

2 インターネットにおける有害情報問題および個人情報保護問題への対策

開発本部 次長 電子ネットワーク担当 政田 十喜雄

1. はじめに

インターネットの急速な普及により、有害情報問題と個人情報保護問題が顕在化しています。政府のミレニアムプロジェクトにより、全国の公立学校へのインターネットの導入が進み、新学習指導要領では情報関連科目が必須となるなど、インターネットの本格的な利用が進められています。インターネットを学校の授業などで使うと、海外で作られた生きた英語を読むことができ、世界中の生の情報を得ることができるといった、非常に有益な側面があります。その一方で、インターネットには、本来18歳未満の青少年には見せてはならないアダルト情報が簡単に見えてしまうといった負の側面があります。この有害情報問題に対しては、各種フィルタリングソフトやサービスが提供されており、ベンダー毎にWebページの格付け（レイティング）を行い、格付け情報（ラベル情報）をデータベース化したラベルデータベースを構築していますが、この格付け作業が、インターネット上の有害情報の急速な増大に十分に追いつけない現状があります。また、Webページの移転や削除が多く、折角集めても約1年で半数のラベル情報（格付け情報）が無効となるという現実があります。上記のラベルデータベースに基づきフィルタリングを行う方法のほかに、有害キーワードを多く含むWebページを自動検出して遮断する方法もありますが、過剰ブロックとなり有益なコンテンツまで遮断されてしまうという課題があります。従って、各団体間でラベル情報を共有したり、情報発信者が自らレイティングを行うセルフレイティングを普及させるなど、何らかの対策が必要となっています。

プライバシーについても頻発する個人情報の漏洩や不正利用に対して（財）日本情報処理開発協会（JIPDEC）の「プライバシーマーク制度」等と連携した技術的解決が強く望まれています。

このため、当協会は（財）インターネット協会（旧電子ネットワーク協議会）と協力して、ラベル情報の共有に向けたレイティングシステムの開発とラベル情報の収集及びWeb情報発信者が自らのページをレイティングするセルフレイティングの推進、また、それらの実験システムとしてのフィルタリングシステムの開発と、インターネットにおけるプライバシー情報管理システムの開発を行って来ました。

2. PICSの重要性

PICS（Platform for Internet Content Selection）はレイティング/フィルタリングの為にWebの国際的標準化団体であるW3Cが策定した国際標準規約であり、これを用いるとフィルタリングに必要なラベル情報の共有などが容易に行えます。PICSはラベル情報の記述、交換、通信プロトコルおよび、受信者側のブロック条件を記述したプロファイルに関する規定を含みます。またラベル情報の問い合わせ先としてのラベルビューロ（PICSに基づくラベルデータベースをいう）およびラベルビューロを参照するための規約も規定しています。ラベル情報は生ものであり、頻繁にページが移転したり、削除されたり、新しいページが作成されたりしますが、PICS標準を用いれば、ある団体がレイティングした結果（ラベル情報）をWeb上でリアルタイムに共有するといったことが可能となります。

またPICSを用いると、情報発信を制限することなく、受信者が主体的に情報を選択することができます。PICSでは、Webページに格付けを行うことと、この情報を元にフィルタリングすることが完全に分離されています。このため、情報に格付けを行うことで情報発信を制限することなく、また、子供や社員などの管理者といった情報の受信者側が格付けのレベルを設定することにより、受信者が主体的に、望ましくないと判断する情報を選択的に受信することができるのが重要です。

2.1 PICSの利用形態

PICSは現在、図1の～のようなところで実際に利用されています。

(1) Webページのセルフレイティング(図中)

Webページ作成者が自分のページを自らレイティングするセルフレイティングは、PICS標準で記述されているため、PICSに対応したブラウザ（マイクロソフト社のInternet Explorerなど）やフィルタリングソフトを用いることによってフィルタリングを行うことができます。

(2) 第三者によるレイティング(図中)

第三者によるレイティング結果をPICS標準で格納する事により、データの相互流通性が高まります。当協会はインターネット協会と協力してPICS標準に基づいてレイティング作業を行っています。

(3)ブラウザからのラベルビューロ参照(図中)

