

平成 11 年度

筑波大学第三学群情報学類

卒業研究論文

題目：Web ページの個人化に関する研究

主専攻 情報工学

著者名 青木裕伸

指導教員 電子・情報工学系 田中 二郎

## 要旨

Web は従来のメディアに比べて情報量が膨大であり、多岐にわたる。情報を整理して利用しやすくする personalization - 個人化 - はユーザをサポートする手段の一つとして研究実装が進められている。しかし、多くの Web パーソナライズの研究は企業や公開側の工夫であって、ユーザから積極的に自分だけの Web 利用を図るための動きは少ない。そこで、本論文ではユーザ側から見た Web のパーソナライジングについて考察し、新システム Webgrep の提案を行った。

WebGrep ではページ巡回の作業の負担を軽減し、編集、レイアウトの機能を取り入れることによって情報の統合的な利用を図ることが可能になる。

# 目次

<b>1</b>	<b>はじめに</b>	<b>3</b>
1.1	ネットワーク上のメディア	3
1.2	Web からの情報取得	4
1.3	本論文の構成	5
<b>2</b>	<b>web パーソナライズの現状</b>	<b>7</b>
2.1	パーソナライズ personalization	7
2.2	ポータルサイト	7
2.3	ユーザ情報解析とその利用	7
2.4	Web 巡回ソフト	10
<b>3</b>	<b>関連研究 事例 Internet Scrapbook および ANATAGONOMY</b>	<b>11</b>
3.1	Internet Scrapbook	11
3.1.1	概要	11
3.1.2	データ獲得	11
3.1.3	マッチング	12
3.2	ANATAGONOMY	12
3.2.1	概要	13
3.2.2	ユーザ情報の獲得	14
3.2.3	レイアウト	14
3.3	両者の比較と問題点	14
<b>4</b>	<b>新システム Webgrep の考察</b>	<b>16</b>
4.1	切取と自動レイアウトによる Web パーソナライジング	16
4.1.1	概要	16
4.1.2	Web ページの切取と編集	16
4.1.3	レイアウト手法	17
<b>5</b>	<b>結論と展望</b>	<b>21</b>

謝辭

22

參考文獻

22

# 第 1 章

## はじめに

### 1.1 ネットワーク上のメディア

World Wide Web はオープンなメディアとして大きく発達しつつある。ネットワーク上での情報伝達やコミュニケーションの手段としては、他に mail、chat などが広く用いられている。これらのメディアは実時間性や解放性などによって特徴づけることができる。

Web とそのほかのメディアと特徴の一つは、Web は誰でも自由に閲覧できるメディアであることである。

もう一つの特徴は、一つの URL を知ることができれば、更新された情報が同一の URL から得られることである。また、リンクを辿ることによって新たな関連サイトを知ることができる。

インターネットで従来用いられてきた情報の入手、分類、整理の手法のひとつとして Netnews がある。Netnews は分野、話題ごとのグループに細分化されており、ユーザは興味のある NewwsGroup だけを選んで購読することができる。しかし、NewsGroup に参加する人間は多く、話題も自分が本当に興味あるものだけとは限らない。参加する人間はかなりコンピュータについて習熟した（または興味を十分に持っている）ユーザが多く、初心者が入りづらい一面もある。基本的には質問や意見交換の場であって、現在の Web のような広範な話題は取り扱えない。

従来用いられてきた手法としてはメーリングリストや掲示板といったシステムもある。これはごく限られた人たちによるメール交換の場で、より親密さや手軽さがある。しかし、NetNews と同じように話題の限定という欠点を持つ。

これらカテゴリを限定した情報源を併用することによってユーザは自分の望む情報を得ることができた。選択によっては多くの情報を管理する必要があったが、情報はあらかじめカテゴライズされたもので、量も Web の膨大な情報量とは比較にならない。

## 1.2 Webからの情報取得

現在、Web ページはネットワーク環境の普及とユーザの拡大によって爆発的な増加を続けている。Lawrence らによれば総ページ数は8億を越えている [9]。さらに、goo.ne.jp の報告によれば1日100万ページのペースで増えている [12] といわれる。

このなかから有用な情報を得るには大きな困難がある。個人の情報収集と処理の能力には限界があり、情報の幅は量に伴って増大するからである。また、Web は個人や企業などによって独自に作られ、全体的な系統や整理区分を作ることが出来ない。

Web における情報取得の問題点は下のようになる。

- 欲しい情報があるのに見つからない
- 欲しい情報が存在しないのに探してしまう
- 欲しい情報がいらぬ情報にまぎれてしまう

情報取得の手段として多く用いられているのは情報検索である。情報検索の手段のひとつとしてサーチエンジンが挙げられる。サーチエンジンは大きくディレクトリ型とロボット型に分けられるが、どちらにも一長一短がある。ディレクトリ型はサーバ側で人間がページを分類し、登録する。これに対してロボット型は一旦すべてのページをクローラと呼ばれるプログラムでサーバの記憶媒体に蓄積し、全文検索をかけることによって目的のページを探す。web 全体に対してカバーする割合は16パーセントと低い [9]。ディレクトリ型のメリットは人間による細かな分類が可能なことであり、手作業で登録するディレクトリ型は情報の精度は高いがサービス側の負担が大きい。ロボット型では検索結果として提示される情報が膨大で冗長なものになりやすい。

