Appropriate Filtering for Education Settings

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	Schools Broadband
Address	Unit 2-4, Backstone Business Park, Dansk Way, Ilkley, West Yorkshire, LS29 8JZ.
Contact Details	Telephone: 011 33 222 333 / Email: info@schoolsbroadband.co.uk
Filtering System	Netsweeper/FortiGate
Date of Assessment	August 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		Fully Compliant. Netsweeper and Schools Broadband are long standing members, supporting the IWF for over ten years, with Council representation.
• and block access to illegal Child Abuse Images (by actively implementing the IWF URL list)		Fully Compliant. The Netsweeper product integrates with the IWF CAIC illegal content list. The IWF functionality is not exposed in the webadmin graphical user interface and cannot be disabled. Importantly, Netsweeper regularly submits URLs discovered by the global systems back to the IWF team, who review the candidate material to decide on inclusion in future updates of the IWF listings. In addition, Netsweeper is also one of the first filtering companies to support the Image Hash List, delivering the most effective and efficient solution for combatting the circulation of child sexual abuse images online. Netsweeper ensures that child-abuse imagery which has previously been identified by the IWF will be identified using Microsoft PhotoDNA and blocked if it appears on a new URL.
• Integrate the 'the police assessed list of unlawful terrorist content, produced on behalf of the Home Office'		Fully Compliant. With access to the CTIRU, Netsweeper uses the UK Home Office's terrorism blocklist to block terrorist content per Government guidelines. Netsweeper integrates the list into its worldwide 500 million user cloud delivery categorising new content and offering unmatched global protection against terrorist and extremist content.

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.		Fully Compliant. Netsweeper has a category called 'Hate Speech'. These sites portray views that are written, verbal, or illustrated, and are intentionally offensive to the public. The intent of these sites is to degrade, intimidate, or incite violent or prejudicial actions against individuals based on race, ethnic affiliation, nationality, gender, sexual orientation, religion, disability, or profession. Advocacy or instructional sites that promote the harming of individuals or groups and encourage or promote peer abuse, videos of physical assaults, written harassment and threats are also included.
Drugs / Substance abuse	displays or promotes the illegal use of drugs or substances		Fully Compliant. Netsweeper has a specific category named 'substance abuse' which is blocked. These sites provide information about or promote the use of prohibited, illegal, controlled, or regulated substances for recreational rather than medicinal use. It can include sites that sell, encourage, or advocate the use of any substance that produces hallucinations, as well as the cultivation, manufacture, and distribution of any intoxicant and related paraphernalia. Informational sites that are clearly intended to provide descriptions of drugs and substances, their negative effects, and addiction potential are not included.
Extremism	promotes terrorism and terrorist ideologies, violence, or intolerance		Fully Compliant. Categories within Netsweeper that block this content include 'Extreme', 'Hate Speech', 'Criminal Skills' and 'Weapons'. Definitions can be found below:

			Extreme - This includes sites that are considered far from normal and are categorised for their degree of intensity. The content features or promotes intentional, direct, and deliberate violence and destruction or the alteration of the human body and other living creatures. These sites may depict or promote torture, self-inflicted harm, mutilation, or other dangerous activities. Images and information that advocate and glorify eating disorders, suicide, death, gore, injuries, or sites that feature grotesque or frightening descriptions are also included. Hate Speech - These sites portray views that are written, verbal, or illustrated, and are intentionally offensive to the public. The intent of these sites is to degrade, intimidate, or incite violent or prejudicial actions against individuals based on race, ethnic affiliation, nationality, gender, sexual orientation, religion, disability, or profession. Advocacy or instructional sites that promote the harming of individuals or groups and encourage or promote peer abuse, videos of physical assaults, written harassment and threats are also included. Weapons - This includes sites that provide information related to the promotion, support, sale, or discussion of weapons and any related device used in combat that can injure or kill, such as guns, knives, or swords. Information on how to build weapons or bombs will also be included in 'Criminal Skills'. Criminal Skills - This includes sites with instructions or methods that promote, encourage, or provide skills considered to be illegal, criminal, violent or harmful to the public, and are forbidden by law. This can include questionable material and sites that promote nonviolent, unethical, or dishonest behaviour such as academic cheating, or software hacking/key breaking. This does not