PICSはラベルビューロを参照するための通信規約も標準化しています。Internet Explorer (IE) はPICS標準を採用しており、IEからラベルビューロを参照して、フィルタリングを行うことができます。

(4)フィルタリングソフトからのラベルビューロ参照(図中)

PICSを採用しているフィルタリングソフトからラベルビューロを参照し、その情報を用いてフィルタリングを実行することが出来ます。当協会がインターネット協会と協力して開発したフィルタリングソフト(SFS)を用いると、ラベルビューロを組み入れた実験システムを構築することが可能です。

(5)ブラウザのプロファイル記述(図中)

PICSは受信者のブロック条件を記述するプロファイル記述についても標準化しています。IEのプロファイル記述を用いることにより、情報発信を規制することなく、受信者が受信情報を選択することが可能となっています。

(6)フィルタリングソフトのプロファイル記述(図中)

PICSを採用しているフィルタリングソフトでは、プロファイル記述とPICS標準のラベル情報とを合わせて運用することにより、情報発信を規制することなく、受信者が受信情報を選択することが可能となります。当協会がインターネット協会と協力して開発したフィルタリングソフト(SFS)はPICSを採用しており、これを用いると、プロファイル記述を利用した実験システムを構築できます。

2.2 レイティング基準について

実際に格付けするためには、PICS標準の上にレイティング基準を策定して行われます。格付けする方法は、大きく二つあり、一つは第三者的な立場で客観的にWebページ・Webサイトを、ある基準に基づいて格付けを行う場合です。これをサードパーティーレイティングといいます。当協会ではインターネット協会と協力して開発したPICSベースのSafetyOnlineというレイティング基準に基づいて格付けを行い、その情報を提供しております。もう一つの方法は、Webページを作成・公開する情報発信者側が、自らの情報の格付けを行うことです。これをセルフレイティング(またはセルフラベリング)といいます。なお、その他のレイティング基準としては、ICRA、RSACiなどがあります。これらの基準は全てPICSに準拠しています。(<http://www.iajapan.org/rating/>)

3. ICRAセルフラベリングツールの開発と普及活動

サードパーティーレイティングによる方式では、急激に増加するコンテンツやアドレスの移転に追いつけない現状があります。また、有害キーワードを多く含むWebページを自動検出する方式では、有益な情報までブロックされるという課題があります。これらの問題を解決しようという方策がセルフラベリングです。

