# The Future Is Bright

*www. The Future Is Bright .net*

# Fabel

# Fact Based Labelling

| | |
|---|---|
| Author: | Joseph Kilcullen BE |
| Date: | 28 September 2005 |
| Version: | 1.0   (A4) |

# Abstract / Executive Summary

A three layered system for implementing censorship for minors is described. The foundation layer is composed of FAct BasEd Labels (FABEL) i.e. labels indicating facts about the content. These labels are located on web servers with the Internet content itself.

The second layer exists on the surfer's computer. It consists of an Age Fact Table which maps/translates the facts to ages that censorship starts at.

The third, and final layer, maps/translates the ages to an appropriate 'National Age Rating Certificate' for the surfer's home country.

After describing the system, possible implementations are discussed and various failure, breakage and hijacking threats are identified, discussed and possible solutions proposed.

# Contents

# Tables

# 1 Introduction

Offensive content, be it violent or adult/sexual is readily available on the Internet. While adults often react in a hostile manner to censorship, it is generally accepted that censorship is desirable for minors i.e. children and teenagers.

Censorship for minors typically consists of an age rating 12, 16 etc. Such censorship ratings can be found on movies and computer games. However introducing censorship for minors to the Internet, a media rich forum, has yet to be achieved. Problems to be solved include:

- Conflicting views on how offensive any given content is.

- Conflicting legal requirements i.e. different regions of the world have different levels of censorship and varying legal responses to illegal content.

- Different languages form barriers for labelling systems while the media, pictures and movies, speak a universal language.

- Any form of censorship will be viewed as a bad thing and will result in hackers attacking the censorship system to fight for freedom of speech.

- It is difficult to identify the scope of the solution i.e. as the number of requirements grow the complexity of the system increases and the likelihood of it being implemented decreases. Also the international cooperation required makes agreement difficult.

# 2 Mission Statement / Objective

The Objective of the Fabel project is to devise a labelling system for Internet media that will implement the Internet equivalent of Age Ratings like those currently displayed on movies and computer games. A full list of objectives follow:

- Children can surf the Internet unsupervised without any fear of offensive content being displayed.

- Parents/Guardians should be able to control what content their children have access to.

- To devise a framework that can be implemented with different technologies i.e. such that it can be extended to television and other media. This is because other media are slowly being absorbed into the Internet e.g. Radio.

- One system acceptable to all so that it will be implemented universally. Not opinion-based systems where different opinions result in different systems that then effectively compete against one another.

- Delivery of National systems i.e. deliver Irish Age Rating Certifications in Ireland, French Age Rating Certification in France etc. rather than delivering an International system that will not satisfy the needs of individual countries.

- The system should not attract unwanted attention from hackers. If the system is used to limit freedom of speech then hackers are likely to destroy the system.

- We need to introduce new technologies to deliver on these requirements while not creating new problems or exasperating existing problems.

- Simple for parents & guardians to manage/control.

# 3  Components of the Solution

This solution consists of the following components:

- International Law component i.e. a set of laws that will be implemented in all participating countries.

- National Law component i.e. each country's own laws that will work with International Law to provide a complete set of laws for that country's requirements.

- Bubbles and Embedded data. These systems are for linking media with censorship data for that media.

- **FA**ct **B**as**E**d **L**abels - Fabel - that are electronic labels that indicate facts about that Internet content.

- Age Fact Tables, Age Rating and Templates

I have chosen to call the system 'Fabel' rather than 'Fable'. The reason for this is that Internet searches will find Fact Based Labelling pages when 'Fabel' is entered as the keyword. If 'Fable' were used then searches would return links to both children's stories and Fact Based Labelling pages.

# 4 The Two Step Censorship Process

For censorship of movies a person will watch each movie and based on a set of guidelines choose an age certification. It is not feasible to censor Internet content in this way. Different countries have different ideas about how offensive content is. Also this labour intensive approach would not be possible with the volume of media on the Internet. Hence a system must be devised to achieve the same result while overcoming the various obstacles.

Here follows a high-level description (low on detail) followed by a low-level description of the censorship process.

## 4.1 *High Level Description*

To achieve automation of the censorship process we split the task into two steps. The split corresponds to:

- Identifying facts about the offensive content.
- Selecting an appropriate age certification, for your country/region, based on those facts.

The Webmaster carries out the first step in the country where the web site is located. This step starts with the offensive content and ends with a list of facts about that content.
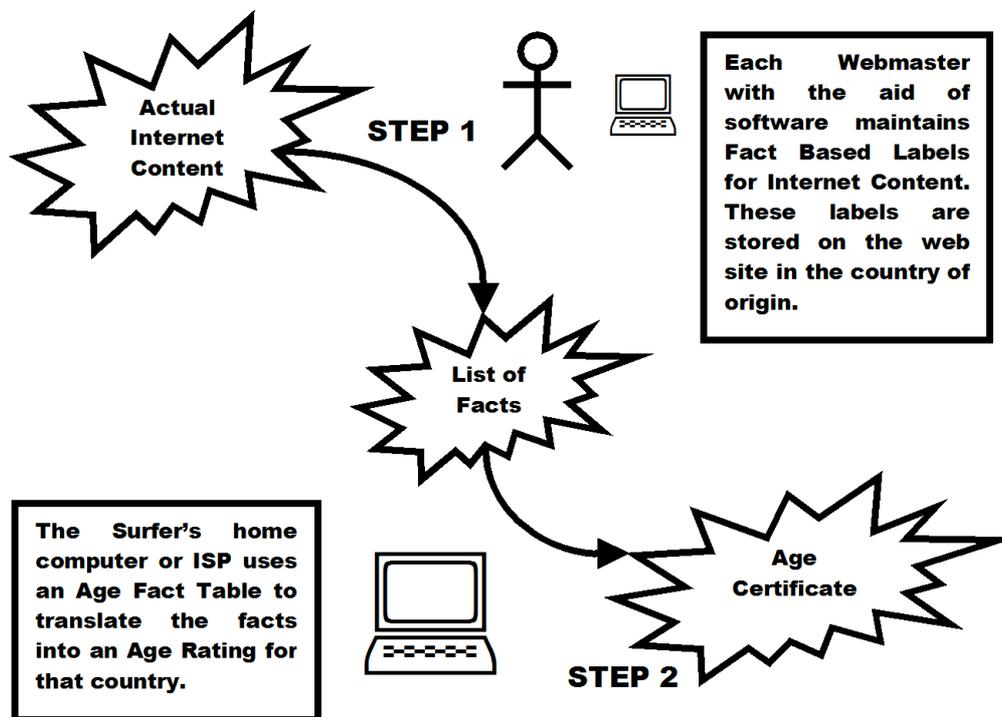


