社団法人 情報処理学会 研究報告
IPSJ SIG Technical Report

2004－DBS－134 (II) (46)
2004／7／14

個人コンテンツからの概念体系の生成とこれに基づくWeb検索のパーソナライゼーション

大島裕明† 小山聡† 田中克己†

個人のコンピュータには，その個人がどのような知識を持っているか，どのような考え方をしているか，ということがわかる情報が含まれている．しかし，それらはコンピュータに利用できるような状態にはなっていない．現在，さまざまな分野でシソーラスのような一般的な概念体系が用いられているが，個人のコンピュータに存在するコンテンツから個人的な概念体系が作成されれば，さまざまな分野におけるパーソナライゼーションが可能になる．本稿では，個人コンピュータに存在する文書とその分類の方法から，個人的な概念体系を作成する手法について提案を行い，作成された個人的な概念体系を用いてウェブ情報検索におけるパーソナライゼーションを行う手法について提案を行う．

Generating the Personal Concept Classification based on Personal Contents and Web Search Personalization

Hiroaki OHSHIMA,† Satoshi OYAMA† and Katsumi TANAKA†

A personal computer has a lot of documents. Those include much imformation that shows what the user is interested in, knows, and so on. However, the computer just has the information and it can not be used automatically. Now, common concept classification like thesaurus is used in many fields, so if the personal concept classification is created automatically based on the personal contents in the personal computer, it will be possible to be personalized in many fields. In this paper, we propose the way to create the personal concept classification from the personal contents and the method of the Web search personalization.

1. はじめに

インターネットが普及し，情報伝達の主要な手段となったことにより，個人はさまざまな情報をコンピュータ上で扱うようになった．それに伴い，メール，仕事に関連する文書，興味のある文書など，さまざまな文書がパーソナルコンピュータで保存，管理されるようになった．当然，個人個人のコンピュータには，その人独自の文書が存在し，もし，他人がそれらの文書を読むと，その人がどのようなことに興味があり，どのようなことを知っており，どのような考えを持っているか，といったことまである程度わかるぐらいの情報が存在している．

一方，それらの文書はコンピュータ上に存在しているだけであり，コンピュータが自動的にそこから得られる知識を利用する，といったことは行われていない．

† 京都大学大学院情報学研究科社会情報学専攻
Department of Social Informatics,
Graduate School of Informatics, Kyoto University

例えば，図1でしめしたように，現在のウェブ情報検索においては，ユーザがすでにどのような情報を知っているか，使っている言葉はどのような考えのもとで使っているのか，ということを検索エンジンに伝えることはできていないのである．

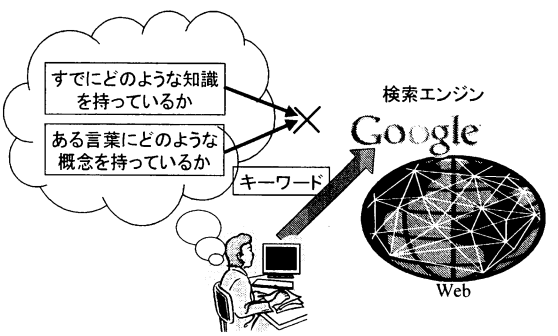

図1 ウェブ情報検索の現状

しかし，コンピュータに存在するさまざまな情報か

ら個人的な概念体系が生成され，自動的にコンピュータが利用可能な状況になることによって，さまざまなアプリケーションにおいて個人の意図や知識を考慮したサービスが提供できるようになると考えられる．先ほどの検索の例でも，個人的な概念体系があれば，ユーザの意図を検索エンジンに伝えることができたり，検索結果を個人に合わせた形で提示できるようになると考えられる．

そこで本研究では，そのような，個人が保有するコンテンツから個人的な概念体系を生成する手法と，そのようにして作られた個人的な概念体系を利用するアプリケーションの一例として，ウェブ情報検索において利用する手法についての提案を行う．

以下，2 章で関連研究について，3 章で個人的な概念体系について，4 章で個人コンテンツからの概念体系の生成について，5 章で個人的な概念体系を用いたウェブ情報検索のパーソナライゼーションについて，6 章で本研究のまとめと今後の課題について述べる．

