Appropriate Monitoring for Schools

June 2016

Provider Checklist Reponses

Schools in England (and Wales) are required "*to ensure children are safe from terrorist and extremist material when accessing the internet in school, including by establishing appropriate levels of filtering*"[1]. Furthermore, the Department for Education published the revised statutory guidance 'Keeping Children Safe in Education'[2] in May 2016 (and active from 5th September 2016) for schools and colleges in England. Amongst the revisions, schools are obligated to "*ensure appropriate filters and appropriate monitoring systems are in place. Children should not be able to access harmful or inappropriate material from the school or colleges IT system*" however, schools will need to "*be careful that "over blocking" does not lead to unreasonable restrictions as to what children can be taught with regards to online teaching and safeguarding*."

By completing all fields and returning to UK Safer Internet Centre (enquiries@saferinternet.org.uk), the aim of this document is to help monitoring providers to illustrate to schools how their particular technology system(s) meets the national defined 'appropriate monitoring' standards. Fully completed forms will be hosted on the UK Safer Internet Centre website alongside the definitions

The results will help schools better assess, in conjunction with their completed risk assessment, if the monitoring system is 'appropriate' for them.

Company / Organisation	e-Safe Systems Ltd
Address	Salford University Business Park, Leslie Hough Way, Salford, M6 6AJ
Contact details	Tel: 08443 443 001 Email: Info@e-safesystems.co.uk Website: www.esafeeducation.com
Filtering System	e-Safe Monitoring Service
Date of assessment	August 2016

System Rating response

Where a supplier is able to confirm that their service fully meets the issue identified in a specific checklist the appropriate self-certification colour for that question is GREEN.	
Where a supplier is not able to confirm that their service fully meets the issue identified in a specific checklist question the appropriate self-certification colour for that question is AMBER.	

[1] Revised Prevent Duty Guidance: for England and Wales, 2015, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445977/3799_Revised_Prevent_Duty_Guidance__England_Wales_V2-Interactive.pdf

[2] https://www.gov.uk/government/publications/keeping-children-safe-in-education--2

Monitoring Content

Monitoring providers should ensure that they:

Aspect	Rating	Explanation
• Are IWF members		e-Safe is a member of the IWF
• Work with CTIRU 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'		e-Safe collaborates with the Home Office and directly with regional Police Prevent teams to identify markers of localised risk. These are incorporated into the dynamic e-Safe Threat Libraries to provide visibility of terrorist related risk at a granular, local level. See e-Safe Service summary section below

Inappropriate Online Content

Monitoring providers should both confirm, and describe how, their system monitors/manages the following content

Content	Explanatory notes – Content that:	Rating	Explanation
Illegal	content that is illegal, for example child abuse images and unlawful terrorist content		e-Safe's sophisticated image and keyword detection technology together with our constantly updated Threat Library captures evidence of a wide range of illegal behaviour, in school and beyond, irrespective of the application or language in use. Child abuse and paedophile activity is often detected from webcam analysis, sometimes on encrypted applications, and illegal static imagery which is not supported by text. In excess of 95% of all child abuse imagery associated with successful prosecutions in the UK has no text associated with it, therefore moving and static image analysis, agnostic of the application, is critical. e-Safe's unparalleled ability to monitor in any language, including any script, ensures that unlawful written content is detected, irrespective of whether it is viewed to screen, entered via the keyboard,