また情報にフィルタをかけて整理することによって情報を取得する方法もある。協調フィルタリングや社会的フィルタリングと呼ばれる手法がある。

直接 Web をブラウジングし、リンクをたどるのも情報取得手段の一つである。ユーザはテキストやリンクから自分の欲しい情報に近づいていく。

各個人が膨大な情報から利用できるように「知識」を取り出す作業はすなわち、情報を個人化することである。この作業を情報の個人化、パーソナライゼーションと呼ぶ [10]。

従来用いられてきた手法の対象は、情報量が現在の Web ほどの膨大なものではなく、ある程度人手によって管理された情報源だった。しかし、そのままでは膨大すぎる情報を持つ Web は、従来そのままの手法では取り扱えない。サーチエンジンによって得た情報源も、年月を重ねるにつれて膨大化し、管理するのが重荷になっていく。情報をユーザに合わせて個

人化するための技術があらたに必要なになっている。そのための手法として、個人化をある程度自動化し、ユーザが意識しなくてもサーバ側で個人に合わせた情報を提示する方法と、またはユーザ自身が自発的な個人化を行い、必要な情報を得るのをサポートする方法の2つのアプローチが考えられる。

個人化をユーザから取得した個人情報からサーバ側で行おうとする動きとして、MyYahoo[5] や MyNetScape[4] といったサイトがある。これらは様々な情報をユーザの趣向に合わせて提示して Web ブラウジングの起点として使ってもらおうというポータルとよばれるサービスである。

しかし、これらの情報のソースはサービス側が提供するものに限られ、受動的な情報しか得られない。レイアウトの仕組みも原始的で稚拙なものであり、必ずしも情報を見易いとは言えない。また、ユーザのプライバシーに関するデリケートな問題をはらんでいる。

また、一方で伝統的な手法として BookMark を利用したサイト巡回による情報取得の方法もある。各ユーザが検索エンジンなどを用いて見つけたページを気に入ったらブックマークに登録し、ネットワークに接続したときにそれらのページを順に回って更新状況をチェックするという方法である。これは従来の情報源に対して行っていたのと近い処理であるが、サイトの内容が時々刻々と変化し、トピックが増減する Web が対象となるので、ユーザの負担は増大する。天気予報、価格情報、ニュースなどの情報は同じページで繰り返し更新されることが多いので、これらのページを周回すれば新しい情報を容易に得ることができる。この手法は、ユーザの知恵を使って情報の個人化を行う作業と言える。ユーザの手間を軽減することでユーザの自発的なパーソナライゼーションを補助することができる。

### 1.3 本論文の構成

2章では 個人化の定義を確認し、Web の個人化を試みた例を幾つか挙げ、利点と問題点について考察する。

3章では、本研究と特に関連の深い研究事例 ANATAGONOMY および Internet Scrapbook について詳しく述べる。本研究のシステムがどのような特性を持つべきか考察する。

4章で、本研究の新システムについて考察する。提案するシステムに必要な要素技術とインターフェースについて述べる。

5章をまとめとし、研究の展望も合わせて述べる。



図 1.1: ポータルサイト Mynetscape



## 第 2 章

### web パーソナライズの現状

#### 2.1 パーソナライズ personalization

パーソナライゼーションとは、「情報の個人化」とは、情報を選択し、整理して、利用可能にし、取得することである。個人化された情報は、扱いやすく、利用もできない膨大な情報に困惑することはなくなる。

#### 2.2 ポータルサイト

MyYahoo [5], MyNetScape などは基本的にユーザ指定型のサービスである [4]。MyYahoo では、ブックマーク、ニュース、天気予報、掲示板、オークションといったサービスを提供している。ユーザはサービスの中からいくつかを選択して、ポータルページに表示させることができる。ユーザの情報はサーバのデータベースと Cookie によって管理されている。Cookie はブラウザに簡単な情報を覚えさせるための技術で、ID やブラウザの状態などを保持するために使われる。ユーザの認証などに便利な反面、ローカルファイルへの書き込みを許すため、セキュリティの穴となり易い面がある。MyYahoo ではユーザ名、パスワード、メールアドレス、年齢、職種といった基本情報のほか、選択したサービス、ブックマークの内容、オークションでの売買履歴と評価、カスタマイズで指定したジャンル、などが個人化された情報としてサーバに保存される。これらの情報はダイレクトメール、ターゲットを絞った広告などに利用されている。

#### 2.3 ユーザ情報解析とその利用

ポータルサイト以外にもユーザ情報を個人化に用いることが試みられている。情報を選別し個人化するには、情報源からの膨大な情報をふるいにかける必要がある。これをフィルタリングと呼ぶ。

ユーザから情報を取得する方法には、Web ログの獲得、フォームを介したユーザからの直接情報、特定のプログラムを用いたシステムの監視などがある。

ユーザから得た情報を興味推定などに使うには、個人の情報のみに基づく方法と個人ユーザよりも全体を見て利用する方法がある。前者を内容に基づくフィルタリング、後者を社会的フィルタリングや協調フィルタリングと呼ぶ [14]。ポータルサイトのカスタマイズは個人情報に基づいて行われている。

