

インターネットにおける 次世代レーティング／フィルタリングシステム

国分明男* 清水昇**

*財団法人 ニューメディア開発協会
〒108-0073 東京都港区三田 1-4-28 三田国際ビル 23 階
E-mail: kokubu@nmda.or.jp

**財団法人 ニューメディア開発協会
〒108-0073 東京都港区三田 1-4-28 三田国際ビル 23 階
E-mail: shimizu@nmda.or.jp

あらまし

筆者らは事務局を務める電子ネットワーク協議会と共に、PICS準拠のフィルタリングシステムの開発を実施し、これに基づき平成9年9月からホームページの格付けデータベースシステムの運用を行ってきた。また、より実効性のある対応をするために、それらの開発・運用経験を生かして昨年、次世代レーティング／フィルタリングシステムの開発を行ってきた。

本報告では、レーティング／フィルタリングシステムを解説し、開発した次世代レーティング／フィルタリングシステムの特長および本年5月から運用サービスを開始した後の状況を述べる。

キーワード

レーティング、フィルタリング、PICS、RSACi、自動化、プロキシ

Next-generation Rating/Filtering System on the Internet

Akio Kokubu* and Noboru Shimizu**

New Media Development Association
Mita Kokusai Bldg. 23F, 1-4-28 Mita, Minato-ku, Tokyo 108-0073
*Kokubu@nmda.or.jp, **shimizu@nmda.or.jp

Abstract

Together with the Electronic Network Consortium, we developed a filtering system that conforms to PICS specifications, and has been operating a database system (a labeling bureau complying with PICS specifications) for homepage rating since September 1997. To create a more effective response to improper information on the Internet, we worked on a next-generation rating/filtering system since last year.

In this paper, rating/filtering technologies are explained in general. The next-generation rating/filtering system we developed is introduced and some of operational results are shown.

key words

Rating, Filtering, PICS, RSACi, Automatic, Proxy

1. はじめに

インターネットは、オープンさと自由さによって成功したネットワークであり、膨大な情報（コンテンツ）と全世界で1億人を超える膨大な利用者とを包含し、現在も発展を続けている。しかし、情報発信が容易であり、発信情報が直ちに世界を駆けめぐる特性を有しているため、チャイルドポルノなどの非合法コンテンツや青少年に有害なコンテンツがネットワーク上を広く流通し、大きな社会問題となっている。この問題を解決するために図1に示すように、インターネット産業界としての自主規制、レーティング／フィルタリングによる技術的解決、ホットラインサービス、法規制から成る包括的な提案が行われている¹。

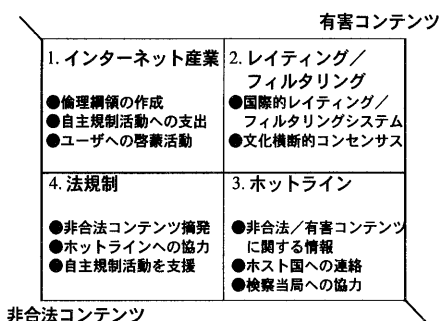


図1 自主規制の国際的システム

レーティング／フィルタリングには、先進的な選択的情報取得・検索機能としての捉え方もあり、電子メールやニュースなどのコンテンツを対象として、さまざまな観点から研究が進められているが²、ここではポルノなどの有害な情報をブロックする機能という捉え方をとする。

筆者らは事務局を務める電子ネットワーク協議会と共に、PICS準拠のフィルタリングシステムの開発を実施し、これに基づき平成9年9月からホームページ・レーティング（格付け）のデータベースシステム（PICS準拠ラベルビューロ）の運用を行ってきた。また、より実効性のある対応をするために、それらの開発・運用経験を生かして昨年

から、次世代レーティング／フィルタリングシステムの開発を行ってきた。

以下では、レーティング／フィルタリングシステムを解説し、次世代レーティング／フィルタリングシステムの特長および本年5月から運用サービスを開始した後の状況を報告する。

