

# О фильтрации школьного Интернета

В обсуждении этой темы предлагаю выделить два направления

1. Реально существующие проблемы предотвращения доступа детей к нежелательной информации в Интернете.
2. Нереально казённая проверка, произведённая прокуратурой.

## 1 Реально существующие проблемы

Вне зависимости от адекватности проверяющих и релевантности результатов проверки (а они были и неадекватны, и нерелевантны, но об этом позже), проблема интернет-фильтрации не надуманная. Действительно, в Интернете полным-полно страниц, посещение которых детьми (детьми!) является как минимум нежелательным, а то и вообще катастрофой.

Для предотвращения доступа детей к нежелательным ресурсам в Интернете разработаны различные методы фильтрации. Рассмотрим один из них — тот, на котором основана фильтрация СКФ, то есть Системы Контентной Фильтрации, созданной на федеральные деньги и внедрённой во всех общеобразовательных учреждениях в приказном порядке. Кстати, именно эта система была установлена и, в строгом соответствии с инструктивным письмом, настроена и в полном объёме функционировала в вечерней школе, когда г. Иванов производил свою эпическую проверку.

Итак, не вдаваясь в технические подробности:

### 1.1 Чёрный и белый список

Существует два основных принципа фильтрации — «чёрный список» и «белый список». «Чёрный список» означает, что есть некий список запретных страниц, на которые заходить нельзя, а кроме этого списка на все остальные страницы можно. «Белый список» наоборот, означает, что есть некий список разрешённых страниц, и больше никуда нельзя, кроме этих разрешённых страниц (как в Северной Корее, Иране, Пакистане и т.п.)

### 1.2 Домены

Структурными единицами Интернета являются домены. Что такое домены, сейчас объяснять не буду, просто представьте себе, что Интернет — это некое пространство, а домены — точки в этом пространстве. Существуют и другие структурные единицы, например, страницы или сайты, но во многих системах фильтрации, и, в частности, в СКФ, именно домены являются единицами фильтрации.

Интернет — принципиально децентрализованная система; децентрализация — один из основополагающих технических и организационных принципов Интернета, одна из основ его надёжности и, хоть и слабая, но всё же гарантия его неподконтрольности. Из-за децентрализации нет абсолютно точной статистики, но статистика, собираемая агентствами интернет-мониторинга, нас вполне устроит. Одно из таких агентств — NetCraft.

По их данным, на весну 2009 года в Интернете было зарегистрировано 250 000 000 доменов.

### 1.3 Категоризация

Чтобы осуществлять фильтрацию, необходимо сначала осмотреть интернет-ресурсы (например, домены) чтобы разделить их на «хорошие» и «плохие». Это называется «категоризацией» доменов и осуществляют эту деятельность эксперты. Эксперты, либо порознь (тогда задействуют механизм голосования экспертов), либо коллегиально,

осматривают домен, и прикрепляют к нему своеобразные ярлыки — «автомобили», «спорт», «путешествия», «порнография», «наркотики», «насилие». Затем администраторы конечного участка в системе фильтрации (например, учителя в классе или родители на домашнем компьютере) указывают, какие категории составляют белый список, а какие чёрный, и включают один из них.

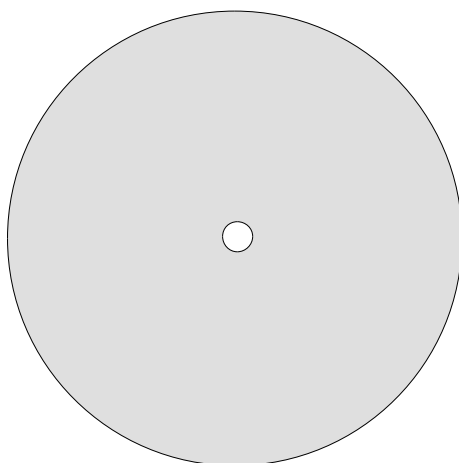
## 1.4 Корень проблем

Итак, на весну 2009 года в Интернете было зарегистрировано 250 000 000 доменов. В то же время в федеральной системе контентной фильтрации (СКФ) осуществлена категоризация 1 200 000 доменов. Миллион двести тысяч — вроде неплохо? Да, пока не поделишь на двести пятьдесят миллионов:

$$\frac{1\,200\,000}{250\,000\,000} = 0,0048 = 0,48\%$$

Меньше половины процента!

