

## 人類学研究支援環境のためのデータモデル構築

本村 康 哲  
関西大学文学部

川上 聰  
神戸大学大学院総合人間科学研究科

川村 清 志  
神戸学院大学地域研究センター

森下 淳 也  
神戸大学国際文化学部

大崎 雅 一  
姫路獨協大学法学部

フィールドワーク支援のためのデジタルアーカイブの構築とその展開の可能性として、より広範囲の資料に対応できるシステムのためのデータモデル構築を、文化人類学が必要とする要件に従って検討する。データをリンクで参照し、生成された複合データを元のアーカイブのデータと同等に扱うことで、システム上でデータを成熟していく、データカルチャリングが実現される。

## Constructing the Data Model for Anthropological Studies on the Digital Workbench

**Yasunori Motomura**  
Faculty of Letters  
Kansai University

**Akira Kawakami**  
Graduate School of Cultural Studies and Human Science  
Kobe University

**Kiyoshi Kawamura**  
Center for Area Research and Development  
Kobe Gakuin University

**Junya Morishita**  
Faculty of Cross-Cultural Studies  
Kobe University

**Masakazu Osaki**  
Faculty of Law  
Himeji Dokkyo University

As the feasibility study of the digital archives for the anthropological field studies, the data model construction for the system applied to wide range of data is examined according to the cultural anthropological needs. By treating the compound data generated with reference to data with the original data in the archive, data may ripen on the system and, so to speak, *data culturing* is realized.

### 1. はじめに

本発表では、フィールドワーク支援のためのデジタルアーカイブの構築とその展開の可能性について検討する。我々は、人類学や民俗学における従来の調査方法論に対する批判を検討し、被調査者や民族誌の読者にも閲覧可能な民族誌資料のデジタルアーカイブのシステムを提案してきた[1]。それは現地との対話を促進し、双方向的な民族誌を作成することを試行したものであった。今回の発表は、前回の文化人類学からのマニフェストを受けて、より広範囲の資料に対応できるワークベンチのためのデータモデル構築をめざし、文化人類学が必要とする要件を満たすシステムのあり方を議論する。このデータモデルは研究作業をワ

ークベンチ上でおこなうとともに、その作業の結果として生み出された解釈をもデータとして登録可能なシステムを目指している。すなわち被調査者や研究者間での対話自体をデータ化していき、各々の視点によって資料が解釈される過程と階層自体をアーカイブ化する方が試みられている。これは、従来のデータモデルが自己完結的で完成品であったのに対して、このモデルは断片的なデータも格納でき、またデータを組織化する過程についても保持し、データが成熟してゆくものである。

### 2. 民族誌の新たな試み

文化人類学ではその主要な成果を示す民族誌について、激しい批判が行われてきた。そこで指摘された調査者と被調査者との非

対称性や民族誌に付与された権威を相対化する試みとして、インターネットによる現地との民族誌資料についての対話や web 上での民族誌の公開が提唱されたり、試験的な試みを実施されたりしている[2]。これらは、インターネットの特質に注目し、それらを利用することで、「民族誌という固定したモノに決して収斂すること」のない、「開かれた民族誌の手法」が提唱されている[3]。

一方、民族誌資料に分類項目を付加して世界規模のデータベースを作り出してきた HRAF<sup>†</sup>は、eHRAF としてその一部の項目が、オンライン上で展開されている。日本においても様々な機関で民俗資料や民具等のデータベースの開発・公開が盛んになってきている。これらは、インターネットの普及にとともに、場所に関わらず大量の情報の提供や検索を簡易化することが目指されている。

しかし、ここに一つの問題がある。これまで web 上に公開された民族誌資料の多くは、研究者によって検証され、きれいに整備された状態で提示されている。つまり、そこで発信される内容は、「著者」である民族誌家の視点によって固定されており、たとえそれらが「単線的なストーリーに沿って配置されるのではなく」でも、「断片的情報を付加してデータベース全体に寄与することができる」かどうかは、疑問と言わざるを得ない[4]。同様に eHRAF においても、元データはすでに発刊された民族誌に他ならず、完成した民族誌を再分解したうえで、項目ごとの分類、階層化を行っているわけである。

本研究が第一に目指すのは、上記の研究が提唱し、部分的には実施されつつあるオンライン上での民族誌資料のデジタルアーカイブの構築である。それに加えて本研究では、そのような「民族誌」が構築される際の視点そのものを相対化する作業を行う

