Appropriate Filtering for Education settings

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

June 2021

Filtering 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”. Furthermore, 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.”

Included within the Scottish Government national action plan on internet safety, schools in Scotland are expected to “have policies in place relating to the use of IT and to use filtering as a means of restricting access to harmful content.”

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

It is important to recognise that no filtering systems can be 100% effective and need to be supported with good teaching and learning practice and effective supervision.

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
Filtering 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.	🟨

.

Illegal Online Content

Filtering providers should ensure that access to illegal content is blocked, specifically that the filtering providers:

Aspect	Rating	Explanation
● Are IWF members	🟩	Netsweeper is a Member of the IWF
● and block access to illegal Child Abuse Images (by actively implementing the IWF URL list)	🟩	Compliant
● Integrate the 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'	🟩	Compliant

Inappropriate Online Content

Recognising that no filter can guarantee to be 100% effective, providers should both confirm, and describe how, their system manages the following content

Content	Explanatory notes – Content that:	Rating	Explanation
Discrimination	Promotes the unjust or prejudicial treatment of people on the grounds of race, religion, age, or sex.	🟩	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
Malware / Hacking	promotes the compromising of systems including anonymous browsing and other filter bypass tools as well as sites hosting malicious content	🟩	Netsweeper has a category for this
Pornography	displays sexual acts or explicit images	🟩	Netsweeper has a category for this
Piracy and copyright theft	includes illegal provision of copyrighted material	🟩	Netsweeper has a category for this
Self Harm	promotes or displays deliberate self harm (including suicide and eating disorders)	🟩	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

Netsweeper integrates and consumes various lists as well as generate it's own database. 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.

Regarding the duration and extent of logfile (Internet history) data retention, providers should outline their retention policy, specifically including the extent to the identification of individuals and the duration to which all data is retained .

Netsweeper will maintain logfiles/history for as long as required or requested by the school.

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.

Filtering System Features

How does the filtering system meet the following principles:

Principle	Rating	Explanation
● Age appropriate, differentiated filtering – includes the ability to vary filtering strength appropriate to age and role	🟩	Users are sourced from the school’s authentication system via grade.
● Circumvention – the extent and ability to identify and manage technologies and techniques used to circumvent the system, specifically VPN, proxy services and DNS over HTTPS.	🟩	Netsweeper has developed a decrypting SSL proxy that inspects all encrypted traffic including DoH.
● Control - has the ability and ease of use that allows schools to control the filter themselves to permit or deny access to specific content	🟩	Schools utilize a Web Administration system that allows schools to manage their policies and content.
● Contextual Content Filters – in addition to URL or IP based filtering, the extent to which (http and https) content is analysed as it is	🟩	Netsweeper has developed an AI system that can inspect content in multiple languages.

streamed to the user and blocked. For example, being able to contextually analyse text on a page and dynamically filter	🟩	
● Filtering Policy – the filtering provider publishes a rationale that details their approach to filtering with classification and categorisation as well as over blocking	🟩	Yes, Netsweeper publishes this information https://helpdesk.netsweeper.com/docs/7.2/Tech_Notes/CNS_Categorization.htm?rhhlterm=categorization https://helpdesk.netsweeper.com/docs/7.2/Quick_Overview/Categorization_Training/Dynamic_Content_Training.htm?rhhlterm=categorization
● Group / Multi-site Management – the ability for deployment of central policy and central oversight or dashboard	🟩	Netsweeper is built with this concept in mind and supports many different models including hub and spoke, central, distributed, etc.
● Identification - the filtering system should have the ability to identify users	🟩	Netsweeper can be integrated with a central authentication system such as LDAP, Azure AD, Novel, Google, Apple, etc, including the ability to statically map the user information.
• 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 filter system block inappropriate content via mobile and app technologies (beyond typical web browser delivered content)	🟩	Netsweeper is deployed as a network service which can detect applications protocols if the Netsweeper is also deployed to perform DPI.
● Multiple language support – the ability for the system to manage relevant languages	🟩	Netsweeper supports over 40 languages.
● Network level - filtering should be applied at ‘network level’ ie, not reliant on any software on user devices whilst at school (recognising that device configuration/software may be required for filtering beyond the school infrastructure)	🟩	Netsweeper is primarily a Network filtering solution with the ability to extend it’s filtering reach via the use of Client Filter software installed on PC, MAC, Google Chrome, IOS and Android devices.

● Remote devices – with many children and staff working remotely, the ability for devices (school and/or personal) to receive school based filtering to a similar quality to that expected in school	🟩	Netsweeper can provide remote filtering if the student/staff devices is connected via VPN to the schools network unless the remote device has the Netsweeper Client Filter software installed which does not require a VPN connection.
● Reporting mechanism – the ability to report inappropriate content for access or blocking	🟩	Netsweeper provides a comprehensive suite of pre-defined reports as well as offer reporting on a custom level. Reporting can be done and managed by the ITC.
● Reports – the system offers clear historical information on the websites visited by your users	🟩	The Netsweeper system logs details of the user activity and then allows reports to be generated including a time line report that shows access over time.

Filtering systems are only ever a tool in helping to safeguard children when online and schools have an obligation to "*consider how children may be taught about safeguarding, including online, through teaching and learning opportunities, as part of providing a broad and balanced curriculum*".[1]

Please note below opportunities to support schools (and other settings) in this regard

Netsweeper offers educational messaging to staff and students that can be used to further inform users of dangers.

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

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