

Web データの個別化と環境適応

清 光 英 成[†] 竹 内 淳 記^{††}

Web データを利用者や利用者グループに特化した表現に再構成する Web の個別化アプリケーションが開発されている。また、利用者が Web データを閲覧している環境に合わせてデータを再構成する研究も進められている。本論文では、Web データの個別化と閲覧環境への適応に関する研究について概観し、主な研究についての分類と比較を行う。特に Web データの個別化研究は、コンテンツを再構成して利用者個別のビューを提供するコンテンツ主体のアプローチと、ページアクセス権限やリンク巡行権限を制御するアクセス権主体のアプローチに分類して議論する。また、今後の Web データの個別化ならびに環境適応研究の展望についても述べる。

Web Personalization and Adaptation

HIDENARI KIYOMITSU[†] and ATSUNORI TAKEUCHI^{††}

The personalization and adaptation of Web data have become very important issues on the Internet Business. In this paper, we survey, categorize, compare and discuss about the proposals on Web personalization and adaptation. First, we introduce researches about personalization of Web data for each user, adaptation of it for user's browsing environment. Then these are categorized and compared in their features. Especially, the researches for Web data personalization can be categorized into contents base approach and access rights approach. We also discuss about some problems and future of Web personalization and adaptation.

1. はじめに

インターネット利用者が爆発的に増加している。その背景として、Web 情報技術の長足の進歩やネットワーク基盤の整備と、Web データを作成するためのソフトウェアの普及をあげることができる。

Mosaic が登場したころ、いわゆる初期の「インターネット」は利用者のページアクセスやリンク巡行を考えると、

- ページコンテンツは誰がアクセスしても不変、
- 設定されているリンクを巡行すれば無条件にリンク先ページへアクセス可能、
- 過去にどのようなページ閲覧やリンク巡行をしたかにかかわらずページ内容とリンク巡行権限は不変、

のような静的な性質があった。その後、CGI、SSI、JAVA Servlet ならびに MS-ASP (Microsoft Active Server Pages) などのサーバアプリケーション技術により、インタラクティブなページを構成できるように

なり、データベースから動的にページを構成するアプリケーションも可能となった。また、JAVA Applet に代表されるクライアントで実行されるアプリケーション技術や Shockwave、Acrobat などのブラウザプラグインの普及により、エンターテイメントとしての利用が可能となったことが利用者の増加に拍車をかけていると考えられる。さらに、Web データを制作するツールも、初心者を対象とした製品が販売されており、高度な技術を要した動的コンテンツ生成も JavaScript などの自動生成機能を用いて特別な知識なしで作成できる環境が整いつつある。

近年、Amazon.com Inc. の 1-click に代表される、商法とアプリケーションを統合したビジネスモデル特許が認められている。また、e-Commerce、e-Business、e-Trade などのインターネット派生語が一般的に使われており、Web ビジネスが社会的に流通するようになっている。このような時流も手伝って、顧客-企業間の商取引や企業-企業間取引を Web 上で実現するためのビジネスアプリケーションを構築する企業が増える傾向にあり、Web とデータベースを連携した Web ビジネス統合開発ツールの普及も進んでいる。

[†] 神戸大学国際文化学部
Faculty of Cross-Cultural Studies, Kobe University

^{††} 神戸大学大学院自然科学研究科
Graduate School of Science and Technology, Kobe University

U. S. PAT. 5960411 (アメリカ合衆国), 特開平 11-161717 (日本)

ビジネス領域だけでなく、個人のホームページを運営している利用者からも Web ページを各閲覧者に個別のコンテンツ構成やリンク構成で提供したいという要求が高まっている。たとえば、サイトを頻りに訪れる閲覧者に対して、他の閲覧者と異なる情報を提供することで、閲覧者の差別化を図るアプリケーションを容易に構築できれば有用である。

このような現状をふまえて、Web データの個別化と環境適応についての研究動向を紹介し、比較と分類を行う。本論文の以降の構成は次のとおりである。2 章と 3 章で Web データの個別化と環境適応のための研究を紹介し、4 章で分類と比較を行う。5 章では問題点と今後の展開を議論する。6 章はまとめである。

2. Web データの個別化

Web データを個々の利用者や利用者グループに適した表現にすることを Web データの個別化と呼ぶ。ビジネス領域では、商取引の顧客-企業間取引や企業-企業間取引は 1 対 1 の個別プロセスで成立しているため、One-to-One の Web アプリケーションを構築して、Web データを個別化する必要がある。また、ページの見せ方を変化させるだけでなく、コンテンツ自体の置き換えや、アクセス権限を個々の利用者に対して個別に制御できる仕組みも重要である。本章では、Web データの個別化に関する主な研究事例を紹介する。

2.1 Web コンテンツの個別化

Kamba らは、利用者が対話的に Web 上の新聞をカスタマイズできる *Krakatoa Chronicle*²¹⁾を開発している。*Krakatoa Chronicle*における紙面上の記事は、複数のニュースサイトに掲載されている記事のトピックやテキストなどの特性と各利用者のプロフィールを解析することで導出され、体裁制御機構を用いて配置される。利用者プロフィールはその利用者が好むトピックや配置を指定することで初期のプロファイルが生成される。利用者は閲覧時にも記事に対する適合度(スコア)を評価でき、プロフィールの更新が可能である。利用者プロフィールの更新には、利用者の要求と表示された記事のスコアを利用者自身が指定することによる明示的適合フィードバックと閲覧時の振舞いをサーバ側で監視することによる暗黙的適合フィードバックを用いている。また、特性の類似したプロフィールを持つ利用者集合をコミュニティととらえ、そのプロフィール集合に対して集約演算を行うことで記事に対するコミュニティのスコアを動的に再計算して即座に反映する機能も有する。