ためのモデルを提示したいと考える。既成の民族誌が構築される際の様々な問題を考慮しながら、現地の視点や他の研究者の視点をどのように反映させていくのかという課題に答えようとする試みである。

### 3. 既存資料の再解釈

本発表が事例として提示する画像は、研究者が撮影した場合と現地の人々が撮影した場合が想定されている。両者のうち、資料としての汎用性が大きな意味を持つのは研究者が資料収集のツールとして画像資料を用いる場合である。今日、研究者は、とりあえず記録しておきたい様々な資料を、デジタルカメラで記録することが可能になった。例えば民具や家屋などの物質文化について、実物を収集できない場合には、記録撮影が重要な参照資料となる。また、調査地などでコピーが不可能な古文書類などの記録においても、画像データが有効である。このような場合に、翻刻する前の一次資料の状態（書体や虫食い状態など）についても画像として記録しておくことは、文献解読のうえで貴重な情報となることは容易に想像がつく。

しかし、このような技術論的な問題だけでなく、写真資料自体を研究対象に据える動きが、人類学、社会学、そして民俗学の周辺で活発になりつつある。デジタル化以前から人類学や民俗学のフィールドワークでは、調査地を記録するための技術として写真の利用が行われてきた。それどころかフィールドワークを推奨し、近代人類学を確立したマリノフスキーの時代、すでに写真は重要な調査ツールとなっている。もっとも、当時の技術では、露光に時間がかかるため、撮影の際には被写体が一定の時間、静止する必要がある。その静止すべき時間を守れない現地人に対して、マリノフスキーが日記のなかで「この野蛮人」と叫んでいたエピソードは、あまりに有名である[5]。

しかし、1980 年代ごろから、画像資料は単なる補助的なデータではなく、文字資料

<sup>†</sup> Human Relations Area Files。世界の主要な社会や民族集団についての文献を網羅的に集め、文献中のデータを細分化しデータベース化したもの。

には残されていない文化的な背景や画像自体の表象の特質を分析、検証する対象として注目されるようになってきた。そこでは、初期の「未開民族」についてのイメージの構成に大きな影響を与えた写真試料やその展覧会の検証が行われている[6]。

また、研究が立ち遅れていた日本民俗学においても、近年、いくつかの注目すべき成果をみることができる。例えば、菊地暁は戦前から戦後における「民俗」事象についての関心が、どのような立場の研究者や写真家から生じたかについて検証している[7]。矢野敬一は、より現地に近い立場からの視点として、写真家熊谷元一の写真集に描き出された郷土の風景の特質について、熊谷自身のライフヒストリーや写真をとりまく同時代の状況を踏まえつつ論じている。熊谷自身は、自らの郷土を対象として写真を取っており、プロの写真家たちが地域の生活や慣習を記録する視線とは異なった特質をもっている。

また、研究者自身が撮影した画像資料にも関心がもたれている。小川直之は、折口信夫がフィールドで撮影した写真を整理し、その特質と今日における資料性について論じている[8]。また、ジャーナリストの佐野眞一は、『宮本常一の写真に読む 日本の昭和』という本において、宮本常一が残した写真資料を再配列することで、その旅の軌跡を明らかにするとともに、彼が生き抜いた昭和という時代の流れを再現することに成功している[9]。ここで佐野の議論を少し追いかけることで、このデジタルワークベンチが目指す方向性について模索してみたい。

宮本常一(1907-1981)は、日本各地をくまなく歩き、各地の民話や口承伝承をはじめとして、その当時の生活の様々な断面を記録してきた。彼は、膨大な著作を発表しているが、フィールドワークや旅のなかで撮影された10万点にも上る写真をも残している。佐野の本は、これらの宮本が残した膨大な写真から、高度経済成長期前後の日本各地の写真を選別し、並べ、解説をつけた

ものであり、各章は写真、写真のキャプション、そして佐野の解説から構成されている。この本の写真は宮本自身によって撮影されたもので、ネガフルム番号、コマ番号、撮影年月日、撮影場所(内容)の四項目に分類したシート表に整理されていると佐野は記しているが、本書の写真のキャプションが宮本自身によるものか否かは明記されていない。

