

民俗学研究のための情報発信

木下 宏揚／能登 正人

I 非文字資料による情報資源と情報流通の管理

はじめに

情報資源の発展は、社会の変化および人間の日常生活に多くの影響を及ぼしてきた。Web 上には膨大な量の情報が蓄積され続けているため、情報資源に関する情報、すなわち情報資源の書誌データを適切に記述し、それを用いて検索出来るようにすることが求められる。

現在、セマンティックWeb や電子図書館を中心としてメタデータが注目を集めている。メタデータは「データに関するデータ」と定義される。これを広義に捉えると、目録情報、索引、事典、辞書、抄録、書評、利用条件など、色々なものがメタデータの範疇となる。ネットワーク上では、情報資源を探し、評価し、利用するというすべての過程でメタデータを必要とするために、情報資源に関する記述（メタデータ記述）が必要不可欠である。メタデータ記述の目的も、資源の管理と提供、保存、知的文化財産権や利用環境など、さまざまであり、目的に応じたメタデータのモデルとメタデータ規則が提案されている。

Dublin Core はインターネット上における情報資源の発見（Resource Discovery）を目的として開発されてきたメタデータで、インターネットや電子図書館における標準的なメタデータ記述として広く認められているものである。Dublin Core は、書誌データをあらわすための15要素と呼ばれる基本項目で定義されている。Dublin Core の特徴は、インターネット上にある多種多様な情報資源に対してさまざまな分野のコミュニティに共通するメタデータ規

則であること、また規則の開発が、図書館や博物館、インターネットといったワーキンググループによって進められてきたことが挙げられる。貴重資料や歴史資料の電子化は、図書館が持つ貴重資料そのものへのアクセスを制限する一方、電子的な複製物によりアクセス性を飛躍的に高めるといふ、「保存とアクセス」の両方の観点から進められてきている。高品質なデジタル化を行うことで、電子データを用いた研究や展示が行えることに加えて、人的災害や自然発生によって破損した貴重な資料の復元が行えるといったメリットが挙げられる。また、図書館や博物館に蓄積された文化財を、どこからでも閲覧可能なことも重要な機能の一つであると考えられる。一方で、電子化された資料がどのように利用されるのか（利用できるのか）、実際に利用しやすくするにはどのようにすべきか、といった点についても十分に考慮されないといけない。単に、資料を電子化しインターネット上で提供するだけでなく、研究利用での利便性の促進や評価、新しい情報技術との組み合わせによる新しい分野での開拓といった活動が求められる。

現在、神奈川大学21世紀COEプログラム「人類文化研究のための非文字資料の体系化」では研究成果を電子化して保存し、研究内容を再構成するために必要な技術や情報の構築を目指している。文化人類学の研究において、民俗調査の際には文化の歴史、遺跡、遺構、遺物、調査方法、ヒアリング内容などを記録するための膨大な数の写真撮影、および資料がある。また研究者は、研究資料の多数を個人的に管理している状態におかれている。そのため、研究

資料の多くは、当該研究者でしか分からない状況にあるといえよう。我々は先祖より贈られてきた文化遺産を後世に伝える義務があると考え。よって、文化的な遺産が永久に失われてしまう前に電子化し、情報資源として発信していく有用性があると判断する。

本研究では、文化人類学的の見地に立って、民俗学調査の際に収集した資料を広く公開する手法を確立することを目指している。一方、情報学的見地からは、非文字情報が持つ情報をメタデータとして整理する方法を確立し、メタデータを用いた情報の流通、利用法を確立することを目的としている。本稿では、非文字情報資源の集まりを捉え、それに対するメタデータ記述の必要性を説く。非文字情報資源の場合、情報資源の粒度はさまざまであり、各々に適したメタデータ記述が求められる。以上のような文化人類学的な分野依存のメタデータのためには Dublin Core に定義されている 15 要素以外の定義が求められる。そのため、Dublin Core を、より詳細で厳密なメタデータ記述方法、dcterms を提案し、考察を行う。

