Network Working Group D. New, N. Borenstein
Internet Draft First Virtual Holdings
June, 1995

KidCode: Naming Conventions for Protecting Children on the World Wide Web and Elsewhere on the Internet Without Censorship

Status of This Memo

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet- Drafts as reference material or to cite them other than as ‘‘work in progress.’’

To learn the current status of any Internet-Draft, please check the ‘‘1id-abstracts.txt’’ listing contained in the Internet- Drafts Shadow Directories on ds.internic.net (US East Coast), nic.nordu.net (Europe), ftp.isi.edu (US West Coast), or munnari.oz.au (Pacific Rim).

Abstract

This memo suggests a "KidCode" naming convention to be used to inform web browsing programs when the material being made available on the web is of a nature that might be unsuitable for children. Some aspects of the convention are also applicable to browsing by FTP, to news reading, and to other protocols as well.

Two separate conventions are specified. One convention is to be used for fixed URL’s, where the permanent URL is itself used to specify the nature of the child-oriented restrictions. Another convention is to be used when the server site desires to use a simpler or pre-existing URL, with restrictions enforced using a naming convention in conjunction with the standard web forms mechanism.

A web site or server is said to be KidCode-compliant when it conforms to one of the KidCode naming conventions for all material that might be considered inappropriate for children. A web client or browser is said to be KidCode-compliant when it blocks access to URL’s that are labelled as inappropriate using the KidCode naming conventions, or when it offers an easy optional mechanism permitting adults to block children’s access in that manner.

The techniques discussed here are open, voluntary, and non-proprietary. They may easily be used in conjunction with explicit "blacklists" of known forbidden sites, and this is a recommended technique for dealing with offensive sites that refuse to adopt these naming conventions.

It is also anticipated that third-party authorities will be desired for the certification of compliance with these techniques. A procedure for declaring such certifications is also described here.

Introduction

With the explosive growth of the World Wide Web [1,2,3], it increasingly appears that one can find virtually anything on the Web. While this is generally viewed as a good thing, many people worry about letting children browse the web, given the growing availability of adult-oriented material.

In response the "anything goes" atmosphere of today’s Internet, there are growing voices advocating government censorship and regulation, in the name of protecting our children. This is an understandable, if possibly misguided response, if the only alternative to censorship is the abandonment of any attempt to permit parents to choose to protect their children from viewing adult materials. In the physical world, however, there is commonly a middle ground. Materials that are permitted to adults are often displayed or sold in special stores or areas that are off-limits to minors.

This document proposes a set of voluntary conventions that will, if generally followed, provide a similar level of protection for children on the Internet to that found in the physical infrastructure of most developed countries. By following these conventions, those who publish adult-oriented materials can choose to permit the automated blocking of minors from the downloading of such materials. Similarly, browser vendors can implement the blocking of such materials from any sites that conform to these conventions.

The naming conventions are described in terms of URL notation for the World Wide Web. However, a closing section of this document describes how the notation might also be used in other contexts.

Overview

The basic idea of the KidCode conventions is to use the URL notation [2] to provide an indication of the suitability of certain materials for children. The conventions provide a range of mechanisms to indicate a minimum age for viewing materials, and to provide a set of textual descriptive category names to indicate the reasons for restricting access.

Two different conventions are established: The fixed URL convention provides strings that may be embedded in a stable URL to specify the nature of the child-oriented restrictions. The dynamic URL convention provides strings that may be embedded in a client response to a web form that specifies age and consent.

The Fixed URL Convention

KidCode fixed URLs are of the following general form:

http://.../.../KidCode.AGE.STR1.STR2.../...

Here "KidCode" is a literal string, while "AGE" is an integer indicating a minimum suggested age for traversal of the link when browsing. It implies that items with this URL should not be permitted to be retrieved by people below that age. An AGE of "0" means the authors explicitly state that they do not believe there is anything offensive at that URL to which parents might object. This is in contrast with a lack of any "KidCode" component, in which the author makes no statement at all about the suitability of the materials for children.

STR1, STR2, etc are optional and indicate the types of possibly-offensive materials that might be seen by accessing this URL. Because the strings are embedded in a URL, they are English words, but may be translated into the local language by the browser interface. This memo provides an initial list of such keywords, which may be extended by general convention and usage. (An IANA registration process might be reasonable in the future.)