宮本の写真は、足を止めたり構えたりせず、ほとんど歩きながら撮り続けられた。汽車に乗ると必ず窓際に座り、車窓に流れる風景にレンズを向けたものであった。人類学者や民俗学者の例にもれず宮本はその時々調査のテーマに従って写真を撮ったであろうが、それ以外にも、その場その場で興味をひかれるものにもレンズを向けていたのであろう。彼の残した写真には、撮ったときの撮影者の意図を超えたものが宿っており、佐野があえて「宮本常一の写真に見る」とはせず「読む」と題したのは、宮本の撮った写真を解釈し、読み解いたということである。

佐野がおこなった写真を読み解く、あるいは解釈するという作業はどのようなものであろうか。これは既存の資料を選別・分類し、閲覧者が独自のコメントおよび解釈等の主観情報を付与するというものである。実は、ここで資料の整理の仕方は、ややオーソドックスであり、そこで読み解かれるものは、宮本の同時代の人々や社会についての感性的な部分が大半を占めているといえる。見方を変えるなら、この本は宮本が残した膨大写真資料を佐野という個人が選択し、再配列することで、ある物語を付与したとみなすことができるだろう。このような再配列におけるキャプションやコメントの重要性は、“Picturing Bushmen”の筆者、Robert Gordon によっても指摘されている[10]。彼は自らの本に1925年の催されたブッシュマンの写真展示会の画像を採録するにあたり、すべてのキャプションを画像本体からは切り離して提示したのである。そのことは、画像にキャプションを付加することがイメージの生産であり、ある方向付

とがイメージの生産であり、ある方向付けを画像資料に与えることになることを示している。本発表は、このような分類が、特定の著者や監修によって行われるのではなく、より広い研究者や場合によっては、被写体や調査対象となった人々の視点を踏まえながら構築することを目指すものである。

既存資料を再解釈すること、つまり既存資料を整理、分類したうえで検証していくことは、個人の能力をはるかに凌駕するものである。このような側面からも、現地との対話による資料の選択に加えて、研究者間での対話と資料の構造化を可能にするデジタルワークベンチの開発は、きわめて急務の作業と位置づけられるだろう。

#### 4. デジタルワークベンチ

フィールドワークで得られる種々の資料は、テキスト・音声・静止画像・動画像などの形態でデジタルアーカイブに格納されている。もし、これらのデータに対して新たな知見を得た場合には、われわれは研究

資料として、何らかの形で、成果として記録・整理しようとするだろう。たとえば、データベースの検索を行った結果、得られたデータ集合を保持したい、グループごとに分類したい、あるいは特定の部分データだけを抽出したい場合等である。このような作業は、資料探索において頻繁に発生するものである。しかしながら、この作業の過程や結果は従来、個人の活動であり、再利用も困難である。デジタルアーカイブは、一次資料の格納を主眼とするものであり、格納されたデータを再利用することや、大量のデータから目的とする情報を効率よく抽出し、体系的に保持することは研究者が解決すべき単なる利用方法と見なされている。

そこで本研究ではこの二次的な利用に着目して、アーカイブに格納されている一次資料から作成される二次資料を統括的に扱うデータモデルを提案する。このデータモデルは、アーカイブからデータをリンクの形で引用し、そのリンクに加えて独自のデ