2. レーティング／フィルタリングとは

基本的な原理を図2に示す。発信者は発信情報にレーティング（格付け）値を加えて、インターネット上で公開する。受信者のシステムはフィルタリングソフトに設定した値（フィルタリング値、設定値）と格付け値を比較して、格付け値が設定値よりも小さければ表示し、大きければ非表示にする。この仕組みでは、発信側で規制するのではなく、受信側で制御することになるので、世界中の多くの人々によって言論・表現の自由と有害情報の規制の両立ができる技術的解決策という捉え方がなされている。

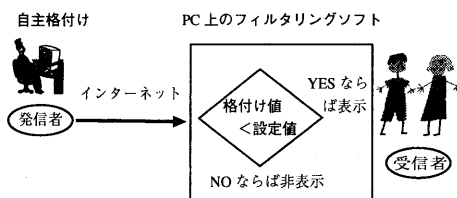


図2 レーティングとフィルタリングとは

PICS (the Platform for Internet Content Selection の略称)³は、WWWの技術規約と標準ソフトを開発するために設立されたWorld Wide Web Consortium (W3Cと略す)が開発した標準仕様の一つであり、フィルタリングサービスの基本モデルとレーティング結果であるラベルの形式とフィルタリングに関する通信プロトコルとを規定している。PICS仕様はW3Cのホームページなどで公開されており、誰でも使うことができる。

PICSに1年遅れて、PICSルールズという標準仕様が追加されている。これは、設定値と格付け値を比較して、格付け値が設定値よりも小さければ表示し、大きければ非

表示にするという単純なフィルタリング機能からくる限界を解決する（例えば、ヌードであっても芸術作品ならば表示できるようにする）ためのものである。利用者の価値観や主観をフィルタリングに反映させることができるようにするため、設定値やポリシーを記述する言語をP I C Sルールズにおいて規定している。P I C Sルールズにより記述されたファイルは、プロファイルと呼ばれる。

P I C S仕様に基づいてフィルタリングを実現するためには、レーティング情報が必要である。発信者レーティング（自主格付け）は、ホームページのHTML記述の中でMETAタグによって、暴力は何点、セックスは何点という形で付加される。サードパーティレーティングは、ラベルの要求に対して応答するラベルビューロと呼ばれる HTTP サーバ上に、発信者レーティングが行われていない URL に対するレーティングの結果であるラベル情報を蓄積したデータベースとして実現される。サードパーティレーティングのデータベースは、異なる価値観に基づき複数存在することが望まれる。受信者レーティングは、受信者の PC 上のフィルタリングソフトに対して受信者が URL リストを入力蓄積することによって実現される。

レーティング情報を規定するためには、レーティング基準（どのような情報にどのような格付けを与えるかの基準）が必要である。その代表例として、図3に示す RSACi (Recreational Software Advisory Council) によるインターネット上のレーティング基準がある。RSACiでは、暴力、ヌード、セックス、言葉のカテゴリがあり、インターネット上のコンテンツが、各々のカテゴリに関して0から4までのレーティング値で格付けできるようになっている。RSACiは、非営利団体RSACによるレーティング基準であるため、米国では発信者レーティングの際の実質的な標準になっており、米国プレイボーイ社のホームページwww.playboy.comでは早い時期から、この基準に基づいてセルフレーティングされていることで有名である。本年5月にRSACの

知的所有権を継承したICRA (Internet Content Rating Association) ⁴における国際的なレーティング基準作成の際にも、この基準に基づいた検討が進められている。

格付け基準例 RSACi

<ul style="list-style-type: none"> ■暴力 0 すべての暴力を制限 1 闘争 2 殺害 3 流血を伴う殺人 4 残忍で過激な暴力 	<ul style="list-style-type: none"> ■ヌード 0 なし 1 露出的な服装 2 部分的なヌード 3 全裸 4 刺激的な全裸
<ul style="list-style-type: none"> ■セックス 0 なし 1 情熱的キス 2 着衣のままの性的接触 3 性的接触の不鮮明な描写 4 性行為の鮮明な描写 	<ul style="list-style-type: none"> ■言葉 0 不快感を与えない俗語 1 穏やかな悪口 2 悪口 3 性的なジェスチャ 4 不快を与える露骨表現

