

## NaMiSEN:ウェブ閲覧履歴への個人嗜好情報付加・抽出ツールの開発

大浦啓一郎<sup>†</sup> 日高 哲雄<sup>†</sup> 森田 哲之<sup>†</sup> 加藤 泰久<sup>†</sup> 奥 雅博<sup>†</sup>

<sup>†</sup> 日本電信電話株式会社 NTT サイバースリユーション研究所 〒180-8585 東京都武蔵野市緑町 3-9-11  
E-mail: †{oura.keiichirou,hidaka.tetsuo,morita.t,kato.yasuhisa,oku.masahiro}@lab.ntt.co.jp

あらまし インターネット上に膨大な数のウェブページが存在する中で、それらのページから個人にとって重要な閲覧履歴を管理し、状況に応じてその閲覧履歴を効率的に抽出し利用者に提示することの重要性が高まってきている。本研究では、閲覧履歴に対して利用者に個人嗜好情報の付加させ、その付加情報を利用して利用者が要求する閲覧履歴を抽出し利用者に提示する方式について検討を行った。また方式の有効性を確認するためプロトタイプの試作を行ったので報告する。

キーワード 閲覧履歴, ブックマーク, 活動履歴, 個人活動, 個人嗜好情報

## NaMiSEN:The development of a personal preference addition and extraction tool to personal web browsing history

Keiichirou OOURA<sup>†</sup>, Tetsuo HIDAKA<sup>†</sup>, Tetsushi MORITA<sup>†</sup>,  
Yasuhisa KATO<sup>†</sup>, and Masahiro OKU<sup>†</sup>

<sup>†</sup> NTT Cyber Solutions Laboratories NTT Corporation 9-11, Midori-Cho 3-Chome  
Musashino-Shi, Tokyo, 180-8585, Japan

E-mail: †{oura.keiichirou,hidaka.tetsuo,morita.t,kato.yasuhisa,oku.masahiro}@lab.ntt.co.jp

**Abstract** With the expanding number of web sites, the use of personal preference information has become increasingly important for the extraction of a user's intended search within his personal web history. This paper examines the addition of personal preference information via user action, and the extraction method obtained by using this information. We also describe a prototype designed for the investigation of the efficiency of these methods.

**Key words** web browsing history, bookmark, active logs, self activation, self preference information

### 1. はじめに

近年のインターネットの普及により、個人が閲覧できるウェブページの数には膨大になってきている。インターネット利用者の多くは、自分のあらゆる活動において有用な情報を得るために検索エンジンなどのツールを用いて自分の活動に資するページを探し利用している。また探索したウェブページが将来有用と利用者が判断した場合に、再利用に備え閲覧履歴を管理することも同時に行われている場合も多い。

最も一般的なウェブページの閲覧履歴を管理し再利用するツールとしてブックマーク機能が挙げられる。本機能により利用者はページ再利用時に効率的に探索できるように活動の種類毎などの名称を付与したフォルダによる分類を行い、URL 保管を行っている。

しかしインターネットの利用用途が多様化し、管理すべき対象が増加するに従って、リストの煩雑化やページの分類が

困難になっていくといった問題がある。また、最近増加している Weblog のようなウェブページにおいては、多種の内容が運営者個人の活動時系列などに従って記述されており、このようなページをある特定部分に興味を持って管理対象とし、ブックマークにより再利用したとしても、以前に管理対象にした部分の特定に時間を要するといった課題もある。

一方で、最近ローカルファイルに格納されているファイルを検索する GoogleDesktopSearch [1] などのローカル検索サービスが開始され、利用者のパソコンに存在する記録から情報を抽出する技術の有用性が注目され始めている。

さらに、パソコン上で行われるすべての個人活動を自動的に記録しておき、その活動から推測されたユーザの興味に従って閲覧履歴を利用者に提示する研究も始められている [2]。

本研究では、閲覧や活動の履歴を活用することに焦点を当て、効率的な閲覧履歴の管理およびユーザ要求に応じた抽出・表示を実現するために、再利用したいページに対して利用者が最小

限の操作で嗜好情報を付加する方法と、付加された嗜好情報を用いて閲覧履歴を抽出・表示する方式について検討した。また有効性を確認するためのプロトタイプである NaMiSEN (波線, Navigation and Mining Service with Efficiency Note) の試作を行い利用検証を行った。

