Home Office Task Force on Child Protection on the Internet

Good practice guidance for the moderation of interactive services for children

November 2005

CONTENTS

> **If you have any comments or queries about this guidance please contact:**
>
> Criminal Law Policy Unit
> Home Office
> 2nd Floor, Fry Building
> 2 Marsham Street
> London, SW1P 4DF
>
> Tel: 020 7035 6981

Foreword by Paul Goggins MP, Chair
Taskforce on Child Protection on the Internet.

I am delighted to launch this good practice guidance for moderation of interactive services for children on behalf of the Task Force on Child Protection on the Internet.

The Task Force was established in 2001 and brings together representatives of the internet industry, mobile phone companies, law enforcement, the children's charities and others, who work together to make the Internet a safer place for children, without diminishing their enjoyment of the exciting opportunities which it offers. Building on this partnership, the Task Force has run several successful education and awareness campaigns. It also assisted the Government in preparing a new offence of meeting a child following sexual grooming, which was introduced in the Sexual Offences Act 2003. This has helped to tackle public concerns about misuse of the Internet by paedophiles. Also in 2003, the Task Force published models of good practice and guidance for the internet industry to consider when providing chat services, instant messaging and web based services, and is working on a kitemark standard for rating, filtering and monitoring software.

This guidance is another important step towards improving child protection standards online. It is intended for use by providers of public interactive communication services aimed at or likely to attract children. This is a rapidly developing area of online activity where children may be at risk from undesirable contact or behaviour and moderation plays a crucial role in making children safer.

The guidance offers advice to service providers for the first time on assessing potential risks to children, deciding whether moderation is necessary and, if so, what kind of moderation. It also sets out good practice in respect of the recruitment, selection and training of moderators and covers the relevant areas of the criminal law.

The document is intended to be of practical help and I believe it is an invaluable guide for any organisation already providing or considering whether to provide interactive services geared toward children. It will also be of interest to parents and carers who wish to know more about moderated services. It represents a substantial achievement on the part of those who have contributed their time and expertise to bring together the best available advice on this subject.

I strongly recommend the guidance and urge providers of interactive services to consider how the recommendations in this document can be applied to their services.

Paul Goggins

Good Practice Guidance for Moderation of Interactive Services for Children: Executive Summary

This guidance has been produced in response to public concern about the safety of children using interactive communication services, such as the Internet. While these services offer huge opportunities for children to communicate and learn, experience has shown that there are some individuals who will use them to contact children in order to "groom" and abuse them. 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 allows a person, or technical filter, to review content posted by users. This document gives a framework to help providers of 'virtual public space' offer a safer environment for children.

The guidance applies to **public electronic interactive communication services** through which individuals can make contact and exchange personal information with other users in a virtual public "space" such as, but not limited to:
- internet chat rooms, message boards, mobile chat services, TV 'text to screen' services, on-line games with chat or messaging facilities, and mobile games with chat facilities.

Purpose of the guidance
The purpose of the guidance is:
- to describe the types of moderation that can be used;
- to inform organisations of all the issues they should take into account when assessing the need for moderation of interactive services, and
- to inform organisations of the issues they should take into account in the recruitment, training and supervision of moderators.

Basic requirements
If you or your organisation are providing, or intend to provide, a public, interactive communication service that is aimed at or likely to attract children, you should:

- assess the potential risk to children, establish if it would be appropriate to use moderation and, if so, decide the form of moderation to use;

- if using human moderation, assess the risk that a child abuser may apply for a position and develop policies for the safer recruitment, training, management and supervision of moderators to safeguard against this, and

- make clear to users whether the interactive service is moderated, and if so, by what means, either human or technical moderation.

The guidance

- Provides information and recommendations for the moderation of public interactive communication services aimed or very likely to attract children in the following areas:

  - information and advice to users;
  - risk assessment;
  - recruitment;
  - training;
  - data security;
  - management and supervision, and
  - escalation procedures

- Gives examples of methods and patterns of behaviour ("grooming") used by child abusers to gain access to children via interactive communication services.

- Considers the relative merits of technical vs. human moderation and concludes that technical moderation has not yet demonstrated the same level of protection as human moderation.

- Provides information about relevant legislation.

BACKGROUND

1.1 Introduction

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 with others, such as online, instant messaging and interactive games.

As technology advances, we are beginning to experience media convergence as the Internet and a range of content services can be accessed through different devices such as personal computers, laptops, mobile phones, game consoles and the TV. This makes it difficult for parents to supervise and monitor their child's use of communication technology, particularly as children's take up of the latest communication technologies will often exceed that of their parents.

Many companies and organisations recognise the challenges faced by parents and are supporting them by providing safety tools and resources such as filtering software and guidance about keeping safe online.