*Figure 1 – High Level Description of Two Step Censorship Process.*

The Webmaster carries out 'Step One' in the country where the web site is located. This step is guided by a software-based questionnaire that helps keep labelling consistent and factual. After completing the questionnaire the software will provide the Webmaster with a list of facts in the form of electronic labels. The software can then be used to 'attach' the labels to the Internet media. Since this first step is carried out the same way in every participating country, the International Law portion of the solution applies to this step. Every participating country will legally require web sites, in that country, to have offensive content labelled with this system.

The end users computer or their ISP carries out the second step. This step is carried out in the web surfer's home country. It starts with the list of facts and ends with an appropriate age certificate for that country. A 'local mapping' or 'Age Fact Table' guides this step. The Age Fact Table specifies the youngest acceptable age certification for each possible fact.

Example: Suppose there are two electronic labels on a picture corresponding to facts X and Y. In the Age Fact Table for this country fact X indicates an age certification of 15 years and fact Y indicates an age certification of 8. Clearly this picture receives a 15-age certificate.

Other countries may have different ages or even a complete ban on this material. This second step corresponds to opinions i.e. the part we can disagree on. Each country/state can have its own rules i.e. it's own opinions. A problem with a complete ban on content is that this will attract hackers who believe they are fighting for freedom of speech. This problem will be discussed in detail in Chapter 7 'Theory Versus Practice – Real World Problems'.

*Postscript*
Note that the final step is carried out on the users computer i.e. this system should not come under attack for limiting freedom of speech, as the system will only be forced on employees in industry, children in schools etc. It will be turned on or off by whoever manages the computer. Home users can choose not to turn the system on, or even choose rating systems other than their National system i.e. Age Fact Table for their country/state.

## 4.2  Low Level Description

When a surfer selects a web page or media file, on the Internet, by specifying a URL. There are three processes that take place to determine whether to censor or not. These processes are:

1. Obtain Fact Based Labels for this content. This happens every time a user selects content on the Internet.

2. Create a Template for this user. This happens once i.e. when the user logs in.

3. Finally the Fact Based Labels are combined with the Template to clear or block this content.

| **Process One** | | | |
|---|---|---|---|
| URL or File (on website) | + | Bubbles or Embedded Data | = | Fact Based Labels (A) |
| **Process Two** | | | |
| Age Fact Table (National) | + | Surfer's Age Rating ( e.g. 12 or 16 etc.) | = | Template (B) |
| **Process Three** | | | |
| Fact Based Labels (A) | + | Template (B) | = | Answer to: Display Content Yes/No? |

*Table 4.1 – The Three Processes that determine whether to censor or not.*

### 4.2.1  Process One. Obtain Fact Based Labels for Content

The two systems, Bubbles and Embedded Data simply retrieve the Fact Based Labels from a website or from the media files. Bubbles works with the Internet while the Embedded Data solution works with the actual media files.

Process one is carried out every time media content is selected. It starts with a URL or a computer file name and ends with the Fact Based Labels for that content.

### 4.2.2  Process Two. Create a user Template.

This process takes place once i.e. when the user logs in or when they turn their computer on.

On the users computer there will be an Age Fact Table. That table will specify the minimum acceptable age for each fact in the Fabel system. The ages specified will correspond to censorship ratings for the surfer's own country. Also the surfer will have an Age Rating specifying what this user can view e.g. 12 Certificate, 16 Certificate etc. To prevent confusion I will refer to this as the 'Surfer's Age Rating'. (See the glossary for a list of Fabel terms.)

Combining the Age Fact Table and Surfer's Age Rating will create a Template. A Template is a list of all the facts in the system with each marked as acceptable or unacceptable for this surfer to view.

Example: A Censorship Age Setting of '12 Year and older' would result in a template that denies access to content with facts suitable for people over 12.

| Fact | | Minimum Acceptable Age for viewing | Template for '12+' | |
|---|---|---|---|---|
| 1 | Violence of type x | 8 | 8 < 12 | Ok |
| 2 | Nudity of type y | 12 | 12 = 12 | Ok |
| 3 | etc. | 8 | 8 < 12 | Ok |
| 4 | etc. | 16 | 16 > 12 | Censor |
| 5 | | 10 | 10 < 12 | Ok |

*Table 4.2 – Sample Template for '12+'*

### 4.2.3   Process Three. Combine Labels and Template

Simply check the list of facts, for facts that the Template indicates should be censored. Here is an example where the surfer can view the content.

| Fact Number | Fact | Template for '12+' |
|---|---|---|
| 1 | True | Ok |
| 2 | True | Ok |
| 3 | False | Ok |
| 4 | False | Censor |
| 5 | False | Ok |

*Table 4.3 – Combining Labels & Template*

### 4.2.4   Sample Implementation

If individual bits (1 or 0) are used to indicate that any given fact is true, 1, or false, 0, then a 2 byte word could store 2x8=16 bits or 16 facts. The fact-based labels for a picture would be a 2-byte word. The template would also be a 2-byte word. To combine them simply flip the template and logically AND it with the facts.

| | |
|---|---|
| 1011,1000,1100,0000 | Facts |
| 1110,1111,1100,0000 | Template |
| 0001,0000,0011,1111 | Template Flipped |
| 0001,0000,0000,0000 | Logical AND of Facts and Template |

*Table 4.4 – Combining Labels & Template  - Sample Implementation*

With a non-zero result there must be facts that this users Template require censorship of. This low-level description appears different because the details obscure the high level design/purpose of the system. The low level description provides the mechanical

nuts and bolts of an implementation while the high level makes it clear how the different countries, different censorship systems can be implemented in one system.