品川らは、利用者プロフィールに基づき、ページ内

に表示するコンテンツを選択する機構を提案している^{46),47)}。プロフィールデータは、利用者の興味を記述したキーワードと、表示詳細度に対する要求を記述した閾値で構成されている。表示するコンテンツの選択は、HTML タグの解析により導出される論理木を用いて決定される。論理木は見出し、段落、表、リストのような構造主体のタグで囲まれた領域を端点とし、構造の上下関係を枝としている。太字、イタリックなどの表示主体のタグはノードの重みとして扱われている。

Baudisch らと Langheirich らは、バナー広告の諸問題を解決する研究を行っている。文献 4) で Baudisch らは、適合フィードバック手法を用いてバナー広告の取捨を議論している。利用者プロフィールは、利用者がバナー広告に求めるキーワードと重みの組を格納している。バナー広告には、バナー画像やリンク先ページの特徴を表現するキーワードが付与されており、利用者がバナー広告をクリックすることでプロフィール中の対応するキーワードの重みが更新される仕組みとなっている。また、利用者がバナーをクリックして閲覧したページの履歴も適合フィードバック資源として利用している。Baudisch らの研究は、利用者にバナー広告の好き嫌いを直接尋ねるため、明示的適合フィードバック手法を用いて Web ページ中の部分的なコンテンツを個別化するアプローチといえる。

一方、Langheirich らは、利用者のバナー広告に対する明示的な要求をプロフィールデータとして利用せず、バナー広告をクリックされたかどうかによって、どのようなバナー広告を Web ページに表示するかを決定する³⁴⁾。また、利用者が入力した検索キーワードや閲覧しているページ URL から最も適合度の高いバナー広告を Web ページに挿入する機構を提案している。Langheirich らのアプローチは、プロフィールデータを記憶しないため、利用者のプライバシーに非侵入的な特徴を持つ。

Ceri らは、電子商取引におけるビジネスルールを用いて Web データを個別化する機構を提案している^{8),9)}。各利用者のプロフィールデータは ID、電子メールアドレス、アクセス履歴、購買履歴などから構成されている。利用者プロフィール中の個人情報には利用者に直接尋ねて収集し、ページの閲覧やページ間の移動、商品の購買履歴はサーバで暗黙的に収集している。Ceri らのビジネスルールは、アクティブデータベースシステムの ECA ルールを模した XML 記法を用いて表現されており、ページに含めるコンテンツの再構成や利用者プロフィールの更新に利用される。

ビデオハイパーメディア研究でも、放送コンテンツ個別化機構の研究開発が行われている。橋本らは、デジタル放送コンテンツを利用者の受信端末で個別化して視聴できる仕組みを提案している^{13),14)}。橋本らの手法では、視聴者の嗜好に合わせたダイジェスト映像を作成するモデル PDMM を導入した番組記述言語 PPML を設計している。PPMLは映像、音声、仮想キャラクタ、字幕スーパーなどの番組素材メディアの同期制御に利用され、表示位置やサイズなどの体裁制御機能も有する。個別化映像の要素となる映像シーンには索引が付与されており、利用者の嗜好情報(たとえば、好きな野球選手など)をプロフィールとして個別化を行っている。橋本らの研究は、Web情報システムにおいて放送型メディアを扱うアプリケーションに利用できる。

2.2 Webサイトの個別化

ここまでで紹介した事例は、Webページのコンテンツを個別化する研究であったが、ページの見せ方を変化させるのではなく、サイトの構造を利用者の好みに合わせて仮想的に構成する研究が行われている。

Atzeniらは、利用者プロフィールを用いて、利用者が興味を持っていると想定されるページを集めた仮想サイトを生成する ARANEUSシステムを提案している³⁾。ARANEUSではまず、利用者プロフィールを元にページを収集して、キーワードとページURL、その注釈情報を属性とするテーブルを生成する。そして、テーブルとして構造化されたデータからビューテーブルを生成し、さらに実体化させたうえで半構造データとしてHTML化している。仮想サイト中のページは、インデックスページからリンクされた収集元のサイトに存在し、ページコンテンツは編集されない。

Perkowitzらは、Webサーバのアクセスログから、利用者のアクセスパターンを抽出し、インデックスページを再構成する機構 *Adaptive Web Sites* を提案している。文献 39) では、利用者のアクセスパターンを抽出するアルゴリズム PageGather を実装して、インデックスページに表示するリンクアンカ候補とその配置順序を決定し、利用者個別の仮想サイトを実現している。

これらの研究は、ページコンテンツを個別化することなく、インデックスページとして表示するサイトの構成情報やリンク構造を個別化するアプローチといえる。

2.3 アクセス権限の個別化

Webデータのアクセス方法をWebデータの作者が主張できるWeb情報システムの研究も行われている。

Kiyomitsuらは、利用者のページ閲覧履歴やリンク巡回履歴に基づいてページへのアクセス権限やリンクの巡回権限を管理している。文献 27), 28) では、ページ作者が定義した巡回済みであるべきリンク集合に含まれるリンクをすべて巡回済みであれば、指定されたリンクを巡回するための権限が発行されるような、リンクに対するアクセス管理機構を提案している。また、ページ作者の意図に反するリンク巡回に対して、ページへのアクセスやリンク巡回を拒否するのではなく、閲覧すべきページや巡回すべきリンクを補充情報として提供する機構も考察している。さらに、文献 25), 29), 50) では、ページに含めるコンテンツを利用者のアクセスしている場所や時刻とアクセス履歴に基づいて動的に変化させる機構と、ページへのアクセス権限やリンク巡回権限を宣言的な言語を用いて実現する方法を議論している。文献 23), 24), 26), 51) では、ページ作者の意図する閲覧方法やリンク巡回を ECA ルールを参考にして XML 表現し、Webコンテンツとアクセス権限を個別化する *Active Web* システムを提案している。