図3 レーティング基準例-RSACi

3. 次世代レーティング/フィルタリングシステム

筆者らは、図4に示す全体図の下で、実験用のレーティング基準を作成し、ラベルビューロにおけるP I C S準拠のサーバソフトを開発し、レーティングデータベースを構築し、P I C S準拠のフィルタリングソフト (Windows95 対応版および Macintosh 対応版 I F S、Internet Filtering System) を開発し、これに基づき平成9年9月からラベルビューロの運用サービスとフィルタリングソフトの配布サービスを行ってきた。

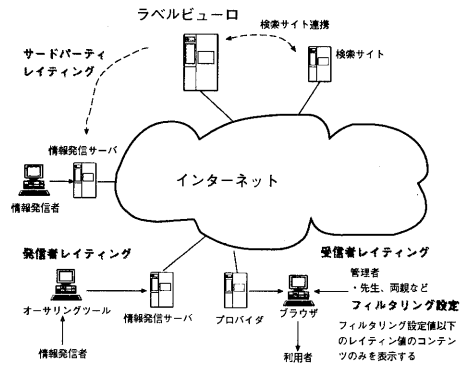


図4 これまでのシステム全体図

このプロジェクトの開発成果である、ラベルビューロのデータベース規模は2万URL

を越え、パソコン搭載のフィルタリングソフトは既に2千以上の教育機関や企業等に利用されている状況にある。しかしながら、データベース構築に当たっては、各ページ毎に目視によりラベル付加作業を実施しなければならないため、日々増加する不適切な情報に十分対処することが困難であった。また、フィルタリングソフトは使用するブラウザやOSのバージョンに依存するため、利用できるユーザー環境に制限があり、学校などのように、多数のパソコンを有する利用者からは、個々のパソコンにフィルタリングソフトをインストールし、設定する作業は煩わしいので解決して欲しいとの要望が出されていた。

このように実際に開発と運用を行うことにより、インターネット上の膨大な情報に対してフィルタリングの有効性をさらに高めるための今後の開発課題も明らかになった。そこで、それらの開発・運用経験を生かして昨年から、次世代「レイティング/フィルタリングシステム」の開発を行ってきた。以下では、図5に示す次世代「レイティング/フィルタリングシステム」の特長および本年5月から運用サービスを開始した後の状況を報告する。

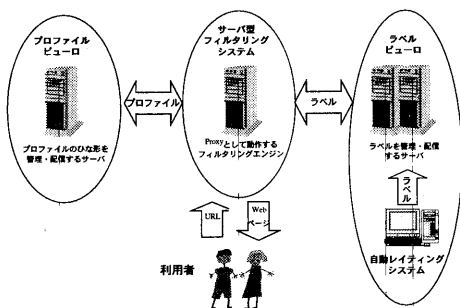


図5 次世代「レイティング/フィルタリングシステム」の構成

3. 1 次世代レイティングシステムの特長

今回開発したシステムにおいては、インターネット上の各ホームページを、テキストのキーワード解析のみならず類似画像認識技術を用いて、効率的にレイティングしてラベルを付加する自動レイティングシステムを開発

すると共に、他のラベルビューロとの間の連携、レイティング基準が異なるラベルの変換機能などを備えたPICS準拠のラベルビューロを開発することにより、日々更新される情報に迅速に対応し、不適切な情報に対するカバー率を拡大できるようにした。

ラベルビューロは、PICS準拠であるため、Internet ExplorerなどのPICS準拠のブラウザやPICS準拠の市販フィルタリングソフトから利用することができる。開発したレイティングシステムの主な特長は、以下の通りである。