2.1はサーバを用いた Web パーソナライゼーションの一般的な構成である。

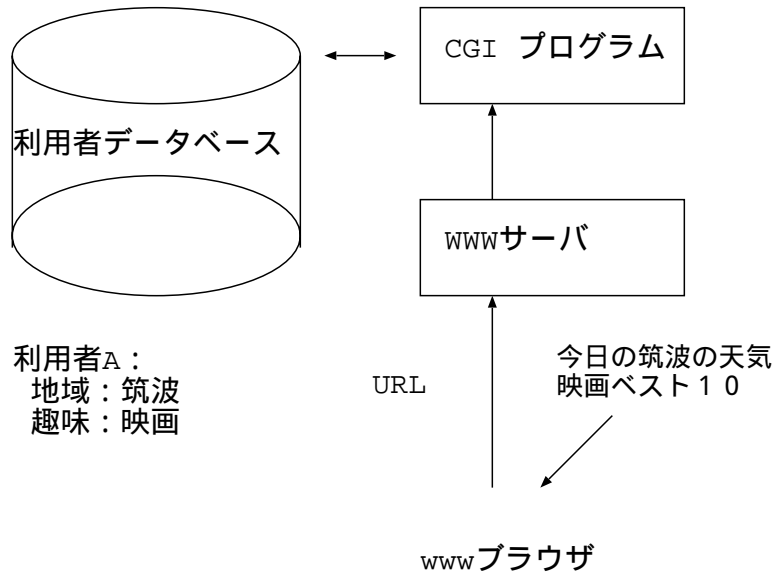


図 2.1: サーバを用いたパーソナライズの略図

ユーザを同定しないログ情報の利用の例として、goo.ne.jp ではユーザの全体の動向を検索ログから検索語のグループ化を行い [12]、情報のニーズを解析したり、トレンドを調べようと試みている。

amazon.com では書籍購入者に関連書籍の推薦やほかの購入者による書評の提示を行っている [7]。その本を買った人がほかにどんな本を買うかを示したり、すでに購入した人がどんな評価を下したかが分かるようになっている。ユーザの関心の個人情報と、それ応じた情報収集をしてできたりリコメンデーション (推薦) を交互に更新し、ほかのユーザの関心の情報との類似度によって情報を提示する [14]。

坂上らは ANATAGONOMY [11] の研究において電子新聞のパーソナライズを試みた。ANATAGONOMY はドキュメントベクトルとユーザプロフィールを用いて予想スコアを算出し、ニュースを自動配列するシステムである。サーバは画面拡大やスクロールの操作によってユーザの関心度を推定し、ユーザの明示的な評価によってこれを補正する。インターフェースは



図 2.2: リコメンデーション amazon.com

JAVA アプレットとして実装されており、サーバ側と連動して実行される。

しかし、これらのユーザ登録やユーザ動向解析を必要とするサービスは、ユーザの情報をサーバ側で大規模に管理する必要がある。そのため、これらの情報がマーケティングなどの重要な資料となる一方、ユーザのプライバシー保護が問題となる。My Yahoo でも登録時などにプライバシーに関するポリシーについて詳細な記述があり、ユーザの不安を取り除こうとしている。ネットワーク上でのプライバシーに関しては Platform for Privacy Preference(P3P)—[16] などで議論が進められているが、個人情報を取得されることによるユーザの不快感は MyYAhoo の例などからも分かるように大きな問題である。

## 2.4 Web 巡回ソフト

Bookmarkなどで管理されたページを順に探訪して更新状況をチェックしたり、サイト構成をローカルディスクに保存することによってユーザの Web 利用を補助するためのソフトは、一般には Web 巡回ソフトやオートパイロットソフトなどと呼ばれ、フリーウェアや製品として世に出ている。研究としては杉浦らの Internet Scrapbook[1]がある。この研究では、ユーザがブラウザ場で指定した Web ページを部分的に切り出して、更新をチェックし保存することによって巡回の手間を省いている。これはユーザ側からの積極的な Web 個人化を補助する試みと言え、本研究と非常に近いコンセプトを持っていると言える。しかし、保存したページの閲覧に際しては Windows の Explorer タイプのインターフェースを踏襲し、収集した情報の閲覧に関しては特に述べていない。

## 第 3 章

# 関連研究 事例 Internet Scrapbook および ANATAGONOMY

### 3.1 Internet Scrapbook

Internet Scrapbook[2] は杉浦らによる Web パーソナライジングに関する研究である。現在はホームページスクラップブックとして製品が発売されている。

#### 3.1.1 概要

Internet Scrapbook はいわゆる Web 巡回ソフトの機能に部分切出しの機能を付加したシステムである。Internet Scrapbook はスクラップページで指定された Web ページから必要な箇所の最新情報を取得する。スクラップページは、ユーザが Web ページの中で必要とする箇所を Web ブラウザ上で例示し、コピーアンドペーストを用いて作成する。一旦登録してしまえば、Web ページの内容が変更されても更新された情報が取得されるので、複数のページにアクセスすること必要は無く、スクラップページの更新要求を出すだけで良い。Internet Scrapbook はパターンマッチングによってスクラップページの更新を行う。例示操作によって作成されたマッチングパターンを用いて、もっともマッチすると思われる部分を抽出する。対話型のインターフェースによってより適切なマッチングを行わせるようにするための学習機能も持っている。

#### 3.1.2 データ獲得

InternetScrapbook では Web ブラウザ上で必要箇所を指定し、コピーアンドペーストを用いてスクラップブックに登録する。ブラウザは Microsoft Internet Explorer および Netscape Navigator をサポートしている。

表示されている Web ページと、タグなどを含んだテキストの HTML 文書は異なるのでカットバッファを用いるだけではスクラップブックの作成に必要な情報を得ることは出来ない。

Internet Scrapbook ではブラウザに固有の API(Application Program Interface) を利用

して Web ページの HTML ソースを取得し、スクラップブックのマッチングパターンと照合する。

### 3.1.3 マッチング

ユーザはブラウザに表示されている Web ページに対して範囲指定をする。マッチングパターンは直接書かずにコピーアンドペーストによる例示操作によって行う。Web ページは順時更新、変更されるので内容の変更によっても残る可能性の高い情報をマッチングのよりどころとする。Internet Scrapbook ではタイトルパターンとタグパターンという 2 つのマッチングパターンを用いている。