In 2003 the Home Office published Good Practice Guidance for the providers of Chat Services, Instant Messaging and Web Based Services, which are aimed at or very likely to attract children. The guidance, when discussing interactive services such as chatrooms, Instant Messaging, Bulletin Boards and Discussion Forums made reference to the fact that some of these services are "moderated" and some are not, see (http://police.homeoffice.gov.uk/operational-policing/crime-disorder/child-protection-taskforce)

Moderation is a process by which the provider of an interactive service e.g. the chat room provider, or the person or company hosting a bulletin board, or interactive game, takes steps to eliminate undesirable or illegal contact or behaviour. The Guide to Web Based Services was supportive of the use of moderation in services for children and young people. It is also clear, however, that moderation is not an absolute guarantee of a child's safety while using a service.

1.2 Aim and scope of the Guidance

Experience has shown that there are individuals who will exploit interactive services to gain contact with children in order to "groom" and abuse them. It is, therefore, important that public interactive services, which are aimed at or likely to attract children address the safety of children using these services.

If you or your organisation are providing, or going to provide, a public, interactive electronic communication service that is aimed at or likely to attract children, you need to do the following:

- assess the possible risks to children and whether it would be appropriate to use moderation and, if so, decide what type of moderation to use;

- consider, if using human moderation, policies for the recruitment, training and supervision of moderators and for the management of users' personal data, and
- make clear to users whether your interactive service is moderated, and if so, whether by human or technical moderation.

1.3   Using the Guidance

The guidance provides details of all the factors that should be considered in determining whether moderation is appropriate and, if so, what it should comprise. It provides a framework, based on current best practice, rather than an absolute model to be followed rigidly irrespective of the circumstances. Interactive services are provided in various and constantly evolving forms and are aimed at different communities.   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 them to judge whether and how far to apply any specific point in the guidance. Where they choose not to, they will obviously want to assure themselves that their decision is justified, given the nature of the service, and that it is consistent with child safety needs.

1.4   Types of interactive services

This document refers frequently to interactive services. It is aimed at **public electronic communication services** through which individuals can make contact and exchange personal information with other users in a virtual public "space" such as but not limited to:

- message boards;
- chat services;
- text to screen;
- TV chat services;
- interactive games with chat or messaging facilities, and
- chat or game services that use location services as part of the location and communication facilities [1].

1.5   Blogging

It is noted that **blogging, mo-blogging and "pseudo blogging"** is still a relatively new phenomenon, elements of which have many of the characteristics of discussion forums and other kinds of spaces where interactive user generated content is created.

Until clearer patterns emerge, it is not possible for us to offer any definitive guidance about *"blogging"* but this may be something to be considered in the future.

---

[1] See Mobile Location Guidance website: http://www.mobilebroadbandgroup.com/social.htm

2.   RISKS TO CHILDREN

2.1 Children's vulnerability

Children and young people **are all vulnerable** due to their young age and inexperience. Many children who have been abused by people they have met on the Internet have come from very loving, stable and positive backgrounds. It is children's lack of experience and trusting nature that is exploited by abusers, especially when children are testing out their identities and being given more freedom by their parents.

Children and young people who are experiencing difficulties at home, school or with friends and those who have disabilities may be particularly vulnerable, and they may seek positive contact, support and friendship through interactive services online. Indeed many children do find new friends on the Internet and sources of support which are positive and offer less complicated relationships than those in their ordinary day-to-day lives.

Experience in the Internet world has shown that some users, including children, behave in more inappropriate and, at times, extreme ways online than they would offline. Some children will engage in behaviour that may place them at risk, such as giving out personal information about themselves and their family; some will act out fantasies by pretending to be someone else; some will test out their sexual identities by engaging in cyber - flirting and communicating in a sexual way, and generally acting out behaviours they would not in real life. In some instances, children and young people may themselves bully, harass and abuse others.

2.2   "Grooming" of children for sexual abuse

Online environments have proved to be very attractive to child sex abusers who have exploited them to contact, "groom" and abuse children. *"Grooming"* is a process by which a child abuser seeks to prepare a child for later abuse. Many child abusers use public interactive spaces to find and meet children. Abusers use a range of techniques to make contact and befriend children. This can include the following examples:

- suggestions that a child leave a public chat room and move to private one-to-one communication such as Instant Messaging;
- asking for personal details: name; address; telephone number; mobile number; name of school or a photograph;
- asking where the home computer is located and/or about parental supervision of Internet use;
- offering the opportunities of modelling;
- meetings with pop idols or celebrities;
- offering cheap tickets to sports or pop concerts;
- offering material gifts including electronic gifts such as games, music or software;
- quick and easy ways to make money;
- seeming eager to meet up offline;

- encouraging a child to share or talk about any difficulties they may be experiencing, such as bullying or difficult relationships, and offering 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 behaviour in the interactive service, and
- telling the child that they know how to locate them, where they live, or where they go to school.

Other"grooming" behaviours may be difficult to identify, since communication that forms part of "grooming" may appear perfectly ordinary and innocent.