Чтобы было нагляднее, я изображу это схематически: большой круг это все домены Интернета, а маленький внутри него — это категоризованные домены, то есть уже осмотренные экспертами и получившие свои ярлычки, «хорошие» или «плохие»:



Кружочек внутри — это осмотренная экспертами и категоризованная часть доменов Интернета, менее 0,5%.

Аналогичные статистические показатели у западных систем фильтрации, а там эту проблему осознали намного раньше и занимаются ею по меньшей мере 20 лет. Запомните, это важно: 20 лет тысячи профессионалов и сотни тысяч добровольцев в западных странах, и уже три или четыре года в нашей стране трудятся над категоризацией доменов, а продвинулись лишь на полпроцента.

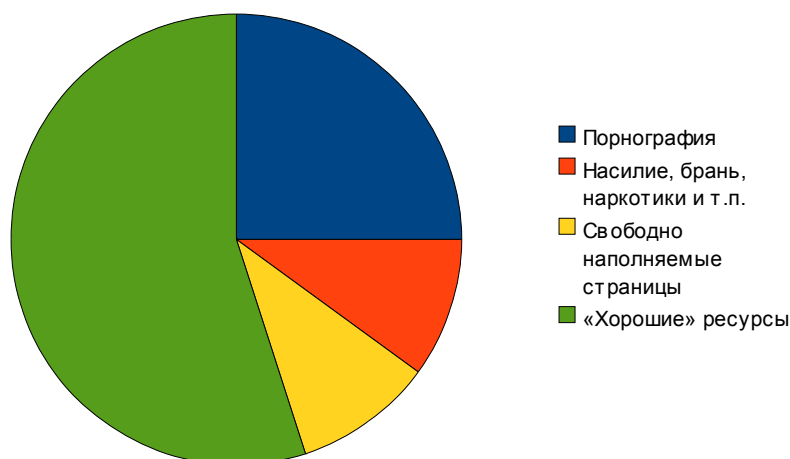
## 1.5 Плохое и хорошее в Интернете

По данным тех же агентств интернет-мониторинга, 20% доменов напрямую посвящены порнографии, ещё 5% содержат порнографию в качестве второстепенных материалов, не менее 10% тематически неприемлемы для детей по другим причинам — одобрительное отношение к употреблению наркотиков, призывы к насилию, побуждение к суициду, нецензурная брань и т.п. И это только тематически определившиеся домены.

А ещё в Интернете существует и ширится практика свободного наполнения страниц, когда содержимое страниц формирует в той или иной степени не держатель домена, а его

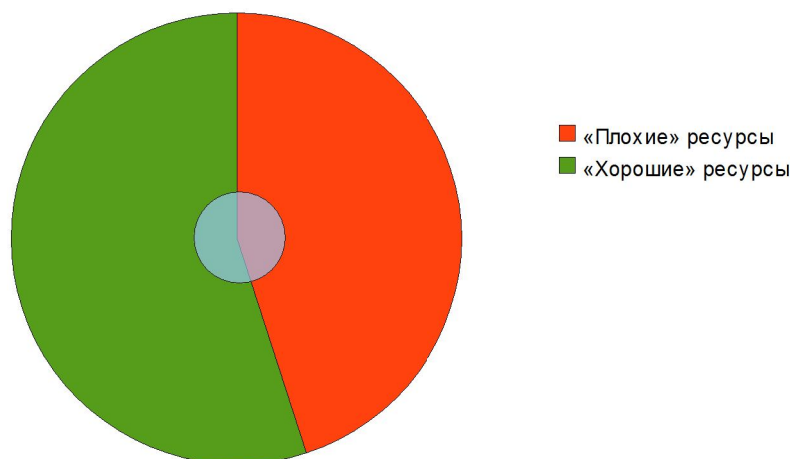
пользователи, то есть посетители страниц — таковы социальные сети, блоги, живые журналы, сетевые дневники, форумы, конференции, чаты и т.п. И даже традиционные интернет-издания со строгой тематикой и многолетней репутацией теперь считают обязательным под каждой публикацией поместить ленту комментариев, в которой читатели могут высказаться о прочитанном/увиденном, поделиться своим личным опытом или видением проблемы. Но также могут поместить в комментарии брань, рекламу порносайта или призывы к насильственным действиям в отношении отдельных представителей власти.