2. 関連研究

すでにいくつかの研究において，個人的な情報からの知識を生成するような研究が行われている．それらについて述べる．

Haystack[1),2)] は MIT が開発した，個人的な情報管理システムである．扱う情報は，e-mail やカレンダー，文書，Web ページなど多岐にわたり，それらを RDF[8)] で管理する．本研究では，Haystack と同様に，さまざまな個人的な情報を管理するが，蓄積された情報そのものをより効果的に利用しようとする Haystack とは異なり，新たな情報を獲得する際に，蓄積された情報を利用することを目的としている．

WorkWare++[3),4)] は富士通研究所が開発した，会社などのグループで用いられるビジネス文書の蓄積と再利用のための情報管理システムである．さまざまな文書が登録され，その登録時には時間などのメタ情報が自動的に付加される．また，人やイベントの情報も同時に管理されている．ユーザは蓄積されたメタ情報を元に，ある研究分野に関してどのような技術が蓄積されているかや，ある事柄を知っていそうな人が誰であるかなどの情報を取得可能である．本研究では，知識としての情報管理を各ユーザが行うとともに，それらを現在の Web の利用のために利用することを目的としており，WorkWare++で行っている，グループによる情報共有や，蓄積された知識の獲得とは異なる．

Hyperclip[5)] は NTT が開発した，知識流通プラットフォームである．ユーザが利用した複数のコンテンツの間の関係を表現することができ，そこで作成された RDF をピア・ツー・ピアネットワークで共有することによって，ある文書と関連する文書を検索することができるようになる．Hyperclip で検索できる文書はピア・ツー・ピアネットワーク上の誰かによってメタ情報が付加されたものである．本研究では，現在の Web にある情報を検索エンジンなどを用いて利用するときに，自分の既得の知識を利用できるようにすることが目的であり，Hyperclip とは目的が異なる．

湯川ら[6)] は，個人が所有する文書に出現する単語の隣接度合いから，それぞれの単語同士の関連度合いを表す概念ベース，パーソナル・リポジトリを個人ごとに作成した．ユーザがコミュニティーのピア・ツー・ピア型システムの他の人が保有する情報を検索するときには，エージェントが検索キーをパーソナル・リポジトリによって拡張し，他人のパーソナルリポジトリ内でどのような情報が検索結果として適当であるかを判断することが可能になる．本研究と同様に，個人が所有する文書をもとに個人の概念を表しているが，その目的がピア・ツー・ピア型のネットワークで共有することであり，Web 情報の自動取得を目的とする本研究とは異なる．

これらの研究は，ある特定の環境やコミュニティーの中で利用可能な知識を作成しようとしている点において，われわれのものとは異なっている．

3. 個人的な概念体系

3.1 概念体系の表現

概念体系とは，上位下位の関係を持った概念の集合のことである．例えば，「犬」という概念に対して，下位関係にある概念には「チワワ」や「ポメラニアン」といったものがあり，上位関係にある概念には「哺乳類」がある．このような概念とその関係をあらわしたものが概念体系である．一般的な概念体系の中で，コンピュータで利用可能なものとして有名なものには，EDR 電子化辞書の概念辞書や日本語語彙大系がある．

厳密には概念体系ではないが，シソーラスも一種の概念体系であると考えられる．シソーラスは語彙の関係を表した辞書で，ある語彙に対して，広義語，狭義語，同義語，反意語などが記述される．シソーラスで有名なものとしては WordNet が挙げられる．

本研究では個人コンテンツから概念体系を生成することを目的とするが，個人コンテンツから概念体系を作成することは非常に困難であるため，概念体系というよりもむしろ，ある語彙に対して個人がどのような考えを持っているか，ということを他の語彙との上位

下位関係を明らかにし，それをあらわすことによって個人的な概念体系とする．

3.2 個人的な概念体系の必要性

個人的な概念体系がなぜ必要かということについて述べる．ある事柄に関して，個人が持つ考えはそれぞれ微妙に異なるものである．例えば，「サッカー」ということについて，人によって最も関心があることは異なり，図 2 で示したように，

