

# 能動的情報資源の概念を用いた 学術情報利用支援システムのアーキテクチャ

千葉 祐<sup>†</sup> 阿部 亨<sup>††</sup> 木下 哲男<sup>††</sup>

<sup>†</sup> 東北大学 情報科学研究科

<sup>††</sup> 東北大学 情報シナジーセンター

**概要** 近年、各種学術情報の電子化が進み、それらの利用を支援するための検索システムが多数提案・開発されている。しかし、これら既存の検索システムを用いたとしても、所望の学術情報を見付け出すまでには、利用者が多くの作業を行う必要があり、利用者へ課せられる負担は依然として大きいままである。これに対し、本稿では、利用者が学術情報を検索する際に行う一連の作業を包括的に支援するシステムのアーキテクチャを提案する。提案するアーキテクチャでは、「能動的情報資源 (AIR)」の概念を用い各学術情報を構造化することで、学術情報自体に能動性・自律性を与えている。これにより、従来利用者へ委ねられていた作業の一部を学術情報側に代行させ、検索に要する作業全体の効果的な支援を図る。

## An Architecture for Academic Information Retrieval Support Systems Based on Active Information Resource

Yuu CHIBA<sup>†</sup>, Toru ABE<sup>††</sup>, and Tetsuo KINOSHITA<sup>††</sup>

<sup>†</sup> Graduate School of Information Sciences, Tohoku University

<sup>††</sup> Information Synergy Center, Tohoku University

**Abstract** Digital academic information circulates widely in recent years. To support the use of it, a lot of academic information retrieval systems have been proposed and developed. However, even if users use those systems, many tasks are left for the users to find the needed information, and thus the users' loads to retrieve academic information are still heavy. To overcome this problem, we propose a novel architecture for academic information retrieval support systems. In the proposed architecture, by employing the concept of active information resource (AIR), each piece of academic information is structured to make it active and autonomous. Consequently, academic information itself can assist the users actively with their retrieval tasks.

### 1. ま え が き

近年、学術論文、貴重図書、史料など、各種学術情報の電子化が進んでおり、これら電子化された学術情報の活用を支援するために、キーワード型やブラウジング型の検索システムが数多く提案・開発され、電子図書館等で利用されている。しかし、現状では、それらの検索システムを用いたとしても、所望の学術情報を見付け出すまでには、情報の取捨選択や、取得した情報を利用した再検索など、多くの作業を行う必要があり、学術情報を

利用する際に利用者へ課せられる負担は依然として大きいままである。

これに対して、本稿では、利用者が学術情報を検索する際に行う一連の作業を包括的に支援するシステムの構築を目指し、それを実現するためのアーキテクチャを提案する。提案するアーキテクチャでは、「能動的情報資源 (Active Information Resource: AIR)」 [1] の概念を用い各学術情報を構造化することで、学術情報自体に能動性・自律性を与えている。これにより、従来利用者へ委ねられていた作業（特に、学術情報の検索に特徴的な「発

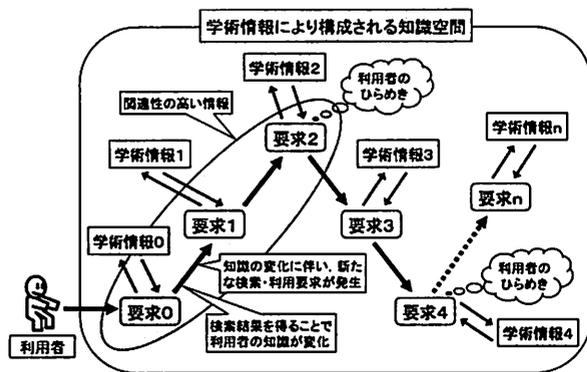


図1 発展検索

展検索)の一部を学術情報側に代行させ、検索に要する作業全体の効率的な支援の実現を図る。

## 2. 学術情報の検索

本章では、学術情報の検索において重要な役割を果たす「発展検索」について説明し、次に、「発展検索」支援の観点から、従来の検索システムの問題を述べる。

### 2.1 発展検索

一般に、利用者が何らかの情報を検索する場合、検索が一回で済むことは稀であり、多くの場合は、図1に示すように、最終的な結果を得るまでに検索を複数回繰り返すことになる。これは、検索で得られた情報を取得することで利用者の知識が変化するためであると考えられる。この結果、利用者に新しいニーズが生まれ、このニーズを満たすために、新たな検索要求の生成、および、新たな検索要求に基づく再検索が、利用者の知識の変化が収束するまで繰り返されることになる。Batesは、要求が変化しながら繰り返されるこのような検索過程を「発展検索 (evolving search)」と名付けている [2]。