Réty⁴¹⁾はリンクをページ対ととらえ、閲覧者のページ閲覧履歴に基づいたページ間の移動制御機構を提案している。Rétyの移動制御機構では、ページ作成者が PROLOG ベースの論理言語を用いて条件付きリンクを実現することができる。

これらの研究は、Webデータの作者がページへのアクセスやリンク巡回のためのデータをメタデータとして付与してアクセス権限の個別化を行うアプローチであるといえる。

3. Webデータの環境適応

Webデータはデスクトップ計算機から携帯情報端末まで、多様な機器からアクセスされ利用されている。デスクトップ計算機と携帯電話のディスプレイでは、表示可能な大きさと解像度は明らかに異なる。また、LAN、電話回線やワイヤレス環境では、通信速度も大きく異なる。このような、端末の表示能力や通信速度などの物理的要因を考慮してWebデータの表現を変化させる要求は自然である。さらに、利用者がWebデータの閲覧形態を指定したり、ページ作者が利用者に見せたい閲覧方式を指定できたりする機構も興味深い研究である。本章では、Webデータを利用環境に適応させる研究を物理的な要因に対するものと、利用

者所望の閲覧形態や利用者の置かれている状況に対するものに分けて紹介する。

3.1 物理的要因に基づく環境適応

Horiらは、Webデータを閲覧する環境に合わせて表示体裁やコンテンツの物理的な量を変化させる機構を、ページ作者による注釈記述によって実現している^{18),53)}。元のデータはXML文書を想定しており、注釈はRDFを用いて外部メタデータとして記述され、XPath/XPointerによってXML文書と対応付けられている。Horiらの手法では、端末として利用される機器の表示能力や通信速度に合わせて、コンテンツを置換したり、画像の解像度や大きさを変化させたりすることができる。特にPDAなどの携帯情報端末に焦点をあて、表示環境や通信環境にWebコンテンツを適応させる研究といえる。

北山らは、端末の能力に応じてWebデータを生成するビジネスアプリケーション構築機構Dharmaを開発している^{16),22)}。Webデータがページとしてあらかじめ用意されているのではなく、表示装置や通信速度に基づいてページに埋め込むデータを構成する動的コンテンツ生成機能を中心に議論している。

マルチメディア情報配信の分野でも、コンテンツの環境適応研究が行われている。Rousseauらはマルチメディア情報を表示するための同期を考慮した閲覧環境の適応技術について述べている⁴⁴⁾。マルチメディア情報を再生するためのメタデータは作者が記述するが、利用者が好みや環境に合わせて再設定することも可能である。さらに、SMILを拡張した同期スキーム記法を提案している。この研究は、コンテンツ表示方法の個別化と環境適応を組み合わせたアプローチといえる。

3.2 閲覧形態と状況に基づく環境適応

リンクをクリックしてページを閲覧する方法がWebデータの表示方式では一般的であるが、文字列コンテンツを音声化してWeb情報の受動的な視聴を可能とする研究が進められている。灘本らは、一般的なclick and readのWebデータの閲覧方式ではなく、利用者があたかもテレビ番組を視聴しているかのようにWebデータを閲覧できる方式を提案している^{15),37)}。灘本らの手法では、利用者が読み上げを行うキャラクタの指定や、ニュース、バラエティなどの番組メタファを指定でき、嗜好に合わせた閲覧形態でWebデータを「視聴」できる。また、Webデータを番組メタファで演出するための番組演出言語Scripting-XML(S-XML)

の仕様記述も行っている。

Kiyomitsuらは、利用者がアクセスしている場所や時刻によってWebデータを変化させる研究を行っている^{23),24)}。KiyomitsuらのActiveWebでは、利用者が実際に存在している位置と時刻、ならびに仮想的な位置と時刻を利用してWebデータの環境適応を議論している。特に仮想位置と仮想時刻は、利用者が移動している状況を想定しており、経由予定地や到着予定時刻を利用したアプリケーションを考慮している。さらに、複数の利用者のアクセス場所や時刻を履歴データとして蓄積することで、Webデータの自動更新も可能である。

4. 各研究の比較と分類

本章では、ここまでで紹介したWebデータの個別化と環境適応に関する主な研究を整理して比較・分類する。

4.1 個別化・環境適応の対象

Webデータの個別化や環境適応を実現するために利用されるデータは様々である。本論文で紹介した主な研究の特徴として、個別化・環境適応の対象、利用されるデータならびにその収集と更新を表1に整理した。

Webデータの個別化は、個別の利用者を対象としたものと利用者グループやコミュニティを対象としたものがある。それぞれを分類すると

- 個別利用者
文献 3), 4), 8), 9), 13), 21), 23)~29), 37), 39)~41), 50), 51)
- 利用者グループ・コミュニティ
文献 21), 23)~26), 29), 50), 51)

のようになる。Webデータの個別化研究は、各利用者に個別のプロファイルデータを利用するものが多い。利用者グループ・コミュニティを対象としたWebデータの個別化は、利用者に個別のプロファイルデータに加えて、プロファイルデータの集合から導出される集約情報を利用している。

Webデータの環境適応研究は、閲覧環境の表示能力、データ処理能力や通信回線といった物理的要因への適応と、利用者あるいは、ページ作者所望の閲覧方法や利用者の状況への適応とがある。それぞれ分類すると、

- 物理的要因
文献 16), 18), 22), 30), 31), 44), 53)
- 閲覧方法・利用者の状況
文献 15), 23)~26), 29), 37), 50), 51)

i-Mode 端末を対象とした画面表示最適化機能が銀行業務アプリケーションとして実用化されている。

表 1 各研究の特徴
Table 1 Features of the introduced researches.