- テレビで観戦すること
- スタジアムで観戦すること
- 実際にプレーすること

などが考えられる．

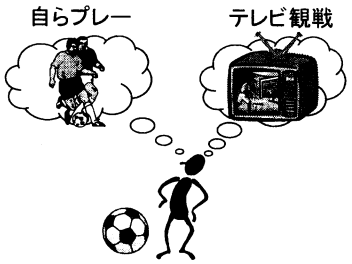

図 2 概念体系の個人差

また，同じサッカー観戦にしても，その対象は

- 高校生のサッカーも含めたあらゆる範囲
- J リーグの特定チームのみ
- ワールドカップのみ
- ヨーロッパ各国のリーグ

など人それぞれであることは簡単に想像がつき，例えば，ヨーロッパ各国のリーグについてだけ興味を示している人が，「サッカー」というキーワードで情報検索を行ったときに，日本の高校サッカーの情報や，日本サッカー協会の情報が結果として返されることには好ましいとは言えない．

このような観点から個人的な概念体系が必要であると考えられる．

3.3 オントロジによる概念体系の記述

オントロジとは，その記述レベルによっていくつかの役割に分けることが可能であるが，基本的には，概念どうしの関係を記述したものである．オントロジの記述に関しては現在，セマンティック Web[7] の分野で盛んに研究が行われており，OWL[9] というオントロジ記述言語が有名である．

OWL では，クラスの記述によって概念体系を記述することが可能である．概念どうしの基本的な関係としてとして上位下位関係があるが，概念体系をオントロジを用いて記述する際には，1 つの概念に対して 1 つのクラスが割り当て，そのスーパークラス，サブクラスが，それぞれ上位概念，下位概念に割り当てることで表現可能である．

この上位下位関係は，基本的に is-a 関係である．OWL では has-a 関係はクラスに対するプロパティで記述される．

OWL で概念体系を記述することの利点の 1 つは，OWL がセマンティック Web のツールの 1 つとして開発されたことであり，それによって，分散されたオントロジを統合するということも十分にサポートされている．

本研究での個人的な概念体系はまさに分散された概念体系であり，これを OWL で記述することによって，将来的に各個人の概念体系を共有したアプリケーションを考えることができる．よって，本研究では個人的な概念体系を OWL で記述することを目標の 1 つと位置づける．しかし，OWL の表現では表現しきれない情報が存在するため，結果としては，OWL によるオントロジと RDF[8] による付加情報という形で個人の概念体系を表現する．

4. 個人コンテンツからの概念体系の生成

4.1 個人的な概念体系を生成する手法

個人的な概念体系を作成する手法には，

- 何もない状態から作成する
- 既存の一般的な概念体系を変化させる

という 2 つの異なる方法が考えられる．

まず，何もない状態から作成する方法について検討する．この手法の場合，一般的な概念体系のように全ての概念を網羅したような概念体系を作成することは事実上不可能であり，ある程度の範囲における概念とその上位下位関係を発見することによって概念体系を生成することが現実的である．

次に，既存の一般的な概念体系を変化させる手法であるが，

- 何もない状態から作成する方法と同じやり方で新しい概念体系を発見し，付加する．
- 概念に利用頻度を付け，評価する．よく使われるものを高スコア，あまり使われないものを低スコアとし，実際に概念体系が利用されるときに，そのスコアも利用する．
- 概念どうしの共起は人によって異なると考えられ，関連する度合いとして共起度を付加する．

といったことを行うことが考えられる．

4.2 概念体系を生成するための情報源

概念体系の生成のためのポイントを

- 概念（の見出しとなる語彙）の発見
- 概念どうしの上位下位関係の発見

ということに絞って，個人が保有するコンテンツから利用可能な情報にどのようなものが存在するかということについて検討を行う．

情報のソースとして考えられるものには，

- 一文書
- 文書内の構造
- 文書に付加されているメタ情報
- 文書間のリンク
- 文書管理方法

といったものが考えられる．

4.2.1 一 文 書

概念の発見のために，文書内の語彙の出現頻度を用いることが考えられる．例えば，TF-IDF 値を用いて特徴ベクトルを作成したときに，高い値を持つ語彙は，その文書における主題となっていると考えられる．また，定型の構文から上位下位関係を見つけることが，自然言語処理の分野で行われている．

