Appropriate Monitoring for Schools

June 2020

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. 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	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 37526880
Monitoring System	FortiGuard™ Web Content Filtering
Date of assessment	April 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		Fortinet is a member
• Utilisation of IWF Hash list to identify the storage or transmission of known child abuse images		Fortinet is currently not implementing this block list but are investigating how it can be implemented in a secure and performant manner.
• 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

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 pre-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 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 http://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. …Additionally Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific pre-configured keywords that are entered in to the search engines. 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. …Additionally Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific pre-configured 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 pre-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.
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. …Additionally Keyword Searches within popular Search Engines can be monitored and logged, and also alerts can be generated if specific pre-configured 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 pre-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.

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

Web filter categorisation is based on a combination of multi language machine learning driven automation and manual review Sites are scanned based on a number of methods:

The system uses language dictionaries to allow support in any language. Sites are scanned based on a number of methods:

- Newly detected domain registrations (Newly Registered)
- Domains which are requested by a user, but which are not rated. Such URLs will go into a queue to be rated based on hit count and the current charge on the system. (Newly Observed)
- New pages on identified popular sites
- Bulk requests from a specific end-customer. Such requests are treated case by case, but we generally offer this as a free service.
- Individual requests received from end-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 with a percentage trust rating deciding if the rating can be directly accepted or if the site needs to be sent for manual review.

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.

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

This is covered below, but to summarise:

A flexible hierarchical search system is used which allows ratings to be given to anything from a top-level domain or an IP address, right down to a fully-specified URL. This allows for example a blogging site such as wordpress.com to have a “Personal Websites and Blogs” rating, whilst individual blogs can have a rating based on their actual content. It also ensures that the entire wordpress domain is not blocked just because a single blogger posts inappropriate content.

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		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 handler can be used on the reporting platform to manage alert management. Additional to this the reporting platform also provides and incident management capability and also the ability to communicate alerts via fabric/API connectors to a third party system.
• 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 has a very comprehensive support for BYOD especially when implemented through Fortinet secure WiFi access points. Personal devices can be recognised as such and automatically assigned to a guest zone with a specific set of policies to avoid infections being passed into the school network, and to prevent access to internal systems. …BYOD policies can be tailored by each individual Education

		Establishment as they see fit to either continue the same level of monitoring & filtering beyond school hours or change the restrictions. No control or monitoring over BYOD devices is available beyond the school location unless Fortinet Client software is installed or unless explicit proxy is configured on the device.
• Data retention –what data is stored, where is it (physically) stored and for how long		All web accesses (as well as any security events such as malware or intrusion detections) can be logged (generally to a centralised customer-owned log server, where log retention is under the complete control of the administrator. The data stored includes used 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 • Andriod Chromebook
• Flexibility – schools ability to amend (add or remove) keywords easily		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. Also the centralized reporting platform provides for consolidated or if the 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 every blocked URL, 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.
• Reporting – how alerts are recorded within the system?		As explained above, 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. FortiAnalyzer include flexible rules for log management including retention times, summarising of older logs, as well as compression and archiving.

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	Carl Windsor
Position	Field Chief Technology Officer
Date	05/03/2021
Signature	C.M. Windsor

AM
April, 2020