Web Filtering Policy
E2BN Protex Web Filtering Services for Schools

Overview

First it must be noted that this policy document is an overview of E2BN policy on filtering Internet access and the Protex system. The systems put in place by E2BN and, in particular, the architecture of the filtering solution, allow academies, schools and Local Authorities with their own local Protex system to vary the policy. The policy outlined here does not negate the need for academies and schools to think through their own policies in this area.

Some would argue that there should be no filtering of Internet content and that exposure to, and education about, the dangers of the Internet must be the way forward. The main argument put forward is that as pupils can access the whole gamut of material on their home computers filtering Internet access at school provides a false sense of security. We reject this argument on two grounds: (1) the argument for pornographic or explicitly violent magazines in a school library has never, to our knowledge, been seriously proposed and we view Internet access in the same way, (2) there are good legal and educational reasons for offering a safe environment for pupils to explore the vast body of information on the internet and preventing access (whether on purpose or by accident) to unsuitable material.

Top of page

URL Filtering

E2BN subscribes to a number of commercial lists of URLs. These are regularly updated and collated into the lists used within the Protex system which are then distributed to all Protex systems overnight.

Clearly even the best URL system has its limitations as unsuitable sites are springing up all the time so it will never be 100% perfect. In fact it is even more difficult to develop a 'perfect' list as not everyone will agree what a bad site is. What is acceptable to one person may be regarded as totally unacceptable by another. Also, it is generally agreed that in education the definition of 'inappropriate' will change depending upon the age of the user.

The URL list is divided into categories which E2BN uses to refine the filter profiles. So, for example, while porn is blocked in all categories (including Staff) other material with an adult audience is blocked in the student profiles but allowed for staff.

The Protex system offers a way to adjust the E2BN Protex central lists to the local requirements. Each Protex server maintains a set of local lists which are added to the standard distributed ones. At its simplest this allows an academy, LA or school with its own Protex server or to add URLs to, or remove them from, the lists provided by E2BN.

The typical terms used in URL based web filtering are 'whitelists' and 'blacklists'. The terms we use when talking about filtering are **Trusted**, **Blocked**, and **ContentChecked** which are more self-explanatory than the typical terms used.

Top of page

Trusted sites

Making a site Trusted has two effects.

Firstly: if the URL would normally have been blocked then this will be overridden and the page will be returned to the client browser. Any subsequent update to the lists will have no effect on a trusted site (for example a trusted site which at some later date is added to the main block lists will still not be blocked - the trusted listing overrides the block list entry.

Secondly: by making a domain or site Trusted you are explicitly trusting any downloadable files it contains as no extension blocking takes place.

Making a site trusted is not to be done lightly as you must really trust the site never to contain or distribute inappropriate material. Parts of sites may be trusted and others not as long as they can be distinguished by URL - for example while mydomain.com may not be trusted you may feel that mydomain.com/education can be.

E2BN has made the decision to trust certain categories of site (.gov.uk; .sch.uk; for example) and some specific sites (e2bn.org is one such others being various education sites,

manufacturer sites, etc.). New trusted sites can easily be added.

Blocked sites

Adding a domain or site to the blocked list bans the site from being viewed. This is simple and straightforward. The only thing to remember is that this, like the Trusted list, will override the main URL lists.

ContentCheck sites

If a site would normally be blocked by the distributed lists but you want your clients to be able to access some or all of the site you have two options. (1) add it to the Trusted list or (2) add it to the ContentCheck lists. You may feel that you cannot fully trust a site but you do not want to ban it completely. Domains or sites added to a ContentCheck list will override a block list but the returned pages' content is analysed and either blocked or allowed depending on the user's filter profile (see below).

Also, files downloaded from ContentCheck sites will be subject to the extension blocking rules in the profile. (i.e. student users will not be able to download .zip files from ContentCheck sites but staff users can - assuming they are using the Staff profile).

All sites that do not appear in any of the active lists are also content checked.

Top of page

Content Filtering

How is the returned page's content checked? This is where the Phraselists come in: the HTML of the page is scanned for various phrases and patterns. There are two types of phrases: those which are either banned or weighted. If a word or phrase in the web pages matches any item in a banned list then the page is blocked. The items in the weighted lists (which are also categorised) all have a numerical value: the items found on the page are totaled to give the page a value which is used to rate the page. Each profile (see Profiles) has a variable called the naughtinesslimit (not our name!) which can be changed to reflect the age group.

Search engines

Searches utilise various other features of Protex which are beyond the scope of this document but include the ability to scan the submitted search and block unsuitable search terms.

Image searches

Image searching is a very powerful tool but some schools have had to ban it because of the nature of some of the thumbnail images displayed to the unwary. E2BN Protex addresses this problem in two ways. Firstly safe search is enforced for all pupil filtering profiles when common search engine are accessed.