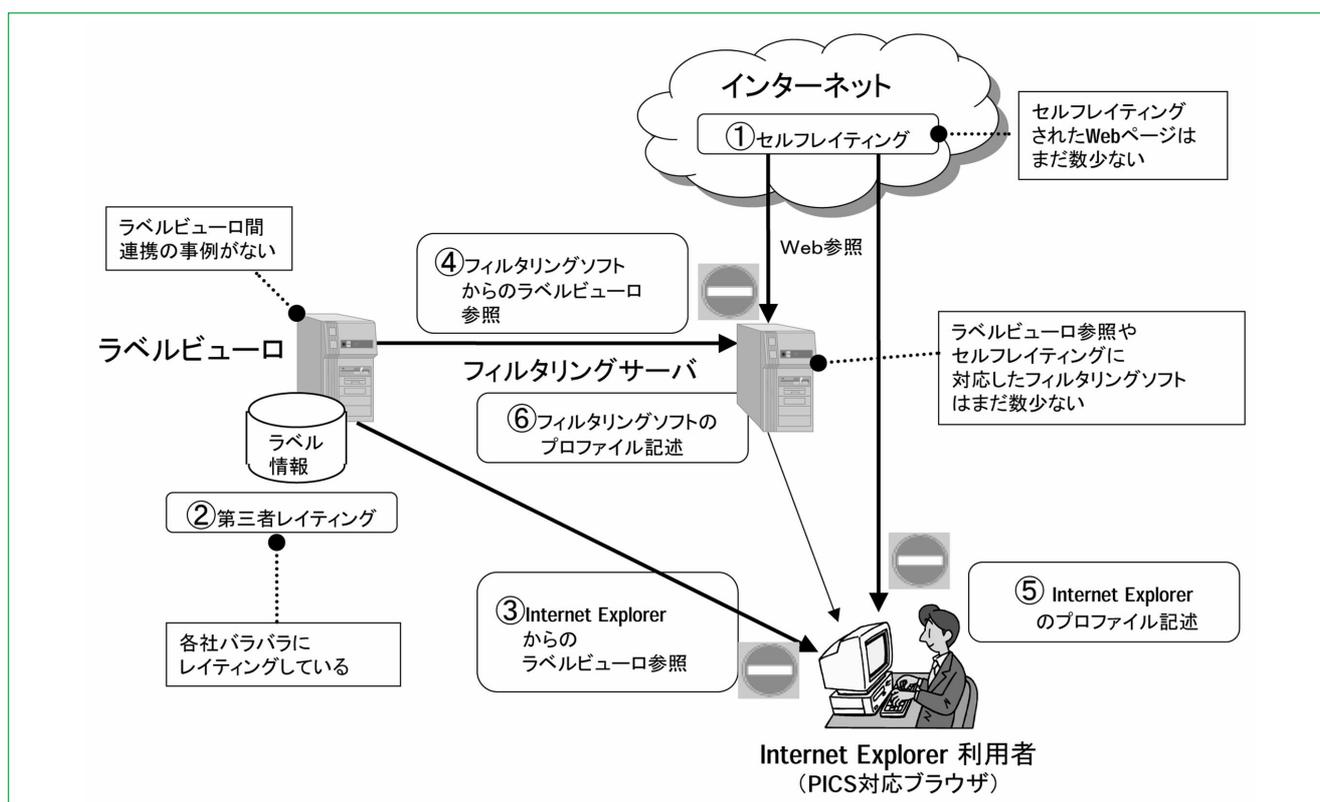


図1 PICSの利用

3.1 セルフラベリングとは

英国の非営利団体ICRA（Internet Content Rating Association）は、従来のRSAC（the Recreational Software Advisory Council）によるレイティング基準RSACi（RSAC on the internet）を見直したICRAラベリング基準（ICRA基準）を開発し、2000年末には同基準に準じたICRAラベリングシステムをWeb上で公開しました。ICRA基準は、基本的に、情報発信者によるセルフラベリングのために作成された基準です。

ICRA基準では、RSACiのように各カテゴリ（ヌード、セックス、暴力、言語）ごとにレベル分けするのではなく、コンテンツを客観的に記述したコンテンツ記述子（「明白な性行為」「人間に対する傷害行為」等）と、そのコンテンツが置かれた文脈を記述したコンテキスト記述子（「明白な性行為」「人間に対する傷害行為」）から成り（記述子の総数は45）、情報発信者が自らのコンテンツに対して、これらの記述子それぞれについて該当する該当しないかをラベル付けすることになります。従来からマイクロソフト社のブラウザInternet ExplorerにはPICS準拠のラベル情報に基づきWebページをフィルタリングする機能（コンテンツアドバイザー）がありますが、今後Internet Explorer 7.0ではICRA基準のラベル情報に対応する予定があります。ICRAはEU委員会から予算を得て開発を進めており、欧州を中心にICRA基

準が浸透していくと予想されます。当協会はインターネット協会と協力し、ICRAと連携して進めることにより、Webページにセルフラベルを付与するためのセルフラベリングツールの日本語版を開発しました。このセルフラベリングツールのサービスは、Webサイト構築者の方のために無償にて提供する予定となっております。（<http://www.iajapan.org/rating/>）

3.2 格付けされたWebページのフィルタリング

セルフラベリングで格付けされたWebページのフィルタリングには、PICSに準拠したフィルタリングソフトウェアやブラウザなどで、受信者が閲覧可能な格付けレベルを設定することにより行います。図2に、ICRAセルフラベリングで格付けされたWebページをフィルタリングする方法を示します。当協会はインターネット協会と協力して、セルフラベリングに対応したフィルタリングソフトウェア（SFS）をフリーで公開しております。また、上述のようにInternet Explorerのコンテンツアドバイザーを使うこともできます。