			accessed or downloaded from a USB or mobile device. Finally, the review of all incidents detected by the specialist behaviour monitoring at e-Safe ensures that evidence of actual illegal content or behaviour is escalated by phone call, in real-time, to a nominated contact.
Bullying	Involve the repeated use of force, threat or coercion to abuse, intimidate or aggressively dominate others		The detection, review and escalation of all safeguarding risk comprises the 3 components mentioned above and outlined in the Summary section below. In respect of bullying behaviour, e-Safe's Threat Libraries contain thousands of bullying related words, phrases, euphemisms and slang, in multiple languages, reflecting culturally specific terms that would be meaningless in English, and localised markers specific to an individual school or region. e-Safe engages with organisations who are specialist practitioners in a particular field of behaviour. This enables e-Safe to source new markers of threats, and create additional markers internally - based on evidence of current trends. In this way the e-Safe provides an unmatched detection capability across a wide range of safeguarding risk – see Summary section below.
Child Sexual Exploitation	: Is encouraging the child into a coercive/manipulative sexual relationship. This may include encouragement to meet		See responses to behaviours above and the summary section below. Child Sexual Exploitation is generally evidenced by imagery, chat, social media and webcam activity, often on encrypted applications such as Skype. The early warning markers of such behaviour are often very subtle and require an expert eye to detect. The multi-lingual team of behaviour specialists at e-Safe provide both the essential expert review and the immediate escalation of genuine risk, 24 x7, 365 days per year.

Discrimination	Promotes the unjust or prejudicial treatment of people on the grounds of race, religion, age, sex, disability or gender identity		See responses to the behaviours above and the Summary section below. Also note that the multi-lingual monitoring can be critical in identifying discrimination and that genuine risk is often masked by large volumes of false positives. The specialist review of all incidents, in all languages ensures that genuine discrimination, even with cultural bias, is identified
Drugs / Substance abuse	displays or promotes the illegal use of drugs or substances		See responses to the behaviours above and the Summary section below. Also note that substance abuse is a behaviour typically associated with a vast number of constantly changing euphemisms and terms, as users attempt to disguise their activity. A new term identified in a London Borough is added to the e-Safe Threat Library to ensure that all e-Safe school and college customers throughout the UK (and internationally) benefit from that marker on the same day. In this way as a term grows in popularity, an e-Safe school or college is automatically prepared and the behaviour visible.
Extremism	promotes terrorism and terrorist ideologies, violence or intolerance		See responses to the behaviours above and the Summary section below. Also note the critical importance of multi-language monitoring and regularly updated Threat Libraries to ensure effective detection of all aspects of extremist risk. In addition, extremism and the grooming of individuals by extremists is often associated with very subtle, and seemingly benign behaviour, occurring over an extended period of time. The specialist monitoring resources employed at e-Safe have both the subject matter expertise and the time to review incidents in context and cross reference with historic activity involving the same user.

			In this way, genuine risk can be identified from a series of low level, seemingly benign incidents that may otherwise go unnoticed. It is also e-Safe's experience that serious extremist risk is more likely revealed by offline activity and use of devices away from school or college. Material is typically passed via USB to read offline and the activity is conducted away from the school or college premises.
Pornography	displays sexual acts or explicit images		See responses to Illegal and Child Sexploitation above and the Summary section below. e-Safe's sophisticated image detection capability will identify moving, static and webcam pornographic material, online or offline, irrespective of the application used. Our experience and evidence of monitoring 500,000+ students and staff in the UK shows that pornographic images and videos are frequently downloaded to school and college devices from pen drives and mobile phones. Sexual acts are performed on webcam and encrypted applications such as Skype. Pornographic material is stored by users on school and college hard drives and central storage areas, hidden in innocent folders. Even if an image is not opened for viewing by the user, e-Safe scans local and central drives to identify stored pornography ensuring that a school or college can be confident such material is always visible and the appropriate intervention made.
Self Harm	promotes or displays deliberate self harm		See responses to the behaviours above and the Summary section below. Mental health related issues such as depression, self-harm and suicide risk represents the single largest category of incidents detected by e-Safe

			across students and staff - at all levels of education in the UK. This is a behaviour which can include image based evidence but certainly is typified in its early stages by subtle markers which require expert interpretation and assessment. e-Safe prides itself on the early warning of self-harm risk which is a direct result of: • the time and effort the team places on ensuring that our Threat Library markers are extensive and current; • the continuous engagement with specialist organisations in this field of mental health; • the subject matter expertise across our highly trained e-Safe monitoring team • the multi-lingual incident review by e-Safe monitoring staff 24 x 7, 365 days per year; • the breadth and quality of our detection capability in identifying offline and offsite behaviour conducted in any language.
Suicide	Suggest the user is considering suicide		See the response to Self-Harm above and the Summary section below.
Violence	Displays or promotes the use of physical force intended to hurt or kill		See the responses above and the Summary section below.

