

# ユビキタス知識環境と知識メディア

## —知識メディアを基盤とする次世代ITの研究—

田中 謙 北海道大学情報科学研究科コンピュータサイエンス専攻 tanaka@meme.hokudai.ac.jp

本拠点ではWeb上の知識のみならず、環境にあまねく埋め込まれた知識担体が動的に生成する知識をも含むユビキタス知識環境の基盤ハードウェア技術と、この環境において、動的に変化する種々の目的に応じて適切な知識を即座に抽出し、目的に応じたフェデレーション(連合体)を即時に形成するソフトウェア技術の研究開発を、知識メディア、量子ナノエレクトロニクス、知的通信の分野の研究者が密に連携することで遂行している。本稿では、ユビキタス知識環境から、目的に応じた知識をいかに探し出し、再利用可能な形式へといかに抽出し、それらの間に自在な機能連携をいかに実現するかという、知識探索、知識抽出、知識フェデレーションの課題に焦点を絞って、これまでの研究成果を報告する。

### ユビキタス知識環境

本プロジェクトでは、来たるべきユビキタス時代の情報環境においては、コンピュータやネットワークといったハードウェアが環境にあまねく存在するだけでなく、その結果として動的に生成され変化する「知識」が環境に遍在するようになると考え、「ユビキタス知識」のビジョンを掲げている。ここでいう知識とは、データやマルチメディア・コンテンツを含んだ文書情報を始め、アプリケーション・ツールや、サーバによって提供される各種サービスをすべて含む。

このような知識は環境の中では何らかの情報機器に乗っかって存在する。このような情報機器を知識担体と呼ぶ。知識担体には、サーバ、PC、組み込み型コンピュータ、PDA、携帯電話、タグチップなどがすでに存在する。今後、スマートチップやスマートダストと呼ばれるような、センサと処理機能と通信機能を一体化した微小知識担体の実現も期待されている。知識担体が微小化されれば、人工物のみならず、生物も含むあらゆる物に埋め込んで、対象物の状態や状況に応じた知識の創成や送受信までを行わせるこ

とができる。このような環境をユビキタス知識環境と呼ぶ。

ユビキタス知識環境は、今日のWebの世界を拡張したものになる。Webの世界では、知識担体として、サーバ、PC、組み込み型コンピュータ、PDA、携帯電話、タグチップなどが連携して知識のネットワークを形成している。工業製品や書籍、衣料品に組み込まれたタグチップにURLアドレスを格納し、この情報をPDAが読み取って、対応するWebページをアクセスすることも実現されている。ユビキタス知識環境では、これに加えて、書籍や、棚、机などの有体物に埋め込まれた各種サイズの知識担体が一過性のネットワークを形成し、相互に知識の連携を行うことにより、「書籍を書棚のあるコーナーから出して、机上に、他の書籍とともに並べた」といった情報を、個々の知識担体内に状況変化のログ情報として蓄積することができるようになる。後日、この書籍を取り出したときに、利用履歴が読み取れたり、あるページを開いた時、「どこで誰がこのページを開き、そのとき周りではどの書籍のどのページが開かれていたか」といった情報が、読み取れるようになる。処理機能やセンサ機能を持った知識担体を有体物に埋め込む場合には、有体物の関連情報やサービス機能を知識担体自体に持たせたり、環境の変化をセンスして提供する情報やサービスの内容を変化させることも可能になる。Webから関連情報やサービスを得る場合にも、知識担体の状態の変化に応じて、アクセスするサイトや要求内容を変えることが可能となる。

### ユビキタス知識ネットワークの要素技術

微小知識担体は数mm角のものから、将来的には数 $0\mu\text{m}$ 角( $1\mu\text{m}=10^{-6}\text{m}$ )のものまで、階層的に用意される必要がある。サイズの違いに応じて、その処理機能、センサ機能、通信機能も異なる。通信距離は知識担体のアンテナのサイズが小さくなるに応じて、よりサイズの大きい知識担体との間の近傍・近接通信とならざるを得ない。