However, even safe search is not foolproof. So, in addition, our system tests the URL of the originating site of each image returned. If Protex finds that it is a site which would be blocked to this user then the returned image is replaced with a blank one. Clearly, if the user clicks on the blank to go to the site it is blocked by the URL filter.

This technique can only be applied to image search engines that include the originating URL in the results page.

Due to the danger of young pupils using unfiltered search engines at home, we recommend the younger students are directed to search engines such as picsearch which are specifically designed for families and to be child friendly.

Top of page

Webmail

Webmail falls into two categories: (1) a web based front end to an email system controlled by the academy, school or the LA and (2) a publicly available webmail system the best known of which are Hotmail and Yahoomail.

Our policy on these is actually very simple although the effects of it may be contentious to some users.

Access to category (1) systems is **trusted**. Assuming the mail system used by a academy, school or LEA is using a .gov.uk or .sch.uk domain name then these are
trusted
by virtue of this domain name. If another domain is being used it can also can be added to the
trusted
list on request.

Public webmail systems (Hotmail, Yahoo) are treated in exactly the same way as other websites.

OK, some explanation and elucidation may be necessary here because email is such an important tool: what does "treated in exactly the same way as other websites" actually mean to a user? For the sake of clarity let us assume you have just opened a new hotmail account. When you access this account you are presented with a web-page with icons linking to an Inbox, Sent Messages, Drafts, etc. on the left hand side and your messages will appear in the main body of the page as a list of from addresses and subject headings. All well and good so far. Let us imagine for a moment a utopian vision where there is no spam mail and no idiots sending mail with swearwords in the subject line. In such a world hotmail would work fine. It will not get blocked by the filtering system because the web-pages generated are innocuous.

Even emails containing swearwords would be visible at this stage as ***only*** the subject lines are presented on the page. Now we see the power on content filtering (as opposed to URL blocking). Let us suppose one of the mails with an innocent subject is actually an email sent to a student from a bully using foul language. When the student tries to open this mail it is blocked by the content filter. Why? Because the web-page containing the mail has been checked by the Content Engine and the phrases used therein have pushed the page over the 'naughtiness' limit assigned to that profile. As you see the profile of the student affects whether the page is blocked or not: a Secondary student will see some mails that for a Primary student would have been blocked.

If people could restrain themselves from using inappropriate language in the subject lines this would all be OK. Good mail gets through, bad mail does not. However, as we all know life is not like that.

First, let's get the simple case out of the way: student 'A' receives a mail with a very inappropriate title which the phraselists weight at over the user's limit. What happens now? 'A' logs into their account and now cannot access their Inbox as the page listing the subject lines in their Inbox has been blocked by the phraselist weighting. What can they do? The best thing would be to go to a member of staff, ask them to log in to A's account so they can access the Inbox. The member of staff notes the address of the sender and then deletes the mail so that the student can access their own account again. The alternative is that the student waits until the evening and does the same thing at home.

Suppose the sender of this mail ('B') is another student at the school (and if hotmail is generally used by staff and students as their main email system this may very well be the case). In this case the teacher above would have their mail address and can talk to them about the mail and how it breached the school's AUP (Acceptable Use Policy). They may even impose sanctions on the user depending upon the content of the mail. Also 'B' is blocked from his/her Sent Messages box while at school for the very same reason 'A' is blocked from their Inbox.

If the sender is not a pupil then clearly it is harder (if not impossible!) to remonstrate with them - but do you really want outsiders sending inappropriate mail to your pupils? And, more to the point, your pupils receiving and reading them? This is exactly the point of filtering the content of the mail - an email communication is being treated in exactly the same way as any other web content.

Now we come to the issue of Spam mail. ***If*** the likes of hotmail and yahoo mail did ***really*** effective spam filtering (the mail was transferred to the Junk folder and then deleted it after a certain time) everything would be OK. However, this is not the case and much of the spam gets through to the Inbox where, by the very nature of much of the spam mail, it throws the content filter over its threshold and the user is blocked from their Inbox. (And, obviously, they also can never access the Junk mail folder at school.)

There is no easy answer to this. E2BN policy is that, given all that has been said above about Internet safety and the nature of the spam causing the block, any further loosening of the filtering policy for Hotmail, Yahoomail, etc. is a matter for individual LEAs or schools to decide.

What could be done? There are only two ways to address this - the choice is yours: (1) make the 'naughtiness' limit higher for each user profile. This will mitigate the webmail problem but will at the same time make access to other unsuitable sites more likely. You are making the whole web-filtering profile looser with all that implies. (2) depending on the exact mail system and how it is structured you **may** be able to add certain URLs to the **trusted** lists to, for example, give unfettered access to the Inbox and all the email it contains - but this would also allow ***any*** attachment to be downloaded. Hotmail seems to work in such a way that you would need to trust *hotmail.msn.com* in its entirety: not something we would do globally.