タイトルパターンはユーザの選択した箇所の前行、先頭行、次行からなる。これは記事のタイトルに注目した手法である。記事の内容が変更されても、文書の構成自体が変更されることは少ない。例えば、「今日の天気」「最新価格表」といった記事のタイトルは通常変更されない。

タグパターンは選択範囲内のタグの種類とそのページ中での順序からなる。ただし、<FONT> タグなどは除外し、見出し <Hn>、リスト <UL>、テーブル <TABLE> 等のレイアウトに関するタグのみに注目する。

更新が行われた時に、タイトル、タグパターン共に完全にマッチすればその箇所を抽出する。完全にマッチする箇所がない場合には、条件を緩和して部分的にマッチする場所の中から、最適と思われるものを抽出する。記事の位置よりもタイトルの方がユーザの例示の意味を明確に表すとの考えに基づいて、基本的にタイトルパターンを優先している。しかし、日付きの更新ログなどでは、タイトルそのものよりも記事の位置が重要な意味を持つ。具体的には、更新があったにも関わらず抽出部分に変更されないような場合にはタグパターンを優先させた候補を選択する。

また、よりユーザの意図を把握するために、アルゴリズムを変更させたマッチング例の候補の中から好みのものを選択させる機能を備えている。選択したアルゴリズムは学習され、その後のマッチングに用いられる。

## 3.2 ANATAGONOMY

ANATAGONOMY[11] は坂上らによる、「パーソナル電子新聞」に関する研究である。パーソナル電子新聞とは、利用者の好みに合わせたニュースを選択し、送信し、レイアウトをカスタマイズするサービスである。ANATAGONOMY は特定のニュースサイトからのニュースをユーザの情報に合わせて表示する。読売 COLiNS[15] と提携して個人向け新聞として運用された。

図 3.1 および図 3.2 がクライアントの実行画面である。クライアントは JAVA アプレット

版とスクリーンセーバー版が用意されている。図 3.1のアプレット版では簡単なレイアウト機能と、ユーザの再評価のためのスクロールバーなどのインターフェースが用意されている。