For further information on this topic refer to the University of Lancashire Cyberspace Research Centre[2] and Childnet International's Chatdanger[3] provides real examples of the '"grooming"' of children.

2.3   The "Grooming" process

Abusers go out of their way to entrap children, either quickly or over a long period of time. They use sophisticated methods which both gain a child's trust and lure them into a world of secrecy. This process isolates children from sources of support such as friends, family or parents. It is often achieved by sending children adult pornography or getting children to send images of themselves, to lower their sexual inhibitions and make the child feel guilty and ashamed. In some cases, the communication towards the child may involve no explicit sexual content. It is aimed at simply gaining the child's trust and confidence. These entrapment processes ultimately have a corrupting effect on children. The end result is that children can, and often do, feel ashamed, guilty and responsible for the communication and abuse that has taken place. Children find it extremely difficult to seek help or tell anyone what is happening to them.

Abusers can be adept at speaking the same language as children and become familiar with their popular culture, hobbies, and interests such as sport, music, celebrities, pop idols, or interactive games. Some abusers will watch a child in a public forum to gather information about their interests over a period of time, so that they can more easily manipulate a conversation with them at a later stage.

Having made contact with one child, abusers may use that child to gain contact with other children who are likely to be using the same online or messaging product, such as instant messaging.

---

[2] "*Cyber Stalking, Abusive Cyber Sex and Online Grooming",* by Rachel O'Connell, Joanna Price and Charlotte Barrow, and "(A Typology Of Child Cybersexploitation And Online Grooming Practices." by Rachel O'Connell).
 http://www.uclan.ac.uk/host/cru/docs/NewCyberStalking.pdf

[3] website www.chatdanger.com

2.4   "Grooming" as a criminal offence

Whilst the "grooming" offence contained in s.15 of the *Sexual Offences Act 2003* can only take place when an adult intends to meet a child with the intention of committing a sexual offence against them, either then or subsequently, other crimes can occur when a person "grooms" a child. For example, s.10 of the Act makes it an offence to cause or incite a child to engage in a sexual activity. This could occur where, for example, a person asks the child to touch themselves or to photograph themselves (the latter may also be an offence under the *Protection of Children Act 1978*). Where a pornographic picture is sent to a child this could also be an offence under s.12, *Sexual Offences Act 2003*. It is important that all moderators are aware of what the law is but they should also be aware that they should not try to investigate incidents themselves as this could prejudice a criminal investigation. Moderators should secure evidence where possible (e.g. transcripts, IP addresses, etc.) and contact the police immediately (See Annex A).

2.5   The Importance of Moderation

In the light of the risks to children described above, moderation is used to try to keep chat and public interactive services safe for children and to provide positive user experience.

Moderation can be used to protect and educate children about safety and how to use interactive services responsibly. It may also help communities take responsibility for running their own virtual space. Communities themselves can give their members the experience of learning how to take responsibility for the conduct of the group as a whole, by agreeing a code of acceptable behaviour and then applying it. Where appropriate, at the right age, and with the right safeguards, children's groups may also learn the same skills in the same way.

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

Although this document is intended for public interactive communication services aimed at or very 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.

Public interactive communication service providers will need to assess for themselves the levels of risk to users of their service and how best to provide an appropriate level of protection for their customers.

3. ASSESSING THE NEED FOR MODERATION

3.1 *What is moderation?*

Moderation is an activity or process whereby a person or technical filter is responsible for reviewing content posted by users. Moderation is usually undertaken according to an agreed set of guidelines or policies to try to ensure users of the service are able to interact safely, responsibly and appropriately. These may be documented, for example, in the service's terms and conditions or "House Rules".

3.2 *Technical or human moderation?*

There are different ways in which the moderation function can be carried out: either entirely by humans, or entirely by software based systems, or by a combination of both.

Technical moderation attempts to filter words and phrases that it has been programmed to identify, telephone and email address formats, profanities and explicit language that may cause offence. However, technical moderation has not yet demonstrated that it can offer the same level of online child protection as human moderation, when used on its own, to combat the sexual "grooming" of children. For example, it can be outwitted by the creative use of combinations of numbers, letters and punctuation marks. Furthermore, software based solutions find it very difficult to pick up and interpret the context within personal communication, for example, the subtleties of "grooming" behaviour.

However, technical interactive solutions which limit a participant's communication to a choice of pre-scripted words and phrases have proved to be effective in significantly minimising risk.

3.3 *Methods of human moderation*

Moderation can be provided in a number of ways, and a combination of these can be applied to interactive services. These are as follows:

- **pre-moderation**: in a pre-moderated service all material supplied by users will be reviewed by the moderator for suitability **before** it becomes visible to other users;

- **post-moderation**: in a post-moderated service, all material supplied by users will be reviewed **after** it becomes visible to other users. The length of time between the material becoming visible and it being checked may vary;

- **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 request** for intervention is made.