Thus, a URL of the form:

http://bookstore.com/KidCode.0/....

might be the root of the children's section of a book store, while the "young adults" section of books that include some bad language might be rooted in the URL:

http://bookstore.com/KidCode.12.profanity/....

and the X-rated section of the book store might be rooted in:

http://bookstore.com/KidCode.21.sex/....

and the really offensive stuff (stories about violent naked doped-out Nazis raping and pillaging while swearing profusely) might be rooted in:

http://bookstore.com/KidCode.21.sex.violence.profanity.nudity.drugs.racism/....

This naming convention allows browsers to enforce restrictions based on age, on categories, or both, or to give the power to make this decision to the parents and teachers who configure the browsers for their children. For example, parents may wish to allow nudity but not violence, regardless of age, or parents may wish to disallow access to URLs marked as having questional religious themes.

KidCode browsers should search each URL accessed for the string "KidCode" (without regard to case) as the leading part of a URL component (i.e. immediately after a "/") and then check the parameters after it for suitability with the configuration. It is recommended that parents, guardians, school officials, etc be given the ability to modify the browser's configuration in this regard, while the children for whom the restrictions are intended should not be able to easily reconfigure the browser.

The formal grammar for KidCode fixed URL naming is given as follows, in the Augmented BNF notation of RFC 822 [4]. Note that the syntactic entity "fixed-KidCode" is designed to be used as either an "hsegement" or an "fsegment" in the formal grammar of RFC 1738. Note also that the syntactic categories "digit" and "uchar" are defined in RFC 1738.

```
fixed-KidCode = KidCode-token "." age-token [category-list]

KidCode-token = "KidCode"    ; case-insensitive

age-token = 1*2digit

category-list = "." category-name [category-list]

category-name = "sex" / "violence" / "profanity" / "nudity" / "drugs"
          / "racism" / "gambling" / 1*uchar
```

Note that the age-token is explicitly required. This means that URLs containing the string "KidCode" without an age specification should not be treated specially by KidCode browsers. This allows, for example, URLs of the form "http://.../about/KidCode.html" and "mailto:help@KidCode.com".

The Dynamic URL Convention

The second KidCode naming convention is intended for use where it is not deemed desirable to embed the KidCode restrictions in the URL. Instead, a web form is generated that requires the browser to return an answer in a predefined format that should be recognized by KidCode browsers. Such a form may be used to ask the browser user to explicitly confirm his willingness to see, for example, sexually explicit material, and returns the indication of consent in a form that can be blocked by KidCode browsers.

In particular, a CGI script can be implemented that prompts for the user's age, using the key word "myage", with appropriate descriptions of what kind of material is being offered behind this URL. The form should include a field "myage" which should be returned by the browser as an integer. It is expected this would be in a form with the method being "get", in order to allow frequent visitors to the site to save it in hotlists.

Thus, for example, if a "privacy curtain" is implemented at the URL

http://adultstuff.com

then that URL will return a web form that requires the user to fill in an integer field named "myage". Of necessity, to continue browsing the user will need to use a URL of the form

http://adultstuff.com/....?...myage=xx...

A KidCode-conformant site or server that does not use the Static URL convention will require the "myage=" return value in a form before returning any age-sensitive material. A KidCode browser should restrict access to such dynamic URLs by checking for URLs of the form "http://.../...?...myage=..." and confirming that the integer after "myage=" is one that the current browser user is permitted to return. Thus, for example, a browser configured for elementary school children would not be allowed to return any age greater than 12. If a student typed in "21" as his age, a KidCode browser that was so configured would never pass the answer to the server.

This convention should naturally be checked for even for URLs on hotlists and in disk caches, since there may be more than one user in a household using the computer and the KidCode mode can be expected to be turned off or modified while the parents or teachers (or older children) are using the computer.

[TIP TO WEBMASTERS AND SERVER IMPLEMENTORS: To reduce the frustration levels, a script which creates a table of contents can have an "KidCode prompt form" on the same page as the table of contents. Entering the proper age can cause a new table of contents to be returned with the string "?myage=..." appended to each embedded URL, allowing the entire document tree to be accessed without further prompting. Of course, with current browsers, each access must invoke a script to assure the proper response has been made. However, there is no reason this cannot be built into a server in the future.]

Dealing with Non-Compliant Sites: Explicit Blacklists

Prior to this specification, in the absence of naming conventions, the only mechanism available for restricting access to offensive material was the construction of explicit blacklists of known offensive sites, to which access could be blocked on a browser-by-browser basis. Such a technique is completely orthogonal to the KidCode naming conventions, and it is recommended that KidCode-conformant browsers should also provide such an explicit blacklist mechanism, to be used to deal with offensive sites that do not follow the KidCode naming conventions. Obviously, however, it is preferable not to have to continually update such a list, so it is hoped that community pressure will persuade most such sites to follow the naming conventions.