4.2.2 文書内の構造

HTML のように，文書内に構造を持つような場合，例えば，タイトルやヘディング要素に記述された内容は，段落要素に記述されていることのより上位の概念が記述されていることが多く；そのような情報から概念の上位下位関係を発見できる可能性がある．

4.2.3 文書に付加されているメタ情報

文書にメタ情報が付加されている場合，概念発見や上位下位関係の発見の精度を高めるために利用できると考えられる．

4.2.4 文書間のリンク

文書どうしのリンクには何らかの意味が存在し，その関係が上位下位関係であると判断することができれば，それらの文書内に含まれる概念やリンク周辺に存在する概念どうしの関係を推測することが可能である．

4.2.5 文書管理方法

個人はさまざまな文書を何らかの形で分類して管理していることがほとんどである．メールはメールソフト内で送信元ごとやタスクごとに分類され，ワープロ文書や論文などは内容ごとに分類されることが多い．その分類の仕方は，個人の意図が反映するものであり，また，ツリー構造上に分類されることが多く，そこから上位下位関係が取り出せる可能性がある．

最近では，個人が Blog サイトを保有してさまざまな Blog 記事を作成するということも増えてきている．Blog では，カテゴリといった形で分類が行われている．現在は，カテゴリの階層構造は作られていないが，将来的に階層構造をもったカテゴリが扱われるようになれば，それもファイルのディレクトリ構造での管理と同様に扱うことができる可能性がある．

本稿の以降では，特にこの文書管理方法に着目して，何もない状態から概念体系を作成する手法について検討を行う．

4.3 文書の管理情報からの概念体系の生成

4.3.1 概　　要

まず，本手法の前提として必要なこととして，個人は保有する文書を，ディレクトリのツリー構造上で管理しているものとする．

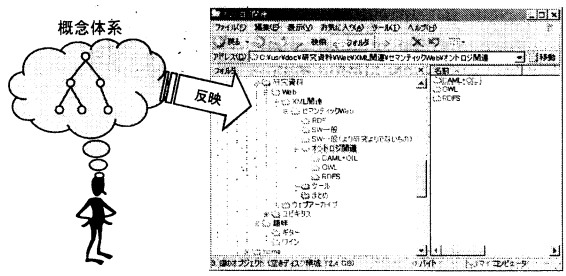

図 3　文書分類のディレクトリ構造の例

図 3 で示したのは，実際の文書分類のディレクトリ構造の一例である．

まず，ある程度のディレクトリ構造が作られるという時点でその人がその分野に関してある程度の数の文書を保有しており，興味があるということがわかる．そして，その構造は，その人なりの文書の分類の仕方であり，それは個人の概念体系の一部を反映しているといえる．

この例の場合，「Web」というディレクトリの下に「XML 関連」と「ウェブアーカイブ」というディレクトリが作成されている．正確に概念として扱えるかどうかは別にして，「Web」という概念と，「XML」や「ウェブアーカイブ」という概念が，上位下位関係を持っていると，直感的に捉えることは可能である．

このような環境から，個人的な概念体系の生成を目指すのだが，その前に認識しておかなくてはならない問題が存在する．それは，

- 個人は必ずしも文書をきちんと分類しない

ということである．これに対する解決法としては，自動分類が考えられるのだが，本稿では，個人が自ら作成した文書の分類のやり方を重要視して概念体系を生

成するため，自動分類で分類された文書群ではあまり意味がない．どのように分類させるか，ということは今後の課題とする

文書がきちんと分類されていたとして，解決すべき問題点として，以下の 3 点が考えられる．

- ディレクトリにつけられている名前が適切ではない可能性がある
- 概念体系と比べると，その階層の深さはせいぜい数階層であり，階層構造としては浅すぎる
- 概念体系では一般的に数十万概念扱うのに対し，文書が分類されているディレクトリはせいぜい数十である