A word about attachments. The same extensions are available for download from webmail as from any other website and will depend upon the profile being used. In particular, this means that staff can download most attachments but students can only download certain acceptable extensions. This includes, for example, all the video and audio extensions but excludes .exe & .zip. But remember that users of **trusted** webmail systems can up and download all file types. So, for example, a student moving files between home and school using zip may have a problem if using Hotmail but would be OK with an approved webmail system.

Top of page

Profiles

The current version of Protex (v5) has 15 standard profiles, including ones for Primary, Middle, Secondary, Sixth Form, and Staff and Walled Garden and Games versions of the three student profiles, and two profiles for public libraries. Each system can be configured to use selected profiles on different ports, locations, and different AD user groups (if using AD Authentication).

The most obvious difference between the profiles is the 'naughtinesslimit' - this is set very low for the PRIMARY profile giving the most restrictive setting and increases on each of the other profiles as the age of the audience increases.

The second important difference is that the student profiles are much more restrictive on the types of file that can be downloaded. This is described in more detail below.

Finally the categories of URLs are slightly different between profiles. For example, while the category "porn" is blocked for all profiles URLs in the category "adult" are allowed for staff but blocked to all students. An overview of the categories that are allowed and blocked for each profile is available for download here .

Top of page

File Extensions

Protex enables the downloading of files from websites to be controlled via their file extension or mime types. By default we do not restrict any of the well known media types (mp3, mp4, mpeg, avi, rm, etc) on any profile. On the student profiles we do block the download of several files & mime types from sites which are not trusted: in particular .doc & .zip files are blocked. A full list of types taken from a student profile can be found here: banned extension list & banned mime types. Note that lines beginning with the # character are not blocked. In the staff profile a very few extensions are blocked while the rest are not.

E2BN had originally blocked the .exe extension but this is now allowed as some sites require .exe files to be downloaded to provide full functionality. We regard this as a security hole and would prefer these files to be blocked so academies and schools must make sure they have other systems in place to prevent viruses and other malware being downloaded and installed by students.

Of the items in these lists please remember that blocking will only apply to sites which are not trusted: if it is important that students are able to download files with these extensions then they will be able to do so if the site is added to the trusted list. If a site is not trustworthy enough (see Trusted sites) to be added to the Trusted list then as a matter of both network security and child protection pupils should not be given free rein to download files from it.

E2BN has, as has been stated above, already added a variety of sites to the Trusted list - both specific sites (bbc.co.uk; sophos.com; e2bn.net; etc.) and generically (.sch.uk; .gov.uk) - which permit all file downloads.

It is worth repeating here that zip files can be downloaded from any site via the Staff profile and it is therefore very important that school systems managers make provision for staff (both teaching and technical) to have access to the Staff profile.

Top of page

Online Games

Another of the benefits of Protex is the ability to ban online games. E2BN had discussions with schools about this decision and it was universally agreed within the test schools than online games should generally be barred to pupils as it is not considered to be an appropriate activity.

Some of the games, as well as being great time wasters, are not suitable for younger students and can also clog up the limited bandwidth. Schools that have their own Protex system can opt to use the 'with games' option where they think it appropriate, either full-time or on a timed basis (e.g. lunch times). To find out more about this please see the Online Documentation .

We have unblocked the audio and video types as we believe the great educational potential here outweighs any possible downside. Clearly these file types can only be downloaded from sites which are not blocked for some other reason.

Top of page

Blogs, etc.

Open blog sites (and other similar online tools) are, in general, not supported by E2BN for school use. We recognise that these are very valuable tools and will do what we can to make sites available on a case-by-case basis consistent with child safety and both our and our member LAs' responsibility to provide a safe environment for students.

Why are they banned? Simply because they are commercial, uncontrolled, globally available web-spaces many of which contain material unsuitable for school viewing. In particular they often contain pictures and images that cannot be filtered just by the textual content of the page.

Can I use them at all? Maybe. While these sites are generally blocked particular sub-domains in some cases it is possible to make particular sites (myblog.blogspot.com) available by adding them to the Trusted or ContentCheck lists. Just send the comment form on the block page and we will consider the request and make it available if possible and appropriate.

These and similar hosting companies sites can only be made available if a unique URL is associated with your account.

Flickr is a special case - because of the structure of the site it has proved difficult to make sub-domains available on a per-account basis.

NB Academies and schools with their own Protex systems are not constrained by this and can modify Protex behavior to suit their own policies on the use of blogs.

Top of page