## 4.3   Hiding Links the Surfer Cannot Open

To help Internet Search Engines, e.g. Yahoo & Google, to hide links a surfer cannot view. We devise the following solution:

*Existing System for Passing Data to Web Sites*
Arguments (information) can be passed to web sites in URLs e.g.

```
http://www.TheFutureIsBright.net/abc.php?a=1&b=2
```

Here the web server receives arguments a and b with values 1 and 2 respectively.

*Alterations Necessary for Search Engine to Receive Surfer's Template*
In the same way the search engines can place a link like

```
http://www.TheFutureIsBright.net/abc.php?template=FABEL_TEMPLATE&a=1&b=2
```

Now the web browser (if configured to do so) can substitute this user's template for FABEL_TEMPLATE. Now the search engines have the user's template and can apply it in their search i.e. the user will not see any links that they cannot open.

*Note*
Chapter 7 includes a discussion of possible ways paedophiles will try to hijack such features for their own purposes and ways these features should be implemented, so as to minimise or prevent such abuses.

# 5 System Components

## 5.1 Legal Components of Solution

In an International context, the diverse requirements make it difficult to make laws for the Internet. However, the split between fact and opinion, inherent in this solution, allows for a similar split in legal requirements. The fact components of the solution coincide with the International law component while the opinion component still functions and can be supported by National Law.

### 5.1.1 International Law

A consistent definition of offensive content should be negotiated. Note that this system will attract more and more attention from hackers as the offensive content definition is broadened to cover more topics. The concept of Democratic Law is critical here. While some people believe Harry Potter books promote the occult, witchcraft etc. Democratic Law will never support censorship of Internet content on either Harry Potter or actual occult/witchcraft web sites. My thinking on the definition of offensive content is that it should cover areas that are clearly offensive i.e. violent and adult/sexual content. Note this is an opinion, so you are free to disagree. Once the offensive content definition tries to encompass other Internet Content, Democratic Law will bring the system under attack.

In France it is illegal to sell Nazi memorabilia. In various regions of the world religious groups will require additional censorship. How are we to resolve the different additions to the basic set that various countries/groups around the world will require? The solution is to create one core labelling set and a number of additional sets, one for Nazi memorabilia, one for topic A, topic B etc. The problem is that only the core set will receive unanimous support. As such additional sets cannot be compulsory. However, we can support these sets with Legally Enforced Honesty i.e. web sites will **not** be forced to use these additional sets, ***but if they do, they must do so honestly/correctly***.

*Components of International Law*
- Legally define offensive content in the context of Fabel.
- Legally require offensive content to be transported via a small set of protocols, including HTTP and FTP.
- Legally require offensive content to be labelled with Fact-Based labels of the core set.
- Legally require a small set number of additional labelling sets to be implemented correctly or not at all. Explicitly state which sets receive this support.

### 5.1.2 National Law

Since Age Fact Tables and accompanying National Age Certificates only apply within their own country, national law can be drafted by individual states to provide for their own needs. The only real concern is that overly restrictive censorship will result in

hackers attacking the system. Such attacks are likely to undermine the system for every country using the system i.e. once the door is open, children and teenagers from all over the world will use it, even if the hackers only intended to promote freedom of speech.

### 5.1.3 High Level Discussion of Legal Requirements.

A question and answer approach is taken here.

**Why seek legal support for Fabel?**

*Without legal support webmasters will not place labels on their web site. Current labelling systems are used on some adult sites. However the diversity of systems creates competition between the systems where we really should be co-operating. Also we need web sites with no offensive content to state just that. Existing labelling systems are rarely used on regular web sites. This undermines the system because these web sites appear as 'unknown' to surfers using labelling. The ambiguity destroys the entire system. Even a labelling system that simply said 'no offensive content here' would grant surfers access to more web sites than existing labelling systems. We need everyone to label with one system. No ambiguity and one system that everyone can use rather than competing opinion based systems.*

**What will happen when we find incorrectly censored web sites?**

*There are two ways that this system can fail. Either the Age Fact Table has a mistake or the labels are incorrect. If the Age Fact Table is incorrect it is up to that country's authorities to identify the problem, rectify it and update the citizens computers accordingly. (The updates solution 'Compel Users to Install Updates' in the eMail Solutions document should help here.)*

*The second failure mode is for labels to be incorrect. On finding incorrect labels the URL should be entered into a system that will inform the country of origin of the problem. The authorities in that country should then pursue the web site proprietors under that country's legislation.*

*Even additional labelling sets like a French Nazi Memorabilia set could be supported with the 'Legally Enforced Honesty' solution i.e. if the set is used then it must be used correctly. See Chapter 7 for more on failure of the system.*

## 5.2 Bubbles and Embedded Data

Adult magazines have front covers that indicate their adult content. Also shops place the magazines 'on the top shelf' or sell them in Adult shops. On the Internet we will need to communicate information about adult content on a number of levels spanning a spectrum from data on individual files, to data on entire web sites. In each case there are varying requirements on what information is required.

Two extremes dictate the development of two systems for linking media content and fact based labels for that content. Here follows a brief description of the two systems and a number of examples to help communicate the needs/requirements of the

systems. Mostly the technical content is supplied as an example or indication of the form/structure of the final solution. The final form of the solution may be completely different or even be developed using one of the existing technologies developed for content labelling.

### 5.2.1   Bubbles

The scale of the Internet means that labelling all content is an unachievable task, unless we simplify and automate the work involved. The Bubbles labelling system represents our opening move in this game.

The first step in this process is to restrict the transport of offensive content to Universal Resource Locator (URL) protocols. Even to a small number of such protocols to start with e.g. HTTP and FTP. Without legal enforcement some people will always move to other protocols, forcing the creation of legislation to stop them. Once the system is in place and working, legislation can be used to support the system i.e. make it illegal to transport content on other protocols.