知識担体間の通信により形成されるネットワークは、Web上の知識ネットワーク構造に加えて、有体物の移動や利用のされ方、他の有体物との動的に変化する近接関係などに応じて、動的に定義される一過性のネットワーク構造を含む。動的に定義される知識担体間のネットワークをフェデレーション・ネットワークといい、その上で定義される知識担体間の機能連携をフェデレーションという。フェデレーションとは、複数のオブジェクトが、相互に協力し合って何らかのタスクを遂行するために一時的に形成される連合体、ないしはその形成を意味する。統合(インテグレーション)と異なり、フェデレーションの場合には、もともと他者との連携を意図して作製されていないオブジェクトの間で一時的な機能連携が形成される。

モバイル端末間のピアツー・ピア通信や、モバイル端末と情報機器間のブルートゥース通信を用いて動的に構成されるアドホック・ネットワークを加えて、Webと、アドホック・ネットワークと、フェデレーション・ネットワークが、ユビキタス知識環境における知識ネットワークを形成する。これをユビキタス知識ネットワークと呼ぶ。

本COE拠点は、このようなユビキタス知識環境を実現するために、それぞれソフトウェア、半導体デバイス、通信を専門とする3グループが密に連携しあって研究教育を遂行している(表-1)。知識メディア・グループは、ユビキタス知識の中から目的に合った知識を自在に「探し出し、相互に「繋ぐ」ことにより、知識の論理的なフェデレーションを自在に形成し新しい知識を創成したり、複雑なタスクを遂行させることができるソフトウェア技術を確立することを目指している。量子ナノエレクトロニクスグループは、単電子デバイス技術を基盤に、センサと通信とプロセッサ機能を持った低消費電力微小知識担体の開発基盤要素技術の確立を目指している。知的通信グループは、低消費電力微小知識担体のための近傍近接知的通信技術と、これとインターネットや携帯電話網との間を繋ぐネットワーク階層の種々の構成要素となる無線および有線の通信技術の確立を目指している。

以下、本稿では、知識メディア・グループの成果を中心にこれまでの主たる研究成果を紹介する。

## 知識メディアを基盤としたユビキタス知識の連携

### ユビキタス知識の自在な連携を実現するフェデレーション

ユビキタス知識のアクセスには、ローカル・アクセスとグローバル・アクセスが考えられる。ローカル・アクセスは近傍のユビキタス知識のみを近傍・近接通信機能を持った携帯電話やPDAを用いてアクセスする。グロー

バル・アクセスは広範囲に分散したユビキタス知識をアクセスする。後者の場合、ユビキタス知識はネットワーク階層によってWeb上に集約され、Web文書中のリスト、Web・アプリケーション、Web・サービスのいずれかの形式でアクセスされる。

ユビキタス知識の連携は、ローカル・アクセスかグローバル・アクセスが可能な知識群の中から、任意の知識群を抽出し、これらを相互に関連付けて機能連携させる技術である。ユビキタス知識のフェデレーションには、個々のユーザが動的に変化する目的に応じてアドホックに定義するものと、分散する膨大な数のサービスや情報が自律的に協調動作を定義するものとの2種類がある。本プロジェクトは、前半3年間で、前者のアドホック・フェデレーション技術を、後半2年間で後者のオートノマス・フェデレーション技術の確立を目指している。