3.4   Undertaking a risk assessment of interactive services

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

The following points are key areas for consideration:

- whether the service is specifically targeted at children and younger users;
- whether the service is very 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;
- 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.

Having undertaken the risk assessment, it is necessary to decide which form of moderation or combination of forms is appropriate.

3.5   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 industry wide agreed definition of a moderator. The role a moderator undertakes will depend on the kind of service offered.

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. 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.

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 both roles of host and moderator.

3.6   Human 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.

In all these cases, moderators may work from home or from an office.

4. INFORMATION AND ADVICE FOR USERS OF MODERATED SERVICES

4.1 It is important for users to be able to identify moderated services. Clear and accessible information should be available to users about what they can expect from the service offered, for example, that the service is moderated, the method of moderation used (human or technical) and how it works.

It is also important for users to be able to contact a moderator for assistance should the need arise. (see recommendations below).

4.2 Recommendations

- Providers of public interactive communication services should provide clear and prominent information to users about the kind of service offered, for example is the chat room moderated or un-moderated? If moderated, what form of moderation is used i.e. technical or human moderation and how does it work?

- Parents need to be advised by interactive communication providers of the importance of communicating with their children about their safety online on a regular and consistent basis, as moderation or other safeguards are not foolproof.

- Children need to be regularly updated by interactive providers of the potential risks to them. This should include what additional measures they can take, for example, tools to block communication or record communication dialogue etc.

- Users of a moderated interactive service should have some means of contacting the moderator for assistance should a concern or difficulty arise while using the service.

5.   RECRUITMENT AND SELECTION OF HUMAN MODERATORS

5.1   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 content and potential contact with sex abusers.

A considerable body of knowledge and evidence now exists. There is general agreement that organisations need to make their professional environments safer for children.

5.2   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 public interactive communication 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 of the moderator, and
- access to personal information about children.

Moderators of sites designed for children and sites which are aimed at or very likely to attract children should be subject to an appropriate Criminal Records Bureau check, be recruited through suitable procedures, and be fully trained and supervised.

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.

5.3   Legislative Developments

There have been considerable legislative developments in recent years to improve the protection of children from harm abuse by those in positions of trust and authority. These include the introduction of new abuse of trust offences in *the Sexual Offences Act 2003*. In 2002 an amendment was made to the *Rehabilitation of Offenders Act (Exceptions) Order* allowing standard disclosures to be made through the *Criminal Records Bureau* on persons in "employment which is concerned with the monitoring, for the purpose of child safety, of communications by means of the internet." [4]

NB. If moderators from outside England & Wales are to be employed 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. For further information contact the CRB Information Line on 0870 90 90 811 or visit their website on www.crb.gov.uk .

---

[4] we understand that further consideration will be given to the range and level of checks which should be available by which changes are made.

5.4   Recruitment and selection for moderators outside the UK

It is accepted that communication technologies operate on a global basis and a number of companies run their operations across a number of different territories. Vetting procedures and standards in other jurisdictions may not match those that exist in the UK. For organisations providing public interactive communication services in these circumstances, it is recommended that, in whatever country they recruit or employ moderators, they use whatever systems are available to them that match the UK's recruitment standards, as closely as possible. Extra care should be taken in following up references and carrying out other checks on a person's background.

The Criminal Records Bureau is not 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 the criminal record as a starting point. However, what constitutes a criminal offence will depend upon the legal framework in each 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. Some countries also operate a *"disqualification from working with children"* system.

The Warner recommendations are UK based and safe recruitment procedures will also vary from country to country.

The Warner Report

Defining the job – Employers should only recruit staff after preparing a job description and a person specification clearly setting the competences (ie skills and personal attributes) and experience required to discharge satisfactorily the responsibilities of the job description.

Advertising - Employers should ensure that all vacancies are advertised usually externally and are open to competition (e.g. by use of websites or the press.)

Use of Public Agencies – Employers should require any agencies used to adopt selection and appointment procedures as rigorous as those for directly employed staff.

Pre-selection information - Employers should require applicants for posts to supply information prior to selection, in a signed document, including: proof of identity; any criminal convictions and whether they have ever been charged with a criminal offence and the outcome; any other relevant information.

Selection Methodology - Employers should use a variety of selection methods as an appointment based upon just one interview by a large panel is more likely to result in a wrong appointment. Selection methods could include written exercises, preliminary interviews, visits to the office, and aptitude tests.

Employers should use preliminary interviews as a standard part of a establishing a fuller picture of the character and attitudes of short-listed candidates.

References – References especially from the current employer, provide key information about the job applicant and employers need to ensure that their procedures enable this to be supplied .Employers should require candidates when applying to provide a full employment history, including periods of unemployment with dates (to the nearest month) and the names and addresses of previous employers.

