Appropriate Monitoring for Schools

May 2025

Monitoring Provider Checklist Responses

Schools (and registered childcare providers) 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". Furthermore, it expects that they "assess the risk of [their] children being drawn into terrorism, including support for extremist ideas that are part of terrorist ideology". There are a number of self review systems (eg www.360safe.org.uk) that will support a school in assessing their wider online safety policy and practice.

The Department for Education's statutory guidance 'Keeping Children Safe in Education' obliges schools and colleges in England to "*ensure appropriate filters and appropriate monitoring systems are in place and regularly review their effectiveness" and they "should be doing all that they reasonably can to limit children's exposure to [Content, Contact, Conduct, Contract] risks from the school's or college's 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	Smoothwall (part of Qoria)
Address	Second Floor, 2 Whitehall Quay, Leeds, LS1 4HR
Contact details	https://smoothwall.com/contact-us
Monitoring System	Smoothwall Monitor
Date of assessment	11/03/2026

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.	

Monitoring Content

Monitoring providers should ensure that they:

Aspect	Rating	Explanation
• Are IWF members		Yes, Smoothwall is a long standing member of the Internet Watch Foundation and implement the IWF CAIC list
• Utilisation of IWF URL list for the attempted access of known child abuse images		The images we see are generally screenshots, and as such aren't suitable for hashing. Hashing is able only to match images with minor changes, and as such cannot hope to match (eg.) an image and that image screenshot in a user's browser.
• Work with CTIRU 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'		Smoothwall works with CTIRU to improve detection in both filtering and monitoring
• Confirm that monitoring for illegal content cannot be disabled by anyone (including any system administrator) at the school		A scheduling option is available when school staff are not available during holidays and/or weekends. Scheduling enables Monitor to be temporarily paused, stopping the system from triggering alerts. This is a hidden feature and must be requested and configured by the school. The feature will only be enabled by Smoothwall where the use case is approved.

Illegal Online Content

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

Content	Explanatory notes – Content that:	Rating	Explanation
child sexual abuse	Content that depicts or promotes sexual abuse or exploitation of children, which is strictly prohibited and subject to severe legal penalties.		Smoothwall Monitor includes the detection of contact with monitored users for sexual purposes. Monitoring looks for signs of grooming and requests for sexual information or images. The "Oversharer" theme alerts in instances where a monitored user might be providing personal information online – their address, full name or phone

			number for example. Monitor also looks for sexual content and references to non-consensual sexual behaviour.
controlling or coercive behaviour	Online actions that involve psychological abuse, manipulation, or intimidation to control another individual, often occurring in domestic contexts.		Monitor's "Vulnerable Person" category aims to identify signs of a person being coerced, including, but not limited to criminal coercion eg county lines.
extreme sexual violence	Content that graphically depicts acts of severe sexual violence, intended to shock or incite similar behaviour, and is illegal under UK law.		The "Sexual Content" category would pick up the creation or searching for such content. The moderation team would act to remove a screenshot if sharing that content would be illegal.
extreme pornography	Pornographic material portraying acts that threaten a person's life or could result in serious injury, and is deemed obscene and unlawful.		The "Sexual Content" category would pick up the creation or searching for such content. The moderation team would act to remove a screenshot if sharing that content would be illegal.
fraud	Deceptive practices conducted online with the intent to secure unfair or unlawful financial gain, including phishing and scam activities.		Monitor's "Vulnerable Person" category aims to identify signs of a person being subject of fraud. Monitor's 'Oversharer' category captures sharing of personal details including addresses, passwords, and financial details, enabling schools to act to protect children at risk of fraud or financial exploitation. **Monitor is not designed as an anti-phishing tool so should not be considered sufficient defence here.**
racially or religiously aggravated public order offences	Content that incites hatred or violence against individuals based on race or religion, undermining public safety and cohesion.		Monitor's "Terrorism and Extremism" category would capture content related to racially & religiously motivated hate and violence.
inciting violence	Online material that encourages or glorifies acts of violence, posing significant risks to public safety and order.		Monitor's 'Violence' category captures acts of violence and the promotion of violence. These are often associated with extremist ideology: Monitor's 'Terrorism/Extremism' category captures the sharing of proscribed organisations and promotion of ideological violence.
illegal immigration	Content that promotes or facilitates unauthorized entry into		Monitor's "Vulnerable Person" category would be triggered if a