### 基盤技術としての知識メディア

知識メディアは、情報知財の再編集と再流通のメディア技術として1987年より著者が研究開発を続けてきたもので、2次元表現メディアのIntelligentPadと3次元表現のIntelligentBoxがある<sup>1)</sup>。IntelligentPadはパッドと呼ばれるカード状の視覚表現を持った機能部品をディスプレイ上で貼り合わせることにより部品合成が定義される。各パッドはスロットと呼ばれる入出力ポートをいくつか持ち、パッドAをパッドBの上に貼り合わせる際に、Bの1つのスロットを選んでAをこのスロットに結合することができる。IntelligentBoxはIntelligentPadのアイデアを3次元表現へと拡張したもので、3次元視覚表現を持ったボックスと呼ばれる部品を用いる。知識メディアは、Web技術の発展と融合して、現在は、Web上のコンテンツとサービスに対する再編集・再流通メディアへと発展している<sup>1), 2)</sup>。

### Web・ビューの関係演算を用いたアドホック・フェデレーション

本プロジェクトでは、前半の2004年度までに、知識メディア技術を基盤とした、ユビキタス知識のローカル・アクセス技術と、グローバル・アクセスにおけるアドホック・フェデレーション技術を確立した。

ローカル・アクセスに関しては、IntelligentPadを分散オブジェクトシステムへと拡張し、ネットワークで繋がった異なるコンピュータ上の2つのパッドの間でhttpプロトコルを用いたスロット結合連携を可能にした。これにより、タグへの読み書きパッドと分析結果の表示パッドをPDA上に用意し、データ解析を行うネットワーク上のサーバのプロキシの役目をするパッドを利用して、これとPDA上の2枚のパッドとをネットワークを介し

ふりがな（ローマ字） 氏名（年齢）	所属専攻・ 職名	役割分担
<b>知識メディアと量子ナノエレクトロニクスへの応用</b>		
(拠点リーダー) 田中 讓 (54)	A・教授	研究統括：連携・探索技術／量子ナノ研究支援環境
岸浪建史 (60)	B・教授	ユビキタス知識技術の大規模工業製品の設計・生産・利用・保守への適用
吉岡真治 (35)	A・助教授	ユビキタス知識メディア技術の大規模システムへの適用
原口 誠 (50)	A・教授	ユビキタス知識探索／知識処理・自己組織化量子集積回路アーキテクチャ
有村博紀 (38)	A・教授	ユビキタス知識の探索発見アルゴリズム／大規模システムへの応用
ZEUGMANN, Thomas (48)	A・教授	ユビキタス知識の探索発見アルゴリズム／大規模システムへの応用
佐藤義治 (58)	A・教授	統計的大規模データ解析に基づくネットワーク上の知識探索
工藤峰一 (45)	A・教授	大規模知識探索と分類に関する研究
雨宮好仁 (56)	C・教授	知識ネットワーク処理・自己組織化量子集積回路アーキテクチャ
<b>量子ナノエレクトロニクス</b>		
(サブリーダー) 長谷川英機 (62)	C・教授	高密度高温動作量子集積回路の実現と IQ チップへの応用
福井孝志 (53)	C・教授	新アーキテクチャに基づく量子集積構造の形成
山本真史 (53)	C・教授	知識処理量子集積回路の高密度集積化・高温動作化技術／大容量スピンメモリデバイス基礎研究
酒井洋輔 (59)	C・教授	プラズマと PLD を用いた CNT, 非晶質 DLC, a-C:F 堆積技術の確立
橋詰保 (47)	E・教授	窒化物半導体ナノ構造結晶成長・表面制御と IQ チップデバイス応用
末岡和久 (38)	C・教授	サブミクロンレベル・スピンメモリの基礎研究
<b>知識担体の微小化と近傍・近接通信技術</b>		
(サブリーダー) 小柴正則 (55)	D・教授	近傍・近接通信と超高速フォトニックネットワーク
小川恭孝 (54)	D・教授	IQ チップ対応近傍・近接通信とマルチアンテナによる高度信号処理
本間利久 (56)	B・教授	IQ チップ対応近傍・近接通信系の数値電磁界解析とその設計最適化
山本 強 (50)	D・教授	アドホック NW とフェデレーション NW のプロトコル
宮永喜一 (47)	D・教授	シリコンベース・スマートチップの試作／知識処理・自己組織化量子集積回路アーキテクチャ