	個別化・環境適応の対象	収集と更新	個別化・環境適応に利用されるデータ
ActiveWeb 文献 23)~29), 50), 51)	アクセス時の状況, 利用者個別, 利用者グループ	明示的 (仮想位置・時刻) 暗黙的	ページ閲覧履歴, リンク巡行履歴, 時刻, 場所, リンク参照元
Krakatoa Chronicle 文献 21)	利用者個別, 利用者グループ	明示的	キーワード, 操作履歴, 適合フィードバック結果
Conditional Linkage 文献 41)	利用者個別	暗黙的	リンク巡行権限の状態
User-Configurable Advertising Profiles 文献 4)	利用者個別	明示的	操作履歴 (バナーのクリック), バナーに付与されているキーワード
One-to-One Personalization 文献 8), 9)	利用者個別	明示的, 暗黙的	個人情報 閲覧履歴, 購買履歴
Annotation-Based Transcoding 文献 18)	表示端末	暗黙的	通信速度, 画面サイズ (接続時のみ有効)
User Adaptable Multimedia Presentation 文献 44)	利用者個別 端末の能力	明示的 (再設定時のみ) 暗黙的	コンテンツ間の同期 (再設定時のみ) 通信速度, 表示能力
Adaptive Web Sites 文献 39)	利用者個別	暗黙的	ページアクセス履歴 (アクセスログから抽出)
ARANEUS 文献 3)	利用者個別	明示的	キーワード
Web 情報の番組化 文献 15), 37)	利用者個別	明示的 (番組作成者が介在)	閲覧方法
パーソナルダイジェスト 文献 13), 14)	利用者個別	明示的 (番組作成者が介在)	嗜好情報

のようになる。

物理的要因に基づく環境適応は、画面サイズや解像度を考慮した表示装置を対象とする研究と通信速度やアクセス形態を対象とする研究がある。紹介した物理的要因への適応に関する研究のほとんどが両方を意識しているので、あえて分類しないが、対象とする環境要素で研究領域を分けることもできる。

閲覧方法に基づく環境適応は、Web データをある種のメタファ上に展開して閲覧することを可能とする。文献 37) では、老人や子供のような、Web ブラウザの操作に慣れていない利用者を意識して、テレビ番組メタファを採用している。文献 15) では、リンク構造を解析することにより、利用者の興味に基づいて導出された経路上のページ群をテレビ番組風に閲覧させるような、個別化と環境適応を複合させたアプローチを採用している。文献 23), 24) は、利用者のアクセス場所や時刻を利用者の状況ととらえて、その状況で最も適切と思われるコンテンツの選択やアクセス権限の発行を行うことができる。

4.2 個別化・環境適応の主導権

Web データの個別化・環境適応研究を、主導権が作者側にあるものと利用者側にあるものに分類することができる。前者はページ作者があらかじめ個別化・環境適応のルールを決めておいて、コンテンツやリンク構造を変化させる方法である。文献 23)~29), 50), 51) の *ActiveWeb* は、ページ作者の個別化・環境適応

に対する意図を反映させるためのルールに従って Web ビューが再構成されるため、個別化・環境適応に利用されるデータは暗黙的に更新される。

後者は、個別化・環境適応の方針を利用者に直接尋ねる方法である。文献 21), 30), 31) では、利用者がページコンテンツの適合度を評価し、適合フィードバック手法を用いてページコンテンツの選定や体裁を変化させる方法を提案している。文献 8), 9) の研究は利用者の個人情報は明示的に入力され、購買履歴や閲覧情報が暗黙的に蓄積される仕組みとなっている。文献 15), 37) では、利用者が Web データの閲覧方法を指定できる。

このように、Web コンテンツのアクセス権や、推奨するページへの誘導を個別化するアプローチは、一般的に主導権が作者側にある。一方、利用者がページの表示をカスタマイズするようなアプリケーションでは、嗜好や要求を利用者に直接尋ねる方式が多い。また、ビジネスアプリケーションの構築を目的とした研究では、個人情報と嗜好や要求を尋ねたうえで、ビジネスルールや適合フィードバック、フィルタリングを適用しながら利用者プロファイルを更新する方法も採用されている。

Web データの個別化研究や環境適応研究を、

- Web データの個別化・環境適応、
- 個別化・環境適応の主導権 (ページ作者・利用者)、
- 個別化・環境適応させるデータ (コンテンツ・ア

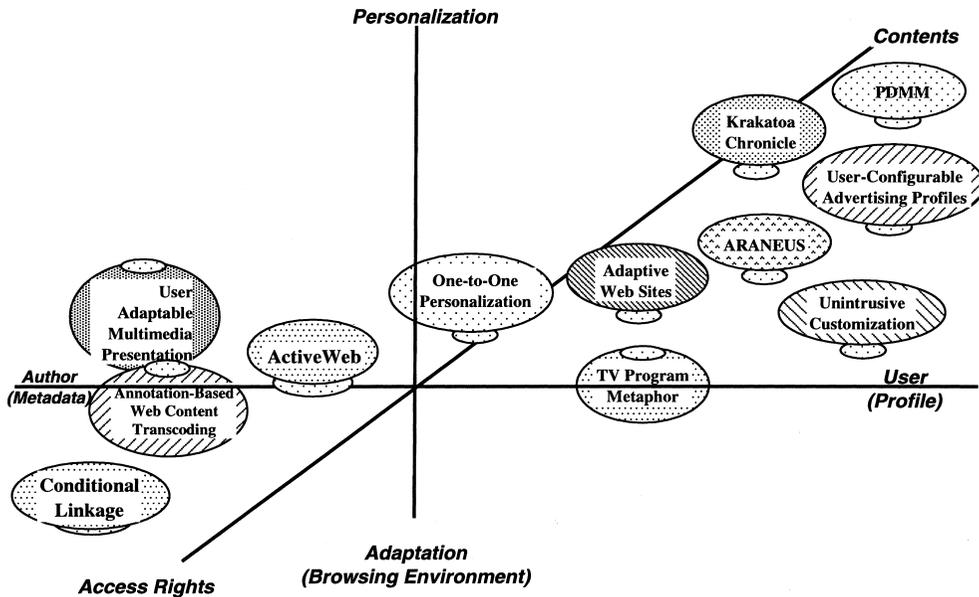


図1 プロファイル・メタデータと個別化・環境適応の分類

Fig. 1 The matrix of research areas.