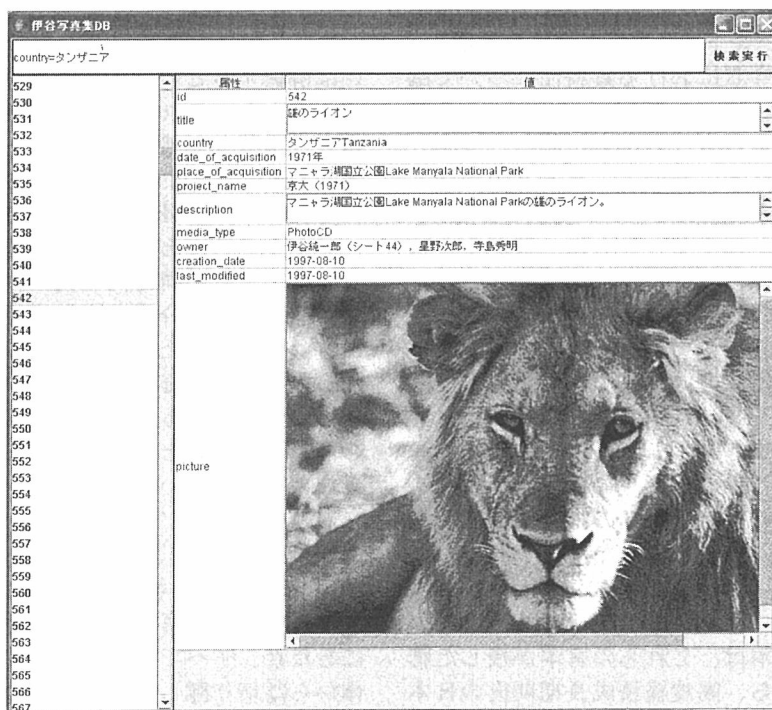


図1:トレイの例

ータを付与し、新しいデータとして登録することで、その研究過程で得られた情報を再利用可能とすることを目的としている。

#### 4.1 システム構成とワークフロー

本システムは、機能として、データプール、トレイ、ワークシートの3つの項目で構成されている。

##### 1) データプール

参照すべきデータのエントリが準備される場所として、データプール（以下、プールと省略）が提供されている。プールへの登録対象となるものは、データベースシステムや URL によって表現されるウェブサイトや、ファイルなどのファイルシステムである。エントリにはデータベースシステムに置かれたテーブルやビューの情報などが保持されており、このエントリによって検索のための足掛かりが与えられる。

##### 2) トレイ

トレイにはプールのエントリから抽出してきたレコードが列挙される。利用者は検索によってプールからレコードの抽出を行う。検索によって得られたレコードはトレイに載せられ、利用者によってブラウズされる。トレイはプールから検索された結果を単に載せるだけではない。利用者がトレイ上のレコードの内容を吟味しながらレコードの取舍選択を自由に行うことができる。図1にトレイの例を示す。ウィンドウの左側がレコードのリストであり、その一つを選んだ時のレコードの内容を右側にブラウズしている様子が見られる。

##### 3) ワークシート

トレイに載せられたレコードから、研究者の望むデータを構成する作業場としてワークシートを提供する。ワークシートでは、トレイに置かれたレコードの引用と独自のデータ項目の作成を行なうことができる。引用の際には、レコードからの部分だけを

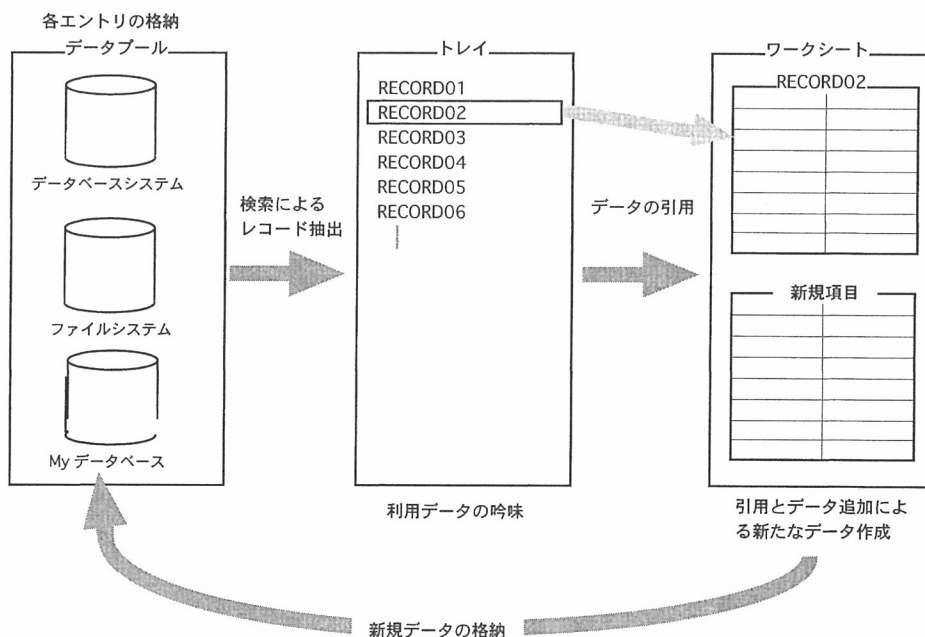


図2:ワークフロー

指定して取り出すことも可能である。

利用者は、プールに置かれたエントリから、レコード単位で取り出して、トレイに入れる。トレイに入れたレコードを吟味しながら、ワークシートに引用し、自らの新しいデータ項目とともに、時にはレコードの部分を取り出す、或いは、複数のレコードを列挙する等して、新しいデータを作り上げる。ワークシートで作られたデータは、プール自身のエントリへと格納され、他のエントリと同等に扱えるようになる。このワークフローを図2に示す。