- 所属 A: 情報科学研究科コンピュータサイエンス専攻  
 B: 情報科学研究科システム情報科学専攻  
 C: 情報科学研究科情報エレクトロニクス専攻  
 D: 情報科学研究科メディアネットワーク専攻  
 E: 量子集積エレクトロニクス研究センター

表-1 推進担当者一覧(平成16年4月現在)

てアドホックにスロット結合させることが可能になった。複数のタグの読み取り値をPDAからサーバに送って分析処理し、その結果をPDA上のパッドで受けて表示させるといったタスクの実行が可能になる。ユビキタス知識のローカル・アクセスを、インターネット上の任意のパッド・アプリケーションやサービスと自在に連携させるアドホック・フェデレーション技術が確立された。

ユビキタス知識のグローバル・アクセスでは、対象となる知識がネットワーク階層を介して、Webページ中のリストか、Web・アプリケーションか、Web・サービスのいずれかの形式で提示されていると仮定してよい。

これらのフェデレーションをアドホックに構築するには、知識抽出技術と知識連携技術の確立を要する。Web上に提示されるユビキタス知識は膨大な数にのぼり、種類ごとに定まった属性を持っていると想定できる。属性のいくつかは特定知識を選び出すのに用いられ、別の属性は刻々更新されるセンサの値に対応する。これらの膨大な数のユビキタス知識は、種類ごとにデータベースにおける関係表と同様な論理的ビューを持つと考えることができる。この関係表形式のビューをWeb・ビューと呼ぶ。知識抽出は、Webページ中のリストやWeb・アプリケーション、Web・サービスからWeb・ビューを定義する