学術情報に対する検索は、論文調査をはじめとする研究活動において行われることが多く、検索時の利用者の知識の変化も非常に活発であるため、新たな検索要求の生成、再検索を行う機会が多い。従って、利用者による学術情報の活用を支援するためには、「発展検索」を支援する機構の実現が非常に効果的であると考えられる。

### 2.2 従来の検索システム

現在、キーワードに基づく検索に加え、Dublin Coreに代表されるメタデータ (書誌情報など) に基づく検索を利用したシステムが提供されている。

このようなシステムに対し、シソーラス [3] を用いることで検索要求を拡張したり、オントロジー [4] を用いて概念間の既知の関係を検索処理に導入する研究が進められている。また、利用者の興味等を記録したユーザプロフィールを検索履歴に基づき随時更新し、そこから得られる統計量を用いることで、利用者の興味の変化 (一種

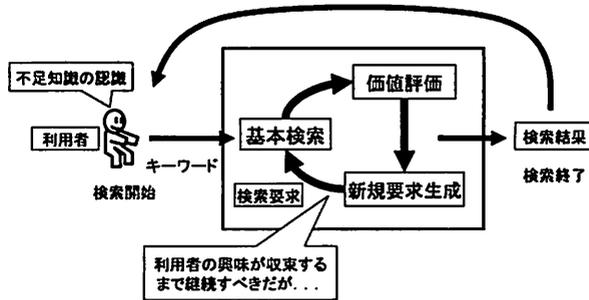


図2 学術情報利用支援システムでの「発展検索」の代行

の知識の変化) を検索処理へ反映する手法も多数提案されている [5], [6]。

これらのアプローチに基づくシステムにより、ある一つの検索要求に対する検索精度を向上させることが可能となる。しかし、そのようなシステムを用いたとしても、ある検索の結果として得られた情報の詳細、あるいは、それにより生じた利用者の知識の変化を、直後の検索要求へ反映することは困難であり、「発展検索」のための諸作業は、依然として利用者に委ねられたままとなる。

## 3. 能動的情報資源を用いた「発展検索」支援

本章では、まず、学術情報利用の効果的支援を図るため「発展検索」をシステム側で代行する手法を提案し、次に、「能動的情報資源」の枠組みに基づく提案手法の実現法について説明する。

### 3.1 「発展検索」の支援手法

本稿では、「発展検索」をシステム側で代行させるために、図2に示すように、次の3つのプロセスをシステム上で順に繰り返す手法を提案する。

- **基本検索**  
各学術情報が持つキーワードやメタデータと検索要求とを比較し、検索要求に合致する学術情報の検索を行う。
- **価値評価**  
検索要求に合致した学術情報中、関連性の高い情報間でグループを構成し、その中で、利用者側の検索履歴や学術情報側の利用履歴などにより、検索結果の順位付け・絞り込み等の取捨選択を行う。
- **新規要求生成**  
選択された学術情報から得られるデータ、および、利用者の知識の変化等に基づき、新たな検索要求を生成する。

これらのプロセスをシステム側で繰り返す場合、基本検索で得られた全ての学術情報を用い新たな検索要求を

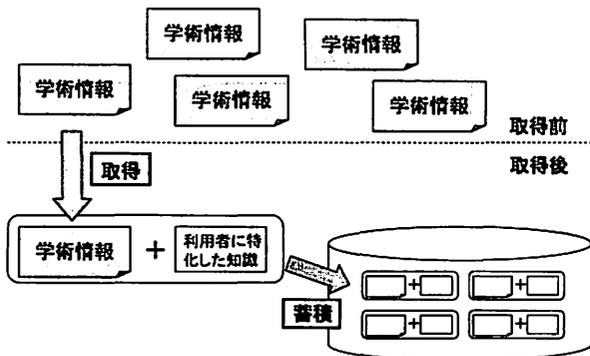


図3 学術情報取得時の知識付加

生成すると、その数は膨大なものとなり効率の低下が予想される。従って、基本検索の結果に対し価値評価を行うことで、利用者にとって価値の高い学術情報のみを選択でき、それらを基に実行される「発展検索」が有効であるといえる。