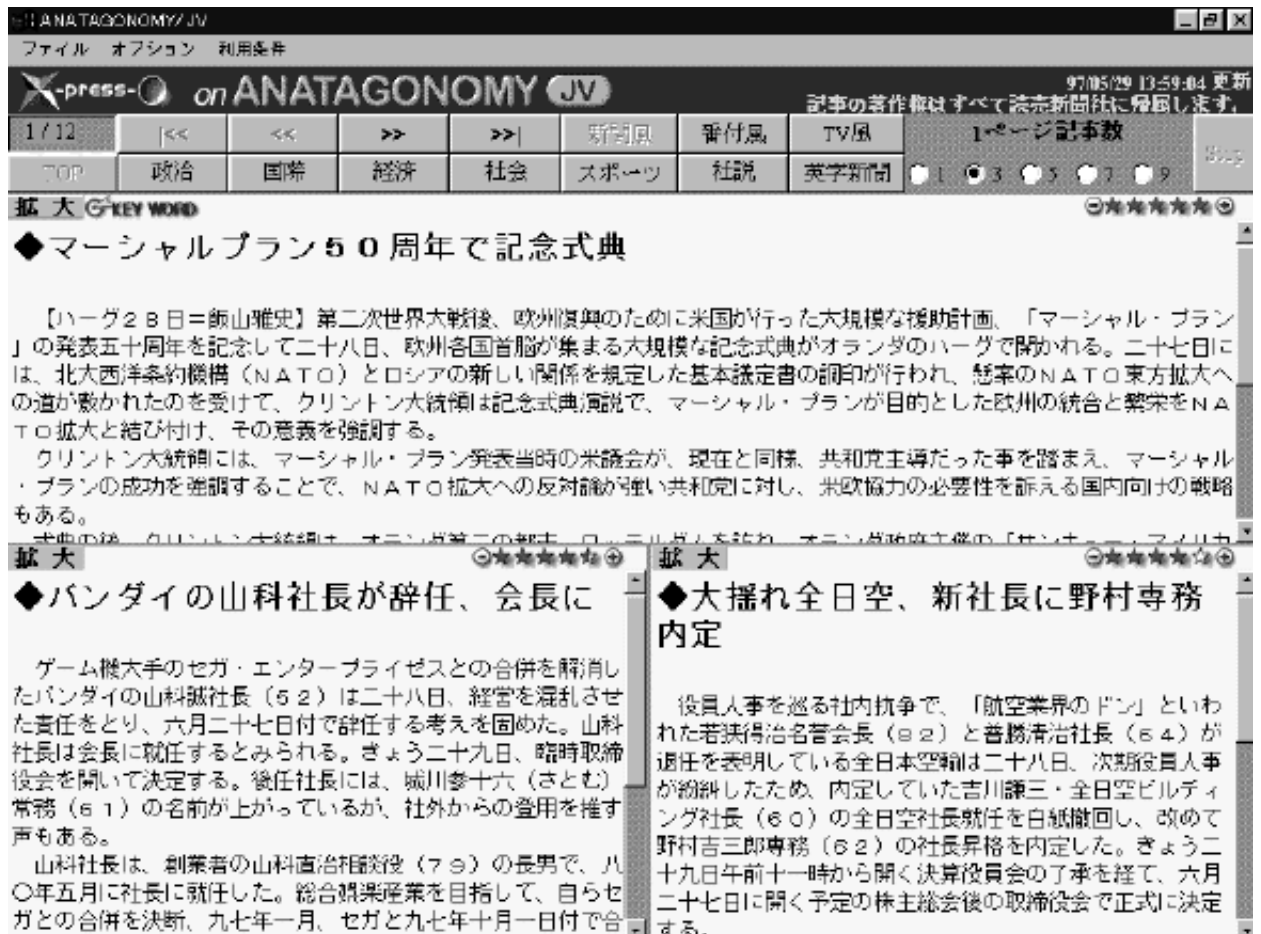


図 3.1: ANATAGONOMY JAVA 版

### 3.2.1 概要

システムは図 3.3のような構成を持つ。ユーザプロファイルにはキーワードと重みのセットが登録されている。学習エンジンはユーザプロファイルに基づいて記事の予想スコアを算出し、ユーザが採点したスコアからプロファイルの更新を行う。ドキュメントベクトルは記事内のキーワードのリストによって記事の特徴を表現する。。システムはユーザプロファイルとドキュメントベクトルを用いてユーザごとの記事の予想スコアを算定する。

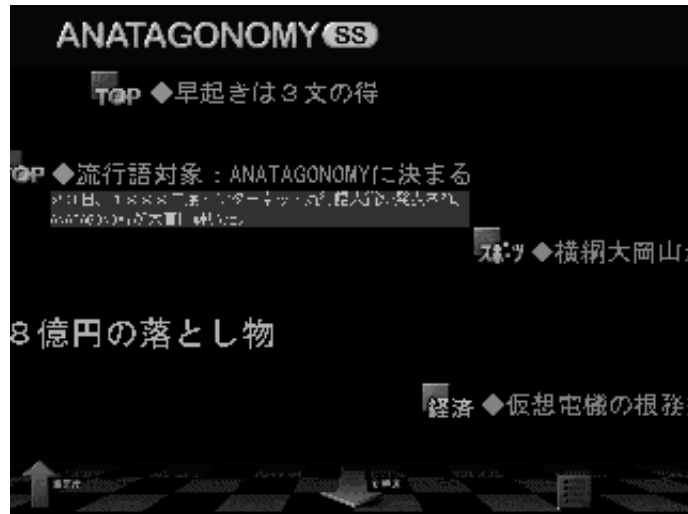


図 3.2: ANATAGONOMY Screen Saver 版

### 3.2.2 ユーザ情報の獲得

ANATAGONOMY は学習エンジン以外からの付加的なユーザ情報を明示フィードバックと暗黙フィードバックの2種類から獲得している。明示フィードバックとは、ユーザが表示クライアントのスコアバーによって記事に直接点数をつけることによってスコアを変化させることである。暗黙的なフィードバックとはスクロールおよび記事拡大の操作に関して、ユーザがどちらかの操作を行ったときには加点する。

### 3.2.3 レイアウト

計算されたスコアに基づいて、各ニュースを記事数に合わせて表示させたり、スクリーンセーバーとして表示させたりすることができる。記事は更新と算出されたそのユーザに対するスコアから優先度をつけてレイアウトされる。ユーザの操作はアプレットを通して監視され、暗黙フィードバックに用いられる。クライアントに付随するスコアバーによって記事の評価をユーザが修正することができる。

## 3.3 両者の比較と問題点

Internet Scrapbook はローカルのみで動作し、サーバを必要としないのに対して、ANATAGONOMY ではサーバを用いてユーザ情報を管理している。。

また、Internet Scrapbook では情報を源を自分の好きなページから取り、自由に切り貼りすることができるが、ANATAGONOMY はニュースソースを限定しており、あくまでニュースのパーソナライズにとどまっている。

