK and the Channel Islands.

The NSPCC Child Protection Helpline is the only free and anonymous way for the public to take action to protect a child.

The NSPCC Helpline also incorporates the following other methods which enable it to reach as many adults as possible:

- Asian Language Helpline – direct: 0800 096 7719;
- email: Helpline@nspcc.org.uk; and
- textphone service for deaf and hearing-impaired callers – direct: 0808 100 1033.

Papyrus
Papyrus is a UK charity committed to suicide prevention, focussing predominately on the emotional well-being of children and young adults
www.papyrus-uk.org

Samaritans
Samaritans is a registered charity, founded in 1953, which offers confidential, non-judgmental emotional support, 24 hours a day, to anyone experiencing feelings of distress or despair, including those which could lead to suicide.

The Samaritans emotional support service is available 24 a day, 365 days a year by telephone on 08457 90 90 90 (1850 60 90 90 in Republic of Ireland) and email at jo@samaritans.org. For more information, visit www.samaritans.org.

Internet safety advice and educational resources

Childnet International
Childnet is a non-profit organisation working with others around the world 'to help the Internet a great and safe place for children'. Childnet provides a range of online safety resources including the Know It All suite of educational resources designed to help educate parents, teachers and young people about safe and responsible use. www.childnet-int.org

Teachtoday
A consortium of service providers in partnership with teachers unions have provided an industry-funded online resource to help teachers learn about new technology and how to teach children to be safe online. www.teachtoday.eu

Thinkuknow
Education and awareness resources provided by CEOP and aimed at children and young people. www.thinkuknow.co.uk.

Appendix C: Contributors to the original 2005 version of the guidance

Main contributors

Chair: Annie Mullins – Global Content Standards Manager, Vodafone

Chris Atkinson – Policy Adviser, NSPCC

John Carr – Internet Adviser, Children's Charities Coalition on Internet Safety/NCH

Julian Coles – Senior Editorial Policy Adviser, BBC Editorial Policy

Sam Devoy Prior – Moderatorsnet

Brian Donnelly – Child Protection Consultant

Ashley Farrell – AOL

Will Gardner – Research and Policy Manager, Childnet International

Phil Hall – Emint

Richard King – Product Manager, Wanadoo UK plc

Tamara Littleton – CEO eModeration Limited

Jasmine Malik – CEO Tempero Limited

Robert Marcus – Director, Chat Moderators

Ewan Macleod – CEO, Neo-one

David Ware – Home Office

Samantha Yorke – Legal Counsel, MSN, Microsoft Corporation

Other contributors

Malcolm Hutty – LINX

Hamish McLeod – Mobile Broadband Group