			necessarily reflect the laws of any particular region or country.
Malware / Hacking	promotes the compromising of systems including anonymous browsing and other filter bypass tools as well as sites hosting malicious content		Fully Compliant. Netsweeper has categories named 'Malware', 'infected hosts', 'phishing', 'viruses' and 'adware'. These categories block websites sites that are associated with this. These are sites containing scripts, or code, that may be ran in a hostile or intrusive manner to a system.
Pornography	displays sexual acts or explicit images		Fully Compliant. Netsweeper has a 'pornography' category which contains URLs that reference, discuss, or display pornographic images, videos, or other sexually oriented material that is created for the purpose of arousing sexual interest. Soft and hard-core pornography, sadomasochism, bestiality, fetishes, erotic stories, adult magazines, sex toys, or any other sexual related products are included.
Piracy and copyright theft	includes illegal provision of copyrighted material		Fully Compliant. Two distinct Netsweeper categories satisfy this requirement, i.e., "Criminal Skills" and Peer2Peer. Criminal Skills includes sites with instructions or methods that promote, encourage, or provide skills considered to be illegal, criminal, violent or harmful to the public, and are forbidden by law. This can include questionable material and sites that promote nonviolent, unethical, or dishonest behaviour such as academic cheating, copyright infringement or software hacking/key breaking. This category does not necessarily reflect the laws of any region or country. Peer2Peer (Torrents included): includes sites that distribute software and facilitate the direct exchange of files between users. Software that enables file searching, sharing, and transferring across a network independent of a central server as well as web-based site.
Self-Harm	promotes or displays deliberate self-harm (including suicide and eating disorders)		Fully Compliant. The Netsweeper "extreme" category blocks sites categorised as self-harm sites, anorexia, bulimia, and other content that prove harmful to children.

Violence	Displays or promotes the use of physical force intended to hurt or kill		Fully Compliant. Violence sites are included in the extreme category which includes sites that are considered far from normal and are categorized for their degree of intensity. The content features or promotes intentional, direct, and deliberate violence and destruction or the alteration of the human body and other living creatures. These sites may depict or promote torture, self-inflicted harm, mutilation, or other dangerous activities. Images and information that advocate and glorify eating disorders, suicide, death, gore, injuries, or sites that feature grotesque or frightening descriptions are also included.

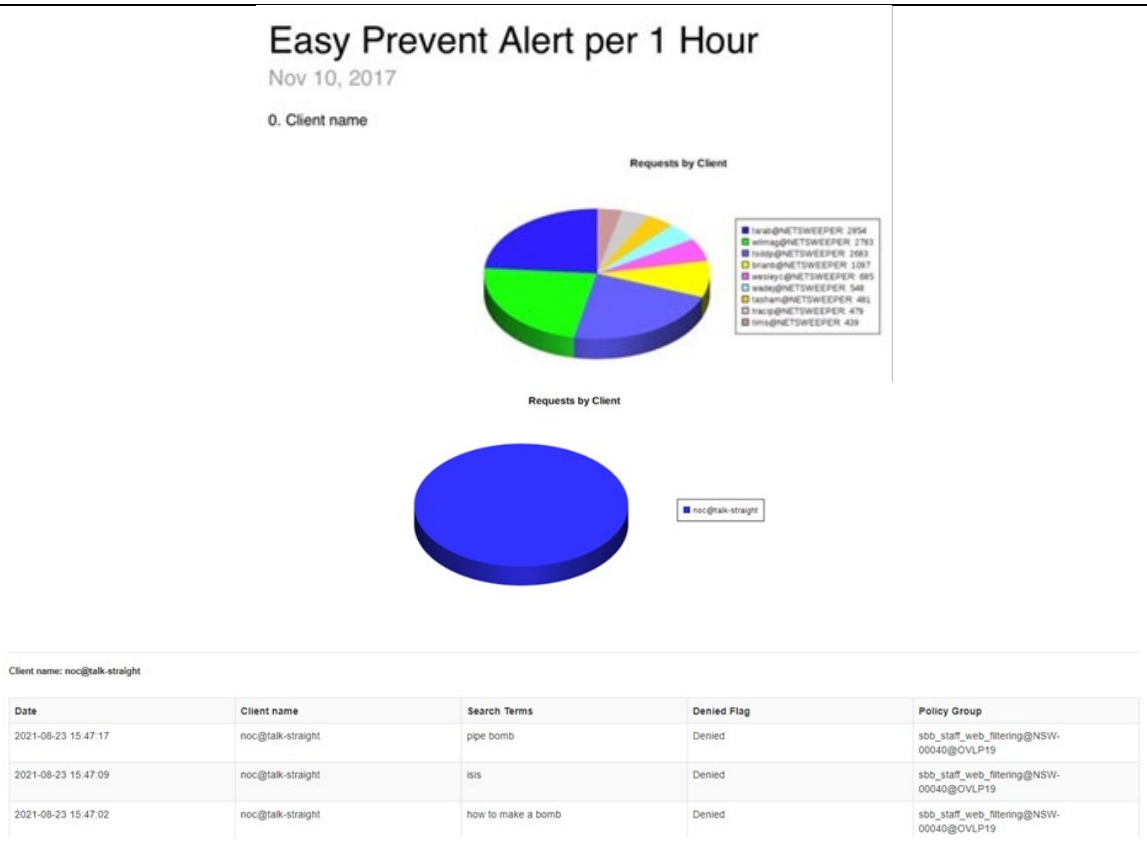
This list should not be considered an exhaustive list. Please outline how the system manages this content and many other aspects:

Netsweeper has provided filtering solutions to the UK Education market for over 15 years and is trusted to protect the networks of over 30% of schools in the United Kingdom. Offering a global collective community experience, Netsweeper resides in over 63 countries, is localised in 30 plus languages, has categorised over 10 billion URLs and is used to filter over 500 million devices worldwide. The web as we know it is consistently changing and by navigating to http://www.netsweeper.com/live-stats/ one can see in real-time the new content Netsweeper is categorising each and every day.

The value we bring to our customers is the Netsweeper collective platform where our customers experience the peer-to-peer benefits of our premium Cloud-based categorisation capabilities; through an intuitive easy to use interface enabling educational administrators to effectively deal with illicit web content on their networks.

Netsweeper offers category-based alerting meaning you can customise and create alerts to be sent to safeguarding members of staff, head teachers, IT personnel. This can be triggered automatically, so for example you can create a 'Prevent' Report as illustrated below.

This can be triggered automatically, so for example you can create a 'Prevent' Report as illustrated overleaf.



Easy Prevent Alert per 1 Hour

Nov 10, 2017

0. Client name

Requests by Client

Requests by Client

Client name: noc@talk-straight

Date	Client name	Search Terms	Denied Flag	Policy Group
2021-08-23 15:47:17	noc@talk-straight	pipe bomb	Denied	sbb_staff_web_filtering@NSW-00040@OVLP19
2021-08-23 15:47:09	noc@talk-straight	isis	Denied	sbb_staff_web_filtering@NSW-00040@OVLP19
2021-08-23 15:47:02	noc@talk-straight	how to make a bomb	Denied	sbb_staff_web_filtering@NSW-00040@OVLP19

In this example an alert has been triggered as these 5 users have tried to access one of the Prevent categories Netsweeper has. You can then drill down and find what time the specific users accessed the content as well as their location

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.

We retain logfile data for two years as per our GDPR policy.

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

Netsweeper defines policies based on the categorisation of URLs. Policies will generally deny selected categories. Policies also have override lists, if a URL would be denied by a category, the lists can be used to amend that decision to be allowed. For example: If the policy denies the 'Social Networks' category, but the administrator wishes to allow Facebook, a simple entry in the local list to allow facebook.com is all that is needed. Users with administrative rights are able to modify their policy (categories and lists) according to their permissions.