## 2. 個人嗜好情報付加, 抽出・表示方式の検討

### 2.1 ウェブ閲覧履歴管理の課題

近年のインターネット利用者の傾向として、ウェブページ閲覧中に再利用したいページを発見した場合、閲覧頻度が高いものはブックマーク機能を用いて管理を行うが、比較的閲覧頻度低いものについてはブックマーク機能を利用しなくなってきている[10]。その理由は、近年のインターネット利用用途の多様化や検索エンジンの高度化などに加えて、ブックマークそのものにおける下記のような課題があると考えられる。

**課題 1** ブックマーク時に閲覧中の文章から視点を外して作業する必要があり閲覧行動が阻害される。

**課題 2** ページ内の興味を持った部分(文章)が特定できない。

**課題 3** フォルダによる分類が面倒である。

**課題 4** ブックマーク時に判断した分類が必ずしも再利用時に要求する分類になっていない。

**課題 5** 管理対象の増大につれリストが煩雑化し表示時の一覧性が損なわれる。

**課題 6** ページタイトルなどの文字列だけでは、ブックマーク時の記憶が想起されない。

上記のような問題を持つブックマーク機能の代替手段として、近年多くの利用者は閲覧経験のあるウェブページにおいても検索サイトなどの検索エンジンにキーワード入力を行い探索を行っている。

しかしブックマーク機能を用いずに、実際に検索エンジンなどにより過去に苦労して発見したページを再度探索しようとした場合、過去に検索エンジンの表示結果から遷移したページからさらにリンクを辿って苦労して発見したページであるなど、過去の探索プロセスを簡単に再現することが困難となり、多くの時間を要する場合がある[2]。また、検索エンジンによる探索は閲覧済みのページに限定して探索を行うわけではなく、その検索結果の一覧には、履歴内を検索したい個人にとって必要のないものが含まれており、検索結果を自分の閲覧履歴に基づきフィルタリングする要望も存在していると考えられる。

そこで、上述したブックマークにおける課題を解決するために、容易に閲覧履歴を管理対象とし、その中から利用者の要求に基づいて抽出する有効な手法が必要であると考えられる。

### 2.2 関連研究

ブックマークの高度化に関連する研究はこれまで数多く行われており、ツールも多く公開されている。より多くの情報をブックマークとして管理し利用する研究として、閲覧した Web ページに辿りついた方法や閲覧回数に着目してブックマークの情報として保持させるもの[5]や、ブックマーク時にコメントを付加できるもの[6],[7]などがある。また、ブックマーク分類などを自動化し、効率的な探索に寄与するための研究として、

ページの内容解析やリンク元の解析等により自動分類を行うもの[8]等がある。ブックマークを閲覧する行動に合わせて視覚化し利用者に提示する研究として、WebFarm[3]は自律的に動作するブックマークアイコンをブラウザ上に配置することで、直感的に内容が理解できる視覚的効果を取り入れたものがある。

さらに、ブックマークの高度化研究以外に利用者の行動履歴を再利用することで効率的な探索を実現している研究として、MemoryRetriever[2]は、そのウェブページに対しての利用者にとっての重要度を、個人の PC 上で行った活動履歴を自動的に取得し解析することにより推測し過去に見たウェブページの中から、その重要度を基に表示順位を決定し利用者に提示するという研究を行っている。

本研究では、ウェブページ中の特定部分が重要であるという利用者の意思表示をブックマーク登録よりも利用者にとって簡易な方法でかつ個人の行動履歴などを含めたできるだけ多くの情報を半自動的に嗜好情報として付加し、付加した情報を活用することで利用者が要求している閲覧履歴の抽出および表示内容を決定することにより、利用者の利便性向上を行うことに焦点を当てている。