両者の特徴を 3.1にまとめた。



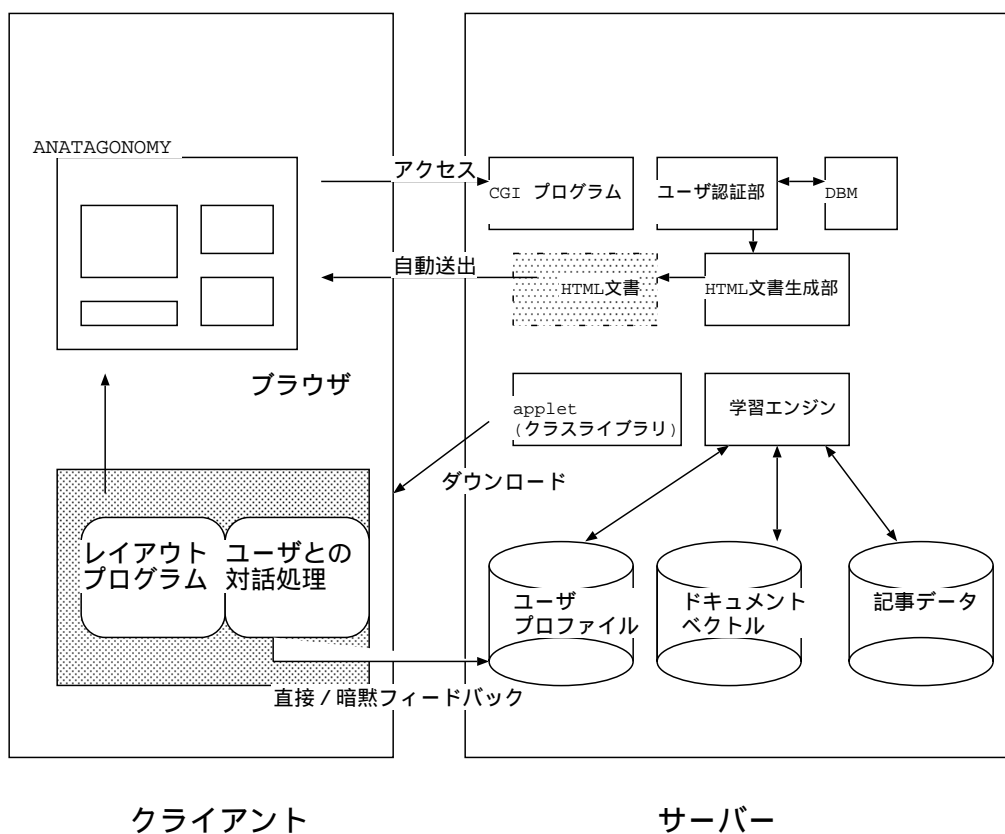


図 3.3: ANATAGONOMY の構成

表 3.1: Internet Scrapbook と ANATAGONOMY の比較

	情報ソース	サーバ処理	レイアウト機能
Internet Scrapbook	全てのページ	必要ない	上下関係のみ
ANATAGONOMY	配信される特定の記事	必要	平面レイアウト

## 第 4 章

### 新システム Webgrep の考察

#### 4.1 切取と自動レイアウトによる Web パーソナライジング

##### 4.1.1 概要

本研究ではユーザの自発的なパーソナライジングを助けるという立場に立っているため、前章で述べた 2 研究のうち特に InternetScrapbook の方針に興味がある。これに、ANATAGON-OMY に備わっているようなレイアウト機能が付加し、編集機能を強化した新システムを提案する。

基本的な機能として、ユーザは好みのページをダウンロードし、ページの任意の部分を指定して、以後、その部分とマッチする部分について WebGrep のみを通して更新状況や内容をチェックすることができる。

新要素としてレイアウト時の切り落としによる直接編集、一覧性を重視したレイアウトを行う。

集めた情報を一覧することによって比較検討などを容易にすることができる。たとえば、同一商品の価格について、複数のサイトの情報を一度に見ることができるサービスをユーザ自身である程度構築できる。情報を自分向けにパーソナライズした利用が容易になる。

##### 4.1.2 Web ページの切取と編集

Web ページの欲しい部分だけを得るためには、まずその部分を指定する方法が必要である。欲しい部分の指定には切り落としを用いる。一旦取ってきたページに対し不要な部分を落としていくことによって指定する。

具体的には、レイアウト指定 GUI で URL を指定すると編集ウィンドウが開き、HTML 文書が表示される。ユーザはソースではなく、ブラウザなどによる実際のレンダリング後のイメージを直接切り取る作業で欲しい部分を指定することができる。

InternetScrapbook の手法では、指定の結果が気に入らない場合ブラウザから何度もコピーアンドペーストを繰り返さなければならない。これに対して、新手法では、ページの表示例を直接操作して取得後のページイメージを例示することができる。指定後のイメージを直接

見て編集することでよりユーザのイメージに近い指定が可能である。

さらに更新されたページとの過去の指定とのパターンマッチの作業が必要である。HTMLはタグによって構造を記述された言語であるからこれを用いてマッチを行う。ユーザが欲しい部分の指定を行う場合、現在のHTML文書に対して指定を用いて例示的に行うことが前提となる。しかし、Webサイト上に置かれたHTML文書は変更されていく。基本的な構成が変わらなければタグの構成も変わらないので、これを基本的なマッチングに用いる。



図 4.1: 元ページ

気象庁02月14日(月)05時発表の天気予報(茨城県)


02月							
日付	14(月)	15(火)	16(水)	17(木)	18(金)	19(土)	20(日)
気温	max	8℃	8℃	7℃	8℃	9℃	10℃
	min	-	-1℃	-4℃	-5℃	-3℃	0℃
天気	☁	☁	☁	☁	☁	☁	☁
降水確率	10/20/20	30	10	10	10	30	40
洗濯指数		40	5時間程度で一応乾く				

図 4.2: 切り落とし編集後

### 4.1.3 レイアウト手法

切り取った Web ページを見易くレイアウトすることを考える。1 画面で集約された情報が得られることが重要である。複数の情報のレイアウトを決めるにはその情報源の個々の重要度や相互の関連が大きな要素となる。Web のパーソナライズの場合、ユーザが指定した個々のページの内容そのものは必ずしも相互に関係しておらず、場合によっては全く無関係な場合もある。つまり、Webgrep で扱う情報は全体の Context 性には乏しいと言える。そこで本研究ではレイアウトのための記事の重要度の評価に、更新の度合いとユーザの意図を直接反映させることにした。

取得したページのレイアウトはレイアウト指定モジュール図 4.3 によって指定される。レイアウト指定部は切り出した部分を配置するための GUI などからなる。レイアウトにあたっては更新情報を重要な要素として評価し、評価によって配置する方法と、固定した位置に特定のページからの情報を置く方法を組み合わせて表記する。