This list should not be considered an exhaustive list. Please outline how the system manages this content and many other aspects

e-Safe Monitoring Service Summary

e-Safe recognises the increasing responsibility that school and college leaders have for safeguarding children and young people in their care; especially with the latest guidance from the Department for Education (Sept 2016) stipulating they should do "all that they reasonably can to limit children's exposure to the risks" from the establishment's ICT system, while putting in place early intervention strategies to help to stop safeguarding risks escalating.

In this context, **e-Safe takes the strain of detecting safeguarding risks** with a unique service that provides highly effective forensic monitoring, not just during term time and school or college hours, but 24 hours a day, 365 days a year**.**

The e-Safe service comprises 3 vital components:

1. **Advanced detection software** - with the capability to monitor both words and phrases, as well as images that are moving *and* static.
   - Safeguarding risk may be wholly or partly written in a foreign language script, a foreign language written using an English keyboard (Romanised), include slang or text-speak variants, or reflect a cultural meaning which doesn't translate to English. e-Safe has a unique ability to detect risks in any language, in any text.
   - The markers of many serious threats are often imagery based, typically moving imagery, on webcam, chat roulette and encrypted applications like Skype. In fact, 95% of imagery associated with child abuse and paedophile activity has no text associated with it at all. With sophisticated image detection technology, e-Safe ensures that static, video and webcam activity - which is not accompanied by text or meta data - is visible too.

2. **Expert interpretation & assessment** - Effective incident review demands time to examine monitoring output and specialist knowledge to interpret the signs. A dedicated team of behavioural experts work diligently to identify the early warning indicators of inappropriate and harmful behaviour. Importantly, the team is multi-lingual with a rich knowledge of different cultures - vital skills for interpreting and assessing the true meaning of words, phrases, slang and text speak in different languages.

   With *e*-Safe all incidents detected are reviewed by the *e*-safe team, categorised into various levels of seriousness and escalated through to the school by phone and encrypted email to a pre-agreed protocol. Depending upon the nature of the incident or issue, the escalation may go to a single individual (e.g. illegal image to safeguarding lead or a Head teacher only) or to a combination/group of contacts (e.g. accessing porn in Year 9 to Year leader & safeguarding lead).

3. **Dynamic Threat Libraries - updated continuously to maintain detection accuracy** - working in collaboration with external partners and schools, our experts update and refine threat libraries on a daily basis to detect emerging behavioural trends at an international, national and local level.
   - The *e*-Safe Threat Libraries comprise tens of thousands of markers across multiple languages
   - Words and phrases are monitored against a series of Libraries covering threats of illegal and inappropriate behaviours, *such as grooming, paedophile activity, child abuse and sexualisation, bullying and harassment, possible self- harm/suicide, HBT, FGM, racism, radicalisation, threats of violence, terrorist activity, trafficking and gang culture.*
   - The Libraries are updated daily to maintain detection accuracy and reflect changes in behaviour trends. New markers are sourced from continuous research by our monitoring experts as well as through close collaboration with external specialist agencies and

schools. For example, the intelligence gathered from our sex offender monitoring work on behalf of UK Police Forces is used to ensure the grooming and paedophile threat library is always up to date.

- Schools also benefit from *e*-safe's global and national library footprint e.g. an issue emerging a day ahead in Australia is added to the relevant Threat Library before the next UK academic day; a marker related to trafficking risk identified in a local region of the UK is added to the Threat Library to aid detection across all schools in that area. This level of responsiveness ensures that schools benefit quickly from monitoring of current and very diverse indicators of safeguarding risk.
- In addition, *e*-Safe maintains and administers bespoke Threat Libraries for individual schools to cope with issues such as gang culture, local slang and requirements for specialist behavioural detection.