### 2.3 個人嗜好情報付加, 抽出・表示方式の検討

2.1 節で述べた、ウェブ閲覧履歴管理の課題を考慮し、閲覧中のウェブページを管理していくための、個人嗜好情報付加, 抽出・表示方法の検討を行う。

まず、個人嗜好情報付加については、閲覧行動を阻害しない簡単かつ一般的な方法で、ウェブページ中の特定部分を管理可能とする必要がある(課題 1,2 に対する検討)。ここでは、利用者によるマウスによるテキスト文字列選択を行わせることにより、ウェブページの特定部分に対する個人嗜好情報を付加することを検討する。その理由として、近年よく利用されている Windows 等のユーザインタフェースにおいて、マウスを用いてテキスト選択する行為を、ウェブページ上の有用なテキスト文字列のコピーを行う場合などに多く利用されていると考えられる。また、利用者が管理対象としたいページと感じたときに読んでいた文章から視点を外さずに、その文章に対してテキスト文字列選択を行うことができ、ウェブページ中の特定部分を管理対象としながら閲覧行動を阻害せず読み進めることができるという効果も期待できる。ブックマーク機能を利用する際に発生するフォルダ等への分類作業については、後述する抽出・表示にその作業を行わせることにより、閲覧行動の中断を防ぐことが可能とすることを考える(課題 3 に対する検討)。

また抽出・表示方式については一覧性を保持する、利用者が抽出したいと考えた時の要求に基づいた分類という観点で実現方式を検討する(課題 4,5 に対する検討)。そこで、利用者のキーワード入力に基づき、個人嗜好情報および閲覧履歴より管理対象の一覧を表示させることを考える。検索エンジンの普及により、キーワードを入力して情報を絞り込んで探索していくという作業は、多くの利用者にとって必要不可欠な動作となりつつあり、ごく自然な操作といえる。抽出・表示時に利用者が思いついたキーワードは、その時の利用者の意図を最も反映したものであると言え、ブックマーク機能における過去のカテゴリ

業が再利用時に要求する分類になっていないという事象もある程度回避できると考えられる。

一覧を表示させる際に、嗜好情報を付加した時の情報をより多く表示させることにより、視覚的に記憶を想起させることも考慮する(課題6に対する検討)。具体的にはテキスト選択をした時に取得するサムネールや、テキスト選択したテキスト文字列を表示することにより、閲覧履歴を管理対象とした時の情報を視覚的にも想起させることを考える。

さらにページ中の特定部分に嗜好情報を付加し、その部分を利用者に提示する方式については(課題2に対する検討)、表示させる際にテキスト選択した部分に選択した時と同様にハイライト表示させることを検討する。これにより、管理対象とするときにウェブページ中のどの文章に対して興味を持ったのかという記憶をすぐに想起させることが期待できる。

前述した検討内容を基に個人嗜好情報付加、抽出・表示方式に求められる要件は下記のようにまとめられる。

- (1) 利用者が閲覧部分をテキスト選択することにより嗜好情報付加(以下、NaMiSEN 選択)を行い、テキスト選択部分(以下、NaMiSEN 選択部分)を保存できること
- (2) 利用者のキーワード入力により NaMiSEN 選択部分の抽出および一覧表示を行うこと
- (3) NaMiSEN 選択部分と NaMiSEN 選択時に同時に取得できる閲覧時間等の情報を利用し、一覧表示の順位を決定すること
- (4) 一覧表示時に NaMiSEN 選択を行った時のサムネールや NaMiSEN 選択部分の表示を行い、視覚的にも記憶想起を容易にするよう配慮すること
- (5) 一覧表示からウェブページキャッシュに遷移した時に、NaMiSEN 選択部分をハイライトし NaMiSEN 選択部分の探索を容易にすること

上記の要件を基にプロトタイプ化した、NaMiSEN について次章に述べる。

### 3. NaMiSEN (波線)

#### 3.1 NaMiSEN の概要

NaMiSEN の利用イメージを図1に示す。利用者はウェブページを読み進め、その中に重要な文章を見つけた際にその文章から視点を外さずに、NaMiSEN 選択を行い閲覧動作を続けていく。また、管理対象から抽出・表示させる際には、NaMiSEN 選択を行った箇所を思い出ししながら、そこに含まれるキーワードを入力することにより利用者の要求に応じた情報を抽出し、一覧を表示させる。表示された一覧にはサムネールなども表示されており、NaMiSEN 選択を行った記憶を視覚的に想起させることを期待している。

試作したプロトタイプの構成を図2に示す。図2中でテキスト選択により嗜好情報を付加する部分を NaMiSEN 選択部、キーワード入力により情報を抽出し表示する部分を NaMiSEN 抽出結果表示部と呼ぶ。キーワード検索を行う全文検索エンジンおよび形態素解析器はそれぞれ namazu [11], chasen [12] を用いた。次節より嗜好情報付加部、抽出結果表示部、また抽出