 Employers should always approach an applicant's present employer; should tell applicants that they reserve the right to approach any previous employer or line manager about a short listed candidate's character and performance before interview; should seek written references on the basis that referees have the job description and person specification and are encouraged to comment frankly on short listed candidate's strengths and weakness in relation to those two documents; and where necessary should explore any aspects of references by telephone with a current or past employer.

Final Decisions - Employers should ensure that those taking a final decision on employment of an employee should have available to them and use all the information about candidates from earlier parts of the selection process; and that they are free to explore areas of doubt and concern to discharge their over riding responsibility to make safe and competent appointments.

Employers should only offer appointments after completing police checks against central government lists and verification of birth certificates and educational / professional qualifications; and **should allow no unsupervised access to children before completion of all checks**

(References ; Warner Report – "*Choosing with Care"* ; Utting Report – "*People like Us")*

5.5   Recommendations for recruitment and selection

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

- the appropriate CRB *(Enhanced Disclosure if available)* check should be undertaken prior to an appointment to a moderation position involving working with children. (It is recognised that delays at the CRB in processing applications, for Enhanced Disclosure prior to appointment, can have significant public and practical implications for the safer provision of services to children.);

- advertisements for such positions should state that a CRB check will be made;

- efforts should be made to adopt safer recruitment and selection procedures based on the *Warner Report – "Choosing with Care".(See the* summary of recommendations which may be relevant for those responsible for recruiting moderators to consider).;

- all prospective employees for moderation positions involving contact with children should be interviewed face-to-face;

- advice on child care recruitment to reflect the importance of safeguarding children (such as the ability to address the applicant's attitude and suitability to work with children) should be sought, for example, from the major children's organisations.

- Contracts, terms of employment or codes of conduct should include:

    - boundaries of personal conduct;
    - prohibition of inappropriate behaviour with children or vulnerable adults who are users of the service, including making arrangements to have personal communication or contact with them;
    - prohibition of use of their moderator screen name outside of employment, and
    - a confidentiality clause, prohibiting the misuse of company information e.g. passing personal information to third parties.

- If moderators move from services not aimed at children to those which are aimed at or very likely to attract children, they should go through the selection processes appropriate for moderators working with children.

- It is accepted that communication technologies operate on a global basis and a number of companies run their operations across a number of territories. Vetting procedures and standards in other jurisdictions may not match those that exist in the UK. For companies providing interactive services in these circumstances, it is recommended that, in whatever

country from which they recruit or accept moderators, they use whatever systems they have available to them to match the UK's recruitment standards as closely as possible.

6. TRAINING OF MODERATORS

6.1 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.

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 possible for the moderator in the particular environment.

Recommendations for training moderators:
Based on your risk assessment some or all of the following may be relevant:

6.2 Role of the moderator

Training should cover:

- 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.

6.3 Escalation procedures

Training should cover:
- the use of an agreed escalation procedure (see 8.4), and
- how, when and why they should refer particular types of incident and to whom a report should be made.

6.4 Recognition of responses:

All moderators should have a reasonable level of awareness of child protection issues, set out below. The depth to which this is necessary will vary with the level of service being provided (a person post-moderating a message board will need a different mix of knowledge from a person moderating a teens' chat room).

6.5 Usage patterns/ behaviour worth investigating further

Experience has demonstrated that there are some behaviour patterns which, while not immediately obvious as signs of abuse, may be worth further investigation (see 2.2).

Case studies can be useful to demonstrate both the ease with which contact can be made and the severity of the harm that can result.

6.6   Child abuse

Training should be provided on:
- the full range of behaviour that constitutes child abuse. The issue of "grooming" should be specifically addressed, including the signs that may warrant intervention (e.g. invitations to meet off-line or requests for personal details (see 1.5 Page 6). Moderators and supervisors should be aware of the relevant law **(see Annex A)**

- 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.

6.7   Vulnerable people

Training should be provided to enable moderators to recognise and 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.

6.8   Bullying/harassment

Training should be provided to enable moderators to recognise behaviour which constitutes bullying and harassment, which in some cases might amount to criminal harassment under the Protection from Harassment Act 1997– (**see ANNEX A**):

6.9   Illegal or Harmful Content

Training should be provided so moderators can be aware of material that is potentially illegal (such as indecent photographs of children) and material that is not appropriate for children, for example, adult pornography.

6.10   Awareness of the risk to children

Training of moderators should seek to raise awareness of the serious risk of harm to children posed by child abusers through the use of interactive services.

6.11   Identifying abusive behaviour

Training needs to address the ability of moderators to identify behaviour which constitutes child abuse and which is possible on the service they are moderating, for example, encouraging a child to share inappropriate images.

6.12 Advising users

Where moderators have direct interactions with users, their training should include the most recent advice on staying safe online so they are able to promote safe practice amongst users.

6.13 Updating the moderation

Training should include advice on what to do if the moderator feels that the procedures and practices they follow seem no longer to be appropriate for the user group or the service that is being delivered, especially where large numbers of children are found to be using a service not designed for them.