4. ラベル情報の作成と

フィルタリングシステムの提供

4.1 日英Webページのレイティング

当協会ではインターネット協会と協力して、これまで日本語Webページのレイティング作業を、半自

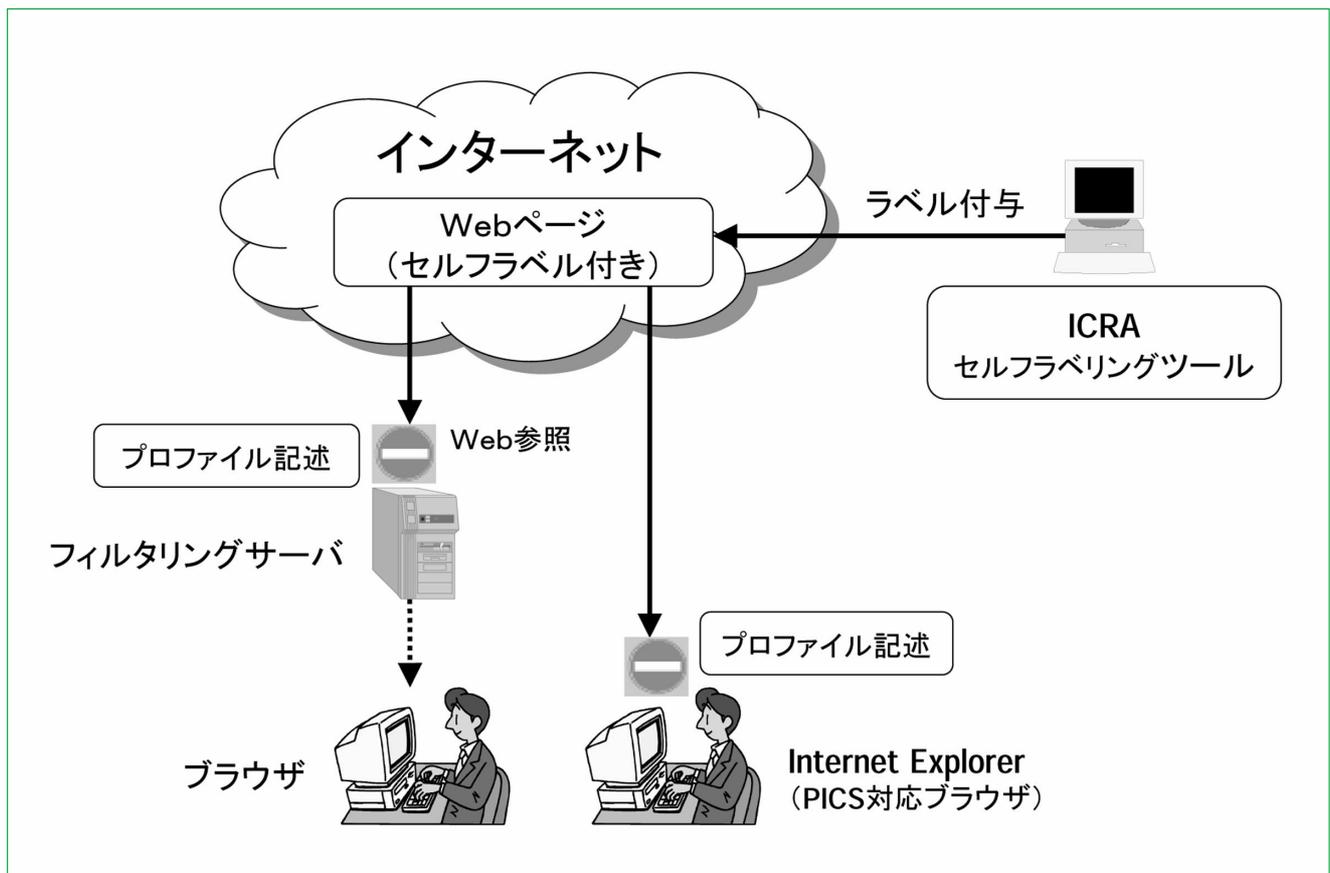


図2 ICRAセルフラベリング

動レイティングツールを利用して日々行っており、これまで50万件のラベル情報（格付けされたURL情報）をラベルビューロとして無償公開していますが、今年度はさらに英語Webページもレイティングの対象に加えました。これらのラベル情報の有効性はWebページの移転や削除により約1年で半減するため、レイティング作業を継続的に進めていく必要があります。半自動レイティングの手順は、まず有害キーワード検索と類似画像検索により自動的にレイティングを行い、次に人手によりレイティング値を確定します。当初、人手のみでレイティングを行っていた場合に比べると作業効率は数十倍に向上しています。また、最終的には必ず人がレイティング値の確認を行っているので、収集したURLに関して、レイティングの誤りは殆どありません。これらのラベル情報は、上述のPICSに準拠しており、異なるシステム間でのラベル情報を相互運用する際に用いることが可能となっています。