More generally: In the event a URL does not contain a KidCode identifier, the browser should be able to be configured such that the URL is either accepted or rejected. That is, the default behavior on an untagged URL should be specifiable. In addition, there may be an exception list of leading URL components that are treated as the reverse of the default. For example, a parent may specify that KidCode sites with an AGE tag of 18 or less is acceptable, and that sites without a KidCode are acceptable, except for http://www.playboy.com and alt.sex. In this event, all URLs starting with http://www.playboy.com and all newsgroups starting with alt.sex would be rejected, unless they specifically included a KidCode entry with an age less than 18.

Definition of KidCode-Conformance

A browser is KidCode-conformant if it implements both the static and dynamic KidCode URL conventions, and makes it easy for adults to configure the browser to manage the access of children who use it, while making it difficult for children to reconfigure the browser to eliminate the restriction.

A web site is KidCode-conformant if any potentially-offensive materials are made available ONLY through URL's that advise KidCode-conformant browsers of the nature of the restrictions.

Verification of KidCode-Conformance

Obviously, it is not possible to enforce this naming convention on the Internet at large. A system such as this one necessarily depends on voluntary compliance. To promote such compliance, and to promote a higher level of comfort on the part of parents and educators, a licensing service to verify and certify KidCode conformance for both web tools and web sites is contemplated by the authors, who may be contacted for additional information.

It is expected that different groups with different sensitivities will wish to develop different "rating authorities" for these purposes. Therefore, a convention is defined whereby a site may be tagged as approved by one or more such authorities.

To determine whether or not a site has been approved by a rating authority, the original URL can be modified to remove everything after "KidCode." and to append "who/" followed by the rating authority name. Thus, to see if the rating authority named "Vatican" has approved the compliance of the site servicing the URL

    http://www.playbeing.com/KidCode.21/Lola.gif

a browser could consult the URL

    http://www.playbeing.com/KidCode.who/Vatican

If that URL cannot be accessed for non-transient reasons (i.e. if the attempt to access it yields HTTP "Not Found" error codes or FTP "could not open" error codes), then the site is not claiming any certification from the Vatican. If that URL does exist, the data inside it give the certification from the Vatican, in a format that is at the discretion of the rating authoritity. Some authorities might choose, for example, to cryptographically sign and date the root URL, and to make such signatures valid for a given period of time only. Others might choose to base their certifications entirely on trademarks, in which case the data returned by the certification URL might be just an empty file or some human-readable information. A KidCode-conformant browser should be able to deduce whether or not a site claims that the desired rating authority has certified the site, and might be able to verify that certification for certain rating authorities.

Applicability of KidCode naming conventions to non-Web protocols

Although they were explicitly designed for use in the World Wide Web, the KidCode naming conventions are applicable in other contexts as well.

For example, FTP may be rendered KidCode-conformant through the use of the fixed URL naming convention in the full path names for files. Thus, an ftp file named "pub/pictures/KidCode.18/whatever" should not be shown to children under the age of 18.

Netnews could be rendered KidCode-conformant through the use of KidCode fixed URL naming conventions in newsgroup names. In the simplest version of this scheme, one might wish to rename the entire "alt.sex" newsgroup hierarchy to be "alt.KidCode.18.sex".

References

[1] Berners-Lee, T. "Hypertext Markup Language (HTML)", draft-ieft-iiir-html-01, June 1993. (expired working draft)

[2] Berners-Lee, T., Masinter, L., and McCahill, M., "Uniform Resource Locators (URL)", RFC 1738, December, 1994.

[3] Berners-Lee, T., Fielding, R. T., Nielsen, H., "Hypertext Transfer Protocol -- HTTP/1.0", draft-fielding-http-spec-0.0, 1994 Nov 28. (working draft)

[4] Crocker, D., "Standard for the format of ARPA Internet text messages", RFC 822, UDEL, August, 1982.

Acknowledgements

The authors wish to thank the members of the Common Solutions Group for their input into the design of this facility.

Security Considerations

Security issues are not discussed in this memo, unless protecting children from offensive materials is considered a security matter, in which case the entire memo is about security.

Author's Address

Darren New
First Virtual Holdings
319-B Crowells Rd
Highland Park, NJ 08904

Email: dnew@sgf.fv.com
Phone: +1 908 572 1135
Fax: +1 908 572 1219

Nathaniel S. Borenstein
First Virtual Holdings
25 Washington Avenue
Morristown, NJ 07960

Email: nsb@nsb.fv.com
Phone: +1 201 540 8967
Fax: +1 201 993 3032