6.14 Concerns arising during training

If the response to training gives a company cause for concern about a particular trainee, the company should be in a position to limit the services that person will moderate to those which exclude children, impose particular supervision, or reconsider their employment.

6.15 Prohibited Activities

Training should address the activities that are prohibited to moderators, for example, unauthorised communication or meeting with service users, together with the reasons for such prohibitions.

6.16 Escalation Procedures

Moderators should be trained in the use of an agreed escalation procedure.

Whether moderation is done internally (e.g. by a content provider) or externally (e.g. by a moderation company), it should be clear who has responsibility for reporting an incident to the relevant authorities, at what stage and in what form. This may involve the police e.g. Virtual Global Task Force (www.virtualglobaltaskforce.com) or the IWF (www.iwf.org.uk).

7.   PERSONAL INFORMATION AND DATA SECURITY.

7.1   Risk management in providing a moderation service is necessary because of the risk of information being misused to contact or maintain contact with a child outside the moderation environment.

If the actual physical environment, including location of computers or storage of data is insecure e.g. in an open access office, there is a risk that unauthorised individuals may gain access to users' personal details, including email, telephone number and address.

If data systems are vulnerable to hacking, or operated by people outside the control of the service operator, there is the potential that the security of users' personal data could be at risk.

Providers must 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.

7.2   Recommendations for data security

- all organisations that collect personal information will need to comply with the Data Protection Act (1998);

- in accordance with the DPA (1998) organisations holding personal data will need to appoint a Data Controller;

- access to personal data about users should be restricted to those authorised by the data controller;

- a record should be kept of who has access to personal data and, where technically practical, when they have accessed it. If IP addresses are recorded, a date and time stamp should be included. Interactive service providers who have employed a moderation company will ordinarily be the data owner (rather than the moderation company) unless agreed otherwise;

- policies and procedures should be devised to address data security in work places. This could include such things as prevention of unauthorised access to systems and physical locations where data is stored and processed, and

- policies and procedures should be devised for situations where moderators work from home, as it is difficult to manage users' personal data in these situations. Providers of moderation services will need to be especially vigilant to comply with the Data Protection Act 1998.

8. MANAGEMENT, SUPERVISION AND ACCOUNTABILITY OF MODERATORS

8.1 Management, Supervision and Accountability

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 for the protection of children.

Professional standards about management, supervision and accountability have evolved, particularly in the last 15 years, following many child abuse inquiries into the systematic abuse of children by staff in professional settings. As far as possible, professional childcare standards should be incorporated into the interactive services environment to ensure the safety and protection of children. Because of the crucial role managers play, they need to be fully informed of the child protection issues involved in the operation of interactive services.

Operators should also be aware of the Good Practice Guidance for Providers of Chat Services, Instant Messaging and Web Based Services.

8.2 Referral of an employee to Protection of Children Act 1999 List

In the UK it has been possible for some time now for employers in organisations involved in working with children to refer an employee, who has placed a child at risk, to a list held by the government. This is known as the *'Protection of Children Act* List' (POCA).

Moderators who harm or place a child at risk during the course of their employment should be referred to the POCA List which contains the names of individuals who have been considered unsuitable to work with children – because in previous employment they have either harmed or placed a child at risk, even if no criminal proceedings have taken place.

It is a criminal offence for those on the POCA list to seek employment with children and it is an offence for an employer knowingly to employ them in such a role, **(see ANNEX B for more details).**

8.3 Recommendations for the Management and Supervision of Moderators

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

- all moderated services should have effective policies and management systems in place for moderators;

- procedures should be in place to ensure moderation practice which fosters awareness of child safety and protection;

- managers should be aware of child protection issues and their responsibilities in respect of protection of children;

- managers should ensure moderators are aware of their responsibilities and of policies and procedures;

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

- If moderators are working at home, management and supervision measures should be carefully considered 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 messenger, video conferencing, telephone contact, to ensure:

  a) the moderator is who they should be;

  b) the work station is set up in a way to keep users' data secure, and

  c) the organisation can monitor what the moderator does.

Managers should supervise the work of moderators so that:

- they get an overall view of the quality and consistency of the moderation being provided;

- 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;

- 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 an employee or former employee is a risk to or has harmed a child, consideration should be given to referring the case to the Department for Education and Skills so that their name can be considered for inclusion on the Protection of Children Act (POCA) 1999 list.

8.4   Escalation procedures

Moderators who look after interactive services intended for children, or where children are very likely to go, should know what to do, whom to refer to, when and how, if they see risky behaviour on the service for which they are

responsible. To do this, a clear escalation policy should be agreed in advance, to avoid confusion and delay.