検索要求の生成に際しては、価値評価により選択された学術情報から得られる各種のデータ（学術情報に予め付与されたメタデータ等）が利用される。従って、書誌情報や引用関係が明確な学術情報（例えば、学術論文など）からは、検索要求の新規生成に有用なデータを取得しやすく、提案手法が特に有効であると考えられる。また、利用者が学術情報を取得した場合、図3に示すように、利用者によるメモや検索時のコンテキストなど利用者に特化した知識を学術情報に付加し、利用者用の記憶領域に蓄積しておく。これら取得済みの学術情報、およびそれに付加された知識を各プロセスで用いることにより、利用者の知識の変化を新たな検索要求の生成へ反映させる。

ここで述べた3つのプロセスを反復することで、過去に取得された学術情報とそれに付加された知識、および、順次獲得される学術情報に基づいた検索が繰り返される。これにより、「発展検索」において関連性の高い検索要求が連続して生成される部分（直観的には、図1中「発展検索」が描く軌跡の直線部分に相当すると考えられる）、例えば、参考文献やキーワードを手掛りに学術情報を芋蔓式に辿る作業などをシステムに代行させることができ、「発展検索」の中でも煩雑・類型的な部分の自動化が可能になると考えられる。

### 3.2 能動的情報資源 (AIR)

前節で述べたように、提案手法では、基本検索の結果に対し価値評価が実行され、利用者にとって価値の高い学術情報を選択される度に、新たな検索要求が生成される。その際、効果の高い検索要求を生成するためには、選択された学術情報毎に、その特徴に応じた検索要求を生成する必要がある。このような機構の実現には、選択さ

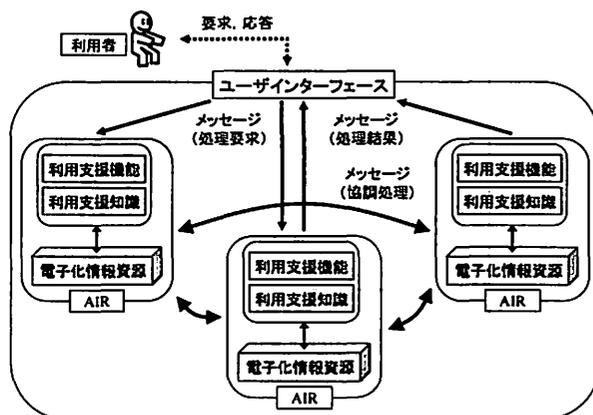


図4 AIR の概念図

れた各学術情報自身が能動的に検索要求を生成するシステムアーキテクチャが有効であると考えられる。そこで、提案手法の実現にあたり、能動的情報資源 (AIR) の概念を学術情報に適用することで、学術情報に能動性を持たせ、各学術情報が自律的に検索要求を生成するアーキテクチャの実現を図る。

AIRは、ネットワーク上に分散した情報資源（電子化された各種データ）の有効活用を図る手法である。AIRの概念を図4に示す。この手法では、利用支援知識（情報資源に関するメタデータや情報資源を操作するための知識など）と利用支援機能（情報資源を操作するための機能）とを情報資源に付加したエージェントとして各AIRが構成される（情報資源のAIR化）。これにより、情報資源自体が能動性・自律性を持つことになり、今まで利用者へ全面的に委ねられていた情報資源に対する操作の一部を情報資源自身が代行できるようになる。例えば、利用者（あるいは他のAIR）からの処理要求メッセージを受信したAIRは、自身が有する利用支援知識・利用支援機能を用い、自身が管理する情報資源に対する操作を代行し、その結果を応答メッセージとして送信する。また、各AIRは、単独で活動するだけでなく、他のAIRと協調し連携・組織化を行うことで、より複雑・柔軟な処理を能動的・自律的に代行することも可能となる。

学術情報をAIR化し、各学術情報に自律性を持たせることで、基本検索、価値評価、新規要求生成を学術情報側に能動的に代行させることが可能となる。この結果、検索の過程で得られる種々の情報を、適切な方法で適時処理に反映させることができ、「発展検索」を効果的に代行できるシステムの実現が期待できる。