Домены со свободно наполняемыми страницами составляют не менее 10% современного интернета и их доля быстро увеличивается. А доля самих свободно наполняемых страниц ещё больше.



Интернет — не библиотека. Интернет — городская площадь, форум, торжище и толковище. Кстати, на тот самый Форум, который древнеримский, детишек вообще не пускали.

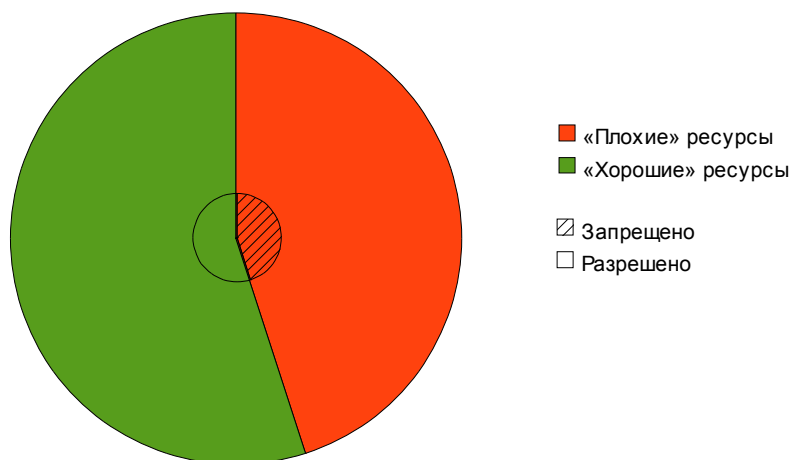
### 1.6 Первая реальная проблема: Выбор «чёрного» или «белого»



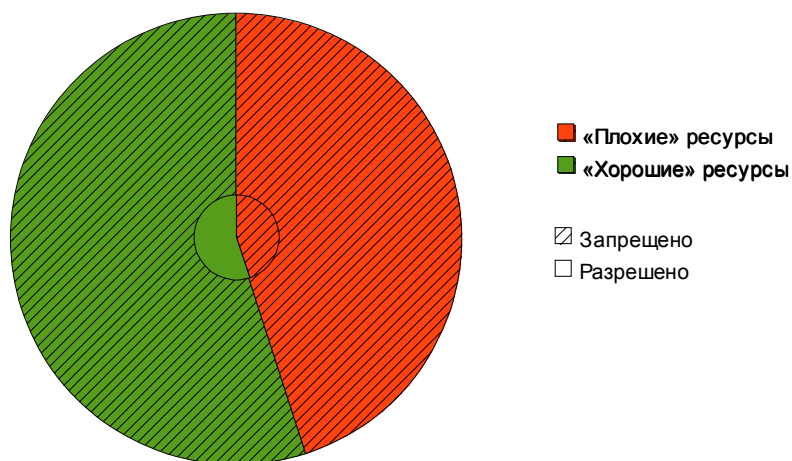
Внутренний кружок опять обозначает, какую часть интернет-доменов уже удалось категоризовать. Только пришлось увеличить его примерно в 10 раз, а то невозможно разглядеть секторы.

Коротко суть дилеммы: Если мы используем «чёрный список», есть реальный и принципиально неустранимый риск, что дети увидят что-то непозволительное. Если используем «белый список», Интернет-пространство сжимается до размеров «дворика под окном», и стоило ли тогда городить огород?

Вот как выглядит применение «чёрного списка»: заштрихованная область — это «плохие» домены, доступ к которым будет заблокирован, незаштрихованная область — это домены, доступ к которым будет разрешён, причём это как «добрые» домены, осмотренные и не осмотренные экспертами, так и «злые» домены, до которых у экспертов пока что попросту не дошли руки, глаза и что там ещё.