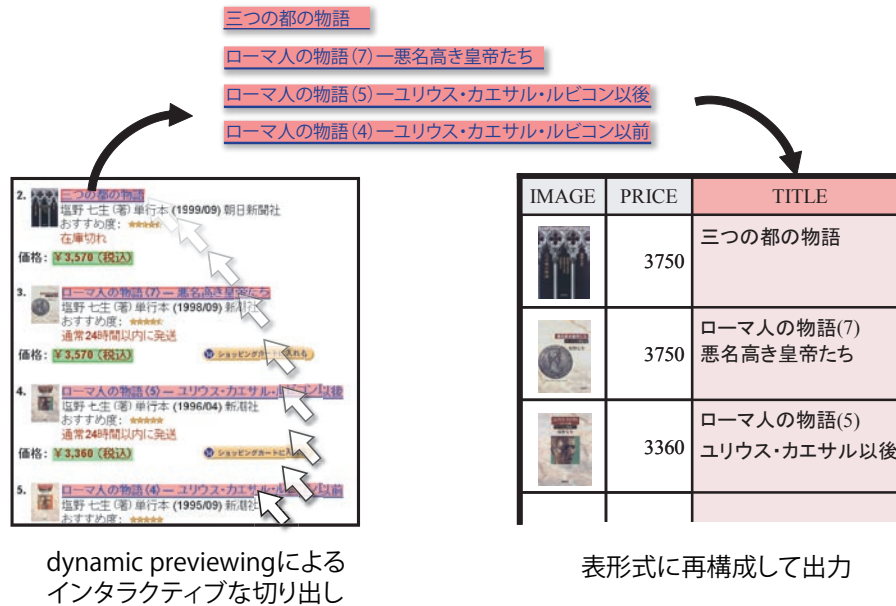


図-1 インタラクティブ・Web・ラッパーを用いたWeb・ビューの抽出

ことにほかならない。複数のWeb・ビューの連携は関係演算を用いて定義することができる。

Webページ文書からのWeb・ビューの抽出定義に関しては、インタラクティブなWebラッパー技術を確立した(図-1)。Webブラウザ・パッドは、マウスを移動すると抽出可能項目を次々と赤枠で囲んでマークする。所望の項目がマークされたときにマウスボタン操作でこれを指定することができる。まず、抽出したい属性の値の第一候補を指定する。この状態でマウスを移動すると、カーソルの直下の項目を第2候補としたときに、2つの候補から一般化操作によって類推される他の候補項目がハイライトされる。抽出したい項目がすべて過不足なくハイライトされたときにマウスボタン操作で抽出指示を行うと、Web・ビュー・パッドが生成され、この属性に対応したコラムがこのパッド内に抽出される。このコラムに任意の属性名を付けることができる。別の属性についても同様の操作を行うことにより、異なるコラムが次々と抽出され、Web・ビューが定義できる。Web・ビュー・パッドは内部的にはこのWeb・ビューを抽出するルールのみを保持し、評価が必要になったときに始めて、対応するWebページをアクセスして表データの抽出を行う。このパッドは各属性に対応するスロットを持つ。

Web・サービスに対しては、WSDL (Web Service Definition Language) を用いたインタフェース記述を自動解析し、スロットとして定義可能な入出力リストをユーザに提示し、ユーザが選んだポートのみをスロットとして持つパッドを自動生成する技術を確立した。このパッドも関係表形式のビューを定義していると考えられるので、これもWeb・ビューとして取り扱う。

このWeb・ビューの評価の際には、入力スロットに相当する属性の値を外部から与える必要がある。

Web・アプリケーションに関しても、IntelligentPad技術に基づいて、Web文書中の任意の要素とそれらの間の入出力関係を自在に抽出することができるC3W (Clipping, Connecting and Cloning for the web) フレームワークを用いてWeb・ビュー・パッドを生成することができる(図-2)<sup>3), 4)</sup>。本プロジェクトで開発したC3Wフレームワークは、先に述べたブラウザと同様のブラウザ・パッドを用いる。Web・アプリケーションを表示して、マウスを移動させながら抽出したい項目を選び、これをパッドとしてドラッグ・アウトすることができる。一連のナビゲーションの中で抜き出した入力フォームや出力文書項目に対応する複数のパッドは、C3Wシート・パッドと呼ばれる特別なパッドの上に並べると、元のナビゲーションにおけるこれらの項目の入出力関係を保持する。各項目にはアルファベット1文字のセル名が自動的に付与され、C3Wシート・パッドには各セルに対応したスロットが自動生成される。別のナビゲーションから複数項目をこのパッド上に抜き出し、2つの独立なナビゲーションから抽出された項目の間に、セル名を用いて関係式を定義し、これらの間に機能連携を定義することが可能である。こうして得られたC3Wシート・パッドはWeb・アプリケーションから抽出されたWeb・ビューを定義している。このWeb・ビューの評価にあたっては、入力スロットに相当する属性の値を外部から与える必要がある。

これらのWeb・ビューを連携させるには関係演算のジョイン演算を用いることができる。Web・ビューに限定を加

えるにはセレクション演算やプロジェクション演算を用いることができる。このフレームワークを用いると、センサを持った多数の微小知識担体が発するデータ群をWeb・ビューとして扱えるだけでなく、これを解析処理するサービスもまたWeb・ビューとして扱うことができる。多数のマイクロ・アクチュエータ群の制御もWeb・ビューを介して行うことができ、これらのWeb・ビューを関係演算で自在に連携させることが可能となる。

WV1(A-SensorID), WV2(SensorID, OutputValue), WV3(Input, Output), WV4(Actuator, ControlValue), V (SensorID, ActuatorID)をWeb・ビューとする。WV1はWebページから、特定のセンサ群のメンバーリストを抽出したもの、WV2はWeb・アプリケーションないしはWeb・サービスから、センサIDを入力し、その現在値を出力とするWeb・ビューを抽出したもの、WV3はセンサ出力値に応じてアクチュエータ入力値を計算するサービスからWeb・ビューを抽出したもの、WV4は指定したアクチュエータに制御値を送るWeb・アプリケーションないしはWeb・サービスから抽出されたWeb・ビューとする。Vはローカルデータベース中で定義したセンサとアクチュエータの対応関係のビューとする。これらに対して、関係演算

$$(((WV1[A-SensorID=SensorID]WV2)[OutputValue=Input]WV3)[A-SensorID=SensorID]V)[ActuatorID=ActuatorID \text{ and } Output=ControlValue] WV4$$