Bubbles are simply groups of URLs that are labelled the same way. The simplest way is for all content, in a bubble/group, to have the same labels. Other more complex bubble types should be developed as required. The examples given below will more clearly indicate the structure and purpose of this system.

### 5.2.2   Embedded Data

Embedded Data simply means that media files will contain Fact-Based labels for their own content. A major problem here is that adding Content Labelling to every media file is almost logistically impossible.

The roll out of Fabel will involve implementing the Bubbles system first, and then using software to transfer the labels stored in the Bubbles system to individual media files. The same software that manages the bubble.def files, described below, will do this. This will overcome the logistics of applying labels to all media files. It is still a mammoth task. It will simply take time for software to be developed, installed and over time all requirements met/implemented.

Finally operating systems can be modified to understand and operate with labels stored in individual media files i.e. embedded data. At this point all media files on the Internet will be labelled i.e. this task is achievable, we simply need to automate and simply this process.

### 5.2.3   Examples

The examples/instances that follow present the requirements that shape the two solutions. Each example is chosen to be as simple as possible. In this way they communicate, more clearly, specific aspects of the systems.

### 5.2.3.1  Instance A – A Child Receives an email.

A child receives an email containing offensive content.

*Discussion*
Here we introduce the idea of embedded data. Basically media files would contain the labels for their own content.

The primary function of Fabel must be to deliver censorship functionality for the Internet. However we are forced, by media being present on our computers to expand the system to our operating systems. We achieve this by embedded data in media files.

*Problems*
This requires all media files on all computers to be modified to store such data. Leaving media files without label data is ambiguous and will undermine the entire system. Conservative users will block access to unlabelled files. This is unworkable, as massive numbers of files will be unlabelled. Just like web sites that are not labelled the ambiguity undermines the entire system.

### 5.2.3.2  Instance B – A Child Creating a Birthday Card.

A child creating a birthday card searches the computer for pictures to put on the front of the card.

*Windows 95/98*
The search finds files in directories belonging to other computer users. However labels embedded in the files deny access.

*Windows XP, Linux etc.*
The operating system denies this user access to other user files. Files in common areas would have labels embedded to determine the outcome.

*Windows 95/98 + older teenager*
An older teenager downloads a software patch to bypass censorship. The younger child has full access to any file on the computer.

*Windows XP, Linux etc. + older teenager*
Older teenager is unable to hack the modern operating system leaving the system secure.

### 5.2.3.3  Instance C – A Web Site With No Offensive Content

A web site with no offensive content requires compliance with Fabel.

*Discussion*
Here we introduce the Bubbles System. Even though individual files will have labels embedded in them, it is not practical for these labels to be used over the Internet. Universal Resource Locators (URL) usually address one file e.g.

    www.TheFutureIsBright.net

```
corresponds to

http://www.TheFutureIsBright.net/index.php
```

However this one file may contain any number of references to other files. To censor this content with embedded data would require all files to be downloaded, have their label data extracted, analysed and then an outcome or result chosen based on user settings and data retrieved. While this can be done it is likely to slow down web browsing to an unreasonable extent.

Instead we will group Internet Content into groups called Bubbles. Every piece of Internet content can be addressed via a URL. All of these URLs start with:

```
http://www.TheFutureIsBright.net/
```

So we can define a group of URLs, a bubble, that all start with this text string. Next we specify one set of Fact-Based labels for all content in this bubble. So a file located at

```
http://www.TheFutureIsBright.net/bubbles.def
```

might look like this:

```
<bubbles>
  <bubble type='uniform'>
    <address>
      HOME/*                                          (1)
    </address>
    <labels type='uniform'>
      <set name='standard'>0</set>                    (2)
    </labels>
  </bubble>
</bubbles>                                             (3)
```

1. The group of URLs covered by this bubble is the home address with anything added onto the end of it i.e. http://www.TheFutureIsBright.net/*

2. The 'set' tag refers to the standard set that must be implemented. None of the optional labelling sets are used here. For example, surfers in France may be blocked from this website as it does not explicitly state that no Nazi memorabilia are for sale here.

3. This tag is to let the software know that the file has been downloaded completely

### 5.2.3.4  Instance D – An Adult Web Site
An adult web site with different types of content requires compliance with Fabel.

*Discussion*
In this case a number of different types of bubble can be used to create the bubbles.def file:

```
<bubbles>
  <bubble type='range'>                                  (1)
    <address>
      HOME/*
    </address>
    <labels type='min' setname='standard'>0000111100001111</labels>
    <labels type='max' setname='standard'>1110001010100101</labels>
    <bubble type='uniform'>                              (2)
      <address>
        HOME/xyz/*
      </address>
      <labels>
        <set name='standard'>01100110</set>
      </labels>
    </bubble>
    <bubble type='amorphous'>                            (3)
      <address>
        HOME/qpw/*
      </address>
    </bubble>
    <bubble type='directory'>                            (4)
      <address>
        HOME/123/*
      </address>
    </bubble>
  </bubble>
</bubbles>
```

There are four bubbles defined here. Bubble (1) covers the entire web site while the remaining bubbles cover subsets of the web site.

1.  This bubble covers the entire web site. As type 'range' a min and max set of labels are supplied. This should provide enough information to determine whether or not to grant access to the site. Only if the user enters the web site will the other bubbles need to be viewed.

2.  Bubble two covers HOME/xyz/* where all content has the same labels.

3.  Bubble three is amorphous meaning the web browser must download all content and look at the labels on individual files.

4.  Bubble four is of type 'directory' meaning 'look in each directory for another bubbles.def file'.

When a user enters a URL that URL will be checked against the bubble addresses. This involves searching bubbles within bubbles until the lowest bubble on a tree is found. The bubble type will then be used to locate the label.

| Bubble Type | Behaviour |
|---|---|
| Uniform | All content has the same labels and those labels are supplied inside the <labels> tag. |
| Amorphous | Web Browsers must retrieve labels from the individual files i.e. use embedded data rather than the bubbles system. |
| Directory | Each directory has its own bubbles.def file. Look there. |

*Table 5.1 – Bubble Types*

Software will be developed to assist webmasters in creating and maintaining bubble definition files.