## 4. 学術情報利用支援システムのアーキテクチャ

本章では、前章で述べた「発展検索」支援手法を実現

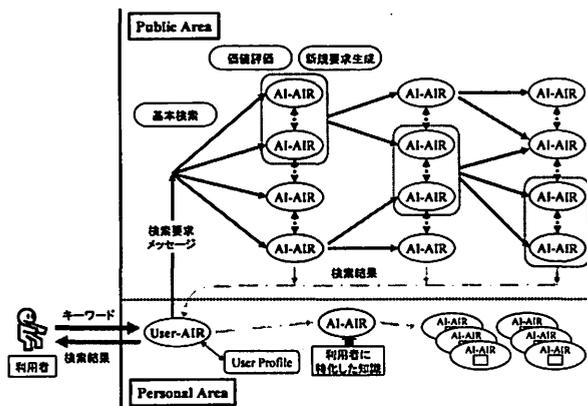


図 5 提案するシステムアーキテクチャ

するためのシステムアーキテクチャを提案し、次に、このアーキテクチャを構成する各 AIR の設計について説明する。

#### 4.1 システムのアーキテクチャおよび動作

提案するシステムアーキテクチャは、図 5 に示すように 2 つの領域から構成される。Public Area は、システムの利用者が共通して利用する領域であり、ここには、検索対象となる AIR 化された学術情報 (AI-AIR) が存在する。基本検索、価値評価、新規要求生成のプロセスは Public Area 内で実行される。これに対して Personal Area は、システムの利用者各々に 1 つずつ用意され、各自のユーザプロフィールや、システムを用いて検索した AI-AIR 等が蓄積される領域である。Personal Area には、User-AIR (利用者とのインターフェース、Public Area と Personal Area とのインターフェース) が存在しており、User-AIR は、利用者からの入力の取得、利用者への検索結果の提示、ユーザプロフィールや獲得済みの AI-AIR など Personal Area 内に存在する情報を Public Area 内の AI-AIR にメッセージとして提供する役割を持っている。

提案するシステムアーキテクチャでは、まず、利用者の入力したキーワードを User-AIR が受信し、1 回目の基本検索の検索要求メッセージが生成され、Public Area 内の各 AI-AIR に送信される。各 AI-AIR は、検索要求メッセージと自身の利用支援知識とを比較し、自身の管理する学術情報が要求に合致するか否かを判断する (基本検索)。要求に合致すると判断した AI-AIR は、関連のある AI-AIR 同士でグループ化を行い、各グループ内で学術情報が持つ価値を比較し重要と判断されるものをいくつか選択する (価値評価)。選択された AI-AIR は、新たに検索要求を生成し、他の AI-AIR へメッセージとして送信することで再度基本検索を実行する (新規要求生成)。

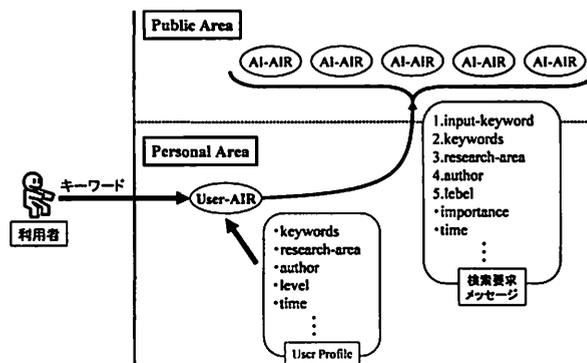


図 6 検索要求メッセージの生成例

この基本検索、価値評価、新規要求生成のサイクルは、検索の終了条件を満たすか、新しい検索要求が生成されなくなるまで行われる。各サイクルにおいて獲得された AI-AIR は、User-AIR を介し、利用者に提示され、それと同時に Personal Area に蓄積される。各プロセスにおいて、Public Area 内の AI-AIR が Personal Area 内の情報 (ユーザプロフィール、獲得済みの AI-AIR) を必要とした場合、User-AIR に要求メッセージを送信し、User-AIR を介して必要な情報を取得する。

#### 4.2 User-AIR の設計

User-AIR は、利用者とのインターフェース、および、Personal Area と Public Area 間のインターフェースとしての役割を担う。利用支援知識としては、表 1 に示すようなユーザプロフィールの内容を持ち、利用支援機能としては以下の機能を持つ。

##### 1 ユーザプロフィールの管理機能

- User-AIR は、利用者毎にユーザプロフィールを作成し、利用者の検索履歴等によりその内容を更新する機能を持つ。また、Public Area 内の AI-AIR からユーザプロフィール参照要求メッセージを受けた場合、必要な情報をユーザプロフィールから抽出し、当該 AI-AIR に送信する。

##### 2 利用者とのインターフェース機能

- User-AIR は、利用者から入力されたキーワードとユーザプロフィールの項目から検索要求メッ

知識の種類	名称
研究分野名	research-area
研究分野におけるキーワード	keywords
研究分野に対する習熟度	level
学術情報にアクセスする権限	authority
一度の検索に費す時間の上限	time
一度の検索で表示する情報の上限	quantity