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		Fully Compliant. Netsweeper is integrated with an existing directory system such as Microsoft AD, Novell LDAP, Apple LDAP, OpenLDAP or Radius Accounting to assign users based on their group or attribute to the correct filtering policy.
• 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.		Fully Compliant. This traffic can be identified alongside standard web traffic. The end user can choose to allow or block this traffic. Netsweeper natively supports multiple tenancy and delegated administration with fine-grained permissions control. This control has two effects, it will simplify the webadmin graphical user interface removing elements of the interface that the user does not have permission for and constraining the access to site data and policies. Nominated individuals will have delegated administration for clusters of sites, and/or individual sites. The ability to manage policies and lists, and report on the associated data will be provided. The Netsweeper collective community numbers over 500 million devices worldwide. This collective together with our technology and human oversight defines URL classification. Netsweeper publishes classification of filtering and categorises on the Netsweeper website as well as a view in real time of new content categorised. This can be found at either http://www.netsweeper.com or https://www.netsweeper.com/live-stats/ Netsweeper's core competency is using our patented techniques to categorise every URL that passes through our deployed systems. Netsweeper is both real-time and employ a hierarchy of data (URLtocategory), with our Category Naming Service (CNS) as a globalmaster database. If any customer anywhere in the world accesses a URL, that URL is submitted to the local policy server, if that policy server cannot

		find a category match, it is automatically submitted to the CNS and looked up there. If the CNS already has the category mapping it is immediately returned to the local system and cached there for future use, a policy decision is then made by the policy server. If neither the local system, nor the CNS has a category match, the URL is submitted to our "Artificial Intelligence" system that will interrogate the content at-and-around that URL, assess the content, detect if it references or contains malware, and assigns one or more categories to the URL into the CNS and then back to the local system, a policy decision is then made by the policy server. The CNS allows us to adapt to trending URLs immediately due to its world-wide scope. If the local system hasn't seen a particular URL yet, then CNS probably has. If the URL has been assigned one or more categories, local systems see immediate responses (subsecond). If the URL is truly "new" then the AI will typically process the content within 20 seconds. The local policy servers can be configured with techniques to minimise the "new URL" wait.
● Control - has the ability and ease of use that allows schools to control the filter themselves to permit or deny access to specific content		Fully Compliant. Netsweeper natively supports multiple tenancy and delegated administration with fine-grained permissions control. This control has two effects, it will simplify the webadmin graphical user interface removing elements of the interface that the user does not have permission for and constraining the access to site data and policies. Nominated individuals will have delegated administration for clusters of sites, and/or individual sites. The ability to manage policies and lists, and report on the associated data will be provided.
● Contextual Content Filters – in addition to URL or IP based filtering, the extent to which (http and https) content is analysed as it is streamed to the user and blocked. For example, being able to contextually		Fully Compliant. Netsweeper has developed an AI system that can inspect content in multiple languages. Webpages are sent to Netsweeper on automatically and are examined through the Netsweeper AI engine which categorises

analyse text on a page and dynamically filter		content based on a specific set of rules. Netsweeper then updates their global CNS systems which in turn automatically updates the Talk Straight Netsweeper Platform.
• Filtering Policy – the filtering provider publishes a rationale that details their approach to filtering with classification and categorisation as well as over blocking		Fully Compliant. The Netsweeper collective community numbers over 500 million devices worldwide. This collective together with our technology and human oversight defines URL classification. Netsweeper publishes classification of filtering and categorises on the Netsweeper website as well as a view in real time of new content categorised. This can be found at either http://www.netsweeper.com or http://www.netsweeper.com/live-stats Netsweeper's core competency is using our patented techniques to categorise every URL that passes through our deployed systems. Netsweeper is both real-time and employ a hierarchy of data (URLtocategory), with our Category Naming Service (CNS) as a globalmaster database. If any customer anywhere in the world accesses a URL, that URL is submitted to the local policy server, if that policy server cannot find a category match, it is automatically submitted to the CNS and looked up there. If the CNS already has the category mapping it is immediately returned to the local system and cached there for future use, a policy decision is then made by the policy server. If neither the local system, nor the CNS has a category match, the URL is submitted to our "Artificial Intelligence" system that will interrogate the content at-and-around that URL, assess the content, detect if it references or contains malware, and assigns one or more categories to the URL into the CNS and then back to the local system, a policy decision is then made by the policy server. The CNS allows us to adapt to trending URLs immediately due to its world-wide scope. If the local system hasn't seen a particular URL yet, then CNS probably has. If the URL has been assigned one or more categories, local