クセス権限)

という3つの軸を用いて各研究を分類すると図1のようになる。それぞれの方向は互いに直交する概念であるため、敵対するものではなく、組み合わせる利用することが可能である。ActiveWeb^{23)~29),50),51)}は、作者が利用者のページ閲覧履歴やリンク巡回履歴に基づいて、ページへのアクセス権限やリンク巡回の権限を発行するためのルールをメタデータとして利用し、Webコンテンツの個別化機能も有するため、個別化するデータはアクセス権限とWebコンテンツである。One-to-One Personalization^{8),9)}は、利用者の要求を直接尋ねたうえで、利用者プロフィールに対してメタデータとしてのビジネスルールを適用するため、主導権が利用者として作者の双方にある個別化研究といえる。

本論文で紹介した各研究が引用している文献を図2に整理した。楕円で示した研究は、表1でプロフィールデータを整理した研究であり、長方形で示した研究は、表1には表れていないが本論文で述べている研究である。破線で囲まれた領域は、各研究の引用傾向と目的により

- 上部 … 環境適応に関する研究^{6),12),16),18),22),44)}
- 左部 … リンク巡回制御を扱った研究^{5),41)}
- 右部 … バナー広告の諸問題を解決する研究^{4),34)}
- 下部右 … 適合フィードバックやソーシャルフィルタリング等を用いたページコンテンツ・サイトビューの個別化に関する研究^{17),20),21),33),39),40),45),48)}

- 下部左 … ビジネスモデルを意識したWeb情報システム研究^{8),9),11),32),38)}

に分類している。

特に破線で囲んでいないが、図2の中央部から左部にかけての研究は、ODMG-93モデル⁷⁾、問合せ処理や大規模トランザクションなどのデータベース技術を参考にしており、図2の右側は、適合フィードバックやソーシャルフィルタリングなどの人工知能技術を参考にしており、各研究は、個別化・環境適応の実現手法により引用傾向が異なるが、図1に示した6つの方向性を組み合わせて利用することが可能であるので、各研究の発展により相互引用されるものと思われる。

4.3 個別化・環境適応のプロセス

Webデータの個別化と環境適応の実現プロセスは、一般的に以下の6過程が考えられる。

- 1) ターゲット利用者の決定
- 2) プロフィールデータの収集、履歴データの照会
- 3) ゴールの導出
- 4) 中間プロセスの生成
- 5) 個別化・環境適応データの生成・配信
- 6) 1)~5)までのデータを蓄積

1)は、利用者の認証と、どのような利用者であるかの特定とに分類することができる。顧客-企業間取引やオークションなどの、ある特定の利用者に固有のビューを提供するWebアプリケーションでは利用者の認証が必要である。文献^{8),9)}では、システムが

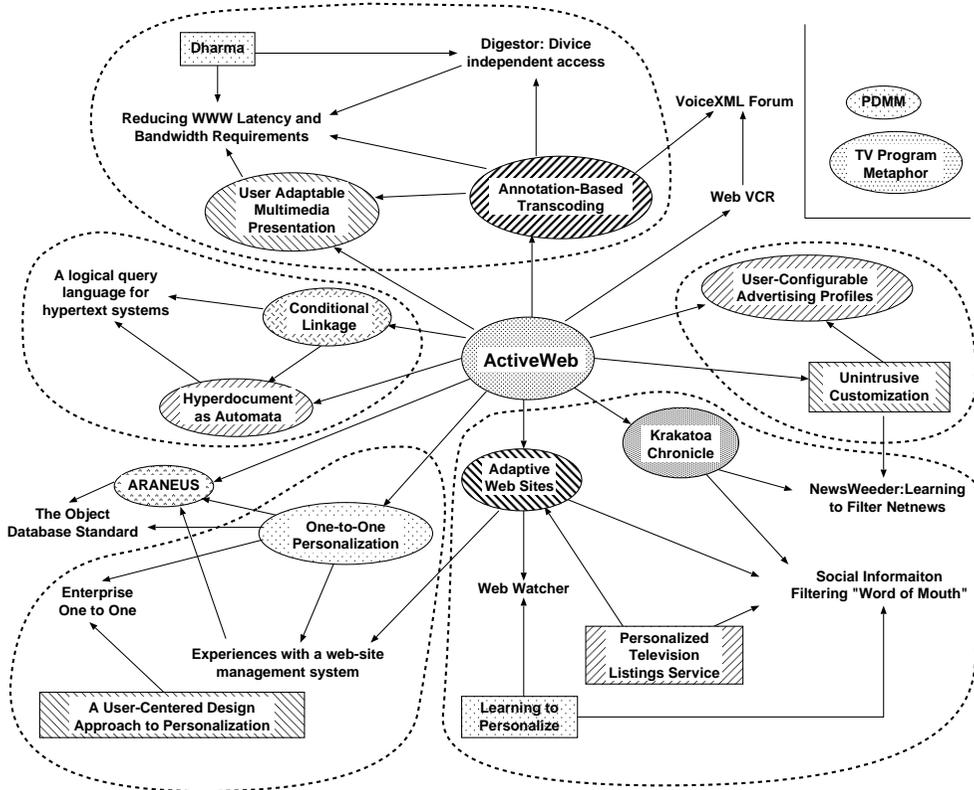


図2 各研究の引用文献
Fig. 2 Reference map of introduced researches.