これらの問題点を踏まえた上で，以下のような概念体系を生成する手法を提案する．

(1) ディレクトリのツリー構造をそのまま仮の概念体系のツリー構造とする．

(2) 仮の概念体系の節ごとにおいて，属する文書を DF 値を基にしたクラスタリングを行い，節内で小さなツリー構造を作成する．そのようにして作られた小さなツリーの節を概念 1 つとする．

(3) 概念に対してそこに属する文書の特徴ベクトルを考慮して，代表すると考えられる語彙を発見してそれを概念の見出しとする．さらに，概念の見出しとならなかった語彙で関連の強い語彙に関する保管情報をメタ情報として付加する．

図 4 がそのようにして生成する概念体系の概念図である．

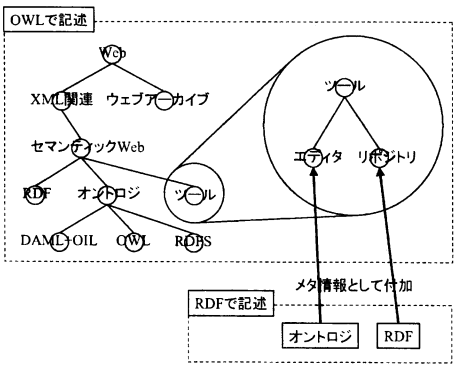

図 4 概念体系の概念図

このように作成される概念体系は，必ずしも is-a 関係のみでは成り立っておらず，ある部分は part-of 関係，ある部分は instance-of 関係になっている．それらは厳密にはクラスに対するプロパティやインスタンスといった形で表現すべきであるが，今回は擬似的な上位下位関係と考え，全ての関係を同等に扱う．

次小節から，細部について述べる．

4.3.2 概念の見出しの決定

先述したように，概念の見出しとしてディレクトリの名前をそのまま用いることには問題がある．そこで，概念の見出しを自動的に決定する手法が必要となる．

まず，ある概念の部分に所属する文書として扱うのは，概念の部分に直接所属する文書とし，下位の概念に含まれる文書は考慮しないこととする．

概念に所属する文書から，語彙の DF 値を取得する．語彙の DF 値は，

$$\text{語彙 w の DF 値} = \frac{1}{\text{全文書数}} \cdot (\text{語彙 w を含む文書数})$$

と定義する．

ストップワードを設定して，一般的な語を取り除くことによって，多くの文書に含まれている語彙の DF 値が高くなることから，DF 値が最も高い語彙をその概念の見出しとすることが可能である．ただし，もし DF 値が最も高い語彙がすでにより上位の概念で使われていれば，その語彙は省くものとする．

4.3.3 概念の細分化

先述したように，概念体系と比べると，ディレクトリの階層構造は非常に浅く，ディレクトリの階層構造を深化させる必要がある．ここでは，ディレクトリの構造は最大限尊重して，仮の概念体系の 1 つの節において，小さな概念体系を発見する．

まず，仮の概念体系の最上位の概念としては，前小節で述べた手法で発見した見出しをもつ概念を置く．

次に，その下位概念の発見であるが，直感的な考え方としては，例えば，DF 値がある程度以上高い語彙 w_1 と w_2 があるとき，それらを含む文書があまり重複していないとすると，w_1 と w_2 がそれぞれ下位概念だと考えられる，というものである．

これを実現する方法として，既存の K-Means のような文書クラスタリングを用いることが考えられ，その場合，文書をあらわす特徴量として，DF 値がある閾値 α 以上の語彙 w_i の TF 値の特徴ベクトルとすることができる．すなわち，$TF(w)$ をある文書に含まれる語彙 w の数とすると，