А так выглядит применение «белого списка»: отсечены все «плохие» домены, осмотренные и неосмотренные экспертами, но зато и из «хороших» доменов доступными являются лишь полпроцента (а на самом деле ещё меньше, потому что обычно эксперты предполагают, что будет применяться именно «чёрный список» и сосредотачиваются на отсмотре и маркировке именно «плохих» страниц, то есть внутренний круг — напоминая, на схеме он в 10 раз больше реального — ещё и сдвинут сильно вправо, в «красную» сторону, что ещё сильнее уменьшает процент доступных «хороших» доменов).



Ситуация усугубляется ещё тем, что экспертным сообществом СКФ поисковые машины не включены в белый список. И у этого идиотского решения есть резоны, особенно принимая во внимание подходы, продемонстрированные нашей прокуратурой: и правда, ведь поисковые машины обязаны точно и беспристрастно выдавать список ссылок на все известные им сайты и страницы, где встречается запрошенная пользователем текстовая строка. А пользователь, гад он этакий, постоянно у этих поисковых норовит машин спросить что-нибудь такое, что не прошло высочайшего одобрения. Однако **практически вся работа в современном Интернете основана на регулярном и систематическом пользовании поисковыми машинами.**

«Выбирай, но осторожно. Осторожно, но выбирай». (Жванецкий)

На практике «чёрный список» даёт превосходные результаты, поскольку, чтобы добраться до незаштрихованной красной области, то есть до незаблокированных «плохих» доменов, ребёнку **одновременно** необходимы: (1) намерение, (2) упорство в его достижении, (3) запас свободного времени, (4) некоторый интернет-опыт и, (5) самое главное — бесконтрольность.

Я бы сказал, что «чёрный список» обеспечивает **разумную безопасность**.

Затруднение заключается в том, что вместо разумной безопасности последнее время стали, по западному образцу, стремиться к безопасности абсолютной. **Абсолютная безопасность абсолютно недостижима**. В частности, гарантированно предотвратить посещение детьми неподобающих страниц в Интернете не поможет даже обесточивание школы — Opera Mini рулит.

### **1.7 Вторая реальная проблема: Проблема ответственности**

Если помните, за 20 лет труда многие тысячи людей на Западе и у нас продвинулись в категоризации доменов аж на целых полпроцента. Очевидно, непосредственно в школе эта задача абсолютно никому не по силам. Значит, придётся пользоваться списками — неважно, белыми или чёрными, — составленными не в школе, а взятыми откуда-то извне. При этом уже категоризовано 1 200 000 доменов — между прочим, это означает, что даже убедиться в правильности категоризации, выполненной посторонними специалистами, у директора нет физической возможности. [ Представим себе, что директор каждый день четыре (!) часа осматривает категоризованные домены, затрачивая на каждый пять минут. Получается 48 доменов в день.  $1\,200\,000 : 48 = 25\,000$  дней  $\approx 70$  лет без выходных и отпусков ]

**Но ответственность за фильтрацию интернета по этим чужим спискам прокуратура желает возложить именно на директора лично!**

### **1.8 Второстепенные проблемы**

**Проблема уровня фильтрации** — из-за одной страницы предполагается отсекается целый домен. В СКФ в качестве единицы фильтрации выбраны домены, однако в пределах одного домена могут располагаться тысячи сайтов и миллионы страниц. Сайты управляются разными людьми, у них разные авторы и разная тематика. Скажем, там тысяча страниц приемлемых для школы без всяких замечаний. Одна страница неприемлемая. Блокировать предполагается весь домен, со всей тысячей приемлемых страниц.

Существуют системы, в которых единицами фильтрации являются меньшие объекты — сайты и даже отдельные страницы. Более того, существуют даже способы подлинной контентной фильтрации, то есть фильтрации по содержанию страницы. Таким образом, уровень фильтрации может быть разным — по доменам, по страницам, по фрагментам страниц.