### 5.2.3.5  Instance E – Porn! What Porn?

In England an incident occurred where a man purchased an adult movie that turned out to be a regular movie. He sued and won!

A webmaster places a set of pictures into a bubble and specifies labels/facts that are true for that set, but facts that might not be true for individual pictures. Software is then used to transfer the Bubble labels into embedded data in the actual files. One picture is then taken from the set and stored separately. Clearly it is now labelled incorrectly.

The problem here is that opponents of the system will try to undermine it with legal challenges centred on the challenge of labelling correctly vast quantities of files and media.

The solution here is to create a 'group' label i.e. a label that states that the picture belongs to a group and the Fact-Based labels supplied refer to the group, not to individual pictures. This group label will assist in transferring labels from the Bubbles system to embedded data i.e. it will reduce the labour involved.

## 5.3   Fact Based Labels

As the name suggests these are labels that indicate facts about the content they are associated with. The intention is to provide enough information for Age Certificates to be assigned to media content while not containing 'opinions'. Opinions being of no value, at the labelling level, as different countries and cultures have different opinions that must be based on facts. Facts form the foundation of the system allowing different cultures to interact on a common level.

Creation of a useful set of facts, for use with Fabel, will be extremely difficult despite our universal agreement on facts. This will be discussed in detail in Chapter 7 'Theory Versus Practice – Real World Problems'

The legal requirements will dictate the nature of these labels.

- Labels indicate facts about the content i.e. straightforward to prove in a court of law.
- Labels should be sufficiently course or fine to allow participating countries to map these labels to their own age certification categories.

## 5.4  Age Fact Table

As the name suggests an Age Fact Table (AFT) simply specifies the minimum acceptable age for viewing content satisfying each fact. The following table is an Age Fact Table with just five facts in it.

| Fact Number | Minimum Acceptable Age |
|---|---|
| 1 | 8 |
| 2 | 12 |
| 3 | 8 |
| 4 | 16 |
| 5 | 10 |

*Table 5.2 – Sample Age Fact Table*

Age Fact Tables will also have a set of 'National Age Rating Certificates'.  For the table above these Certificates would be 8, 10, 12 and 16. Since each country can create their own AFT the Age Rating Certificates will be for their country. Hence a '12' rated movie in France could contain different content to a '12' rated movie in Ireland, as the AFT's are different for the two countries.

Each country gets to create their own Age Fact Table and associated National Age Rating Certificates.

# 6 Theory Versus Practice – Real World Problems

In practice this system will only succeed if we identify the various ways the system will come under attack, and then alter the system to deflect or minimise the impact of these attacks. The sections that follow categorise the types of attack this system will be subjected to and possible solutions.

## 6.1 'Censorship of All Content' v. 'Age Ratings for Minors'

*Democratic Law*
**Democratic Law is the concept that laws become more difficult to enforce as the portion of the population that support them declines. For example enforcement of drug laws is difficult because people who buy drugs are opposed to anti drug legislation. They want to be able to buy drugs.**

The concept of 'Democratic Law' means that laws are in fact 'defined behaviours that the population agrees is unacceptable' e.g. driving on the wrong side of the road, murder etc. This means that enforcement of the law is the government, on behalf of the people, punishing individuals who go against the population's defined acceptable behaviour.

The concept of Democratic Law means that our ability to enforce laws weakens as the portion of the population against those laws increases. Technologies have a major impact on this, as technologies impact productivity and efficiency levels and as such the economics & feasibility of any given law enforcement attempt. Also individuals who wish to break the law can use productivity and efficiency gains to further their objectives.

The impact that this has on Internet censorship attempts includes:

- Either extreme, zero or total censorship, will not gain public support and will result in the system coming under attack.

- For technological reasons attacks on the system will break the system for everyone i.e. a small leak in a real world dam can be plugged relatively easily, while an IT dam is considered to be breached if there is any leak. This is because high levels of productivity will allow vast quantities of fluid to flow through the crack, no matter how small that crack is.

- The international nature of the Internet means that an inability to enforce the rule of law in some regions of the world will significantly threaten attempts to legislate the Internet i.e. the leak in our technological dam will exist in foreign countries that we have no control over.

Generally speaking every system will have flaws, failings. In this case our objective must be to minimise the problems while maximising the effectiveness of the system. Once we over step 'the line' Democratic Law will bring the system under attack. Unlike less technologically advanced systems (only 70% of population in favour of

smoking ban in Ireland, but the system still works) a tiny portion of the population can wreck havoc e.g. Virus writers. This, in a nutshell, is the problem. The high levels of productivity associated with technology means that a tiny portion of the population in a region of the world where we cannot enforce our laws will break our technological dam.

The solution, that I propose, attempts to have our cake and eat it. By dividing the system into compulsory and optional, labelling and the use of labelling, we undermine the need to attack the system i.e. if it is optional for surfer's to turn on/off censorship then there is no need for anyone to hack the system. Why bother? Anyone who does not like it can simply turn it off! The alternative is to force the system on people and have the effectiveness of the entire system destroyed by hackers.

Effectively we are creating a framework that can be utilized by parents/guardians and industry to control content access via their IT infrastructure. Attacks on the system will be a minimum as there is no need to attack a system that people can turn off. Optimistic as this may appear, this does actually create problems:

- Inability to censor the Internet, in any way, allows access to highly offensive content including child pornography and excessively violent content.

- Regions of the world that require censorship other than extreme content will not be able to censor such content as this will bring the entire system under attack.

- As an optional system we need parents/guardians to intervene in some way to turn it on. Usually parents will be less computer literate than their children and some parents may not be aware of the content the open system will grant their children – thinking it similar to television, which clearly Internet content is not.