文書 d の特徴量 $= \{TF(w_1), TF(w_2), ..., TF(w_i)\}$

となる．この特徴量を用いて，K-Means で 2 分割する，ということを繰り返すことで，概念体系を深化させることが可能である．

4.3.4 メタ情報の付加

概念の細分化によって，扱える概念の数はある程度多くなるが，それでも一般的な概念体系と比べるとは

るかに少ない概念しか生成することはできない．実際に概念体系が用いられるときには，語彙のマッチングによるので，ある概念に関連する語彙をメタ情報として付加することで，より多くの語彙から概念のマッチングが可能となる．

このメタ情報の付加によって，より個人差が現れると考えられる．

ある概念に対して付加される語彙は，その概念に所属する文書のから求められる語彙の DF 値と，文内の語彙の共起度から求める．

語彙の DF 値についてはやはりストップワードによって一般的な語を排除する．DF 値がある閾値以上の語彙から，DF 値と共に関連する語としてメタ情報とする．

共起度については，ある語彙 w_i と概念の見出し語 w_h との共起度を

$$\text{語彙 } w_i \text{の共起度} = \frac{1}{w_h\text{が現れる全文数}} \cdot (w_i \text{と} w_h \text{が共に現れる文数})$$

と定義し，この値がある閾値以上であれば，メタ情報とする．

いずれの場合も，概念の見出しに使われているものやより上位の概念の見出しに使われているものを除外する．

5. 個人的な概念体系を用いたウェブ情報検索のパーソナライゼーション

5.1 概　　要

新しい情報を得るために，ウェブ情報検索で検索エンジンを用いることは非常に一般的となっている．例えば，「サッカー」について何か情報を得たいと考えた場合，検索エンジンに対して検索キーワード「サッカー」を入力し，検索結果を得る，といったことがしばしば行われる．そのようなときに利用される検索エンジンで有名なものとしては，Google や Yahoo!が挙げられる．

Google で検索を行った場合は，検索結果は PageRank によって順序付けが行われる．PageRank によって高くランク付けされるページは，他の多くのページからリンクされているページであり，それは，社会的にみて有名なページであると言える．例えば，「サッカー」というキーワードで高いランクになっているページには「日本サッカー協会」などがある．

Yahoo!では，カテゴリや登録サイトで「サッカー」が含まれているものがリストアップされる．登録サイトで結果が得られるような場合，そのページがどのようなカテゴリーに属するかということを知ることができる．

これらの検索結果におけるランク付けやカテゴリ分類は，社会の一般的のものであるということができる．

しかし，そのような一般的な基準によりランク付けは，ユーザにとってのランク付けとは異なっていたり，一般的な分類も，ユーザにとっての物事の分類と異なっていたりする．よって，検索結果の上位には，ユーザにとって無駄な情報がいくつも含まれていると考えられる．

そのような無駄な情報を取り除いた残りの結果をある程度有意な文書であると考え，次に行うことができるのが，その文書の内容の判断である．

ある概念に対して，より概要的な内容であるのか，詳細な内容であるのか，ということが概念体系を使えばある程度判断することが可能である．

また，検索語についても拡張を行うことができる．ある語による検索結果の件数が多すぎる場合や，件数が少ない場合に，それぞれ別の処理が考えられる．

概念体系をさまざまな場面において利用可能であることを，以下で述べる．

5.2 不要情報の除去

まず，不要情報の除去について述べる．

検索エンジンから返される検索結果のうち，

- ユーザにとって既得の情報
- ユーザの興味からずれている情報

という 2 種類の情報はユーザにとって不要な情報である．そこで，これらを取り除く方法について述べる．

5.2.1 既得情報の除去

まず，ユーザにとって既得の情報の除去であるが，これは，すでに保有している文書と，検索で得られたページの特徴ベクトルの類似度を計算することにより，その類似度が非常に高く，ある閾値以上であった場合に，既得であると判断することによって除去可能であると考えられる．