によりビューを定義する。このビューを評価することによって、Webを介してアクセスできるセンサ群とアクチュエータ群をローカルデータベースが定義する結合関係に従って関連付け、間にセンサ出力をアクチュエータ入力に変換するサービスを介してこれらを連携結合するフェデレーションが定義できることになる。

図-3は、関係演算を用いたWeb・ビューの連携定義を3次元知識メディア環境で実現したものである。ユネスコが公開している世界遺産に関する情報から、遺産名称、国名、登録年からなるWeb・ビューと、遺産名称と写真からなるWeb・ビューを抽出し、これらをジョイン演算で結合し、3次元のデータベース可視化フレームワークを用いて可視化した。



図-2 C3Wフレームワークを用いたWebページからの項目の抽出とそれらの機能連携(上段は企業の株価情報サービス,下段は為替変換で,Cセルを'='B'と定義することにより,ドル建ての株価を円建てに変換する連携が構築されている)

### ルックアップ・サービス

ユビキタス知識のローカル・アクセスとグローバル・アクセスをアドホックに繋ぐためには、ユーザがローカル・アクセスで得た知識をグローバル・アクセス可能なリポジトリに直接操作で簡単に登録するための技術開発が必要である。本プロジェクトではWikiサーバの技術をベースに、ドラッグ・アンド・ドロップでパッドの登録・削除・検索・再利用が可能なPiazzaと呼ばれるリポジトリ技術を確立した。PiazzaはWebページと同様にページを新しく定義でき、URLで各ページをアクセスすることができる。各ページには合成パッドをいくつも並べて登録可能で、登録はパッドをローカルな環境からPiazzaのページにドラッグ・アンド・ドロップするだけで完了する。各ページに登録されているパッドは複製をローカル環境へドラッグ・アウトして利用することができる。

### ユビキタス知識の探索技術

ユビキタス知識の探索技術に関しても、統計的データマイニング、パターン認識、類推理論、半構造データのマイニング、計算論的学習理論など多角的立場から研究を進めている。紙面の都合上、研究内容の一部の概略を紹介する。

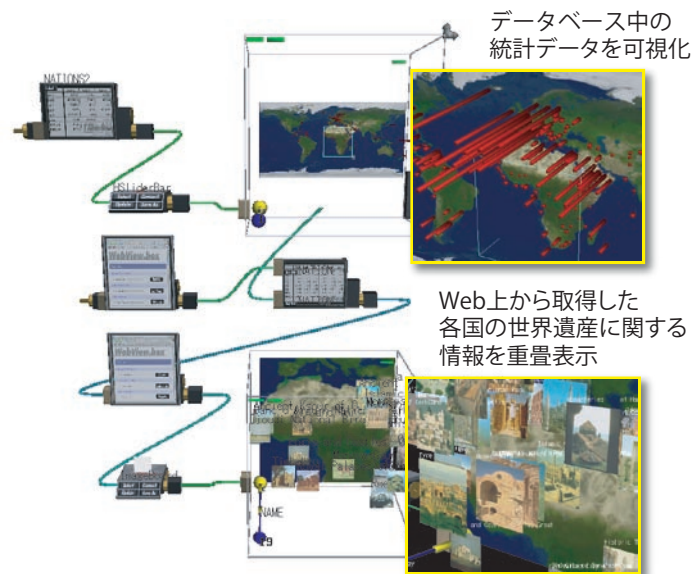


図-3 UNESCOのWebページから抽出された世界遺産に関するWeb・ビューを組み合わせて得られた関係ビューを3次元可視化表示(左端中央の2つの要素がWeb・ビュー)