4.2 ラベル情報のシームレスな配送と利用

当協会およびインターネット協会ではラベル情報を作成すると共に、他団体との相互協力によってラベル情報を収集し、PICSを用いてリアルタイムにラベル情報を共有することを可能としました。図3の様に配送サーバを多段配置し、フィルタリングシステムへラベル情報をリアルタイムに配信することがで

きます。配送サーバとラベルビューロは、インターネット協会のホームページより無償でダウンロードして、図4の様なシステムを構築して利用できます。Internet Explorerのコンテンツアドバイザーからラベルビューロを参照したり、各種フィルタリングソフトからラベル情報を利用することにより、フィルタリングシステムを容易に構築することが可能となりました。

4.3 フィルタリング実験システムの提供

通信・放送機構（TAO）の学校インターネットプロジェクトでは、学校へインターネットを急速に普及させ、かつその利用環境を整える活動を行っていますが、当協会がインターネット協会と共に開発したフィルタリングシステムはそのプロジェクトの実験システムに採用されています。

4.4 今後の展望

当協会はインターネット協会と協力して、配送サーバおよびラベルビューロを提供し、PICS標準を普及させることにより、ラベル情報の共有化を進めていきます。これにより、各ベンダーの提供するフィルタリングソフトが捕捉する有害URL件数の増大に資することを目標とします。また、海外機関との連携を進める事により、海外Webページのラベル情報の充実化も図っていきます。また、セルフラベリン