Here again our system comes under attack. To some it will be unacceptable that the system be optional. Others, such as many regions of the United States, will be delighted as this will allow the system to be implemented in spite of freedom of speech laws i.e. it will not actually censor the Internet, it will simply allow individuals not to view content before it is displayed on screen. Like choosing not to view a blue movie because it has an '18' cert on it. It's about choice not censorship.

Though this solution addresses a number of major concerns. It is not enough. My preferred approach to additional requirements is that as much as possible should be achieved with this system. Then countries cooperating, through this system, can pool resources and technologies to identify and block other content in a unified manner. As such a task requires considerable design/research effort I do not intent to include these requirements in this documents objectives i.e. addressing child pornography, legitimate censorship and illegal content is not part of this system. This system is intended to address Age Rating censorship for teenagers and minors.

In the next section an example is given of the failure of this system for actual censorship of excessively violent material.

## 6.2   Failure, Breakage, Abuse and Responses to Same

Before discussing the major task of creating a core labelling set, we need to understand the various failure modes and abuses, the system will be subjected to. Each time we identify a failure mode or abuse of the system we must attempt to adjust the solution to resolve the problem without breaking the overall solution. As the title suggests there are three major categories of problems:

- Failure – where the system does not succeed for some reason.

- Breakage – where people try to bypass or break the system deliberately.

- Abuse or Hijacking occurs where individuals use Fabel for purposes other than its design objectives e.g. individuals using the system to find offensive content.

### 6.2.1   Failure Types and Responses/Contingency Plans

There are three types of failure. The first two and responses to them are:

- A problem with the Age Fact Table. Where an AFT is incorrect it is up to that country's government to create a new AFT and distribute it to end-users.

- An incorrectly labelled website. The country that the website is located in should pursue the proprietors under that country's legislation. (It would actually help the system if adults surfing the Net for porn flag content they find not to be labelled correctly.)

Of all the solution components the International Law component is the most likely to fail. Other solutions either depend on technology or National Law. The International Law component can fail in two ways:

- No law present in one or more countries.

- Laws present but not enforced.

Where no laws are present, participating countries may try to create 'second opinion labelling' or block all content from such countries. Second opinion labelling means a third party will label the content. Clearly the logistics of this task makes this solution impossible to implement i.e. it is only achievable on a small scale. This leaves blocking all content, another unpleasant solution.

Another failure is where countries agree to adhere to the system but do not actually enforce the law. Logically we approach this by asking 'why would someone do this?' The answer is that countries, like individuals, act in their own self-interest. As such we can identify ways that countries will not cooperate by asking how it would be in a country's self interests not to cooperate. Unfortunately the answer is, where a country has a child prostitution tourism industry. It is preferable that children's charities work with such countries to resolve such problems.

### 6.2.2  Breakage Types and Responses/Contingency Plans

Suppose countries producing extremely violent material place labels, to that affect, on such content. Then other countries censor this content at the ISP level i.e. they try to block this content in a compulsory manner. Web sites producing such content will experience a drop in sales. In response to this they may move to other countries that do not require labelling and/or support the creation of 'label stripping websites'. Yes existing systems for bypassing censorship will also be used if/when Fabel is implemented. Such systems are already well documented so I will not discuss them here.

*Label Stripping Websites*
**As the name suggests these web sites will download labelled web pages for you and replace their labels with a 'no offensive content' label. Such websites will not only allow access to extreme content they will also allow teenagers and minors to gain access to anything. Viruses and other Malware could even modify people's Home Address and bookmarks to pass these URLs through such a website. This would remove all censorship labels, as all content would be filtered through the Label Stripping Website.**

The existence of Label Stripping (or Label Filtering) websites constitutes a Breakage as minors will use such websites to bypass the system.

### 6.2.3  Hijacking/Abuse of Fabel

Unfortunately people are always going to try and use systems to further their own objectives. Even if their own objectives are the opposite of the system they are hijacking/abusing e.g. Catholic Church and Paedophiles. Possible hijackings include:

- Paedophiles may try to identify children surfing the Net by requesting a surfer's Template. Browsers should be configurable so that Templates are only transmitted to specific websites, rather than all website. Parents could specify a Children's Charity that would provide a TPV type list of trustworthy sites. When other search sites are used the Browser could be configured to download each link's labels and disable links accordingly. Yes slow but this will work.

- Websites that are aimed at children may have labels to help confine children to that portion of the Internet. However paedophiles may create websites to try and gain access to children. Just as children's charities require background checks on staff, children's websites are likely to have a standard created for them similar to the Chatterbox standard for Chat Rooms. Like Chatterbox such websites should be audited and a TPV solution used to verify adherence to the standard.

- Finally Fabel may be used to help people find offensive content. When the Age Ratings have very course grading e.g. 12, 18 etc. this is not a problem. With Fact Based Labelling individuals may use individual facts to find content. Again this is not a problem if the facts correspond to content at the lower end of the offensive scale. However use of Fabel to find more offensive content would exasperate existing problems with violence associated with

such content. Clearly this form of abuse requires us to modify our design. The modification required is referred to as Regrouping (described below).

- Our choice of labels will be constrained by possible abuses e.g. a label indicating the age of individuals in the content would result in surfers searching for content with the youngest age allowed by the system. Then websites will place labels of the youngest possible age on content with even younger people in it. Basically paedophiles will place the youngest allowable age on child porn and claim legitimacy in whatever country they find that will not prosecute them. This leads to the idea of 'Ascertainable Facts' described below.

## 6.3  Label Regrouping

Existing National Age Rating Certificates are grouped facts i.e. an 18 cert indicates that one or more of the '18 cert' facts is true. Fabel, in its final form involves all National Age Rating Certificates being broken down into individual facts and then 'regrouped' into groups that are of use to all participating countries.