佐藤義治は、各次元の値が2値の多次元データの分布に対し、統計学的に定義されたある距離構造を定義することにより、距離構造を保存しつつこの分布を球面上の分布へと写し、球面上の確率分布を用いた判別関数を提案した。この手法をWebページの中の画像が広告か否かの判別に適用するために、HTMLに含まれる1,555個の言葉やフレーズ(最大2語)を属性として、この語が含まれる、含まれない2値で測定したものをデータとして用いて実験したところ、誤判別率は約5.5%となり、従来の線形判別関数の誤判別率26%を大きく改善する結果が得られた。

工藤峰一は、ユビキタス知識の各々に付与された記述内容と、知識連携の「目的」や「意図」に関する記述内容とのパターン・マッチングによる検索手法の研究を行っている。このマッチング処理では、付随情報の記述の多様性を吸収しつつ内容のマッチングを判定する必要がある、構文的な適切さと内容の整合性を同時に評価することが必要となる。工藤はこの問題に対して新しい手法を提案し、Web上の飲食店に関する評判情報から意図した評判情報を収集する実験に適用し、良好な結果を得た<sup>5), 6)</sup>。

2人は、これらの手法を拡張してユビキタス知識の選択に応用することを目指している。

### ユビキタス時代のアドホック性と自律性

現在、「ユビキタス」がブームであるが、ユビキタス知識という観点から基盤技術を再検討すると、これらの[ユビキタス]システムと異なり、対象知識の範囲も目的も限定されないオープンな状況を考えなければならず、種々の目的に応じたフェデレーションが即時に形成

される技術が指向されなければならない。誰もが現場で即座にユビキタス知識のフェデレーションを定義できるアドホック・フェデレーション技術と、個々の知識自体が自立的に協調して特定の目的のためにフェデレーションを形成するオートノマス・フェデレーション技術が特に重要となる。そこには常に人が関与するという点も重要である。本プロジェクトの前半3年間では、知識メディア技術を基盤にすることにより、人の関与を容易にし、Webに知識メディア技術を適用してWeb・ビュー抽出技術を確立し、アドホック・フェデレーションをWeb・ビューの合成演算(関係演算)として実現する技術を確立した。後半の2年間では、タブルスペースに知識メディア・アーキテクチャを適用することにより、オートノマス・フェデレーション技術を確立することを目指している。

#### 参考文献

- 1) Tanaka, Y.: Meme Media and Meme Market Architectures – Knowledge Media for Editing, Distributing, and Managing Intellectual Resources, IEEE Press – John Wiley(2003).
- 2) Grieser, G. and Tanaka, Y.(ed.): Intuitive Human Interfaces for Organizing and Accessing Intellectual Assets, LNAI 3359, Springer (2005).
- 3) Tanaka, Y., Fujima, J. and Ohigashi, M.: Meme Media for the Knowledge Federation over the Web and Pervasive Computing Environments, LNCS 3321, Springer, pp.33-47(2004).
- 4) Tanaka, Y., Ito, K. and Kurosaki, D.: Meme Media Architectures for Re-editing and Redistributing Intellectual Assets over the Web, International Journal of Human-Computer Studies, Elsevier, Vol.60 No.4, pp.489-526(2004).
- 5) Nakamura, A. and Kudo, M.: Mining Frequent Trees with Node-Inclusion Constraints, 電子情報通信学会研究技術報告 COMP2004-44 (2004).
- 6) 長谷川博之, 工藤峰一, 中村篤祥: 構造と内容に基づくWebページからの評判抽出におけるパターンの構成法, 電子情報通信学会第16回データ工学ワークショップ予稿集(to appear).

(平成17年3月9日受付)