以下、1.2 でメタデータの基礎概念を述べ、1.3 で機械処理可能なメタデータ技術応用のセマンティック Web とオントロジーの関係を述べ、1.4 で非文字情報を発信する際に適した提案概念、dcterms と Application dcterms について述べる。また、1.5 で今回付与した非文字資料情報の考察を述べ、最後に 1.6 で、まとめと今後の課題について論じる。

メタデータの基礎概念

本節では、書誌情報のように、情報資源を組織化・管理し、利用に提供するための情報資源の性質について述べる。また、情報資源発見のためのメタデータ規則として開発された Dublin Core について述べる。

1——メタデータの目的と種類

計算機技術の発達には計算能力の向上だけでなく、さまざまな応用技術を生み出した。飛躍的に増加し続ける情報資源を管理するための新しい概念として発展してきたのがメタデータである。メタデータは

「データに関するデータ」と定義され情報の内容を表現すると同時に、情報を構造化するものである⁽¹⁾⁽²⁾。メタデータは情報資源管理と情報資源発見を目的としており、応用は基本的にその対象範囲を広げること、現在、研究など多くの分野を対象を広げている。メタデータは、ある情報に対する付加的なデータで、データを意味付けするために重要な役割を果たす。これまでのメタデータは各個人や組織、団体など限られた範囲内で限られたデータを扱うために利用されてきた。図書館や博物館は、館内あるいは同様の館同士、たとえば図書館同士の整理、管理、検索を効率よく行うことを目的として独自の目録システムを構築し利用していた。ところが近年の情報電子化は対象を選ばず急速に進んでおり、同時にこれまで独自に運用してきた情報同士を交換したり、横断的に検索したいという要求が生じてきている。各組織が独自のメタデータのまま交換しただけでは、内容に関する情報が正確でないばかりか、意味がまったく違ってしまふ可能性もある。そこで近年、メタデータを国際的に共通化、標準化する動きが活発になっている。メタデータを見る際に大きく分けて、情報資源を組織化・管理する提供側（管理者側）の視点と情報資源を探し、利用するという利用者側の視点がある。図書目録や書誌データは主に前者の視点で作られてきており、検索システムが扱うメタデータは後者の視点から決められるといえる。図書館と博物館などではメタデータ記述に関する要求は異なる。ネットワーク上の情報資源と従来の冊子体資料でも要求は異なる。こうした要求を満たすには、複数のメタデータ規則を目的に合わせて適切に利用すること、異なるメタデータを統合的に利用することなど、メタデータの相互利用性が求められる。そのため、記述対象や応用目的に合わせて情報資源の性質を記述するためのメタデータ規則は決められている。また、メタデータ規則そのもの、およびそれに関連する情報を登録し、人間にも機械にも提供するサービスの必要性があると考え。情報資源の組織化・管理し利用に提供するための情報資源の性質を表すものを以下に示す。

1. 情報資源の内容に関する記述かつ情報資源を探

表 1.1 Dublin Core の基本 15 要素

要素名 (英語)	識別子	識別子定義および説明
タイトル (Title)	Title	情報資源に与えられた名前
作成者 (Creator)	Creator	情報資源の内容の作成に主たる責任を持つ実体
主題およびキーワード (Subject and Keywords)	Subject	情報資源の内容のトピック
内容記述	Description	情報資源の内容説明
公開者 (Publisher)	Publisher	情報資源を公開することに対して責任を持つ実体
寄与者 (Contributor)	Contributor	情報資源の内容に何らかの寄与をした実体
日付 (Date)	Date	識情報資源のライフサイクルにおける何らかの事象に対して関係付けられた日付
資源タイプ (Type)	Type	情報資源の内容の性質もしくはジャンル
形式 (Format)	Format	情報資源の物理的形態ないしデジタル形態での表現形式
資源識別子 (Resource Identifier)	Identifier	与えられた環境において、情報資源への一意に定まる参照
出处 (Source)	Source	現在の情報資源が作り出される源になった情報資源への参照
言語 (Language)	Language	当該情報資源の内容を表すために用いられた言語
関係 (Relation)	Relation	関連情報資源への参照
時間的・空間的対象範囲 (Coverage)	Coverage	情報資源の内容が表す範囲あるいは領域
権利管理 (Rights Management)	Rights	情報資源に含まれる、ないしは関する権利に関する情報