結果表示部の中で、履歴の表示順位を決定するために検討した適合度計算アルゴリズムについてそれぞれ詳説する。

#### 3.2 NaMiSEN 選択部

詳細イメージを図3に示す。

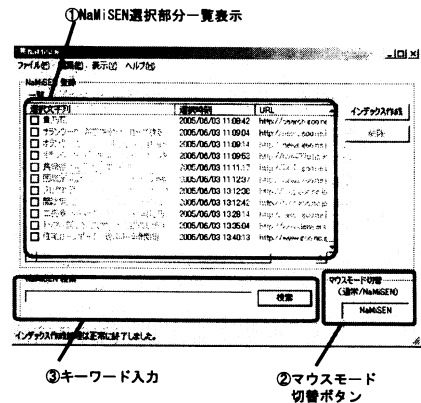


図3 NaMiSEN 選択部

利用者はウェブブラウザによりホームページを閲覧しながら、重要だと感じた部分や文章に対して NaMiSEN 選択を行いながら閲覧を進めていく。

利用者が NaMiSEN 選択を行うことにより嗜好情報として、URL、NaMiSEN 選択部分文字列、ウェブページ閲覧時間、NaMiSEN 選択時のページサムネール、ウェブページキャッシュの取得をシステムが自動的に行う。

また、利用者に対して NaMiSEN 選択により嗜好情報が登録されたことをフィードバックするため、NaMiSEN 選択を実施した際に図3中の① NaMiSEN 選択部分一覧表示部分に選択内容を追加し更新する。さらに、ウェブページ中のテキスト情報を NaMiSEN 選択を行わずに、コピーなど他の用途でマウスによるテキスト選択を利用する場合に、②のマウスモード切替ボタンを押下することにより NaMiSEN 選択が行えるモードと通常のマウス操作を行うモードを切替えることができるよう考慮した。

#### 3.3 NaMiSEN 抽出結果表示部

詳細イメージを図4に示す。利用者は NaMiSEN 選択部分に含まれる単語やそのウェブページに含まれる単語を想起しながら、図3の③や図4の④にキーワードを入力し NaMiSEN 選択部分の抽出を行う。

抽出結果として、図4の⑤に示すとおり、検索エンジンによりヒットした NaMiSEN 選択を行ったウェブページの情報である NaMiSEN 選択時サムネール(⑤-1)、ページタイトル(⑤-2)、URL(⑤-3)などがそれぞれ表示される。サムネールを用いることにより NaMiSEN 選択時の記憶を視覚的に想起できる可能性を考慮した。またこれらにはリンクが設定されておりサムネールの拡大表示(⑤-1のリンク)、ページキャッシュ(同⑤-2)、ウェブページ(同⑤-3)などが関連付けられている。利用者はそのリンクをクリックすることによりそれらの情報を閲覧することが可能となる。

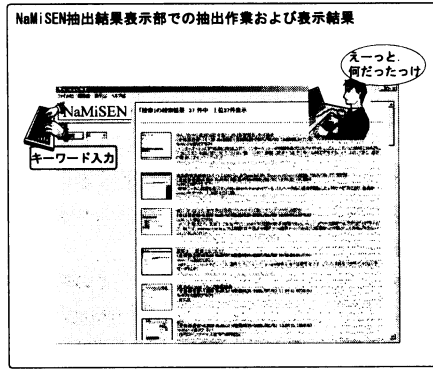
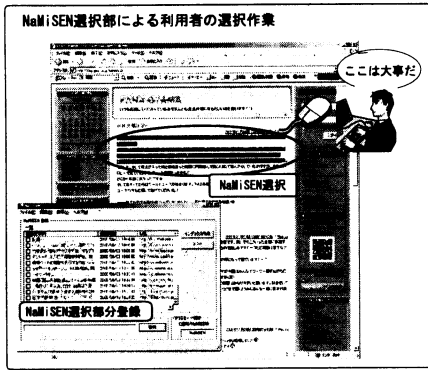