and people smuggling	a country, including services offering illegal transportation or documentation.		user generated content indicating they had been victims of trafficking or similar.
promoting or facilitating suicide	Material that encourages or assists individuals in committing suicide, posing serious risks to vulnerable populations.		As with Self Harm, suicidal ideation and discussion of or researching suicide related material is covered by the "Vulnerable User" theme. If a risk to life is suspected, the DSL will receive a phone call straight away – 24/7/365.
intimate image abuse	The non-consensual sharing of private sexual images or videos, commonly known as "revenge porn," intended to cause distress or harm.		Monitor's "Vulnerable Person" category would be triggered if a user generated content indicating they had been victims of intimate image abuse, or were in a position where they might be abused.
selling illegal drugs or weapons	Online activities involving the advertisement or sale of prohibited substances or firearms, contravening legal regulations.		A number of risk categories will (depending on circumstance) be triggered by someone engaging in these activities, including "Violence" and "Vulnerable Person"
sexual exploitation	Content that involves taking advantage of individuals sexually for personal gain or profit, including trafficking and forced prostitution.		Monitor's "Vulnerable Person" category would be triggered if a user generated content indicating they had been victims of Sexual Exploitation.
Terrorism	Material that promotes, incites, or instructs on terrorist activities, aiming to radicalise individuals or coordinate acts of terror.		Monitor's "Terrorism and Extremism" category covers all aspects of this type of material, including up to date flags for proscribed organisations.

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
Gambling	Enables gambling		Gambling would be classified as "Vulnerable Person"
Harmful content	Content that is bullying, abusive or hateful. Content which depicts or encourages serious violence or injury. Content which encourages dangerous stunts and challenges; including the ingestion, inhalation or exposure to harmful substances.		Monitor would pick up this type of risk in a number of different areas, depending on whether the monitored user is perpetrator or victim, and the type of harmful content.

Hate speech / Discrimination	Content that expresses hate or encourages violence towards a person or group based on something such as race, religion, sex, or sexual orientation. Promotes the unjust or prejudicial treatment of people with protected characteristics listed in the Equality Act 2010		Monitor's "Terrorism and Extremism" category covers all aspects of this type of material.
Malware / Hacking	promotes the compromising of systems including anonymous browsing and other filter bypass tools as well as sites hosting malicious content, including ransomware and viruses		**Monitor is not designed as a cyber security tool, and as such is not considered a mode of malware protection**. However, a student engaged in illegal activity may be picked up as "Vulnerable Person".
Mis / Dis Information	Promotes or spreads false or misleading information intended to deceive, manipulate, or harm, including content undermining trust in factual information or institutions		Qoria (and by extension Smoothwall) partners with Newsguard, and along with our own signatures for misinformation, use their expertise to guide our classification.
Pornography	displays sexual acts or explicit images		The "Sexual Content" category would pick up the creation of or searching for such content.
Self Harm and eating disorders	encourages, promotes, or provides instructions for self harm or eating disorders		Material encouraging Self Harm is covered by the "Vulnerable User" theme. If a risk to life is suspected, the DSL will receive a phone call straight away – 24/7/365.
VAWG	Promotes or glorifies violence, abuse, coercion, or harmful stereotypes targeting women and girls, including content that normalises gender-based violence or perpetuates misogyny.		This would be captured under the "Violence" risk category. Smoothwall's team are particularly aware of VAWG and would raise the risk level appropriately

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

Our risk categories are broad, and constantly expanding. However, it should be noted that many risks noted here are the types of risk which also include consumption of material found online. While Monitor will pick up users searching or discussing these types of material, it is vitally important to pair monitoring with high quality, real time content filtering.