#### 4.2 データモデル

このようなシステムを実現するための試みはいくつかあるが[11-14]、ここでは、ワークシートを構成するモデルとして、データを引用する為の拡張された URI と追加するデータ項目の集合によって構成されるデータ構造を用いる。

拡張された URI とは、プールに置かれた利用者の求めるデータが特定できる識別子を指す。例えば、関係データベースシステムであれば、レコードを特定できる ID になる。また、更にレコードの部分データを抽出する場合には、特定のフィールド名を加える。より細かな内部を特定するならば、取り出されたデータ内の範囲指定子を更に加えたものとなる。これによって、利用者の求めるものを引用することができるよう拡張する。

利用者が、既存データの分類や系列化を表現できる構文を用意する。拡張された URI の列に対して、順序付けやグループ化を表現できる構文である。

利用者が独自に付与するデータは、項目名と値からなる組を基本とする。これらは、上の引用と自在に配置できるようにして、特定の意味付けが表現できるようにする。

このようなデータ構造を格納できるシステムとしては、XML が有効である。そのため、プロトタイプシステムは、XML による自由形式のデータ構造を用い、XMLDB に格納している。

システムの元となるプールに置かれるものとして、外部に存在するデータベースシステムやファイルシステム（Web 情報も含む）を想定している。システム自身がデータを全て保持する立場を取った場合、利用範囲が著しく狭まると考えたからである。寧ろ既存のシステムと有機的に統合できることがアーカイブされたデータを活かすことになると考え、外部のリソースを扱うシステムという立場に立った。

#### 4.3 システムの要件

上記のデータモデルも用いて、前節で述べたデジタルワークベンチを実現する要件を挙げる。

1) 元データの改変は行わない。公開されたアーカイブの資料は一般のデータとは趣が異なり、重要な事実を表す記録と見ることができる。これを書き換えることは、いわば出版された印刷物を書き換えるのと同じことになる。もし、かりにアーカイブ内のオリジナルデータの改変を考えることは、記録に対する異論として、別途、付与するべきであるという立場を取る。

2) 外部リソースは引用のみでシステムにコピーしない。外部リソースからのデータをコピーして取り込み、自在にその見方を変えることは、著作権の問題が発生する場合もあり得る。記録の源を尊重するという立場からも、常に引用する形でデータを扱うことにする。効率の問題は発生するが、これを前提にシステムを構築することにした。これによって、1)の条件も常に満足できることになる。

3) 利用者の作成したデータもまた、プールに登録可能である。利用者がこのシステムに基づいて作業したデータもまた、他の外部リソース同様、プールに公開し(publish)、必要に応じて、リソースとして提供できる。この形で研究資料が公開されることによって、冒頭で述べたように我々の目指している目標である、調査者と被調査者の対称性が確保され、開かれた民族誌がシステム上で実現される。

#### 4.4 検索のカスケード

このように既存のデータを引用して生成された新しいデータは利用者のエン트리としてプールに保存され、公開可能なものとして、外部リソースと同等に扱える。この場合、利用者の作成した複合データには、引用するデータのリンクしか保持されていない。この複合データにリンク先のデータに含まれるキーワードの検索命令を与えた場合、システムは複合データのリンクをたどり、元の外部リソースに検索命令を伝えて、結果を得る必要がある。往々にして複合データのプールエントリーは沢山の外部リソースへのリンクを保持していると考えられるので、システムは複合データのエントリーにあるデータ一つ一つに対して、外部リソースのデータへの検索を実行することになる。リンクに基づく複合データを作り続けていけば、複合データへの一つの検索から、データからデータへと検索命令が雪崩のように発生する。これが検索のカスケードである。

この多量の検索命令をどう効率良く実行するかということはこのシステムの根本的な問題であり、なかなか解決できない困難である。例えば、プールの複合データのエントリー自身に含まれるリンクの外部リソースの情報を追加して、個別の複合データを検索するより先に、外部リソースへの一括検索を生成して、ヒットした引用データとデータを照合する形で検索の最適化を図ることを検討している。

#### 5. おわりに

デジタルアーカイブやデータウェアハウスと呼ばれるシステムで利用されるものとして、データマイニングという技術がある。これは膨大なデータ集合から、データ相互の関連を見いだすものである。この技術は既存のデータから計算機によって解を導き出すことを目的としたものである。