図1 利用イメージ

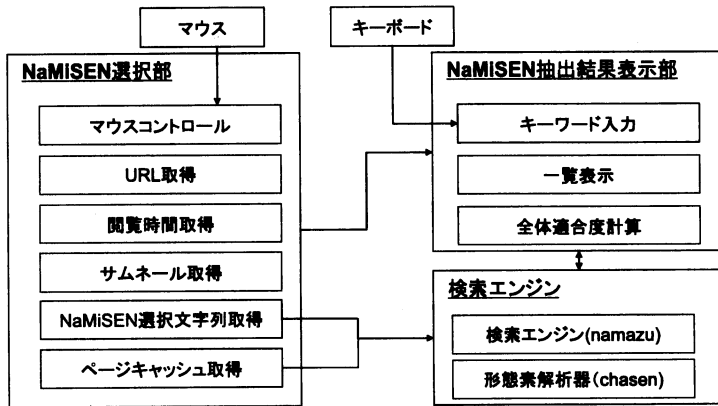


図2 NaMiSEN プロトタイプ構成

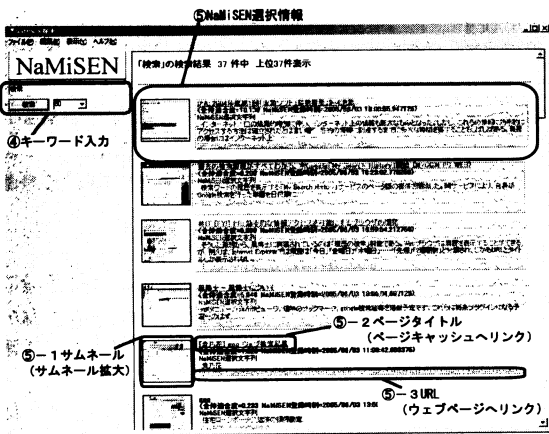


図4 NaMiSEN 抽出結果表示部

NaMiSEN 選択情報の表示順位は適合度順に表示しており、算出方法は次節で述べる。

### 3.4 適合度算出アルゴリズム

2.3 節の要件 (3) で述べた、抽出結果表示を利用者の要求に近づけるために、取得した嗜好情報を活用し表示順位を決定するアルゴリズムについて検討する。

ここで利用者は本システムにより、閲覧履歴を抽出するために入力すべきキーワードについて下記の情報が与えられていると仮定する。

- (1) NaMiSEN 選択部分に含まれる単語を入力する
- (2) NaMiSEN 選択したウェブページに含まれる単語を入力する
- (3) (1) を入力するほうが表示順位が高い

上記のような仮定の下、入力されたキーワードに対する、NaMiSEN 選択文字列空間 ( $W^n$  とする) と NaMiSEN 選択したウェブページの文字列空間 ( $W^c$  とする) をそれぞれ検索エンジンによって検索した結果の TF-IDF 値 [9] と、NaMiSEN 選択時の当該ウェブページの閲覧時間を利用し表示順位を決定することを考える。これにより、 $W^n$ 、 $W^c$  それぞれの空間におけるキーワードとウェブページ間の距離、および上記の情報が与えられた利用者の要求を考慮した表示順位を示す値を得ることができると考えられる。

そこで次のようなアルゴリズムにより適合度を算出することを考える。ウェブページ  $P_i$  に対する NaMiSEN 選択したウェブページの閲覧時間を  $T_i$ 、利用者が入力したキーワードを  $q$ 、NaMiSEN 選択を行った箇所数を  $m$ 、 $W^n$ 、 $W^c$  をそれぞれ検索した時に得られる TF-IDF 値を表示順位に反映させる度合をあらわすパラメータを  $\alpha_n, \alpha_c$ 、当該ウェブページを閲覧した時

間の長さを表示順位に反映させる度合をあらわすパラメータを  $\alpha_t$  とする。

【適合度算出アルゴリズム】

(1)  $W^n$  での TF-IDF 値  $d_i^n(q)$  ( $i = 1, \dots, m$ ) を求め、最大値で割ることにより 0-1 正規化し  $d_i^{n*}(q)$  とする。

(2)  $W^c$  での TF-IDF 値  $d_i^c(q)$  を求め、最大値で割ることにより 0-1 正規化し  $d_i^{c*}(q)$  とする。