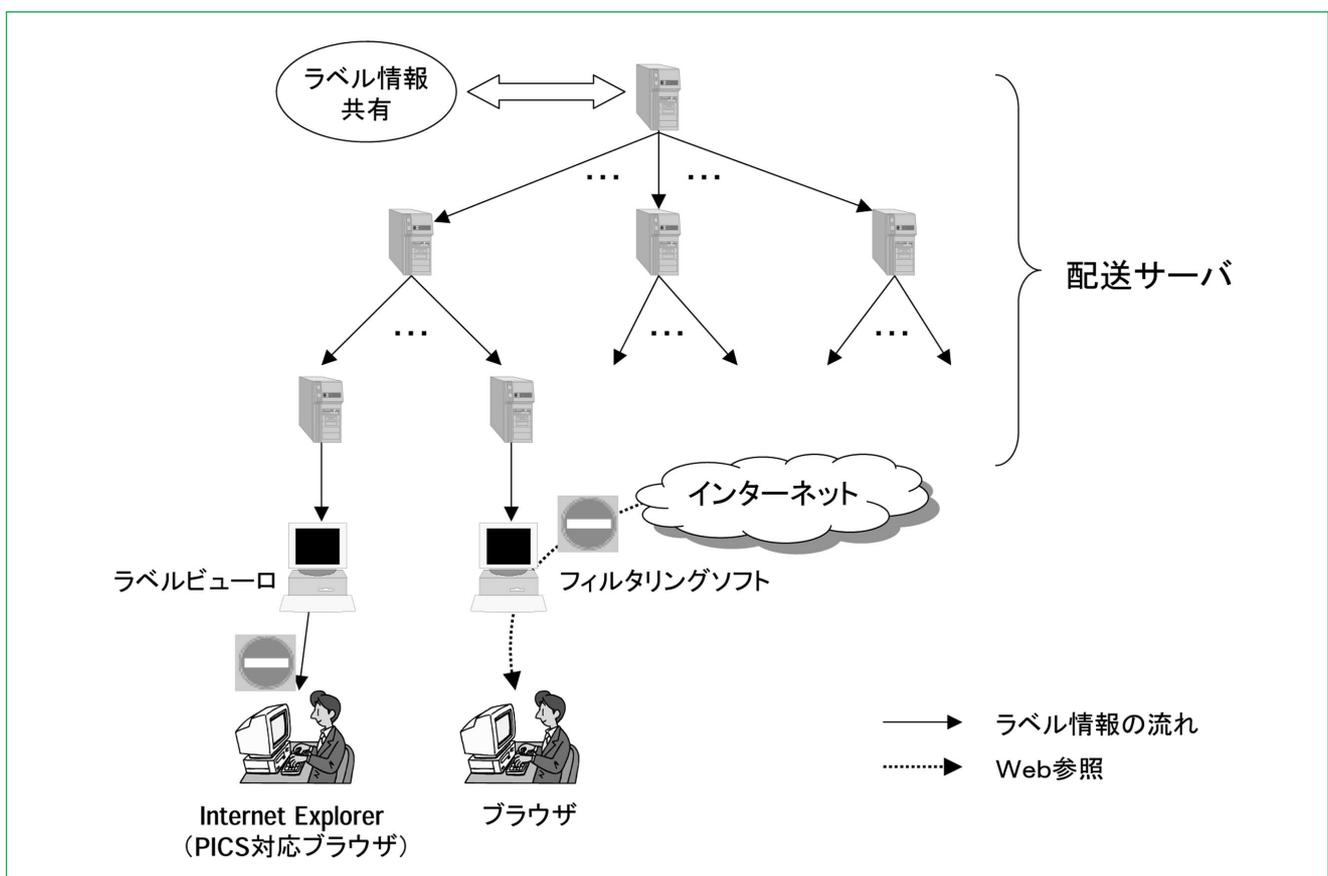


図3 配送サーバの多段構成

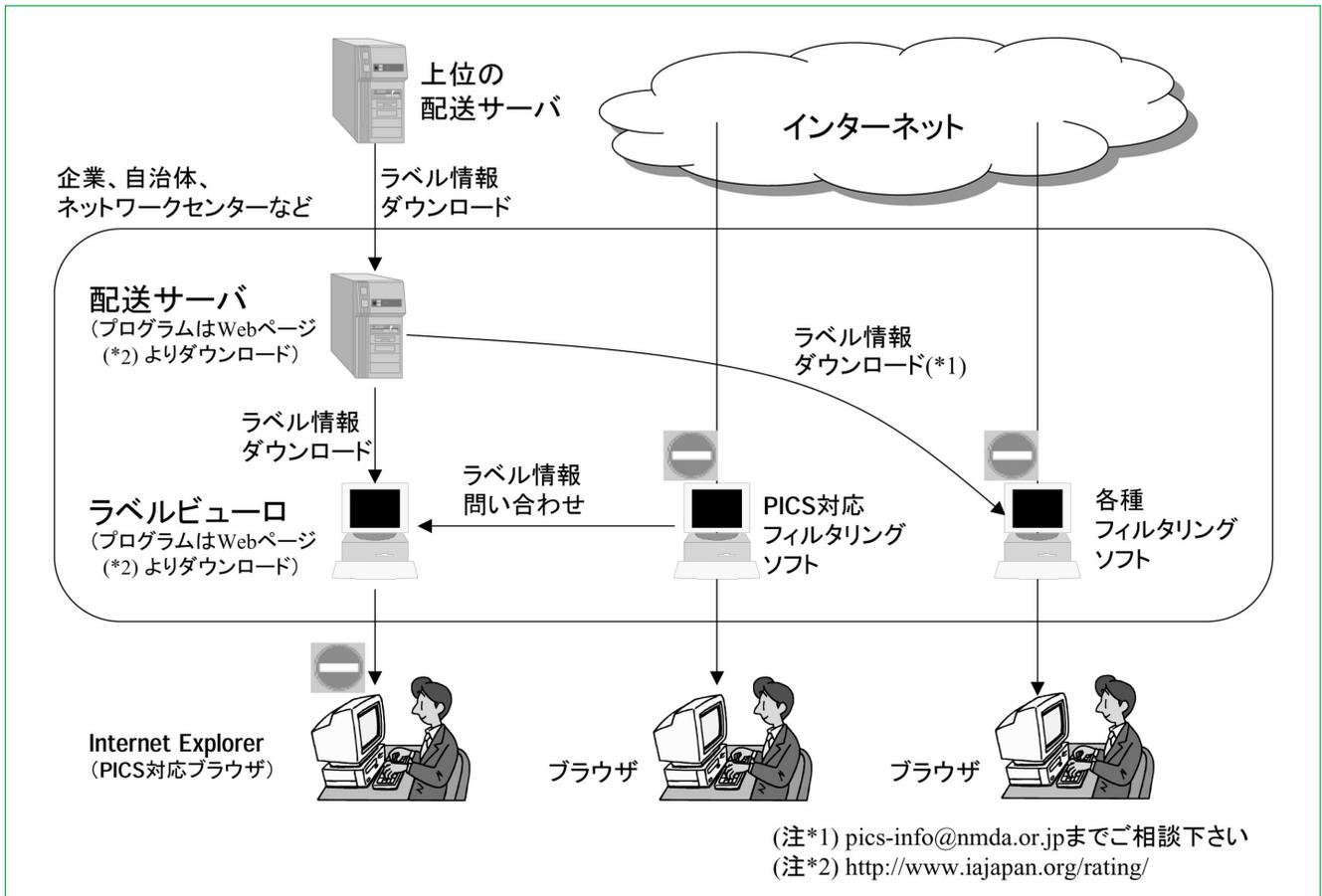


図4 配送サーバとラベルビューロの利用イメージ

グを普及促進していきます。これにより、ラベルが付与されていないWebページのみをレイティングすれば良くなり、レイティングの効率化と高品質化を促進することが出来ます。

5. インターネットにおける「プライバシー情報管理システム」の開発

5.1 背景

オンラインショッピングに限らず、インターネットにおける取引においては少なからず利用者の個人情報収集しますが、これらの収集された個人情報が利用者の知らない間に他に流用されたり、不正に使われたりする問題が数多く発生しています。こういった問題に対して一つの側面から改善する方法として、収集された個人情報の利用目的や開示範囲を情報収集の前に予め利用者に提示して同意を得るようにし、個人情報が適切に取り扱われるようにするインターネット上のプロトコル（P3P）の開発がW3Cにより進められています。当協会はインターネット協会と協力して1998年からP3P技術を使った研究を行い、1999年に「プライバシー情報管理システム」、2001年に「ポリシーウィザード」を開発・公開し、P3Pの国内への普及活動、およびW3Cへの技術フィードバック等を行っております。

5.2 P3Pとは

P3Pとは、Platform for Privacy Preferences Project（プライバシー情報取扱いに対する個人の嗜好を支持する技術基盤）の略で、W3Cが開発中のインターネットを含むネットワーク上のプライバシー保護を目的とした技術標準です。