8.5 Recommendations for Escalation Procedures

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

- a clear escalation policy should be established in advance;

- moderations should be familiar with the escalation policy and should know what to do, who to refer to, when and how, if they observe risky behaviour on the service;

- moderators should be familiar with the service's *'house rules'* or terms and conditions or other published code of conduct for users;

- moderators should be particularly familiar with how the service's house rules etc… apply to behaviour which is risky for children, for example, an attempt to publish personal information or to arrange a face-to-face meeting with a stranger;

- moderators should know what to do if a user breaks the house rules etc and what level of sanction to apply if it is their responsibility, or who to refer to;

- moderators should know at what stage an incident should be referred to their supervisor or manager and in what form;

- consideration should be given to how any potential evidence of a crime should be captured and for how long it should be stored, subject to the Data Protection Act 1998;

- whether moderation is done internally (e.g. by a content provider) or externally (e.g. by a moderation company) it should be clear who has responsibility for reporting an incident to the relevant authorities, at what stage and in what form. This may involve the police, the IWF or children's charities;

- the person responsible for reporting an incident to the relevant authorities should know to whom to report the incident and how to do it (see section 6.16), and

- it may be necessary to arrange out-of-hours contact if the moderated service is available outside office hours.

The Criminal Law affecting personal interactions in interactive services.

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 does not only apply to issues such as distributing illegal material, but also, for example, to behaviour which may cause harm because it amounts to a course of harassment. Inciting someone to commit an offence is also no less an offence simply because it is done through a computer or mobile phone. 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 will be helpful. No-one using an interactive service should be under the illusion that the criminal law does not bear on what they do.

Communications Act 2003
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, or persistently makes use of a public electronic communications network.
These provisions allow prosecution of nuisance or silent telephone callers in the online world but could apply to similar behaviour committed on any public communications network. The offences carry a penalty of a maximum of 6 months' imprisonment and/or a level five fine (£5000).

Protection from Harassment Act 1997
The Protection from Harassment Act 1997 was introduced primarily to tackle stalking but the offence of harassment 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 will catch the most serious cases where behaviour is so threatening that victims fear for their safety. It carries a penalty of a maximum of 5 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 will 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 5 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.

Other public order laws designed to deal with online behaviour may also be applicable to behaviour in an electronic interactive service, depending on the circumstances.

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 5 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 concerning the defence to "making" an indecent photograph of a child provided by s 46 of the Sexual Offences Act 2003 is available on the CPS website, www.cps.gov.uk .

Sexual Offences Act 2003

Section 10: Causing or Inciting a child to engage in sexual activity.

Section 10 makes it an offence for a person to cause or incite a child to engage in sexual activity. This is a very wide provision and 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. What amounts to a "sexual" activity will be decided by the court but section 78 of the Act defines "sexual" in such a way that the circumstances and motives of an offender may be relevant.

The offence does not require any element of coercion although it is possible that this may occur in some situations. The offence is committed even where the child apparently consents to performing the act. There is no requirement that the offender receives sexual gratification from the act although this is

likely to occur in the vast majority of cases, simply that the offender intentionally causes of incites the activity.

The offence has a maximum penalty of fourteen years imprisonment

Section 12: Causing a child to watch a sexual act
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 or sends a child indecent images over the internet.

The offence does not require any element of coercion, though it may be a factor in some cases. The offence is committed even where the child apparently consents to watching a sexual act. In order for an offence to be committed, the adult must act for his own sexual gratification. This ensures that, adults showing children sex education material, either in a school or other setting, will not be liable for this offence.
The offence has a maximum penalty of ten years imprisonment.

Section 15: Meeting a child following sexual "grooming"
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.

The section is intended to cover situations where an adult establishes contact with a child and gains the child's trust so that he can arrange to meet the child for the purpose of committing a "relevant offence" against the child (essentially this means sex offences). The contact with the child may take place through communications on the Internet, but equally, it could for example, be through meetings, letters, text messages or telephone conversations. The Police may become aware of the contact between the offender and the child by a number of means, for example, reporting by the child, or by concerned parents/teachers.
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 initial communications between the adult and child may have a sexually explicit content, for example, conversations about sexual acts he would like the child to engage in or sending the child indecent images. However, this

need not be the case. Prior communications could for example, involve an adult giving a child music lessons or running a youth club the child attends, an adult serving sweets to a child in a sweet shop, meeting incidentally through a friend, or chatting about innocent subjects.

It is for prosecutors to prove the intent of the adult to engage in unlawful sexual behaviour with the child on the occasion of the meeting or on a subsequent occasion. Proof could be derived from the communications between the adult and the child before the meeting, for example, from conversations about the nature of the sexual activity that is planned. Such evidence might be obtained by examining the contents of e-mails or letters which have been sent or received, or from the transcripts of chat room conversations which might have been logged either on an individual's computer or on the computer of an internet service provider. Evidence may also be drawn from other circumstances, such as the adult travelling to the meeting with ropes, condoms and lubricants.