(3)  $T_i(q)$  の最大値で割ることにより  $T_i(q)$  を 0-1 正規化し  $T_i^*(q)$  とする。

(4)  $r_i(q) = \alpha_n d_i^{n*}(q) + \alpha_c d_i^{c*}(q) + \alpha_t T_i^*(q)$  なる  $r_i(q)$  を適合度とする。

上記のアルゴリズムにより適合度を計算し、NaMiSEN 抽出結果表示における表示順位を決定する。

#### 4. 利用検証

プロトタイプの有効性を確認するため、利用検証を行った。3.4 節で述べたパラメータを  $\alpha_n = 10, \alpha_c = 1, \alpha_t = 0.1$  とし、NaMiSEN 選択部分に含まれる文字列がキーワードとして入力された際の表示優先度を高く設定し、主に NaMiSEN 選択部分が一覧表示結果に与える有効性の確認を行った。利用者の興味に基づき約 100 ページについて NaMiSEN 選択を行いながら閲覧し、いくつかのキーワードを入力して表示結果を確認した。

また、表示結果のそれぞれが利用者が求めているページかどうかを主観的に判断し表示順位 5 位までについて 3 段階 (○ △ ×) で評価を行った。

結果の例を表 1, 2 に示す。表の項目の「順位」は表示順位、「ヒット部分」は、キーワードが  $W^n, W^c$  のいずれにもヒットした場合は both,  $W^c$  のみにヒットした場合は  $W^c$  と記述し、「 $r_j(q)$ 」には、算出した全体適合度を記している。

表 1 キーワード「ゴルフ」の表示結果

順位	ヒット部分	主観的判断	$r_j(q)$
1	both	△	11.094
2	both	△	11.094
3	both	○	11.094
4	both	△	11.094
5	both	△	11.094

表 2 キーワード「ワールドカップ」の表示結果

順位	ヒット部分	主観的判断	$r_j(q)$
1	both	△	10.979
2	both	○	8.738
3	both	○	7.438
4	$W^c$	△	2.058
5	$W^c$	○	0.800

表 1 を見ると、 $r_j(q)$  の値がすべて同じになっている。これは、同じウェブページの「ゴルフ」という単語を含む文字列に利用者が NaMiSEN 選択を行い、表示においても同じウェブページが表示順位 5 位までを占めたことが原因である。

また表 2 を見ると、「ワールドカップ」を含む文字列を

NaMiSEN 選択した結果以外にも、利用者が主観的判断により有効だと判断したウェブページが第 5 位に存在する。このウェブページを分析すると、NaMiSEN 選択を行った部分の近傍に当該単語が存在することがわかった。

キーワードに対する平均抽出数は 7 個でおおのこのキーワードにおいて、表 3 のようになった。

表 3 キーワードに対する抽出数

キーワード	抽出数
ゴルフ	12
岩手 そば	7
地震	4
ワールドカップ	8
地図	4

システムの反応速度については、検索ボタン押下後すぐに結果が表示される。しかし本プロトタイプでは抽出・表示を行う前にインデックスを作成する必要があり、NaMiSEN 選択数と NaMiSEN 選択時に保存されるウェブページキャッシュのサイズに依存し時間がかかることがわかった。

#### 5. 考 察

利用検証の結果を見ると、利用者は抽出・表示時において NaMiSEN 選択を行った部分に含まれる単語を主としてキーワードとして選択していることがわかった。しかし NaMiSEN 選択した部分の近傍のキーワードを選択し当該ウェブページを要求している場合もあり、このようなウェブページを利用者の要求とみなして表示順位をできるだけ上位にするためには、 $\alpha_n, \alpha_c$  のパラメータ設定の検討と同時に、NaMiSEN 選択部分の近傍に存在する文字列を適合度計算アルゴリズムに反映させるためのアルゴリズムの改良について検討する必要がある。

分類の観点から見ると、NaMiSEN 選択の総数 100 に対して、1 つのキーワードに対して平均 7 件の NaMiSEN 選択が抽出される。主観的判断から、抽出表示時における利用者が入力したキーワードによりある程度分類が行われていると考えることができる。

表 1 の結果では、上位 5 位がすべて同一のウェブページに付与された NaMiSEN 選択であり、表示結果が重複している部分が存在し一覧性を損ねているため、表示結果の URL を基にしたグルーピング方式の検討が必要である。