		systems see immediate responses (subsecond). If the URL is truly "new" then the AI will typically process the content within 20 seconds. The local policy servers can be configured with techniques to minimise the "new URL" wait.
• Group / Multi-site Management – the ability for deployment of central policy and central oversight or dashboard		Fully Compliant. We provide a single pane of glass service where policies can be shared across multiple schools
• Identification - the filtering system should have the ability to identify users		Fully Compliant. Netsweeper sits in the core of the network and configured to use a transparent proxy. Netsweeper is integrated with an existing directory system such as Microsoft AD, Novell LDAP, Apple LDAP, OpenLDAP or Radius Accounting to assign users based on their group or attribute to the correct filtering policy. For guest/wireless networks Netsweeper can utilise Radius Accounting packets which are generated by the Wireless Access Controller to identify when a user has authenticated. Users can either be identified by the agents which can be used on various platforms such as Apple, Google & Windows, or be identified through a captive portal for devices that do not wish to have an agent installed.
• 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)		Fully Compliant. Netsweeper is deployed as a network service which can detect applications protocols if the Netsweeper is also deployed to perform DPI. As the Talk Straight Netsweeper deployment in an inline network filter, we filter on all traffic passing through ports 80/443. We do not have to add additional complexity for our customers, to implement filtering on mobile and app content. For full Layer 7 application control to facilitate full Mobile and app content inspection/filtering we provide this service through our Hosted FortiGate platform.

● Multiple language support – the ability for the system to manage relevant languages		Fully Compliant. Importantly, Netsweeper can categorise pages not only in English, but also in multiple languages. Netsweeper currently supports over 30 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)		Fully Compliant. All web-based traffic that traverse the network will have a full audit trail and be identifiable via the Active Directory source used by the trust. Users at each site will be identified via Active Directory. As a user logs into their chosen device this can vary between; Laptops, Desktops, Mobile devices, iPads and more, their details are identified by the platform and provide the correct level of filtering based on the group that user is in. These groups will be based on the role of the user and the specific site. If a user was not to authenticate on the network, as the filtering product is a network level content filter, we will ensure that all devices, authenticated or not, have a default level of filtering authorised by the trust/site. No software is required to enable authentication on devices. For example, devices that the customer does not want the Netsweeper agent installed, we can force a Captive Portal on those devices. The captive portal enables the user to authenticate using their chosen directory credentials and forces the filtering of that specified user depending on their Active Directory group.
● 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		Fully Compliant. As part of the Netsweeper solution we can facilitate the authentication and filtering of all remote/BYOD devices on the school network. Using the captive portal, we are able to give users specific filtering on any device on the school network. Furthermore, all information on these devices is tracked and monitored. This facilitates the ability to fully report on all user information for every device on the school network. For devices that are used from home, we can also facilitate the same experience as the users get in school at home. The user will be

		provided with a proxy port which will automatically be removed and enabled when the user brings the device back onto the school network. Full filtering and reporting capabilities will be facilitate for devices which are not on the school network.
• Reporting mechanism – the ability to report inappropriate content for access or blocking		Fully Compliant. Reports can trigger emails; thus, a scheduled report can be considered an alert if the report contains data (if the report contains no data take no action).
• Reports – the system offers clear historical information on the websites visited by your users		Fully Compliant. Netsweeper provides a very flexible reporting tool. Out of the box several predefined quick reports are available, these can be adapted or removed as desired. Quick reports are typically graphical and provide visibility of for example "top 10 web sites" (there are many different quick reports). The reporting tool has the concept of demand reports and scheduled reports. Demand reports are typically one-off reports for a specific demand. For example, "Can you tell me what web sites were accessed today between 10am and noon?" Scheduled reports run on a defined schedule. Scheduled reports are useful for informative infographics. For example, top 10 web sites visited this week, top 10 web sites denied this week. Reports can be graphical (pie charts, bar-charts), or detailed (tabular text), or combined. Graphical reports can be multi-layered, allowing for interactive reports where further detail can be discovered by drilling down. There are various export options (image, PDF, CSV, etc.)

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

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

Netsweeper can also be used for tracking usage of subscription-based learning tools. Using the Netsweeper reporting system schools can identify if they any students are using specific educational resources and the frequency. Netsweeper can also provide custom messaging to students when they attempt to access specific content. For example: Netsweeper can provide an information page on hate speech, drug abuse, viruses, etc rather than just simply denying them. This allows schools to message specific topics and provide students with more information on how and where to obtain help.

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	David Tindall
Position	Chief Executive Officer
Date	14th September 2021
Signature