This first table represents the current Fabel design without Regrouping. Shown are three countries with their National Age Rating Certificates and corresponding facts. This system is open to abuse as surfers can search for individual facts.

| Country X | 18 | | | | | | 12 | | | 10 | | 8 | | | |
|-----------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Country Y | 18 | | | | | 16 | | | | 12 | | | | |
| Country Z | 21 | | 18 | | | | 16 | | | 12 | | | 10 | | |
| Fact | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |

*Table 7.1 – Before Regrouping of Facts.*

Regrouping involves grouping of facts that can be swapped for one another without any National Age Rating Certificate being altered by the swap. Confusing? Just look at the next table.

| Country X | 18 | | 12 | | 10 | | 8 | |
|-----------|----|----|----|----|----|----|----|----|
| Country Y | 18 | | 16 | | 12 | | |
| Country Z | 21 | 18 | | 16 | | 12 | | 10 | |
| Fact | A | B | C | D | E | F | G | H | I | J | K |

*Table 7.2 – After Regrouping of Facts.*

The following table indicates the changes in the facts used by these systems.

| Regrouped Facts | Individual Facts |
|---|---|
| A | 1, 2, 3 |
| B | 4, 5, 6 |
| C | 7, 8 |
| D | 9 |
| E | 10, 11 |
| F | 12 |
| G | 13 |
| H | 14, 15 |
| I | 16 |
| J | 17 |
| K | 18 |

*Table 7.3 – Fact Groups*

In both systems surfers can search for websites where fact 'K' (Number 18) is true. However surfers searching for fact '15' in the grouped system cannot be guaranteed to get a page where that fact is true. They are only guaranteed that either fact '14' or '15' is true. Similarly a search for fact 'A' cannot guarantee any one fact in that group being true, only that one or more of these facts are true.

Since we only want to deny searches for extreme content from working, all we do is group facts associated with this content into the oldest age rating group. In this way such content will not be easier to find using Fabel. Some may say that we want this material distinguished so we can block it. My preferred approach is a 'Yin & Yang' approach i.e. both ends of the spectrum. By grouping such content with other '18' content, countries that do not censor such content will be happy because that content will not be easier to find. Countries that do want to censor it will block such content at the ISP level by blocking access to their IP addresses. Any International agreement should simply be for such websites to have their IP addresses registered centrally so countries that want to block it will be provided with a list of their IP addresses. Yes existing ways of bypassing such blocks will still exist. However, existing solutions to those problems will still work.

## 6.4   Concerning the Creation of a Core Label Set

Clearly the most challenging part of the solution is the creation of a Core Label Set that will satisfy the difficult requirements identified so far and any additional requirements that present themselves in the future. My preferred approach is to create a website for participating countries to interact with one another.

### 6.4.1   www. The Future Is Bright .net

It is likely to be a prolonged and difficult process. Existing labelling set and their design criteria have already categorised the types and levels of content. This should help. The real challenge is the collaboration required to reach a consensus on facts, fact groups and their mapping to National Age Rating Certificates. It is conceivable

that other mappings may be required i.e. systems that allow parents more choice/control.

Development plans for www.TheFutureIsBright.net include the following:

- The creation of Projects within the website e.g. 'Fabel', 'Spam' or any other area where technology is considered to be harmful to children.

- The ability to create Forums within each Project e.g. a 'Requirements' forum for debates about the problem. And 'Design' forum where technical participants will have debates on possible solutions.

This system will be further developed so that Fabel Labelling sets can be created through the website. The idea is that a 'Fabel – Core Label Set' project could be created. Then participants from various countries would propose motions in the various debating forums. After debating the issues votes on the motions could be used. In this way participants would slowly agree on 'facts' that the system should use. Once a new fact is added to the system participants could specify an age, in their own Age Fact Table, for that fact. As the process proceeds the website will identify possible 'Fact Groups' based on the facts agreed by participants, so far. With the assistance of a website participants can interact and even test/assess the system as they build up a list of facts. Any problems with facts, fact groups etc. should be clear and these problems can be debated in the forums.

### 6.4.2   The Four Labels Set

It is preferable that all websites are labelled, as ambiguity about content will destroy the system. Surfers using labelling will not wish to enter unlabelled sites as they appear as 'unknowns' on the system. Since labelling sets are likely to be detailed, in their final form, a gradual rollout may help.

The first and simplest labelling set contains four labels i.e. 'no offensive content', 'offensive content', 'mixed' and 'dynamic'. Dynamic would be used by search engines that respond to your Template. If such a label set is rolled out first, then the bulk of the Internet, websites with no offensive content, will be labelled as such. The objective of this set is to get websites with no offensive content labelled as such. If the system is designed correctly such websites will not need to update their labels when new label sets are released i.e. the 'no offensive content' label will remain while the other labels will be replaced by more detailed labels.

This can be achieved by having forward compatible versioning of label sets i.e. this Four Label Set will be Version 1.0 of the core labelling set. Each new version will keep some of the previous set and optionally break/replace some of the previous set. Implemented correctly each label set version release will result in a portion of Internet websites being updated rather than every website being updated every time. Though this appears to be a rollout issue it is described here as the most basic labelling set.

### 6.4.3   Ascertainable Facts, Ascertainable Opinions and Transition Porn

*Ascertainable Facts & Ascertainable Opinions*
**In the context of Fabel, Ascertainable Facts are facts that can be determined or discovered definitely from the media content. For example, the level of nudity present is ascertainable from pictures/movies. However the actual age of individuals depicted in that content cannot be ascertained simply by looking at that media.**

**In the context of Fabel, an Ascertainable Opinion is an opinion that can be formed from the media content alone i.e. we can retrieve sufficient data from the media to form that opinion.**

*Transition Pornography*
**Legally Transition Pornography does not exist. Legally pornography is either legal or illegal i.e. adult porn or child porn. With Transition porn it does not matter whether or not it is legal. It looks illegal and a lack of evidence either way undermines our ability to deal with it. Transition porn is pornography where due to our inability to determine the age of individuals depicted, pornography cannot be clearly determined to be adult porn or child porn. Transition, in this case, refers to the grey area between legal and illegal.**

**Personally, I use this term o refer to 'dubious' content that is presented as legal adult content.**

One issue to be addressed is different ages of consent in different regions of the world. Throughout Europe alone the ages of consent range from 13 to 17 ([www.InHope.org](www.InHope.org)). The ages specified in Child Pornography Legislation, in Europe, vary from 16 to 18. If we attempt to place the age of individuals into our facts list the system will fail for the following reasons:

- Paedophiles will search for content labelled with the lowest possible age.

