Appropriate Monitoring for Schools

UK Safer Internet Centre
www.saferinternet.org.uk
brought to you by
SWGfL Childnet IWF

June 2021

Monitoring Provider Checklist Reponses

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 they "should be doing all that they reasonably can to limit children's exposure to the above 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	Netsweeper
Address	Suite 125-126 Pure Offices, 4100 Park Approach Thorpe Park, Leeds, United Kingdom, LS15 8GB
Contact details	Lou Erdelyi, lou.erdelyi@netsweeper.com
Monitoring System	Netsweeper
Date of assessment	August 18, 2021

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	🟩	Netsweeper is a Member of the IWF
● Utilisation of IWF URL list for the attempted access of known child abuse images	🟩	Compliant
● Work with CTIRU 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'	🟩	Compliant

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	🟩	Netsweeper has a category for this
Bullying	Involve the repeated use of force, threat or coercion to abuse, intimidate or aggressively dominate others	🟩	Netsweeper has a category for this
Child Sexual Exploitation	Is encouraging the child into a coercive/manipulative sexual relationship. This may include encouragement to meet	🟩	Netsweeper has a category for this and works with IWF and enforces this content
Discrimination	Promotes the unjust or prejudicial treatment of people on the grounds of race, religion, age, sex, disability or gender identity	🟩	Netsweeper has a category for this
Drugs / Substance abuse	displays or promotes the illegal use of drugs or substances	🟩	Netsweeper has a category for this
Extremism	promotes terrorism and terrorist ideologies, violence or intolerance	🟩	Netsweeper has a category for this
Pornography	displays sexual acts or explicit images	🟩	Netsweeper has a category for this
Self Harm	promotes or displays deliberate self harm	🟩	Netsweeper has a category for this
Suicide	Suggest the user is considering suicide	🟩	Netsweeper has a category for this
Violence	Displays or promotes the use of physical force intended to hurt or kill	🟩	Netsweeper has a category for this

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

Real-time content filtering ensures students always have the best protection. Netsweeper's AI-based web content categorization platform is the industry's most accurate and effective solution to classify online content with over 90 categories and 50 languages.

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

The Netsweeper AI-powered platform provides on-the-fly categorization for all content that ensures detection of new and emerging safeguarding threats that is very accurate. Industry-leading database of over 3 billion URLs with real-time dynamic updates to education-specific categories including hate speech, weapons, cyberbullying, and substance abuse. Categories are populated dynamically by AI – they are not static lists. Categorization occurs in over 47 languages. Every Netsweeper deployment globally contributes to the platform with over 150 million URLs categorized every day. Should an over blocking occur, Administrators are able to easily update the system on their own or report such occurrences directly to Netsweeper for further analysis.

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. Further situations may warrant additional capability, for examples boarding schools or community based access	🟩	Users are sourced from the school's authentication system via grade.
• Alert Management – how alerts are managed – if schools manage system alerts or support/management is provided	🟩	Netsweeper can generate an alert but it's the school's responsibility to act upon that alert.
• 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 and how data is managed. Does it monitor beyond the school hours and location	🟩	Netsweeper can monitor any device that is connected physically or wirelessly to the Netsweeper. Furthermore, Netsweeper can expand it's monitor capabilities by allows BYOD devices to install a small lightweight client that applies policies regardless of location.
• 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 as per policy and GDPR. Netsweeper does not have access to the data, but rather the school has access. Netsweeper can

	🟩	be configured to backup data if requested but it's the school's policy that determines or dictates this.
• Devices – if software is required to be installed on devices, the monitoring system should be clear about the devices (and operating systems) it covers	🟩	Software does not need to be installed but optionally can be. All major operating system platforms, Windows, Mac, IOS, Android and Chrome are supported.
• Flexibility – schools ability to amend (add or remove) keywords easily	🟩	Yes, this is completed via the Web Administration system.
• Group / Multi-site Management – the ability for deployment of central policy and central oversight or dashboard	🟩	Yes, this is managed, configured, and monitored via the Web Administration system
• Monitoring Policy – How are all users made aware that their online access is being monitored? Is any advice or guidance provided to support schools?	🟩	This is done using the schools Internet Access Acceptance Policies (IAAP) that students and parents need to sign.
• Multiple language support – the ability for the system to manage relevant languages?	🟩	Netsweeper supports over 40+ 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 can be generated in multiple way depending on the school's needs.
• 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.	🟩	Netsweeper can monitor any device that is connected via VPN without the need to install any software. If no VPN connection is used, then client filter software is requiring which is available for all major operating system platforms.
• Reporting – how alerts are recorded within the system?	🟩	Alerts are recoded as an event type and are visible to the ITC.
• Harmful Image detection – The inclusion or extent to which visual content is discovered, monitored and analysed (eg Image hash)	🟩	Netsweeper works with IWF and Microsoft and can detect and identifying CSAM images.

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

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	Lou Erdelyi
Position	CTO
Date	August 18, 2021
Signature	Lou Erdely