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 and vulnerability. This will in turn define which alerts are prioritised and responded to. Further situations may warrant additional capability, for examples boarding schools or community based access		Monitor has several age group settings which are applied during onboarding. This alters alert thresholds and settings relevant to the age group chosen. Age group settings can be applied to specific Groups of users, allowing for granular control of Monitoring sensitivity.
• Alert Management – how alerts are managed – if schools manage system alerts or support/management is provided		Monitor allows fully customisable alert settings for the site or groups within the site. Alerts can be tailored for each Safeguarding user of the system, allowing them to chose the Groups they receive alerts for, and the severity at which they will be notified
• Audit – Any changes to the monitoring system are logged enabling an audit trail that ensures transparency and that individuals are not able to make unilateral changes.		Users are not able to perform any actions in the UI that would cause concern. For example they are not able to delete alerts from the database. Adding and removing users via the user management tool is audited.
• BYOD (Bring Your Own Device) – if adopted by the school and the system includes the capability to monitor personal mobile and app technologies (ie not owned by the school), how is this deployed and supported and how data is managed. Does it monitor beyond the school hours and location		Monitor does not currently fully cater for BYOD. However if a student logs into their school Chrome account on a BYOD that activity would be Monitored
• Data retention –what data is stored, where is it (physically– ie cloud/school infrastructure) stored and for how long. This should also include any data backup provision		Data is stored on our secure servers for a period of 15 months and then permanently deleted. Monitor's integration with all widely-used safeguarding record keeping systems allows Safeguarding users to automatically copy data across, providing longer term storage

● Devices – if software is required to be installed on devices, the monitoring system should be clear about the devices (and operating systems) it covers		Monitor supports Windows, MacOS, iOS and ChromeOS. Customers are informed of this during the sales and onboarding process
● Flexibility - changes in keywords (addition or subtraction) can be easily amended according to an agreed policy		Schools have the option to feed back into the moderation system. The AI and human moderation components are part of a carefully calibrated system where new sources of alerts are added by our professional analysts.
● Group / Multi-site Management – the ability for deployment of central policy and central oversight or dashboard		Requirements for monitoring across all sites within a group of schools will be discussed during a customer's onboarding. Where centrally-managed policies are required they can be easily mirrored across sites, and central users can be given a variety of levels of visibility over their sites. Monitor's reporting tools are suitable for single sites and large groups of schools, and automatically display information on all sites the user has access to
● Harmful Image detection – The inclusion or extent to which visual content is discovered, monitored and analysed (e.g. Image hash).		Harmful images may be detected by our moderators, and in addition, users can opt for “Cloud Scan” which identifies a number of categories of harmful image stored in Schools’ online services. Lack of local image hash detection is the reason for orange, rather than green.
● Identification - the monitoring system should identify users and devices to attribute activity (particularly for mobile devices) and ensure the application of appropriate configurations for individual users.		All monitoring is attached at a user level, and while device information is captured, it is not used for aggregation. MAC address and operating system are captured, enabling schools to differentiate between school and personal devices.