表 1 User-AIR の利用支援知識

セージを生成する機能を持つ。具体的には、図 6 に示すように、ユーザプロフィールから必要な項目を抽出し、入力されたキーワードに付加する。また、検索要求が持つ重要性として importance の値に初期値 1.00 を設定し、検索要求中の項目の優先順位を決定する。生成された検索要求メッセージは、Public Area 内の AI-AIR に送信される。

- User-AIR は、基本検索により Public Area で獲得された AI-AIR を受信し、それらを利用者に提示する機能を持つ。また、受信した AI-AIR を、Personal Area に蓄積する機能を持つ。
- 獲得された AI-AIR へ利用者がコメントを付加する場合、User-AIR は、利用者からの入力を受信し、当該 AI-AIR の新しい利用支援知識（利用者に特化した知識）として付加する機能を持つ。

### 3 Public Area と Personal Area とのインターフェース機能

- Public Area 内の AI-AIR は、基本検索、価値評価、新規要求生成を行う際、必要に応じて User-AIR へメッセージを送信し、Personal Area 内の情報を要求する。このメッセージを受信した User-AIR は、Personal Area 内のユーザプロフィールおよび獲得済み AI-AIR から情報を抽出し、これらを必要としている Public Area 内の AI-AIR に応答する機能を持つ。
- User-AIR は、検索結果として Public Area から獲得された AI-AIR に、検索時のコンテキスト（検索要求に含まれる情報等）を選別し、新しい利用支援知識（利用者に特化した知識）として付加する機能を持つ。

#### 4.3 AI-AIR の設計

AI-AIR は、利用者による「発展検索」を支援するた

知識の種類	名称
AIR の名前, 作成者	title, creator
AIR の属する研究分野名	research-area
AIR の持つ特徴 (キーワード)	keywords
参照している AIR	reference
参照されている AIR	referenced-by
類似している AIR	similar
参照している AIR の優先度	priority
類似した AIR の類似度	similarity
他の AIR からのアクセス履歴	access
AIR にアクセスする権限	authority

表 2 AI-AIR の利用支援知識

めに、基本検索、価値評価、新規要求生成のサイクルを能動的に代行する。そのため、AI-AIR は、表 2 に示す利用支援知識と以下の利用支援機能を持つ。

#### 1 基本検索機能

- 基本検索は、Public Area 内の AI-AIR が、User-AIR もしくは他の AI-AIR からの検索要求メッセージを受信した場合に開始される。AI-AIR は、メッセージの項目と自身の利用支援知識の項目とを比較し、自身の管理する学術情報が要求に合致しているか否か判定する機能を持つ。また、検索要求メッセージを受信可能な（学術情報が要求に合致する）場合、応答メッセージを検索要求メッセージ送信元の AIR に送信する機能を持つ。
- AI-AIR は、検索要求メッセージを受信可能と判定した場合に、自身のアクセス履歴（利用支援知識の一部）を更新する機能を持つ。

#### 2 価値評価機能

- AI-AIR は、検索要求メッセージを受信した AI-AIR 間で、相互の関連に応じてグループを構成する機能を持つ。グループの構成法は、まず、検索要求メッセージの送信元の AI-AIR が、応答メッセージを返した AI-AIR 間の項目の関連により、グループの分類とマネージャとなる AI-AIR の決定を行い、それらを応答メッセージを返した AI-AIR に送信する。次に、これを受信した AI-AIR が、自身が属するグループのマネージャにグループ化受理メッセージを送信する。マネージャは、グループ化受理メッセージを返した AI-AIR を 1 つのグループとして認識する。
- AI-AIR は、自身の管理する学術情報が利用者に対し有する価値を算出する機能を持つ。この価値は、Personal Area 内の情報（ユーザプロフィール、獲得済みの AI-AIR の利用支援知識）と自身の利用支援知識の項目とを比較することで算出される。
- AI-AIR は、自身の価値を他の AI-AIR の価値と比較する機能を持つ。各グループでは、各 AI-AIR の持つ価値が比較され、価値の高い AI-AIR が選択される。各グループのマネージャは、選択された AI-AIR を価値評価の結果として User-AIR に送信する。