(1) テキストと画像を自動レイティング

- ホームページ上のテキスト情報に対して、情報カテゴリ毎の重み付きキーワードリスト（現在、不適切語は約6千語、適切語は約2万6千語）と突き合わせを行い、分類することにより、レイティング値を決定する。目視による確定作業を経て、キーワードリストを更新することにより、自動化の程度を高めることができる。
- ホームページ上の画像情報を効率的にレイティングできるようにするため、類似画像データベース（現在、約2千画像）の検索に基づきレイティング値を決定し、画像情報をサムネール化（閲覧性を高めるために元の画像を小さくし、表示すること）し、目視により効率的にレイティング値を確定する。確定画像を類似画像データベースに登録することにより、自動化の程度を高めることができる。
- ラベルデータベースの陳腐化を防ぐため、データベースに蓄積されているラベル情報に対応するURLを定期的にアクセスし、更新日時やサイズを調べて、変更があれば、自動レイティングの対象とすることができる。
- 自動レイティング対象のリストアップは従来と同様に、主要検索サイトにおける検索結果のクリッピングやロボット検索などに基づいて行うことができる。

(2) ラベルビューロ間の連携

全世界の膨大な数のホームページを、単一のラベルビューロで網羅することは不可能である。本ラベルビューロは、レイティング値の問い合わせがフィルタリングソフトから行われたとき、そのホームページのラベルデータが存在しない場合は、他のPICS準拠のラベルビューロに代行して問い合わせを行い、レイティング値を返すことができる。その際に、レイティング基準（どのような情報にどのような格付けを与えるかの基準）が異なるラベルビューロであれば、当協会のレイティング基準にラベルを変換することができる。

なお、レイティング基準としては、国際的に広く使われている米国の非営利団体RSACの基準（RSACi）をベースとするレイティング基準（SaftyOnlineと呼ぶ）を使用している。

(3) 高速・高信頼性の実現

ラベルビューロは、複数台のUNIXマシンにより構成されている。これらの中で負荷分散処理を行っているために、快適なレスポンスが得られる。また、1台のUNIXマシンが故障しても、他のUNIXマシンが代替処理を行うので、可用性が高くなっている。

3.2 次世代フィルタリングシステムの特長

学校などからの要請に基づいて、サーバ型フィルタリングシステム（SFS）を開発することにより、学校のように多数のパソコンを有する利用者や、NetscapeなどのPICS未対応ブラウザ利用者におけるフィルタリング機能のインストールを容易にした。主な特長は、以下の通りである。

(1) プロキシによりフィルタリング

- 学校などにおいてフィルタリングを実施しやすく、プロキシサーバによりフィルタリングを行う。このことにより、フィルタリング機能の設定や管理を個々のパソコンごとに行う必要がなくなる。
- SFSはJavaで記述されており、JDK1.1.6が動作するコンピュータ環境であれば、稼働させることができる。

OSとして、Linux、Solaris、Windows NTで動作することを確認済みである。

- ホームページ閲覧のみならず、ファイル転送、ネットニュースにも対応している。

(2) プロファイルビューロを用意

SFSは、PICSルールズ対応であるので、先生や親などのアクセス管理者が、SFSにフィルタリング値やポリシーを記述したプロファイルを設定しなければならない。しかしながら記述が容易とは言い難いので、ニーズや管理者のスキルに合わせた3通りの記述方法が可能にしている。自由度は低いけれども最も容易な方法として、プロファイルの共有とインストールのためのプロファイルビューロを用意されており、そこには、世代別、用途別のひな型プロファイルが登録され、アクセス管理者がダウンロードして使用できるようになっている。