Monitoring 24/7, 365 days a year

To effectively fulfil statutory safeguarding duties, school leaders have to maintain visibility of ICT use whenever and wherever it is deployed. School devices are regularly taken off site by students and staff for use at home during evenings, weekends and holidays. In fact, our evidence shows that 1/3 of incidents happen offline – often in the evenings, weekends and holidays

Most traditional monitoring software solutions, and all filtering solutions that offer a degree of keyword detection, are focused on online activity, usually on site. This means that the significant volume of incidents that happen offline - in the evenings, weekends and holidays - are invisible to the safeguarding team.

With e-Safe, incidents that happen out of hours are reviewed 24/7, 365 days a year - and those requiring immediate intervention are escalated in real-time to ensure effective intervention, protection and support for the individual, and minimal reputation risk for the school.

Focused on finding the early markers of harmful and inappropriate behaviours, however subtle

The new DfE guidance states that you must have early visibility of markers of harmful and inappropriate behaviours, so that intervention strategies can be put in place to stop the risk from escalating.

The problem is, the markers associated with the range of safeguarding risks can often be incredibly subtle. The content is likely to be very benign: the obvious markers of risk will be absent and the safeguarding threat easily overlooked.

Self-administered solutions place an increasing burden on a school's resources if the required level of safeguarding is to be attained. Even with highly accurate detection technology, high volumes of false positives are produced and every incident must be checked - in context with historic activity. The word 'suicide' is a classic example. Applying weighting to 'suicide' creates vast volumes of false positive incidents due to school project work, music lyrics, etc. The resulting pages of incidents presented are numerous and, to make matters worse, the risk of suicide, and other extreme behaviours, is rarely indicated by a user typing the actual word 'suicide' or a known variant.

The specialist team at e-Safe has the time and necessary skill to assess the severity of incidents and - importantly - to distinguish between the genuine issues (requiring intervention) and the false positives. This not only removes the burden other systems place on non-specialists, it ensures the very serious incidents are not lost under a blanket of false positive and less serious data.
The high volume of false-positives associated with traditional software monitoring solutions often forces schools to adopt a 'top10' approach, focusing on incidents by volume. In an attempt to manage the volume of activity that comes through, markers can be removed. It is vital that experienced and trained specialists make these decisions based on a robust volume of evidence.

At e-Safe we recognise that the ICT environment is a rich source of behaviour markers and believe, therefore, it is important that all incidents are reviewed. Typically, the more serious behaviours do not occur in volume and a 'top 10' approach or the removal of a marker can easily result in the complete loss of visibility of a serious safeguarding risk. Large volumes of benign incidents easily mask the single subtle marker of life threatening behaviour, or illegal activity, and at e-Safe <u>all</u> incidents are reviewed by experts skilled in behaviour monitoring.

Each team member at e-Safe shares a passion for safeguarding and

- *Holds a degree in at least one of the following areas: Child Psychology, Criminology, Forensic Science and Computing Forensics.*
- *Has experience of working with and supporting young people and adults in a variety of behaviour related situations e.g. bullying and harassment, grooming, child abuse, mental health, offender management.*
- *Has on-going training across the range of inappropriate or harmful behaviours.*
- *Is DBS checked and security vetted to NPPV (Non Police Personnel Vetting) level 2 or 3*

e-Safe provides accurate baseline measurement that makes the assessment of intervention strategies straight forward
The DfE (and Inspectorate) requires that when intervention plans are put in place, the effectiveness is measured and used to refine future intervention strategy.

e-Safe provides analysis of actual and genuine behaviour to illustrate the effectiveness of interventions as the baseline changes over time.

The trouble with the reports typically provided by traditional software-based monitoring solutions is the significant volume of false positives that are included in the data, as they cloud the genuine underlying incident baseline.

The monthly, termly and annual reports and analysis provided by e-Safe only reflect incidents that have been reviewed by our specialist team and require intervention. Meaning the baseline of behaviour and safeguarding risk is accurate, enabling the leadership team to correctly assess the effectiveness of their interventions and plan future intervention strategies accordingly.