Webサイトには個人情報の取扱いについて記述したプライバシーポリシーを掲載している所が増えていますが、P3Pの技術を使うことによって、プライバシーポリシーを、標準化されたマシンリーダブルな形式（XML形式）で記述することができます。利用者側では、P3P対応のクライアントツールやブラウザなどを使うことにより、個人情報収集画面においてWebサイトのプライバシーポリシーを参照したり、予め登録しておいたプリファレンス（どのようなサイトに、どのような使用目的で、どのような個人情報を提供して良いかといった個人の嗜好を記述した規則）とポリシーとを照合することができます。最近のP3P仕様の動向としては、2000年6月にニューヨークにてW3C主催でP3Pの発表会が行われ、当協会はインターネット協会（旧電子ネットワーク協議会）と協力してデモンストレーションを行いました。2000年12月にP3Pバージョン 1.0のCandidate Recommendation（勧告候補）が公開され、この半年後にはP3P仕様の一部を採用したMicrosoft社のInternet Explorer 6のリリー

スもありました。このあと、W3Cでは2002年9月にワーキングドラフト、2002年1月にはP3Pバージョン1.0のProposed Recommendations(勧告案)が公開されました。最終的にW3C勧告となるのは2002年前半ではないかと見られます。

P3Pに準拠したシステムとしては、<http://www.w3.org/P3P/implementations>にあるように十数社がサーバ側ツール、あるいはクライアント側ツール、開発ツールなどを開発しています。日本では、当協会のシステムの他、NECが、P3P対応のサーバ構築ツールを開発しています。また、P3Pに対応したWebサイトはhttp://www.w3.org/P3P/compliant_sitesにあるように既に数十サイトがあり、これも増加中です。