● Monitoring Policy – How are all users made aware that their online access is being monitored? Is any advice or guidance provided to support schools?		Smoothwall provides assistance to customers in informing their users about monitoring with Smoothwall Monitor, including plain-language explanations for children about monitoring.
● Mobile and app content – mobile and app content is often delivered in entirely different mechanisms from that delivered through a traditional web browser. To what extent does the monitoring system operate across mobile devices and app content. Providers should be clear about the capability of their monitoring system to monitor content on mobile and web apps and any configuration or component requirements to achieve this.		Mobile apps are challenging to monitor directly due to Android and IOS sandboxing. Apps on desktop operating systems (Windows, Macos) are fully monitored.
● Multiple language support – the ability for the system to manage relevant languages?		Smoothwall Monitor is used across the UK, US and Australia. Monitor fully supports English, Spanish, and contains a number of keywords from many commonly-used languages.
● 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?		Alerts are categorised on a scale of 1 to 5, initially by AI, then a human reviewer. Alerts are then sent according to theme and severity. Almost all events will trigger an email, some higher level events will trigger a phone call to the Safeguarding Team
● Remote monitoring – with many children and staff working remotely, the ability, extent and management for the monitoring of devices (school and/or personal). Included here is the hours of operation together with the explicit awareness of users. Monitoring should focus on school-owned and managed devices. When shared devices are used, schools must ensure users log in individually. This allows monitoring systems to apply restrictions and configurations based on user profiles, improving the safeguarding process.		Monitored devices are actively monitored 24/7/365, whether the device is in-school or elsewhere. All activity on a monitored device will be analysed. Only school issued devices and accounts are supported by Monitor. Smoothwall provides assistance to schools in making users and parents aware of monitoring

● Reporting – how alerts are recorded within the system?		Alerts are recorded separately to capture information. All alerts are available in the portal and can be searched, linked through to associated screen captures. Permanent storage should be in the school's record management system. A full set of reports over time showing alert types and levels is available within the Monitor dashboard.
● Safeguarding case management integration – the ability to integrate with school safeguarding and wellbeing systems to better understand context of activity		Monitor is capable of API integration with all major safeguarding platforms in the UK.

Pro Active Monitoring - how any pro-active monitoring support is provided including if any automation is utilised and the safeguarding capability of the organisation's teams.

Smoothwall Monitor only supports pro-active monitoring. Smoothwall believes this is the only way monitoring implementations can be successful. Automation is used to support the monitoring team in weeding out captures which are not harmful, and presenting the moderation team with the data in the most efficient way. The moderation team are all Smoothwall employees, fully DBS checked, and have the support of counsellors and the HR team. None are on a zero hours contract. The moderation team do not make decisions on the outcome, they are there to make sure you don't see false positives. As such, they are not trained safeguarders per se, rather operatives trained in understanding what they are seeing and whether a DSL or other safeguarder would need to be alerted.

Where illegal material is screenshotted, this will be deleted, and a log of the deletion kept for investigative purposes. School alerts will include a note about the deleted image.

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

As part of the wider Qoria group, Smoothwall offers a huge range of products and support for schools, including best in class Monitoring, Classroom Management and Student Wellbeing tools. Additionally Smoothwall offers training and resources to promote safety in UK schools, including a school branded "hub" for parents and students

How does your monitoring system identify and respond to activity involving Generative AI technologies (e.g. AI prompts, content creation, or platform use)?

In your response, please explain how your system captures or analyses user interactions with Generative AI tools; to what extent logs or alerts reflect potential safeguarding risks associated with AI-generated content (such as harmful prompts or inappropriate use of image and text generation); any known limitations—whether technical, privacy-related, or device-specific—that may affect your

system’s ability to monitor such activity; and what guidance you provide to schools to support their understanding and management of Generative AI-related risks.

In the main, the risks propagated by AI are similar to risks generated by interactions with other people, only more intense. As such, the current scheme of risks in Monitor has seen no need for expansion. In terms of the monitoring itself, Monitor is capable of capturing conversations in both desktop apps, and anywhere on the web. This includes all major AI chatbot tools. We have seen a significant rise in risk signals from these sources, in line with what we would expect. This has led to changes in our detection and moderation processes to cope with the demand, and increased training for moderators in spotting dangerous AI usage. While chat is the poster-boy for AI, there are ways for AI generated text and images to be presented to students outside of chat conversations, so monitoring just chatbots is not sufficient. Additionally, it is essential that your monitoring is paired with high quality filtering that can detect harmful text, images and video in real time, as AI generated content is often ephemeral.

Monitoring 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	Gabi Walshaw
Position	Vice President, Product for Early Detection and intervention
Date	12/03/2026
Signature