利用者に与える ID や利用者の e-mail アドレスをプロフィールデータとして保持し、認証に利用している。一方、利用者の閲覧方法や閲覧環境に基づいてページビューを変化させるアプリケーションでは、利用者個人の認証ではなく、アクセスに利用している機器や通信環境、ある特定の操作を行ったかどうかの検証が行われる。

2) は、個別化・環境適応を行うためのデータを収集する工程といえる。利用者プロフィールの収集には、利用者に個人情報や嗜好データを直接尋ねる明示的な方法と、ページ滞在時間、スクロール操作の有無など、利用者の行動を監視することで取得する暗黙的な方法がある。

3) は、2) で収集されたプロフィールデータや履歴データを処理して、Web データにどのような操作を行うかを決定する工程といえる。文献 21), 34), 39) では、ソーシャルフィルタリングや適合フィードバック技術を用いて、個別化ビューに含めるコンテンツと体裁を決定している。

4) は、3) で決定したビューを、どのように生成するかをプログラムする工程であるといえる。実際に

は、CGI, SSI, MS-ASP, JavaServlet などにおけるモジュール選択あるいはプログラムの動的生成と、スケジューリングを行う。

5) は、4) で選択・生成されたプロセスを適用して Web ビューの生成を行い、利用者に配信する工程である。

文献 8), 9) は、1)~3), 6) を XML ベースで行い、プロフィールデータの更新にビジネスルールを適用するアプローチである。文献 23)~29), 50), 51) では、1)~6) の方法をプログラムに埋め込むのではなくページ作者がルールを定義することによって明示的に指定できる。文献 19) は、1)~6) をビジネスモデル的に扱い、XML 環境での実現を目的としている。

5. 問題点と今後の展望

大規模な Web 情報システムを構築するためには時間的・金銭的コストの問題がある。そのため、個々の Web アプリケーションの制作と複数のシステムコンポーネントを統合、再利用する総合的な開発機構が必要である。今日では ColdFusion のような Web 情報システム開発ツールが市販され、データベースシステム

ムとの連携も比較的容易となっている。

一方、利用者はシステムとの対話が俊敏であることを期待している。そのため、応答時間が短くなければ利用者あるいは顧客を逃してしまうことになるだろう。データベースシステムなどの大規模システムはその資源の豊富さが価値として評価されるため、ビジネス用途などの接続に限っては利用者が待ってくれる。しかし、一般的な Web 情報システム利用者の大半は娯楽や購買などの個人利用を目的としている。すなわち、特別な希少価値や特典がコンテンツに含まれていなければ、「待たされるサイト」は利用者から敬遠されるようになることが予想される。

一般に、システムのパフォーマンスは時間とともにコスト対効果が向上するが、現状ではハードウェアの進歩よりも利用者の増加ならびにアクセス数の増加が勝っている。応答時間の問題に対処するには、ページコンテンツの軽量化やベースシステムである Web サーバソフトウェアの高速化、サーバ側アプリケーションモジュールの最適化が必要である。ハードウェアによる高速化とともにソフトウェアの最適化は、大衆を対象とした大規模 Web アプリケーションでは重要な課題といえる。

6. ま と め

Web データの個別化と環境適応に関する研究を紹介し、整理・分類した。Web データの個別化研究には様々なアプローチがあるが、コンテンツやその表示方法を個別化して利用者に提供するものと、ページへのアクセス権限やリンクの巡行権限を個別化するものに分類した。さらに、個別化に利用されるプロファイルデータを整理し、個別化の主導権に関する傾向をまとめている。Web データの個別化と環境適応に関する研究には、6つの方向性があるが、互いに直交する概念であるため、様々な組合せによる応用が可能である。

Web データの環境適応は、PDA や携帯型 PC だけでなく、i-Mode などを含む携帯情報端末での応用が期待される。将来 IMT-2000 などの高速接続が可能となるが、携帯情報端末の表示装置の物理的制約は解消されないため、有用な成果が期待される研究分野といえる。

また、今後の普及が予想されるデジタル放送の個別化研究も紹介した。特に、橋本らの研究は、Web 情報システムにおいて放送型メディアコンテンツを利用した個別化と利用者所望の閲覧環境に適応させる拡張が期待できる。

Web データの個別化・環境適応の問題点を示したが、

アプローチをかえて、利用者が待ってくれるシステムを構築することもこれらの問題を解決する鍵となる。はなはだ抽象的ではあるが、痒いところに手が届く便利なサイトを構築すればよいのである。そのため、魅力的なコンテンツを提供し続けることや頻繁なページコンテンツの更新が必要である。また、利用者の絞り込みを行い、特殊なコミュニティや接続方法に特化したサービスの提供も有用である。これらは、紹介した主な研究を応用することで実現可能である。Web 情報システムの普及と発展にとって、ある利用者に特化した Web データや利用者所望の閲覧形態をサポートできる Web データの個別化技術や環境適応技術がますます重要になることが予想される。

動的コンテンツをサーバで生成して利用者に配信する Web 情報システムでは、プロキシキャッシュとブラウザキャッシュによる版の不整合という問題がある。Web システムでは重要な更新が行われたときに更新データのプッシュや能動的更新も可能である。しかしながら、利用者へ個別な Web データを確実に送信できる機構の開発は重要な課題である。

最後に、本論文では直接言及しなかったが、*Communications of the ACM* の 2000 年 8 月号に Personalization 特集がある。特に文献 10), 17), 32), 35), 36), 42), 43), 48), 52) は本論文で紹介した事例も含まれるが、参照されることをお勧めする。

謝辞 本論文の一部は、日本学術振興会未来開拓学術推進事業における研究プロジェクト「マルチメディア・コンテンツの高次処理の研究」(プロジェクト番号 JSPS-RFTF97P00501)、ならびに文部省科研費「分散型ハイパーメディアからの構造発見とアクセス管理」(課題番号 12680416)による。ここに記して謝意を表します。