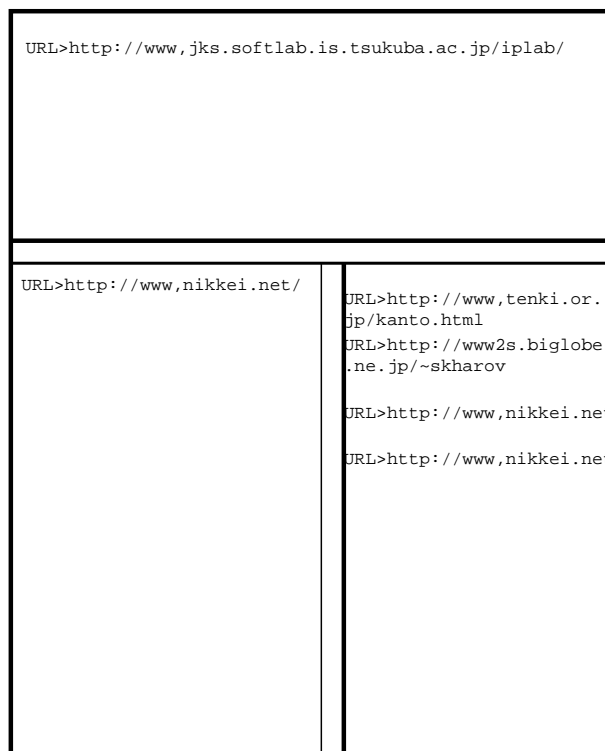










図 4.3: レイアウト指定 GUI

レイアウトは評価と指定に基づく平面分割 [3] によって行う (図 4.4)。集約された情報を利用することが本システムの最終目的であるが、閲覧にあたって、ユーザが登録する情報源が多くなると、レイアウトし切れないページができてしまう。そこで優先度の低いものに関しては更新状況のみを表示する。

また、スクロール等の操作も極力行わなくて良いレイアウトを行う。アニメーションズームを取り入れ、フォーカスの当たっているフレームを一時的に大きく表示する(4.5)。

URL><http://www.jks.softlab.is.tsukuba.ac.jp/iplab/>

気象庁02月14日(月)05時発表の天気予報(茨城県)

		02月						
日付		14(月)	15(火)	16(水)	17(木)	18(金)	19(土)	20(日)
気温	max	8℃	8℃	7℃	8℃	9℃	10℃	10℃
	min	-	-1℃	-4℃	-5℃	-3℃	-4℃	0℃
天気								
降水確率		10/20/20	30	10	10	10	30	40
洗濯指数			40	5時間程度で一応乾く				


  

<p>URL&gt;<a href="http://www.nikkei.net/">http://www.nikkei.net/</a></p> <p style="text-align: right; font-size: small;">最終更新時刻: 2000/02/14 07:34</p> <ul style="list-style-type: none"> <li>▶ 中田、移籍後初ゴール・ペルージャと引き分ける</li> <li>▶ 中田、自慢いられる4位・W杯スケート最終日</li> <li>▶ 日本、3-0で白星獲得・サッカーアジア杯予選</li> <li>▶ 「編者方は退屈しない」平井さん、地上とTV交信</li> <li>▶ さっさと雪まつり閉幕・入浴200万人突破</li> <li>▶ スタービー生みの親、C・シールドの死去</li> <li>▶ 赤子の会社員身辺紹介の件暴論起訴・黒田法務</li> <li>▶ 無効ゴルフ会費種を赤りつける - 信託受託で2人判明</li> <li>▶ 「話し相手が見つかった」、女性誌の白黒記者が復活</li> <li>▶ 社会人が4番とめる・ラグビー日本選手権</li> <li>▶ 大沢が5輪代表に参進・東京国際マラソン</li> <li>▶ 男子は奇襲、女子田中が優勝・スピードスケート全日本シユート</li> <li>▶ 丸山、69で2打差2位・米ビュイック招待第3日</li> </ul> <div style="border: 1px solid gray; padding: 5px; margin-top: 10px; text-align: center;"> <p>中田、移籍後初ゴール・ペルージャと引き分ける</p> </div>	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 60%;">URL&gt;<a href="http://www.tenki.or.jp/kanto.html">http://www.tenki.or.jp/kanto.html</a></td> <td style="width: 40%;">02/10 FRI 15:15</td> </tr> <tr> <td>URL&gt;<a href="http://www2s.biglobe.ne.jp/~skharov">http://www2s.biglobe.ne.jp/~skharov</a></td> <td>02/10 FRI 09:15</td> </tr> <tr> <td>URL&gt;<a href="http://www.nikkei.net/">http://www.nikkei.net/</a></td> <td>02/02 WED 15:15</td> </tr> <tr> <td>URL&gt;<a href="http://www.nikkei.net/">http://www.nikkei.net/</a></td> <td>01/31 MON 12:30</td> </tr> </table>	URL> <a href="http://www.tenki.or.jp/kanto.html">http://www.tenki.or.jp/kanto.html</a>	02/10 FRI 15:15	URL> <a href="http://www2s.biglobe.ne.jp/~skharov">http://www2s.biglobe.ne.jp/~skharov</a>	02/10 FRI 09:15	URL> <a href="http://www.nikkei.net/">http://www.nikkei.net/</a>	02/02 WED 15:15	URL> <a href="http://www.nikkei.net/">http://www.nikkei.net/</a>	01/31 MON 12:30
URL> <a href="http://www.tenki.or.jp/kanto.html">http://www.tenki.or.jp/kanto.html</a>	02/10 FRI 15:15								
URL> <a href="http://www2s.biglobe.ne.jp/~skharov">http://www2s.biglobe.ne.jp/~skharov</a>	02/10 FRI 09:15								
URL> <a href="http://www.nikkei.net/">http://www.nikkei.net/</a>	02/02 WED 15:15								
URL> <a href="http://www.nikkei.net/">http://www.nikkei.net/</a>	01/31 MON 12:30								