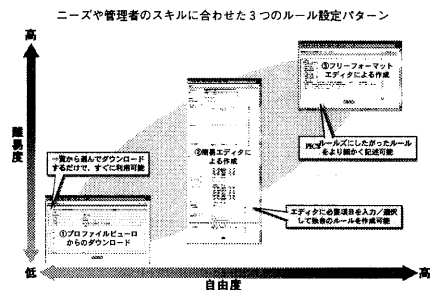


図6 フィルタリングのルール設定

(3) 利用者に合わせて、様々なレイティングの活用が可能

SFSでは、①学校の先生などのSFS管理者がレイティングしたもの、②第三者がレイティングしたもの（当協会によるものなど）、③発信者が自らレイティングしたもの（例えば、米国のプレイボーイ社のページなどが実施しているもの）の3つが利用可能である。複数のレイティングが実施されている場合には、上記①、②、③の順で優先してフィルタリングを行う。

3. 3 フィルタリング機能の実現方法⁵

今回開発のシステムでは、以下の3種類の利用方法によってフィルタリング機能を実現することができる。

(1) SFSのインストール

SFS(サーバ型フィルタリングシステム)を、学校やプロバイダの主ネットワーク管理者が中継サーバにプロキシとしてインストールすることにより、フィルタリング機能を実現する方法。SFSは当協会が無償配布しているが、SFSの稼働環境が主としてUNIXであり、サーバへの組み込みの際に技術的サポートが必要と考えられるため、オンライン配布は行わずメールによる申し込みにより対応している。また、ブラウザのプロキシ設定変更を阻止するためのツール(リダイレクタ)も配布している。

(2) 公開プロキシの利用

これまでプロキシを経由せずにプロバイダに接続しているユーザの場合、当協会が提供しているインターネット上の公開SFSをブラウザのHTTPプロキシとして設定することによるフィルタリング機能の実現方法。この方法により、PICS対応でないブラウザ(Netscape Navigator)を用いて、一般ユーザが簡単にフィルタリング機能を実現することができる。

(3) ラベルビューロへの直接問い合わせ

PICS対応のInternet Explorerをブラウザとして使用して、プロキシを経由せずにラベルビューロに直接問い合わせをすることによってフィルタリング機能を実現する方法。ブラウザのコンテンツアドバイザーにラベルビューロを設定する。但し、Internet Explorerの現行バージョンはPICSルールズ対応でないので、フィルタリングポリシーなどが記述されたプロファイルを設定することはできない。

4. 運用サービス開始後の状況

(1) 自動レイティング作業

開発した自動レイティング機能を用いて、主としてわが国の不適切な情報を定期的にレイティングし、データベースを大規模化した。既に25万URL以上が登録されており、1千URL/日程度の目視による確定効率でレイティング作業を継続中である。既に登録されているURLをアクセスし、更新日時やサイズを調べて、変更があれば、自動レイティングの対象とすることができるので、サンプリング的に更新状況を調べたところ、10%~30数%のURLが更新されている。

自動レイティングと目視による確定作業との関係は、テキストではカテゴリにも依存するが、比較的辛めに自動レイティングされる傾向がある。しかし、見落としを防ぐことができるなどのメリットがあるので、支援ツールとしての実用性は高い。画像については、類似画像データベースの規模が小さいので、補助的にしか使われておらず、今後、大規模化して実用性を高める必要がある。

(2) SFSの利用状況

本年4月までの運用と比べると、今回開発したプロキシサーバを用いた県全体や市全体での公立学校への導入など大規模な利用が始まっている。例えば、本年秋の新学期から新潟県スクール防犯連絡協議会に加盟の主要プロバイダに接続する県内ほぼ全域で、小中学校からインターネットにアクセスする際にはSFSをプロバイダにインストールすることによってフィルタリング機能が働くようになっている。このような大規模利用に対応するために、プロキシサーバの性能向上や信頼性向上などの多くの改善を続けている。

参考文献

1. Self-regulation of Internet Content, Bertelsmann Foundation, Sept. 1999
2. www.ee.umd.edu/medlab/filter/resources.html
3. <http://www.nmda.or.jp/enc/w3c/> (和訳)
4. <http://www.icra.org/>
5. <http://www.nmda.or.jp/enc/rating/>