NaMiSEN 選択の方式として採用したマウスによるテキスト選択においては、ウェブページ中のリンクが設定されている部分やリッチコンテンツなどにおいて、情報が付加できないもしくは付加しにくい部分が多く存在するため、嗜好情報付加方式の再検討が必要である。

利用者が入力するキーワードが、NaMiSEN 選択を行った近傍の文字列にであることも多く見られるため、適合度算出において、これらの文字列を反映させるアルゴリズムについて検討する必要もあると考えられる。

## 6. おわりに

### 6.1 まとめ

本稿では、ウェブページの閲覧時に個人嗜好情報を容易に付加し管理対象とする方式と、それらの情報を利用者の要求に基づいて抽出し、わかりやすく表示するための検討を行った。また、適合度算出アルゴリズムにより抽出時の表示順位を決定し、できる限り利用者の要求に沿ったウェブページを上位に表示できるように考慮した。

また、プロトタイプである NaMiSEN を試作し、利用検証を行うことにより検討した方式の有効性の確認と課題点の抽出を行った。

本プロトタイプをブックマーク機能と比較した際のメリットとしては、下記が挙げられる。

- キーワードにより管理対象を抽出し嗜好情報に基づいて順位を決定し表示させるため、一覧性を損なわずに情報を表示することができる。

- キーワードにより分類して表示させるため、利用者が当該ウェブページを管理対象としたいときに分類作業が発生しない。

- 閲覧を行っている箇所視点を外すことなく管理対象とするための作業を行うことができるので、利用者の閲覧行動を阻害せずにウェブページを管理対象とすることができる。

- ウェブページ中の特定部分（文章）を管理対象とし、再閲覧時に強調して表示させることによりウェブページを表示した後の文章の特定に時間を要しない。

- 一覧表示時に文字列のみではなくサムネールなどを表示し利用者の NaMiSEN 選択時の記憶を視覚的に想起させることで、直感的な探索が可能である。

### 6.2 今後の課題

今後は探索効率向上のために、適合度計算アルゴリズムの妥当性や最適パラメータ値の検討、および表示している情報の有効性の検証などについて多くの被験者による実験により行っていく。

また 5. 章で述べた適合度計算アルゴリズム改良、URL によるウェブページグルーピング方式、NaMiSEN 選択方式、NaMiSEN 選択部分の近傍の文字列を定義し適合度算出に反映させるアルゴリズムなどの検討を行う予定である。

## 文 献

- [1] GoogleDesktopSearch <http://desktop.google.com/>
- [2] 森田哲之, 日高哲雄, 倉恒子, 藤本強, 加藤泰久 "Memory-Retriever ~ 履歴を利用したローカルファイル検索システム", 電子情報通信学会 2005 年総合大会講演論文集, D-8-11, pp83, 電子情報通信学会, 2005
- [3] 日野洋一郎, 中島伸介, 小山聡, 田中克己 "WebFarm: 動物メタファを用いた Web ブックマーク再利用機構", DEWS2003 7-P-06, 2003
- [4] 中島伸介, 黒田慎介, 田中克己 "閲覧履歴を反映したコンテキスト依存型 Web ブックマーク", 情報処理学会論文誌: データベース, Vol.43, No. SIG05, 情報処理学会, 2002
- [5] H. Takano, and T. Winograd, "Dynamic Bookmarks for the WWW", Proceedings of the ninth ACM Conference on Hypertext and Hypermedia, 1998
- [6] Yasuyuki Kashiwagi, "Sleipnir Version 1.66", <http://sleipnir.pos.to/>

- [7] W.S.Li 他, "A System for Personalizable Web Information Organization, Sharing, and Management", 8th International World Wide Web Conference, 1999
- [8] Yoelle S. Maarek 他, "Automatically organizing bookmarks per contents", Computer Networks and ISDN Systems, Vol.28, No.7-11, pp.1321-1334, 1996
- [9] 北 他, "情報検索アルゴリズム", 共立出版, 2002
- [10] 山崎潤一郎, "ブックマーク「お気に入り」が廃れたワケ", <http://arena.nikkeibp.co.jp/col/20030212/103753/>
- [11] Namazu Project "http://www.namazu.org/doc/manual.html"
- [12] 松本 他, "形態素解析システム「茶筌」version2.3.2 使用説明書", <http://chasen.aist-nara.ac.jp/hiki/ChaSen/>, 2003