- Commercial websites will identify this market and attempt to supply this market with individuals below the lowest acceptable age with the lowest acceptable age placed in the labels. Since this is legally Child Pornography and yet no proof exists, then this qualifies as Transition Pornography.

Clearly the issue of age must be addressed. Here are some ideas:

- Where the rule of law can be trusted, in our own country, content can be presented and appropriate evidence of age presented to the authorities.

- Outside of our legal jurisdiction a more realistic approach of 'if it looks like child pornography it is child pornography' should be adopted i.e. use Ascertainable Opinions.

Basically we will serve up our own population with content supported by the law and you can server up your own population with whatever you want. On an international basis content should be assessed on its appearance, Ascertainable Facts and Ascertainable Opinions only, no 'just take our word for it' non-sense.

In theory this sounds great, in practice no technologies exist to deliver on these requirements in a meaningful manner. Content from one website can always be 'bounced' off other websites that are permitted to 'export' content. This will prevent censorship type Firewalls from working on a National or International level.

This, in fact, is addressing problems outside of Fabel's requirements. However we need to address this because labels in the core label set should not assist paedophiles or businesses wishing to make money from them. Put simply we cannot have labels that indicate age or any other factor that can be hijacked by paedophiles or any other group of people. Debates, assessing possible facts should consider this possibility. No doubt other 'check list' types questions will need to be checked for each fact e.g. 'will this fact be grouped with a sufficient number of other facts to deny the possibility of searched for this fact'.

## 6.5  Implementation / Roll Out

Should government and industry accept the technical, legal and political aspects of the solution then the challenge of creating a core labelling set should be addressed first. The website approach described above is my preferred approach.

If negotiations to create a Core Label Set take place, then software companies are likely to develop software in anticipation of acceptance of the system. Developing software before a core labelling set is agreed is risky. There is no guarantee that such a set can be agreed. Without Regrouping of facts the system may not be welcomed by participating countries, and may not be implemented. Also problems not predicted in this document are likely to appear. The system's design may or may not respond to additional requirements favourably.

In assessing this solution a thorough set of software requirements is likely to be documented. It is preferable that such specifications be published to help companies developing software for the system.

In the event that everything works and sufficient countries seek implementation then the system will be rolled out. This is likely to take a considerable period of time. The Bubbles system will link URL's to labels first and over time file formats may be modified to include labels, allowing embedded data to be added.

Some remaining points concerning rollout:

- Parents are usually less computer literate than their children. As such effort should be made to make the system easy to use and configure.

- Where aspects of the solution appear overwhelming or overly labour intensive, IT productivity solutions should be sought or division of labour solutions like having webmasters label their own websites.

- Since considerable work will be required the system should be rolled out in a large number of small steps e.g. introduce the 'The Four Labels Set' first and then introduce other labels in phases.

## *6.6   Conclusion*

Though a number of features of this solution distinguish it from previous attempts to censor the Internet. This solution is still lacking in a number of key areas. Should this system be accepted and adopted it will merely constitute one brick in the wall of solutions currently required by the Internet and IT industry in general.

Problems not identified in this document are likely to appear as we attempt to build and implement the system. Interactions/uses of the system not previously predicted are also likely.

Finally other requirements not addressed here may impact the industry's interpretation of this solution. Seeking to create a portion of the Internet for children, for industry, for different religions etc. may propel labelling in directions not predicted here. However, the discussions presented here concerning abuses of the system, using division of labour, breaking problems into fact and opinion and others will undoubtedly impact further efforts in this area. This work may simply prove to be a foundation for further work.

# Glossary

### *Age Fact Table*

An Age Fact Table specifies the youngest acceptable age certification for each possible fact. Here is a mock up of an Age Fact Table with five facts/fact groups.

| Fact Number | Youngest Acceptable Age |
|---|---|
| 1 | 8 |
| 2 | 8 |
| 3 | 16 |
| 4 | 12 |
| 5 | 10 |

*Table Glossary.1 – Age Fact Table*

### *Bubbles*

A system for grouping URLs with the same Fabel labels.

### *Democratic Law*

The concept that laws become unenforceable as support from the population declines.

### *Embedded Data*

In this content, embedded data refers to the idea that Fabel labels can be embedded in media files.

### *Fact Based Labels*

Labels for Internet media that indicate facts about the media content.

### *National Age Rating Certificates*

Any given nation's age ratings for movies, computer games etc. e.g. '12', '16' indicating content suitable for 12 year olds and 16 year olds respectively.

In the following table the first three rows indicate the National Age Rating Certificates for countries X, Y and Z.

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Country X** | 18 | | 12 | | | 10 | | | 8 | | |
| **Country Y** | 18 | | | 16 | | | | 12 | | | ▓ |
| **Country Z** | 21 | 18 | | | 16 | | | 12 | | 10 | |
| **Fact** | A | B | C | D | E | F | G | H | I | J | K |

*Table Glossary.2 – National Age Rating Certificates*

### *Surfer's Age Rating*

The censorship setting for this surfer e.g. a 13 year old in Ireland may have their setting set to '12' meaning content suitable for individuals of 12 or more years will be displayed. Content requiring older age ranges e.g. 18 or over, will not be displayed.

## *Templates*

A surfer's Age Rating is combined with their Age Fact Table to identify which facts can be true and still allow this surfer to view the content.

| Country X | 18 | | 12 | | | 10 | | 8 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Fact | A | B | C | D | E | F | G | H | I | J | K |

| Country Y | 18 | | | 16 | | | 12 | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Fact | A | B | C | D | E | F | G | H | I | J | K |

| Country Z | 21 | 18 | | 16 | | 12 | | 10 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Fact | A | B | C | D | E | F | G | H | I | J | K |

*Table Glossary.2 – Templates, coloured area, for a '12' certificate.*