参 考 文 献

- 1) VoiceXML Forum Version 0.9 Specification, VoiceXML Forum. <http://www.vxml.org/specs/VoiceXML-0.9-19990817.pdf> (Aug.1999).
- 2) Anupam, V., Freire, J., Kumar, B. and Lieuwen, D.: Automating Web Navigation with the WebVCR, *Proc. WWW9* (1999).
- 3) Atzeni, P., Mocca, G. and Merialdo, P.: To Weave the Web, *Proc. VLDB'97*, pp.206-215 (1997).
- 4) Baudisch, P. and Dirk, L.: User-configurable advertising profiles applied to Web page banners, *Proc. 1st Berlin Economics Workshop* (1997).
- 5) Beerli, C. and Kornatzky, Y.: A logical query

- language for hypertext systems, *Proc. European Conference on Hypertext*, pp.67–80 (1990).
- 6) Bickmore, T.W. and Schilit, B.N.: Digestor: Device-independent access to the World Wide Web, *Proc. WWW6*, pp.1075–1082 (1997).
 - 7) Cattell, R.G.G.: *The Object Database Standard ODMG-93*, Morgan Kaufmann Pub. (1994).
 - 8) Ceri, S., Fraternali, P., Maurino, A. and Paraboschi, S.: One-to-One Personalization of Data-Intensive Web Sites, *Proc. WebDB'99*, pp.1–6 (1999).
 - 9) Ceri, S., Fraternali, P. and Paraboschi, S.: Data-Driven, One-To-One Web Site Generation for Data-Intensive Applications, *Proc. VLDB'99*, pp.615–626 (1999).
 - 10) Cingil, I., Dogac, A. and Azgin, A.: A Broader Approach to Personalization, *Comm. ACM*, Vol.43, No.8, pp.136–141 (2000).
 - 11) Fernandez, M.F., Florescu, D., Levy, A.Y. and Suciu, D.: Catching the boat with Strudel: Experiences with a web-site management system, *Proc. SIGMOD '98*, pp.414–425 (1998).
 - 12) Fox, A. and Brewer, E.A.: Reducing WWW latency and bandwidth requirements by real-time distillation, *Proc. WWW5*, pp.1445–1456 (1996).
 - 13) 橋本隆子, 白田由香利, 飯沢篤志: パーソナルなダイジェスト作成モデル(PDMM)の検討, 情報処理学会研究報告, Vol.2000, No.69, pp.519–526 (2000).
 - 14) 橋本隆子, 白田由香利, 真野博子, 飯沢篤志: TV受信端末におけるダイジェスト視聴システム, 情報処理学会論文誌: データベース, Vol.41, No.SIG3(TOD 6), pp.71–84 (2000).
 - 15) 服部多栄子, 沢中郁夫, 灘本明代, 田中克己: リンク構造とディレクトリ構造を用いた複数文書の受動的視聴, DBWeb 2000, 情報処理学会シンポジウムシリーズ, Vol.2000, No.14, pp.129–136 (2000).
 - 16) 広瀬紳一, 北山文彦, 久世和資: 多種端末向けWebアプリケーション構築システム Dharma: ビューオブジェクト生成とHTML生成機構, 情報処理学会第57回全国大会文集分冊3, pp.392–393 (1998).
 - 17) Hirsh, H., Basu, C. and Davison, B.D.: Learning to Personalize, *Comm. ACM*, Vol.43, No.8, pp.102–106 (2000).
 - 18) Hori, M., Kondoh, G., Ono, K., Hirose, S. and Singhal, S.: Annotation-Based Web Content Transcoding, *Proc. WWW9*, pp.197–211 (1999).
 - 19) 石川 博, 太田 学: XMLに基づくEビジネスモデルの記述言語について, 情報処理学会研究報告, Vol.2000, No.69, pp.439–446 (2000).
 - 20) Joachims, T., Freitag, D. and Mitchell, T.: Webwatcher: A tour guide for the world wide web, *Proc. 15th International Joint Conference on Artificial Intelligence*, pp.770–775 (1997).
 - 21) Kamba, T., Bharat, K. and Albers, M.C.: The Krakatoa Chronicle: An Interactive Personalized Newspaper on the Web, *WWW Journal*, Vol.1, No.1, pp.159–170 (1995).
 - 22) 北山文彦, 広瀬紳一, 久世和資: 多種端末向けWebアプリケーション構築システム Dharma: システム概要とアプリケーションオブジェクト, 情報処理学会第57回全国大会文集分冊3, pp.390–391 (1998).
 - 23) Kiyomitsu, H., Takeuchi, A. and Tanaka, K.: ActiveWeb: XML-Based Active Rules for Web View Derivations and Access Control, *Proc. ITVE 2001*, Australian Computer Science Communication, Vol.23, No.6, pp.31–39, IEEE Press (2001).
 - 24) Kiyomitsu, H., Takeuchi, A. and Tanaka, K.: Web Reconfiguration by Spatio-Temporal Page Personalization Rules Based on Access Histories, *Proc. SAINT 2001*, IEEE-CS/IPSJ, pp.75–82, IEEE Press (2001).
 - 25) 清光英成, 竹内淳記, 田中克己: アクセス履歴と利用者の位置に基づくコンテンツの再構成, 情報処理学会研究報告, Vol.2000, No.44, pp.17–24 (2000).
 - 26) 清光英成, 竹内淳記, 田中克己: アクティブルールを用いたWebページの動的再構成, 情報処理学会研究報告, Vol.2000, No.96, pp.383–390 (2000).
 - 27) Kiyomitsu, H. and Tanaka, K.: Navigation-Dependent Web-Views: Changing Web Pages and Link Structures By User Navigation History, *Proc. ICSC '99*, Vol.LNCS-1749, pp.421–426 (1999).
 - 28) 清光英成, 田中克己: Webリンクの巡行に基づく動的なリンク活性化とアクセス管理, 情報処理学会論文誌: データベース, Vol.42, No.SIG 8(TOD 10) (2001).
 - 29) 清光英成, 竹内淳記, 田中克己: 閲覧履歴に依存したWebコンテンツの動的生成方式, DEWS2000 論文集 CD-ROM (2000).
 - 30) 小池雄一, 小向孝典, 神場和成, 古関義幸: パーソナライズ情報配信基盤PIDEMの提案, 情報処理学会第57回全国大会論文集分冊3, pp.386–387 (1996).
 - 31) 小向孝典, 小池雄一, 神場和成, 古関義幸: パーソナライズ情報配信基盤PIDEMにおける配信方法多様化の実現, 情報処理学会第57回全国大会論文集分冊3, pp.384–385 (1996).
 - 32) Kramer, J., Noronha, S. and Vergo, J.: A User-Centered Design Approach to Personal-