How does it work?
The *e*-Safe application is securely installed on the school computer device(s), in the cloud (e.g. Google Chrome), or on servers controlling such as thin client and virtual desktop environments (VDI) - to monitor a user's online (Internet) and offline activity, both in school and away from the school network. Each computer device can be uniquely identified in its environment, and the user is allocated to a specific group to aid granular monitoring and reporting e.g. staff, student, year group, vulnerable or site specific.

When potentially inappropriate behaviour is flagged to the team at *e*-Safe via the application, this is what happens:

- The incident is captured as a screen shot
- A screen shot is uploaded to an *e*-Safe server along with the user id, machine name, time and date stamp.
- The incident is reviewed by the specialist team at *e*-Safe, who are all DBS cleared and vetted to NPPV (Non Police Personnel Vetting) level 2 or 3.
    - Incidents requiring intervention are identified and escalated to nominated contacts using the pre agreed escalation/reporting protocol.
    - Serious incidents (e.g. child abuse imagery, grooming, life threatening behaviour) are reported by telephone, directly to the specified safeguarding contact(s) in real-time.
- An encrypted report is produced, including supporting material where appropriate, and sent by email - stating the user id, machine id, time/date of the incident captured and a narrative of the incident.

The data captured by e-Safe in the course of monitoring users is protected during transmission and at rest:

- All data transmitted between a school/college device & the e-Safe server is encrypted (256 AES)
- The e-Safe servers are located behind a secure firewall at a UK based ISO27001 accredited data centre
- Access to the servers is password controlled
- Access to the e-Safe application on the servers is password controlled
- Data is validated upon receipt to prevents code insertion attacks.
- Database access is password protected and the data 'at rest' in the database is encrypted to provide another level of database obfuscation

Providers should be clear how their system does not over block access so it does not lead to unreasonable restrictions

The e-Safe monitoring service is geared to ensure that the 'blocking' of content is appropriate and does not restrict legitimate activity. Blocking can be applied to:

- Content e.g. an image blocked but not written content

- Document, file or URL blocked entirely based on pre-defined criteria e.g. it is on a published banned list, known to be unlawful, contains inappropriate material such as pornography or radical comment etc,
- Applications e.g. Chatroulette, Social Media Gaming, Tor browser
- Device e.g. attempts to access an inappropriate file/content held on a pen drive

Blocking can be applied at school, user group, device group, user and device level as the school or college requires.

Monitoring System Features

How does the monitoring system meet the following principles:

Principle	Rating	Explanation
• Age appropriate – includes the ability to implement variable monitoring appropriate to age. This will in turn define which alerts are prioritised and responded to		e-Safe monitoring settings are variable at the software level. However, the unique incident review service delivered by e-Safe behaviour specialists ensures that all incidents are reviewed. The escalation protocol agreed with the school or college determines the prioritisation and escalation process. The burden of review and prioritisation is eliminated. The content and context of the behaviour is key to determining safeguarding risk.
• BYOD (Bring Your Own Device) – if the system includes the capability to monitor personal mobile and app technologies (ie not owned by the school), how is this deployed and supported?		e-Safe monitors BYOD and can be completely platform independent without the need to install an e-Safe client package or app on the device. The e-Safe client can monitor use of school or college services from within a thin client, enhanced terminal services, virtual desktop or similar environment. A user connecting this way via a personal computer, mobile phone, , iPad, iPhone, Smartphone, Android tablet

		etc is monitored for all online activity. Schools and colleges with a Google for Education account automatically monitor BYOD activity along with all school/college owned devices within the Google Chrome environment. The e-Safe Chrome monitoring app is imported into the GfE account and provides full image and keyword monitoring. With the owner's consent it is possible to install the e-Safe client to a non-school or college owned device e.g. Windows laptops, laptops & mobile phones, Apple MAC, & from December 2016 any Android device. Monitoring can be restricted to school site only and the e-Safe client auto disabled when the device is removed from site.
• Data retention –what data is stored, where and for how long		The following data is captured by e-Safe in the course of monitoring users: • The user login id (not the name of the user) • The date and time stamp an incident occurred • The id of the device (& serial numbers of various components within the device) that the user was logged in to at the point the incident occurred The data above is held on dedicated servers located in an ISO 27001 accredited UK data centre.