図 4.4: 表示部 GUI

URL>http://www.jks.softlab.is.tsukuba.ac.jp/iplab/

気象庁02月14日(月)05時発表の天気予報(茨城県)

02月							
日付	14(月)	15(火)	16(水)	17(木)	18(金)	19(土)	20(日)
気温	max	8℃	8℃	7℃	8℃	9℃	10℃
	min	-	-1℃	-4℃	-5℃	-3℃	-4℃
天気	☁	☁	☁	☁	☁	☁	☁
降水確率	10/20/20	30	10	10	10	30	40
洗濯指数		40	5時間程度で一定乾く				

URL>http://www.nikkei.net/

最終更新時刻: 2000/02/14 07:26

- 中田、移籍後初ゴール・ペルージャと引き分ける
- 田畑、白鷺いすれも4位・W杯スケート最終日
- 日本、3-0で白星奔進・サッカーアジア杯予選
- 「無重力は温泉みたい」毛利さん、地上とTV交信
- さっぽろ雪まつり閉幕・入出200万人超す
- スノーピー生みの親、C・シュルツ氏が死去
- 米子の会社員射殺仲介の作業員起訴・鳥取地検
- 無効ゴルフ会員権を売りつける・詐欺容疑で2人逮捕
- 「話し相手が欲しかった」・女性監禁の依頼容疑者が供述
- 社会人が4強占める・ラグビー日本選手権
- 大伏が五輪代表に前進・東京国際マラソン
- 男子は寺尾、女子田中が優勝・スピードスケート全日本シヨート
- 丸山、69で2打差2位・米ビュイック招待第3日

中田、移籍後初ゴール・ペルージャと引き分ける

URL>http://www.tenki.or.jp/kanto.html 02/10 FRI 15:15  
URL>http://www2s.biglobe.n.jp/~skharov 02/10 FRI 09:15  
URL>http://www.nikkei.net/ 02/02 WED 15:15  
URL>http://www.nikkei.net/ 01/31 MON 12:30

↑  
FOCUS

図 4.5: フォーカス変更によるズームング

## 第 5 章

### 結論と展望

Web は従来のメディアに比べて情報量が膨大であり、多岐にわたる。情報を整理して利用しやすくする personalization - 個人化 - はユーザをサポートする手段の一つとして研究実装が進められている。しかし、多くの Web パーソナライズの研究は企業や公開側の工夫であって、ユーザから積極的に自分だけの Web 利用を図るための動きは少ない。

そこで、本論文ではユーザ側から見た Web のパーソナライジングについて考察し、新システム Webgrep の提案を行った。

WebGrep ではページ巡回の作業の負担を軽減し、編集、レイアウトの機能を取り入れることによって情報の統合的な利用を図ることが可能になる。

今後は実装を進め、評価を行っていく予定である。現在、閲覧時のズームングアニメーションについて考察している。また、Chat や News といった他のメディアも含めてパーソナライズの対象として含めていく研究も行いたい。

## 謝辞

本研究を進めるにあたり、叱咤激励しつつ指導して下さった田中二郎教授に深く感謝します。また、NEC CC メディア研究所の古関義幸氏には研究に関して重要な助言を頂きました。筑波大学工学研究科 三浦 元喜氏には研究の内容から進め方まで多大な助言を頂きました。田中研究室の皆さんの暖かいサポートに感謝致します。



## 参考文献

- [1] 杉浦 淳, 古関義幸, Internet Scrapbook: 例示プログラミングによる Web ブラウジング, インタラクティブシステムとソフトウェア V 日本ソフトウェア科学会 WISS '97.
- [2] Sugiura A. and Koseki Y., "Internet Scrapbook: Creating Personalized World Wide Web Pages", Extended Abstracts of CHI "97 pp.343-344,1997.
- [3] 杉山公造, グラフ自動描画法とその応用 - ビジュアルヒューマンタフエース - 計測自動制御会, 1993.
- [4] <http://my.netscape.com/>
- [5] <http://my.yahoo.co.jp/>
- [6] 田中 一男 , ポータルサイト技術の最新動向, ACM SIGMOD 日本支部第 13 会大会, 1999.
- [7] <http://www.amazon.com/>
- [8] <http://celes.softlab.is.tsukuba.ac.jp/ssr/private/ppt/kishi99/ref.html>
- [9] <http://www.wwwmetrics.com/>
- [10] 神場 知成 小池 雄一 古関義幸 , 情報のパーソナライゼーションとその記述方式 , 人工知能学会誌 Vol.14. NO.6 ,1999
- [11] 坂上 秀和 神場 知成 古関 義幸 , パーソナル電子新聞 ANATAGONOMY の開発と評価 , インタラクティブシステムとソフトウェア IV 日本ソフトウェア科学会 WISS '96.
- [12] 梶谷浩一 , サーチエンジンの最新動向, ACM SIGMOD 日本支部第 13 会大会, 1999.
- [13] 通信白書 平成 11 年 郵政省  
<http://www.mpt.go.jp/policyreports/japanese/papers/99wp/99wp-0-index.html>
- [14] 福原知宏 強調フィルタリングに関する研究動向 1998  
<http://bandits.aist-nara.ac.jp/tomohi-f/Docs/>

[15] 読売 COLiNS  
<http://pnews.cplaza.ne.jp/>

[16] P3P  
<http://www.w3.org/P3P/>