5.3 P3Pポリシーウィザードの提供

WebサイトをP3P対応にするためには、プライバシーポリシーをP3P(XML)で記述しなければなりません。これをエディタなどで作成するのは難しい問題があります。そこで、当協会はインターネット協会と協力して、P3P「2000年12月勧告候補版対応」に準拠したP3Pポリシーウィザードを開発し、<http://www.nmda.or.jp/enc/privacy/>にて提供しています。P3Pポリシーウィザードとは、WebサイトをP3Pに対応させるために、P3P(XML)を簡単に作成するツールです。Webサイト作成者は、このP3Pポリシーウィザードを使い個人情報の利用目的や開示範囲を選択する簡単な操作により、ウィザードが作成したP3Pを自分のWebサイトに組み込むだけで、そのWebサイトはP3P対応となります。組み込む方法には、(a)HTMLファイルの<meta>タグにエディタなどを使ってコピーする方法、(b)P3Pのwell-knownロケーションといわれるトップのドメイン名からの/w3c/ディレクトリ、例えば<http://www.nmda.or.jp/w3c/p3p.xml/>といった場所にファイルを作成する方法、あるいは(c)httpヘッダに入れるためにCGIをプログラミングするか、Webサーバによってはある特定のディレクトリにファイルを作成する方法などがあります。一番良く使われる方法は、(b)ですが、固定個所にファイルが置けない場合は(a)となり、あるいはパナー広告会社などは(c)の方法を使うのが一般的と思われる。

5.4 プライバシー情報管理システム

当協会はインターネット協会と協力して、P3Pワーキングドラフト1999年10月版、および2000年5月版に準拠したプライバシー情報管理ツールを実験用に提供しております。プライバシー情報管理ツールは、主にP3のWebサイトを構築するP3Pオーサリングツールと、Webブラウザ側で使うP3Pユーザエージェントから成ります。具体的にはP3Pオーサリングツールを使って作成したP3P対応のWebサイトに

アクセスすると、P3PユーザエージェントがP3Pを受けた場合は、予めユーザが登録しておいた個人情報の利用目的や開示範囲を超えるような利用要求が来た場合には、オンラインの取引を中止するというデモが可能となっています。また、本システムでは、これらの機能のほか、ユーザが登録するための個人情報設定ツール、P3Pユーザエージェントの稼動状態を見るモニター、WebブラウザからP3Pユーザエージェントを強制利用させるリダイレクタ、およびたやすく改竄できるHTML形式のプライバシーポリシーの改竄チェックを行うサイトビューロ機能なども提供しております。これらの機能を元に国内のオンラインショップなどを使った実証実験も行い、W3Cへの技術的なフィードバックも行いました。

5.5 今後の展望

2002年度には国内の法整備(個人情報保護法案)が期待されること、あるいは日本情報処理開発協会(JIPDEC)のプライバシーマークなどの普及などから、国民の個人情報保護に関する問題意識は年々高まっており、加えて、近いうちにP3PもW3C勧告となる予定であり、インターネットにおけるP3P技術の普及が益々期待されます。また、P3Pはクライアント側、サイト側が共に相乗効果で普及させていくべき特徴があります。

このような中、サイト作成者にとっては負担がかかるP3P対応作業を、より安価に容易に行えるツールを日本語版で、また最新のP3P仕様に対応して提供しなければなりません。クライアント側も前述のとおり、現状のブラウザではP3Pの有効な機能を使っていないこともあり、より一層の強化が望まれます。一方、人が読むプライバシーポリシーとマシンが読むP3Pポリシーの適合性に関する問題の解決として、国内では既に実用化されているJIPDECのプライバシーマーク制度との連携、相互保証の検討、などが重要になってくると考えています。この他にも、Webサイトに提供した自分の個人情報を事後に修正・削除したり、その使用を差し止めたりできるオプトアウト機能、国内特有の個人情報(住所や氏名のふりがな、生年月日における元号など)のデータスキーマ基準の開発と普及などの必要性が考えられます。

このように、P3Pの今後の立ち上がりとともに、インターネットをより安全にかつ、快適に利用できる環境を国内でも整えていくことが必要です。