し出すために利用されることが主であるもの (Descriptive Metadata)

2. 情報資源の保存やアクセス制御など管理方法に関するもの (Administrative Metadata)
3. 情報資源の物理的・論理的内部構造に関するもの (Structural Metadata)
4. 情報資源の利用に必要な技術的要件に関するもの (Technical Metadata)

Dublin Core は情報資源の内容記述が主である。一方、研究資料など、書誌情報の電子化を指向しているメタデータの場合には、これらすべての性質を含んでいる。また、メタデータに関するデータ、すなわちメタ・メタデータや異なるメタデータ規則を横断的に用いるためのメタデータ規則間の対応関係なども情報資源の記述として挙げられる。

メタデータ規則の構成要素は以下のように一般化して捉えることが出来る。

1. 記述対象属性の定義
2. 記述対象の属性値の定義
3. メタデータ構造の定義
4. 省略可能性や繰り返し回数など、記述対象属性

ごとに決められる制約

5. メタデータ記述構文
6. 実際の対象情報資源に対して、どのように記述すべきか内容を抽出し、メタデータを記述するかに関する指針

以上の観点から見ても、メタデータ記述に当たっては、記述対象をどのように捉えるかが重要な視点になるといえよう。

2— Dublin Core

メタデータを記述する語彙にはさまざまなものがあり、⁽³⁾ 個人が自由に作成し、付与することが出来るが、それらを記述する代表的なメタデータ規則に Dublin Core がある。Dublin Core はメタデータ規則の開発を進める組織 DCMI (Dublin Core Metadata Initiative) によって提示された。1994 年の WWW の国際会議での議論から、インターネット上でのリソースの発見を目的としたメタデータの必要性が認証され、1995 年にアメリカ・オハイオ州の Dublin で最初のワークショップが開かれ、色々な分野に共通に適用出来るメタデータ記述の要素の必要性が合意された。これは、情報資源と電子

化データを意識せずに扱うように設計されたもので、現在、電子データやWeb情報のメタデータとして標準化が進められている。また、Dublin Coreは多様な分野で作られるメタデータを相互利用できるようにすること、分野の違いを超えて情報資源を検索できるようにすることを目標としている。換言すれば「色々な分野の情報資源を見つけ出すための共通の枠組み」と捉えることができる。

Dublin Coreは表1.1に示すように15要素で構成されている。この15項目は最も基本的なものに限定されており、すべての要素は省略可能かつ繰り返し可能と規定している。項目を階層的に細分化することも可能で、細分化された項目をsubelement、それぞれの項目の意味付けをQualifierと呼ぶ。Qualifierは組織や分野によって異なるが、各組織が全く独立に規定するのではなく、Dublin Coreで基本的な部分を規定している他⁽⁴⁾、同じ目的や分野を持つ団体グループとなってQualifierを規定するように議論が行われている。現在、図書館、博物館などのグループでそれぞれの詳細なメタデータを標準化する作業が行われている。なお、Dublin Coreは要素と、その意味情報(Semantics)を規定するもので、構文(Syntax)や実装を規定するものではない。現在の技術においては、RDF⁽⁶⁾を利用してメタデータを意味付けし、XML⁽⁷⁾などを利用して実装を行うのが一般的である⁽⁸⁾。