The intended "relevant offence" does not have to take place for the offence to be committed. It is sufficient for the adult to travel to meet the child with the intent to commit a "relevant offence" against the child. The adult might intend to commit the "relevant offence" on that occasion, or on a future occasion. An example of the latter would be where a person communicates with the child over the internet, expressing his intention that they engage in sexual activity. He then arranges to meet the child for the first time in a public place, with the intention of meeting her again at a later date in private, at which point he plans to have sex with her. In this example a section 15 offence would have been committed at the point at which the adult sets out for the first meeting.
Either the meeting or at least part of the travel to the meeting must take place in England, Wales or Northern Ireland. However, the adult's previous meetings or communications with the child can have taken place anywhere in the world and it would also be possible for the person to intend to engage in sexual activity with a child in another jurisdiction.

In some cases it might be appropriate to charge a person with an attempt to commit the offence rather than the offence itself. For example, where an undercover policeman takes the place of the child at the meeting in a covert operation, the defendant could be charged with attempting to commit the offence, assuming the necessary intent could be proved. The attempted offence has the same penalty as the offence itself.
The offence has a maximum penalty of ten years imprisonment.

Risk of sexual harm orders (RSHOs)
Sections 123 to 129 of the Sexual Offences Act 2003 provide for a new civil preventative order, the risk of sexual harm order (RSHO). This is a new 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.

A chief officer of police may apply for a "risk of sexual harm order" in respect of a person aged 18 or over if it appears to the chief officer that there is reasonable cause to think it is necessary, and that person has on at least two occasions done one of the acts listed. These are:

   (a) engaging in sexual activity involving a child or in the presence of a child;

   (b) causing or inciting a child to watch a person engaging in sexual activity or to look at a moving or still image that is sexual;

   (c) giving a child anything that relates to sexual activity or contains a reference to such activity, and

   (d) communicating with a child, where any part of the communication is sexual.

It is not necessary for the defendant to have a prior conviction for a sexual offence. The court can make an order if it is satisfied that it is necessary for the purpose of protecting children generally or any individual child from harm from the defendant. The order entitles the court to prohibit the defendant from doing anything described in it. The minimum duration of an order is 2 years. The order is intended as a preventative measure to deter unlawful or harmful sexual activity with, or conduct towards, a child. Breach of an order, without reasonable excuse, is a criminal offence that is tried either summarily or on indictment with a maximum penalty on indictment of five years imprisonment.

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.

The criminal law affecting individuals considered unsuitable to work with children.

Protection of Children Act 1999

All child care organisations (as defined by the Act) have a statutory duty to refer names for possible inclusion to the Protection of Children Act list of those individuals considered unsuitable to work with children. **Any other organisation *may* refer names for possible inclusion on the list.** This applies where a worker in a regulated position has been dismissed, resigned, or moved away from work with children, on grounds of misconduct which harmed a child or placed a child at risk of harm

A child care organisation is defined in the Act as an organisation:

▪ which is concerned with the provision of accommodation, social services or health care services to children or the supervision of children;

▪ whose activities are regulated by or by virtue of any prescribed enactment, and

▪ which fulfils such other conditions as may be prescribed.

Unless they fall into this category through some other role, it is unlikely that companies offering interactive services will be classed as child care organisations. They may nevertheless make referrals to the list in the circumstances set out.

Regulated positions include

▪ a position whose normal duties include caring for, training, supervising or being in sole charge of children, and

▪ a position whose normal duties involve unsupervised contact with children under arrangements made by a responsible person.

Once a referral has been received, the decision as to whether to include a person on the list is made by the Secretary of State. There is an avenue of appeal to the Care Standards Tribunal. In order for the Secretary of State to go on to confirm a person on the POCA list he must form the opinion that:

- the employer reasonably considered the individual to be guilty of misconduct (whether or not in the course of his employment) which harmed a child or placed a child at risk of harm, and

- that the person is now unsuitable to work with children.

Details that should ideally be sent with a referral

- Full name, date of birth of the individual;

- National Insurance number (if available);

- Confirmation that the individual occupied a child care post or "regulated position";

- Full details of the alleged misconduct;

- Detailed explanation about how - by his/her misconduct - the individual harmed a child or placed a child at risk of harm;

- Details of the investigation carried out to date - and their conclusions - including copies of all relevant papers (including statements, notes of interviews, minutes of meetings and minutes/notes of disciplinary hearings) and details of the organisation's disciplinary procedure;

- Details of the action taken against the individual - has he/she been suspended, dismissed or transferred from a child care position etc;

- Information of any police involvement (or the involvement of any other agency);

- Details of proposed further action - i.e. dates for disciplinary hearings, timetables on further investigations etc;

- Any other information considered relevant to the circumstances of the alleged misconduct.

Contact Address:

Children's Safeguarding Operations Unit (POCA),
Ground Floor Area E,
Mowden Hall, Staindrop Road,
Darlington,
DL3 9BG

01325 392 030

Project Team

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 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, Chatmoderators

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