**Проблема места фильтрации** — на протяжении трёх лет с регулярной настойчивостью от нас требуют установить программу ПКФ на каждый компьютер, а между тем это совершенно лишено смысла, так как фильтрация, в точности совпадающая с рекомендуемой фильтрацией ПКФ, осуществляется нашими провайдерами — связистами Ростелекома, по контракту с Министерством образования. Даже если организация подключена, минуя «президентский» интернет, несложная настройка внутреннего DNS-сервера опять-таки обеспечивает точно такую же фильтрацию, как и ПКФ с рекомендуемыми настройками.

На этом фоне особенно любопытно, что под прокурорскую «раздачу» попала именно та организация, где все требования министерства об установке и настройке ПКФ были выполнены безукоризненно, слово в слово. А единственный исполнительный информатик в городе получил за это взыскание.

**Проблема преодолимости.** Любая фильтрация, организованная техническими средствами, преодолевается другими техническими средствами на счёт «три».

## **1.9 И в заключении он сказал...**

Мы, технари и информатики, привыкли, что Интернет — это территория свободы. Нам нелегко принять тот факт, что свобода эта ограничивается, и уж тем более нелегко самим возводить колючую проволоку.

Постулаты:

1. Фильтрация интернета для взрослых пользователей не осуществляется. Здесь «взрослые» — это совершеннолетние сотрудники школы.
2. Для технических целей Интернет должен быть доступен в полном объёме.
3. Фильтрация в случае постоянного присутствия наблюдателя-педагога не требуется. В отсутствие учащихся фильтрация не требуется.
4. Во время выхода в интернет на уроке по заданию следует включать политику «чёрного списка» или отключать фильтрацию совсем.
5. Поисковые машины должны быть включены в белый список в обязательном порядке, по крайней мере Яндекс с включённым «семейным поиском» и Google с включённой «строгой фильтрацией».
6. Руки прочь от Википедии! Википедия должна быть доступна любому пользователю, включая детсадовцев, в полном объёме. То же касается других энциклопедий общего назначения (неотраслевых).
7. Домены, включённые в «Каталог образовательных ресурсов сети Интернет» (<http://katalog.iot.ru/>), подготовленный институтом «Информика» по заказу Министерства образования, должны быть включены в белый список в полном составе.
8. В случае введения в школе политики «белого списка» решение о включении в список новых доменов должен принимать только директор школы на основании письменных заявлений от желающих.
9. Претензии о некачественной фильтрации по базам федеральной СКФ следует адресовать экспертному сообществу. Претензии о наличии доступа к ресурсам, признанным в установленном законом порядке экстремистскими, следует адресовать провайдеру, поскольку предотвращение доступа к таковым ресурсам является обязанностью оператора связи (Ростелекома), а не конечного потребителя (школы). Сотрудники школы не имеют возможности заниматься категоризацией доменов или поиском недостатков фильтрации.

Дополнительно ещё замечу, что, по моему мнению, протоколы взаимодействия клиентов фильтрации с серверами СКФ должны быть открыто опубликованы, чтобы независимые разработчики получили возможность создавать альтернативные программы фильтрации с опорой на ту же федеральную СКФ. База адресов, пополняемая энтузиастами-общественниками и поддерживаемая на бюджетные деньги, не может быть технически доступна только одной компании в стране (и в мире).

## 2 Нереально кафкианская проверка

Для начала позволю себе самоцитирование. В беседе сквозь Интернет с одним выпускником я описал проверку, насколько смог себе представить по рассказу учительницы информатики:

**Петрович**

Теперь обещанное о визите прокурора.

Летом пришёл к ним в вечернюю школу прокурорский работник с характерной и редкой фамилией Иванов. Проверять фильтрацию интернета. Следи за маразмом:

Вечерняя школа, все до единого учащиеся — совершеннолетние.

Лето. Ученики на каникулах.

Но если бы и были, напоминаю, они все совершеннолетние.

Школа закрыта на ремонт.

Компьютерный класс опечатан.

И обесточен.

Компьютеры зачехлены.

Учительница информатики в отпуске.

Иванов требует отозвать из отпуска учительницу информатики.

Причём немедленно, пока он тут.

И вскрыть класс.

И включить компьютеры.

Садится за один из них и в Яндексе набирает «эмо».