一方、ここで提唱したデータモデルを用いたシステムでは、利用者によって既存のデータから読み取られた新たな知見をアー

カイブへと追加・格納することによって、システム内でデータが培養され、増殖してゆくことを目的としている。これは、既存のデータに人間が新たな意味を与えて構造化してゆく過程であり、これをデータカルチャリングと呼びたい。データマイニングでは計算機の力を借りて膨大なデータの中から相関性を発見するのに対して、データカルチャリングでは、人間が全ての過程に関与し、構造化し、ひとつの解へと到ることを目的としている。類似した技術として、データファーミングという技術もあるが、その目的は、データマイニングを支援するものに過ぎない。

様々なデジタルアーカイブの構築がなされており、その多くはデータの蓄積を目的としたものであるが、ここで提唱したシステムでは、本来のアーカイブの目的のひとつである利用を主眼としたシステムであるということも出来る。

このシステムは、人類学におけるいくつかの困難な状況を打開したいという現状から始まったものであるが、勿論、このシステムの利用に関してデジタルデバインドなどの問題点があることは我々も認識しており、人類学が直面する課題の全てを解決できるものとも考えてはいない。むしろ、我々の目指すものは、情報資源の平準化やそれらを使いこなすためのメディア・リテラシー、情報の取得に必要な言語の共有、コンテンツへの関心の高さといった問題をクリアしたうえで実現されるものかもしれない。

だが、それらの課題は、研究者間での民族誌の再構築に関わる共同作業の場において試験的な試みを広げていき、随時、調査対象であった人々による参画を求める形で進めていきたい。インターネットやマルチメディアの利用が飛躍的に拡大しつつある今日、これらの実践的な作業のなかでこそ、遵守すべき指針や原則も確定されていくのではないだろうか。そのような試行錯誤の過程自体が、カルチャリングなのであり、双方向的な文化の構築するための試みであると位置付けておきたい。

## 参考文献

- [1] 本村康哲, 川上聡, 川村清志, 森下淳也, 大崎雅一, 「人類学における研究支援環境としてのデジタルワークベンチ」, pp.219-226, じんもんこん 2003.
- [2] 太田好信『トランスポジションの思想——文化人類学の再想像』世界思想社、1998、大村敬一「本質主義的な記述を超えるために——『社会・個人』の二元論への状況的認知論からの挑戦」『文化人類学研究』3、76-100、2002.
- [3] 湖中真哉「民族誌の未来形へ——オンライン民族誌の実践から」『民博通信』102、6-8、2003.
- [4] 飯田卓「異文化をプロデュースする——テレビ番組と民族誌」『民博通信』102、2-5、2003.
- [5] マリノフスキー.B.『マリノフスキー日記』谷口佳子訳、平凡社 1987.
- [6] Elizabeth Edwards *Anthropology and Photography 1860-1920*, Yale University Press, 1994.
- [7] 菊地暁『柳田國男と民俗学の近代——奥能登アエノコトのアルケオロジー』吉川弘文館, 2001.
- [8] 小川直之「画像資料と民俗学」  
[http://www2.kokugakuin.ac.jp/frontier/bulletin1/bulletin1\\_13.pdf](http://www2.kokugakuin.ac.jp/frontier/bulletin1/bulletin1_13.pdf).
- [9] 佐野眞一『宮本常一の写真に読む 日本の昭和』, 平凡社, 2004.
- [10] Robert J. Gordon *Picturing Bushmen: The Denver African Expedition of 1925*, Ohio University Press. 1997.
- [11] S.Ueshima, K.Ohtsuki, Q.Qian, H.Oiso, K.Tanaka, J.Morishita, "Incremental Data Organization for Ancient Document Databases," *Proc. 4th International Conf. on Database Systems for Advanced Applications (DASFAA '95)*, pp.457-466, 1995.
- [12] S.Ueshima, K.Ohtsuki, T.Sugiyama, J.Morishita, "Hierarchical Graph Model: Collecting Objects and Generating Multiple Views for Semi-structured Data," *Advanced Database Systems for Integration of Media and User Environments'98*, Ed. by Y.Kambayashi et.al.(World Scientific Ltd) pp.203-208, 1998.
- [13] 上島伸一, 大月一弘, 杉山武司, 森下淳也, 「階層構造グラフを用いた半構造化データの構造化手法」, 情報処理学会論文誌 Vol.39, 構造化手法」, 情報処理学会論文誌 Vol.39, No.4, pp.857-867, 1998.
- [14] S. Abiteboul, Peter Buneman, Dan Suciu, "Data on the Web: From Relations to Semistructured Data and Xml," Morgan Kaufmann Pub. 257p, 1999.