すなわち，すでに持っている文書の特徴ベクトルを F_i，検索結果の文書の特徴ベクトルを R_j とし，類似度をコサイン類似度によって計算すると，

$$Similarity(F_i, R_j) = \frac{F_i \cdot R_j}{|F_i| \cdot |R_j|}$$

となる．この値がある閾値 ω 以上であれば，その文書を検索結果から除去する．

5.2.2 興味からずれている情報の除去

本研究で生成する概念体系は，概念に対して見出しと，メタ情報によって関連のある語彙が付加されており，一種の特徴ベクトルとなっている．そこで，各概

念の特徴ベクトルと，検索結果の文書から得られる特徴ベクトルの類似度を計算して，どの概念の特徴ベクトルとの類似度もある閾値以下であれば，ユーザにとってあまり興味のない文書ということで除去可能である．

5.3 情報の特性判定

同じ事柄について述べた文書でも，

- より概要的な文書
- より詳細なことを説明した文書

という違いは存在する．これは，文書に含まれる概念と検索キーワードとして用いられた語彙から得られる概念との関係によって推測することが可能である．ここでは，以下のような式を用いて概要か詳細かを判断する．

$$\sum(\text{上位概念}\cdot\text{関連度}) - \sum(\text{下位概念}\cdot\text{関連度})$$

上位概念には，メタ情報として付加されている語彙も含め，関連度とはDF値や共起度を元にしてメタ情報に付加されている値のことである．この値が正であればより概要的な情報であり，負であればより詳細な情報であると判断可能である．

5.4 検索キーワードの拡張

検索キーワードの拡張には，検索結果をより絞り込むものと，検索範囲を広げるものと2通り考えられる．基本的にはどちらにおいても検索キーワードに

(1) 検索キーワードが含まれる概念の他の語彙を付加する
(2) 検索キーワードの上位概念の語彙を付加する
(3) 検索キーワードの下位概念の語彙を付加する

ということを行う．

検索結果が大量である場合は，検索結果をより絞り込むためには，既存のキーワードに対してAND条件で上記の拡張を行う．基本的には(1)(2)(3)の順で絞込みを行うことが考えられ，この順に絞込みの力が強くなると考えられる．

検索範囲を広げるためには，OR条件で上記の拡張を行う．この場合，(1)(3)(2)の順で行うのが順当と考えられる．

6. まとめと今後の課題

本稿では，個人がコンピュータ上に保有するさまざまな文書から個人的な概念体系を作成する手法について提案を行った．さらに，それを用いたアプリケーション例として，ウェブ情報検索のパーソナライゼーションの手法について提案を行った．今後は本手法の評価とシステムの実装を行う必要がある．

謝　辞

本研究の一部は，平成16年度科研費特定領域研究(2)「Webの意味構造発見に基づく新しいWeb検索サービス方式に関する研究」(課題番号：16016247，代表：田中克己）および21世紀COEプログラム「知識社会基盤構築のための情報学拠点形成」によるものです．ここに記して謝意を表すものとします．

参 考 文 献

1) E.Adar, D. Karger and L. Stein, "Haystack: Per-User Information Environment", Proc.1999 Conference on Information and Knowledge Management, pp.413-22, 1999.
2) Haystack: http://haystack.lcs.mit.edu/
3) 片山佳則，小櫻文彦，井形伸之，渡部勇，津田宏，セマンティックグループウェアWorkWare++とKnowWho検索への応用，情報処理学会 研究報告「情報学基礎」No.071, 2003.
4) 内野寛治，津田宏，松井くにお，WorkWare:WEBを用いた文書の時間順整理の試み，情報処理学会研究報告「情報学基礎」No.051, 1998.
5) Hiroyuki Sato, Yutaka Abe and Atsushi Kanai, "Hyperclip: a Tool for Gathering and Sharing Meta-Data on Users' Activities by using Peer-toPeer Technology", WWW2002 Workshop on Real world RDF and Semantic Web applications, 2002.
6) 湯川高志，吉田仙，桑原和宏，パーソナル・レポジトリに対するピア・ツー・ピア型協調検索機構の提案，電子情報通信学会 信学技報 AI2001-48, 2001.
7) Semantic Web ホームページ (W3C): http://www.w3.org/2001/sw/
8) W3C Resource Description Framework (RDF) (W3C): http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/
9) OWL Web Ontology Language Reference (W3C): http://www.w3.org/TR/owl-ref/