Яндекс выдаёт кучу ссылок.

Первая ссылка на Википедию.

Иванов встромяет указующий пёрст в эту строчку и изрекает:

«Вот видите, дети у вас могут попасть на деструктивный сайт».

**Собеседник**

и вы его не убили?

**Петрович**

Господь помиловал меня. Я там не был.

Далее Иванов проделал то же самое со словами «готы», «свастика», «бей евреев» (идиот! надо — «бей жидов», какой же антисемит напишет «евреев»), «как изготовить бомбу».

Становящемуся на глазах знаменитым русскому юристу Иванову в довольно резких тонах предложили осмотреть школьную библиотеку на предмет обнаружения свастики, каковая там имеется в каждой второй книжке о войне, а также в Большой Советской Энциклопедии, Новой Российской Энциклопедии, Энциклопедии Кирилла и Мефодия, Энциклопедическом словаре школьника и рекомендованных министерством образования учебниках истории. Юрист с достоинством ответил, что указаний о проверке школьных библиотек ему не поступало.

**Собеседник**

ппц, и как дальше жить с такими идиотами?

**Петрович**

На основании этой чудо-проверки советник юстиции (!) с характерной и редкой фамилией Иванов (заметь, не какой-нибудь старый маразматик; сказали, лет тридцати он, вряд ли сорока) выносит представление об устранении нарушений и наказании виновных.

**Собеседник**

и как он это обосновал?

**Петрович**

Законом о противодействии экстремизму.

**Собеседник**

жжжжееесть

**Петрович**

Вот это точно можно отправлять куда угодно, хоть на Баш, хоть на {sh}It happens, хоть на сайт к Медведеву. Потому что пц.

Ну, имеется, конечно, некоторое нагнетание в рамках литературной обработки, но суть дела передана верно. Итак:

1. Проверка фильтрации летом, в отсутствие учеников, сходна с наложением взыскания за свисающие с потолка провода во время ремонта, выполняемого электриками на каникулах, на том основании, что дети (которых нет) могут получить электрический удар.
2. Как я убедился, программа фильтрации установлена и настроена в строгом соответствии с указаниями министерства образования. Программа работает, проверки на правильность работы, описанные в том же инструктивном письме, отрабатываются программой безупречно. Таким образом возникает вопрос, что же проверял господин Иванов? Реально выполняемую фильтрацию? Тогда вопрос — как школа должна была проверить эту самую фильтрацию? По методике министерства? Так и было сделано!
3. Где опубликована применённая господином Ивановым методика проверки, кто и когда её разработал, кто и когда утвердил?
4. На основании какого юридически значимого документа господин Иванов объявил деструктивным сайтом всемирную свободную энциклопедию «Википедия»? Следует ли нам сообщить.wikisoобществу, что они официально признаны в России деструктивным сайтом? И в какой именно части России?
5. Чем помешали господину Иванову безобидные эмо и древние готы?
6. Отличает ли господин Иванов профашистскую литературу от энциклопедии или он не собирается утруждать себя столь сложной мыслительной деятельностью?
7. Следует ли одолжить господину Иванову бензинчика для разжигания на городской площади костра из книг, внутри которых содержится упоминание или изображение свастики — древнего мистического символа, осквернённого в двадцатом веке одними преступниками и преследуемого теперь другими преступниками?
8. В курсе ли господин Иванов, что найденная им ссылка по запросу «бей евреев» является заголовком статьи на сайте газеты «Комсомольская правда» — зарегистрированного органа массовой информации, — а сутью статьи является осуждение антисемитизма, как одной из форм экстремизма?
9. Известно ли господину Иванову, что ДПНИ — движение, которое он, по моему мнению, справедливо считает экстремистским, — является официально зарегистрированной организацией и выиграло несколько судебных процессов в разных регионах страны, как по обвинению в экстремизме, так и по закрытию изданий и сайтов?
10. Ну и, наконец, последний вопрос: Почему, обнаружив недопустимые ресурсы, господин Иванов **скрыл их адреса** от технического персонала и руководителя организации, тем самым лишив нас возможности исправить обнаруженные недостатки хотя бы в этой малости?