Appropriate Monitoring for Schools

May 2023

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 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	Fortinet, Inc. (which is the manufacturer (not the supplier) of Fortinet network security products and related services)
Address	899 Kifer Road, Sunnyvale, 94086 California, United States
Contact details	+44 20 3752 6880
Monitoring System	FortiGuard Web Content Filtering
Date of assessment	June 2023

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		Fortinet is a member
● Utilisation of IWF URL list for the attempted access of known child abuse images		The list is part of FortiGuard Web Filtering Service
● Work with CTIRU 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'		The list is part of FortiGuard Web Filtering Service
● Confirm that monitoring for illegal content cannot be disabled by the school		Illegal content filtering/monitoring cannot be disabled without local admin rights.

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		There are a number of categories that block illegal content, activities and object, covered in the following categories. These categories can be both blocked and/or reported/alerted upon. **Category – E.G. Drug Abuse, Extremist Groups, Hacking, Pornography, Explicit Violence, Explicit Violence**

Bullying	Involve the repeated use of force, threat or coercion to abuse, intimidate or aggressively dominate others		**Category – Explicit Violence** This category includes sites that depict offensive material on brutality, death, cruelty, acts of abuse, mutilation, etc. These categories can be both blocked and/ or reported/alerted upon. This combination with the ability to monitor HTTPS and use Application Control to inspect communication channels such as Instant messaging if permitted. Abusive and bullying content can be alerted to a safeguarding team based on a pre-defined or user-defined list of words whilst monitoring the typical social media applications. Reporting on this activity is also provided and can be scheduled and delivered to the safeguarding team.
Child Sexual Exploitation	Is encouraging the child into a coercive/manipulative sexual relationship. This may include encouragement to meet		**Category – Child Sexual Abuse** Websites that have been verified by the Internet Watch Foundation to contain or distribute images of non-adult children that are depicted in a state of abuse. Information on the Internet Watch Foundation is available at https://www.iwf.org.uk/ This combined with the ability to monitor HTTPS and use Application Control to inspect communication channels such as Instant Messaging if permitted.

Discrimination	Promotes the unjust or prejudicial treatment of people on the grounds of race, religion, age, sex, disability or gender identity		**Category – Discrimination** Sites that promote the identification of racial groups, the denigration of subjection of groups, or the superiority of any group. These categories can be both blocked and/or reported/alerted upon
Drugs / Substance abuse	displays or promotes the illegal use of drugs or substances		**Category – Drug Abuse** Websites that feature information on illegal drug activities including drug promotion, preparation, cultivation, trafficking, distribution, solicitation, etc. These categories can be both blocked and/or reported/alerted upon.

Extremism	promotes terrorism and terrorist ideologies, violence or intolerance		**Category – Extremist Groups** Sites that feature radical militia groups or movements with aggressive anti-government convictions and beliefs. Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific user-defined keywords that are entered in to the search engines. These categories can be both blocked and/or reported/alerted upon.

Gambling	Enables gambling		**Category – Gambling** Sites that cater to gambling activities such as betting, lotteries, casinos, including gaming information, instruction, and statistics. These categories can be both blocked and/or reported/alerted upon.
Pornography	displays sexual acts or explicit images		**Category – Pornography** Mature content websites (18+ years and over) which present or display sexual acts with the intent to sexually arouse and excite. **Category – Nudity and Risque** Mature content websites (18+ years and over) that depict the human body in full or partial nudity without the intent to sexually arouse. These categories can be both blocked and/or reported/alerted upon.
Self Harm	promotes or displays deliberate self harm		**Category – Explicit Violence** This category includes sites that depict offensive material on brutality, death, cruelty, acts of

			abuse, mutilation, etc. Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific user-defined keywords that are entered in to the search engines. These categories can be both blocked and/or reported/alerted upon. This combined with the ability to monitor HTTPS and use Application Control to inspect communication channels such as Instant Messaging if permitted. Self harm content can dynamically be alerted to the safeguarding team based on user-defined lists of words whilst monitoring the typical social media applications. Reporting on this activity is also provided and can be scheduled and delivered to the safeguarding team.

Suicide	Suggest the user is considering suicide		**Category – Explicit Violence** This category includes sites that depict offensive material on brutality, death, cruelty, acts of abuse, mutilation, etc. Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific user-defined keywords that are entered in to the search engines. These categories can be both blocked and/or reported/alerted upon. This combined with the ability to monitor HTTPS and use Application Control to inspect communication channels such as Instant Messaging if permitted.
			Suicide content can dynamically be alerted to the safeguarding team based on user-defined lists of words whilst monitoring the typical social media applications. Reporting on this activity is also provided and can be scheduled and delivered to the safeguarding team
Violence	Displays or promotes the use of physical force intended to hurt or kill		**Category – Explicit Violence** This category includes sites that depict offensive material on brutality, death, cruelty, acts of abuse, mutilation, etc. These categories can be both blocked and/or reported/alerted upon.

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

General categorisation is based on an automated engine which has been developed in-house and which has evolved over more than 15 years since its initial conception. The system uses language dictionaries to allow support in any language. Sites are scanned based on a number of methods:

- New pages on identified popular sites
- URLs which are requested by a user, but which are not rated. Such URLs will go into a queue to be rated based on popularity.
- Individual requests received from customers or users. These requests can be received in a number of ways (see below) and may be either request to rate an unrated site, or requests to change the rating of a site.