#### 3 新規要求生成機能

- 価値評価によって選択された AI-AIR は、基本検索の際に受信した検索要求メッセージを書き

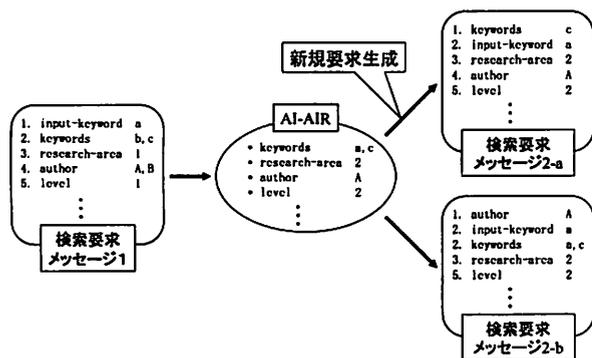


図 7 新規要求生成の例

換え、新たな検索要求メッセージを生成する機能を持つ。書き換えは、AI-AIR の利用支援知識と Personal Area 内の情報を用いて、検索要求メッセージ中の項目の優先順位と項目の値を換えることで実現する。書き換え規則は種々考えられるが、例えば、図 7 に示すように、受信した検索要求メッセージと利用支援知識とで一致した項目の優先順位を 1 とし、他の項目を利用支援知識の項目に書き換える等がある。

- 検索要求メッセージと自身の利用支援知識に一致している項目がある場合は、一致している項目の優先順位等から、importance の値を再計算する機能を持つ。この importance の値が、新たに作成された検索要求メッセージに付加される。
- AI-AIR は、新規要求生成によって作成した新たな検索要求メッセージを他の AI-AIR に送信することで、基本検索を開始する機能を持つ。なお、新たな検索要求メッセージ内の importance の値がしきい値を下回る場合、メッセージは送信されずに破棄される。基本検索を開始した AI-AIR は、User-AIR に検索の継続を告げるメッセージを送信する。

## 5. まとめ

本稿では、利用者による学術情報の活用を支援するために、学術情報の検索において重要な役割を果たす「発展検索」に着目し、これをシステム側で代行する手法を提案した。提案手法では、複数のプロセス（基本検索、価値評価、新規要求生成）を繰り返すことで、参考文献やキーワードを手掛りに学術情報を芋蔓式に辿る作業など、「発展検索」の中でも煩雑・類型的な部分の自動化を図っている。

さらに、この提案手法を効果的に実現するためのシステムアーキテクチャの提案を行った。提案したアーキテクチャでは、各学術情報を AIR として構成することによ

り、各プロセスを学術情報自身に能動的・自律的に代行させている。このため、学術情報の検索に際し、従来利用者に課せられていた負担の大幅な軽減が期待できる。

現在、本稿では設計が不十分であった

- ユーザプロフィールや獲得済の AI-AIR から、利用者の知識の変化を推測し、検索処理に効果的に反映する手法
- 利用者の活動を効果的に支援するための検索結果の表示法

に関し、具体的なアルゴリズムの検討を進めている。また、これと並行して、エージェントフレームワーク ADIPS / DASH [7] を用い、提案アーキテクチャに基づいたプロトタイプシステムの実装を進めている。このプロトタイプシステムは、各学術情報がエージェント型の AIR として実装されたマルチエージェントシステムとして構成されている。今後、このプロトタイプシステムを用いて評価実験を行い、提案手法およびアーキテクチャの有効性を検証する予定である。

## 文 献

- [1] 木下, “分散情報資源活用の一手法 — 能動的情報資源の設計—,” 信学技報, AI99-54, pp.13-19, 1999.
- [2] M. J. Bates, “The design of browsing and berrypicking techniques for the online search interface,” Online Review, Vol.13, No.5, pp.407-424, 1989.
- [3] 大山 他, “大規模学術情報データベースに適した情報検索システムの開発,” 信学論, Vol.J84-D-I, No.6, pp.658-670, 2001.
- [4] P. Weinstein et al., “Agent-based digital libraries: Decentralization and coordination,” IEEE Commun. Mag, Vol.37, No.1, pp.110-115, 1999.
- [5] F. Liu et al., “Personalized Web Search for Improving Retrieval Effectiveness,” IEEE Trans. Knowledge Data Eng, Vol.16, No.1, pp28-40, 2004.
- [6] 土方 他, “情報推薦・情報フィルタリングのためのユーザプロファイリング技術,” 人工知能学会誌, Vol.19, No.3, pp.365-372, 2004.
- [7] “Dash — Distributed Agent System based on Hybrid architecture —,” <http://www.agent-town.com/dash/index.html>.