		The retention period is determined by the school or college, or in the instance of illegal/criminal activity, the Police.
• Flexibility – schools ability to amend (add or remove) keywords easily		A school or college can request new terms to be added to the e-Safe Threat Library on demand. Upon receipt of a request by email or phone, the new terms will be added within 15 minutes by the e-Safe monitoring team. Importantly, e-Safe is proactively updating the Threat Library daily with new words, terms and euphemisms based on input from 3rd party organisations and our own behaviour experts. Removal is a critical decision and with software based monitoring solutions administered by schools and colleges internally, it is often performed to reduce false positives at the risk of making the single serious incident or threat invisible going forward. The e-Safe monitoring team shoulder the monitoring workload, reviewing all incidents and eliminating the false positives, thereby eliminating the need for a school or college to consider keyword removal.
• Monitoring Policy – How are all users made aware that their online access is being monitored? Is any advice or guidance provided to support schools?		Safeguarding risk is not confined to online activity. Our analysis of monitoring over 500,000 UK students and staff in primary, secondary and further education shows that nearly 1/3rd of all serious incidents occur offline, away from the Internet. e-Safe monitors behaviour 24 hours per day, 365 days per year, irrespective of whether the

		user is online or offline, in school/college or offsite. e-Safe monitors behaviour of students and staff in accordance with the school/college statutory requirements, its AUP, Code of Conduct and employee contracts. The behaviour e-Safe identifies and escalates to a school or college will often reveal the need to amend policies and contracts – e-Safe staff will prompt school and college leaders of a need to review policies based on the monitoring evidence. e-Safe will support and guide schools in particular interventions and also recommend where appropriate 3rd party specialists to assist.
• Multiple language support – the ability for the system to manage relevant languages?		e-Safe uniquely supports <u>all</u> languages, including script based languages at the technical level. The e-Safe Threat Library contains culturally specific terms across all behaviours, including foreign language slang variants. The e-Safe monitoring team is multi-lingual. All three components are critical for effective monitoring of safeguarding risk in the UK's diverse, multi-cultural society.
• Prioritisation – How are alerts generated and prioritised to enable a rapid response to immediate issues. What operational procedures are in place to facilitate that process?		See Summary section above. • Incidents requiring intervention are identified and escalated to nominated contacts using the pre agreed escalation/reporting protocol either immediately, same day, or weekly.

		• Serious incidents (e.g. child abuse imagery, grooming, life threatening behaviour) are reported by telephone, directly to the specified safeguarding contact(s) in real-time. • Non-life threatening or illegal incidents are escalated by encrypted report which includes the narrative explanation from the e-Safe behaviour specialist of the incident and the reason for escalation • The protocol for incident reporting and escalation is determined by the school or college. The seriousness of an incident, the type of incident and the user involved will determine who receives the escalation.
• Reporting – how alerts are recorded within the system?		See How does it work? in the Summary section above.

Please note below opportunities to support schools (and other settings) with their obligations around Keeping Children Safe in Education?

Visit www.esafeeducation.com & view or download the report - Key facts about Keeping Children Safe in Education

PROVIDER SELF-CERTIFICATION DECLARATION

In order that schools can be confident regarding the accuracy of the self-certification statements, the supplier confirms:

- that their self-certification responses have been fully and accurately completed by a person or persons who are competent in the relevant fields
- that they will update their self-certification responses promptly when changes to the service or its terms and conditions would result in their existing compliance statement no longer being accurate or complete
- that they will provide any additional information or clarification sought as part of the self-certification process
- that if at any time, the UK Safer Internet Centre is of the view that any element or elements of a provider's self-certification responses require independent verification, they will agree to that independent verification, supply all necessary clarification requested, meet the associated verification costs, or withdraw their self-certification submission.

Name	Mark Donkersley
Position	Managing Director
Date	31st August 2016
Signature