In general, initial rating is done by the automated rating system for easily identifiable content using supervised machine learning. Content not readily identifiable is passed to our multilanguage human reviewers. Malicious content (viruses, exploits) is not rated using this system (more details below) because such sites have legitimate visible content.

Ratings may also be obtained from third-party feeds, including feeds from governments or other organisations, containing such content as extremism or sexual violence.

Requests to change the rating of an already categorised URL will always be dealt with by a human, to ensure that the request gets the highest level of care and attention.

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		Rating can be varied depending on a user group profile, and different users can be added to groups depending on their age. Categories are not directly linked to age, due to the subjective nature of deciding what is appropriate. Rather, it is left to the system administrator to decide which categories make sense for the different age groups.

• Alert Management – how alerts are managed – if schools manage system alerts or support/management is provided		Event handlers can be used on the reporting platform for alert management. Additionally, the reporting platform provides an incident management capability alongside the ability to communicate alerts via fabric/API connectors to third party systems.
• 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.		All Fortinet products involved with web traffic monitoring automatically log any changes made to the system. These logs can then be sent to a log analysis tool, such as FortiAnalyzer™, for the generation of alerts when certain actions have been performed and reporting on any actions performed.
• 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		Fortinet can monitor all destinations users visit on BYOD devices and log them for further analysis. When combined with FortiAuthenticator's SmartConnect function

		this monitoring is extended, allowing BYOD devices to be monitored as if they were school owned devices. Please note that the above only applies to onsite devices. The monitoring of devices off the school premises requires an endpoint agent (FortiClient) to be installed on the device.
• 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		All web accesses (as well as any security events such as malware or intrusion detections) can be logged (generally to a centralised customerowned log server) where log retention is under the complete control of the user. The data stored includes user ID (assuming that users are individually authenticated), URL visited, the ID of the security device which handled the transaction, as well as low-level addressing information and a timestamp.

□ Devices – if software is required to be installed on devices, the monitoring system should be clear about the devices (and operating systems) it covers		Client Software is not necessary to monitor devices when on School Premises and connected to the school network via Wired or Wireless. Client software is only required if Monitoring and Filtering is required when 'off network'. This is in the form of FortiClient™ software and is available for the following operation systems:

		• Windows Desktop (XP or newer) • Apple Mac • Apple iOS • Android • Chromebook
□ Flexibility - changes in keywords (addition or subtraction) can be easily amended according to an agreed policy		The Fortinet web filter allows complete freedom to add, modify and remove keywords to be checked in web pages. Wildcard matches can be used for more flexible searching and thresholds can be applied to block only if a certain word appears multiple times.
□ Group / Multi-site Management – the ability for deployment of central policy and central oversight or dashboard		Fortinet provides a centralized management platform to allow policy to be consistent across multiple locations. The centralized reporting platform also provides for consolidated or, if preferred, autonomous reporting for the multiple locations.

• Monitoring Policy – How are all users made aware that their online access is being monitored? Is any advice or guidance provided to support schools?		In general, such communication would be done via training or security awareness sessions. Note that if a user is blocked after accessing a website which belongs to a blocked category, a customisable block page will be displayed which can contain an explanation for the block, as well as information on whom to contact for more information, or a link to a support website.
• Multiple language support – the ability for the system to manage relevant languages?		The Fortinet web filtering solution is completely multilingual,

		both in the automated rating system, used for the majority of website rating, as well as for the human rating team which contains skills in all major languages to allow for detailed verification of page ratings.

• 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?		Any blocked access will generate a log message, which may also generate an alert. A log analysis tool, such as FortiAnalyzer™, can generate alerts based on these logs. Such alerts may be generated for blocked URLs, or rules can be specified to limit alerts to specific categories or individual users or groups of users, or even a time of day. The alerts can then be sent by a number of means including email, SMS, SNMP traps.
• 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.		Fortinet can protect school-owned off-site devices with an endpoint agent installed in the following ways. FortiSASE™ – Cloud delivered security inspection. This is available for managed devices (FortiSASE agent) or unmanaged devices (Secure Web Gateway). The latter is particularly suitable for providing webfiltering security to Chromebooks. FortiClient™ can provide local protection or enforce a VPN connection with

		centralised protection. Logs from FortiClient will be uploaded automatically for reporting and alerting. FortiClient™ can synchronize webfilter policy with the FortiGate™ for a seamless approach either on premise or remote working
• Reporting – how alerts are recorded within the system?		All detections of forbidden or suspicious accesses will be logged, either locally in the FortiGate (for very small installations) or to a centralised log manager, such as FortiAnalyzer.
• Harmful Image detection – The inclusion or extent to which visual content is discovered, monitored and analysed (eg Image hash)		Images today are only detected by their URL. We don't do analysis of the images themselves.

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.

Fortinet can provide pro-active monitoring in two ways.

1. Scheduled reports – FortiAnalyzer can be used to ingest log feeds from a multitude of Fortinet and third-party products, and provide scheduled reports to DSLs on user activity from that data.

2. Alerting – During the ingestion of logs, described above, FortiAnalyzer can send an alert to DSLs based off of a multitude of criteria. For example, attempts to access blocked websites (individual attempts or multiple attempts in a given timeframe), attempts to use certain applications or applications of a certain type to evade the monitoring platform, searching specific words/phrases in a search engine, etc.

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 selfcertification 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	Ben Wilson
Position	VP Product Management
Date	October 16, 2023
Signature	