- ization, *Comm. ACM*, Vol.43, No.8, pp.45-48 (2000).
- 33) Lang, K.: NewsWeeder: Learning to Filter Netnews, *Proc. ICML '95*, pp.331-339 (1995).
- 34) Langheirich, M., Nakamura, A., Abe, N., Kamba, T. and Koseki, Y.: Unintrusive Customization Techniques for Web Advertising, *Proc. WWW8* (1998).
- 35) Manber, U., Patel, A. and Robison, J.: Experience with Personalization on Yahoo!, *Comm. ACM*, Vol.43, No.8, pp.35-39 (2000).
- 36) Mulvenna, M.D., Anand, S.S. and Büchner, A.G.: Personalization on the Net using Web Mining, *Comm. ACM*, Vol.43, No.8, pp.123-125 (2000).
- 37) 灘本明代, 服部多栄子, 近藤宏行, 沢中郁夫, 草原真知子, 田中克己: Web 情報の番組化のためのオーサリング機構, 情報処理学会研究報告, Vol.2000, No.10, pp.99-106 (2000).
- 38) Peppers, D. and Rogers, M.: *Enterprise One to One: Tools for Competing in the Interactive Age*, Currency-Doubleday (1997).
- 39) Perkowitz, M. and Etzioni, O.: Towards Adaptive Web Sites: Conceptual Framework and Case Study, *Proc. WWW8* (1998).
- 40) Perkowitz, M. and Etzioni, O.: Adaptive Web Sites, *Comm. ACM*, Vol.43, No.8, pp.152-158 (2000).
- 41) Réty, J.-H.: Structure Analysis for Hypertext with Conditional Linkage, *Proc. Hypertext '99*, pp.135-136 (1989).
- 42) Riecken, D.: Personalized Communication Networks, *Comm. ACM*, Vol.43, No.8, pp.41-42 (2000).
- 43) Riecken, D.: Personalized Views of Personalization, *Comm. ACM*, Vol.43, No.8, pp.26-28 (2000).
- 44) Rousseau, F., Garcia-Macias, J.A., de Lima, J.V. and Duda, A.: User Adaptable Multimedia Presentations for the WWW, *Proc. WWW8* (1998).
- 45) Shardanand, U. and Maes, P.: Social information filtering: Algorithms for automating word of mouth, *Conference on Human Factors in Computing Systems—CHI '95* (1995).
- 46) 品川徳秀, 北川博之: ユーザの視点に即した仮想 WWW ページの動的生成による閲覧支援, 情報処理学会研究報告, Vol.99, No.61, pp.425-430 (1999).
- 47) 品川徳秀, 北川博之, 河田 純: ユーザプロフィールに基づくビューページの動的生成による WWW 閲覧支援, 情報処理学会論文誌データベース, Vol.41, No.6, pp.22-36 (2000).
- 48) Smyth, B. and Cotter, P.: A Personalized Television Listings Service, *Comm. ACM*, Vol.43, No.8, pp.107-111 (2000).
- 49) Stotts, D. and Cabarrus, C.R.: Hyperdocuments Automata: Verification on Trace-Based Browsing Properties by Model Checking, *ACM Trans. Information Systems*, Vol.16, No.1, pp.1-30 (1992).
- 50) 竹内淳記, 清光英成, 田中克己: アクセス履歴の集約とメタルールに基づく Web コンテンツのアクセス管理機構, 情報処理学会研究報告, Vol.2000, No.96, pp.315-322 (2000).
- 51) 竹内淳記, 清光英成, 田中克己: アクティブルールに基づく Web 個別化・環境適応システム: ActiveWeb の実装, DEWS2001 論文集 CD-ROM (2001).
- 52) Wells, N. and Wolfers, J.: Finance with a Personalized Touch, *Comm. ACM*, Vol.43, No.8, pp.31-34 (2000).
- 53) IBM 東京基礎研究所: XML オーサリングツール構築技術: FREEDOM. <http://www.trl.ibm.co.jp/projects/freedom/>

(平成 12 年 12 月 20 日受付)

(平成 13 年 3 月 28 日採録)

(担当編集委員 遠山 元道)



清光 英成 (正会員)

1994 年図書館情報大学図書館情報学部卒業。1996 年同大学大学院図書館情報学研究科図書館情報学専攻修士課程修了。1998 年奈良先端科学技術大学院大学情報科学研究科博士後期課程中退。同年神戸大学経済学部助手。2001 年同大学国際文化学部講師, 現在に至る。データベースシステム研究, Web 情報システム研究, Web データの個別化・環境適応研究に従事。IEEE Computer Society, ACM 各会員。



竹内 淳記 (学生会員)

2000 年神戸大学工学部情報知能工学科卒業。現在同大学大学院自然科学研究科修士課程在学中。Web 情報システム研究, Web データの個別化・環境適応研究に従事。