セマンティックWebの概要

近年、Webコンテンツに意味情報を付与することにより、Webの有用性を飛躍的に高めようとするセマンティックWebが、注目されており、この意味情報を表現するための方法の一つとしてオント

ロジー(Ontology)が用いられている。セマンティックWebは、オントロジーの記述に適したRDF記述を採用している。本節では、これらセマンティックWebの概要について説明し、その現状および課題について述べる。

1—セマンティックWeb

情報資源管理はこれまでデータベースの研究が中心であったが、インターネットとWWWの登場により、WWW上の資源情報を管理するという方向が新たに生まれてきた。セマンティックWebはWebの創始者であるTim Berners-Leeが1998年頃に提唱し始めた技術である。Web上にある文書などのセマンティック、すなわち意味情報を取り扱う技術である。Web関連の標準化団体W3Cによる仕様策定や研究開発が進められ、技術の一部を実用化した企業も登場している。

2—セマンティックWebとRDF

HTML(HyperText Markup Language)は主に人が読み書きを理解するための文書の体裁を記述するメタデータ仕様であったのに対して、RDF(Resource Description Framework)は文書内容の意味情報を記述するメタデータ情報である⁽⁶⁾。RDFはセマンティックWebの最大の特徴であり、基盤技術である。RDFは宣言文(Statement)から成り、主語(Resource)と述語(Property)、そしてその目的語(Propertyの値)の三つの組みで構成される。以下のようなStatementをRDFで表す場合を図1.1に示す。

RDFの設計思想は意味情報を機械処理することを志向しており、処理そのものには曖昧性がないように設計されている。しかし、意味自体には曖昧性、

RDF形式での記述例

```
<rdf:RDF>
<rdf:Description about="http://www.himoji.jp/">
<s:publisher>Systematization of Nonwritten Cultural Materials for the Study of Human Societies
</s:publisher>
</rdf:Description>
</rdf:RDF>
```

多様性、断層性があり、これらを意味情報に特有な性質を取り扱う仕組みとして、階層構造のフレームワークが検討されている。これらのフレームワークにより、Web 自体を巨大な知識システムとして取り扱えるようにしようというのがセマンティック Web ビジョンである⁽¹¹⁾。現在、RDFで記述したメタデータがどのような意味情報を持つか規定する RDF Schema、その上位の機構として OWL (Web Ontology Language) の標準化が進められている⁽¹²⁾。今後は、セマンティック Web フレームワークのどの階層までを取り込むかはアプリケーションの目的、将来の拡張性を考慮して決める必要がある。

3——セマンティック Web とオントロジー

WWW によって、インターネットの上に文書レベルでの流通と共用の基盤が確立されたといえる。セマンティック Web はそれを意味情報を付加するまで高めようといった取り組みである。現在、セマンティック Web ではオントロジーの議論が盛んに行われている。これは XML によって構造を持つ文書レベルでの共用が可能となった WWW において、さらに語の意味情報の共用を可能にするための取り組みであるといえる。セマンティック Web などのオントロジーの取り組みによってネットワーク上での語の意味情報の共有が進むことが期待されている。

4——セマンティック Web の今後の課題

1.3.3 で述べたようにオントロジーの概念はセマンティック Web を支える中核基盤の一つであるといえよう。今後セマンティック Web が普及するための今後の課題として以下のようなことが考えられる。

1. オントロジーの作成

セマンティック Web の応用メリットを享受するには膨大なオントロジー記述が必要と考えるが、それをどのようにして作っていくかという課題がある。またセマンティック Web が目標とする複数オントロジー並存の情報交換を実現するためには、オントロジーの相互交換、統合などを容易にするようなオントロジー構築の方

図 1.1 RDF 記述例

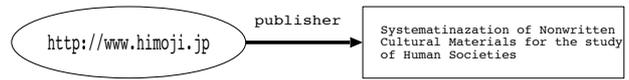
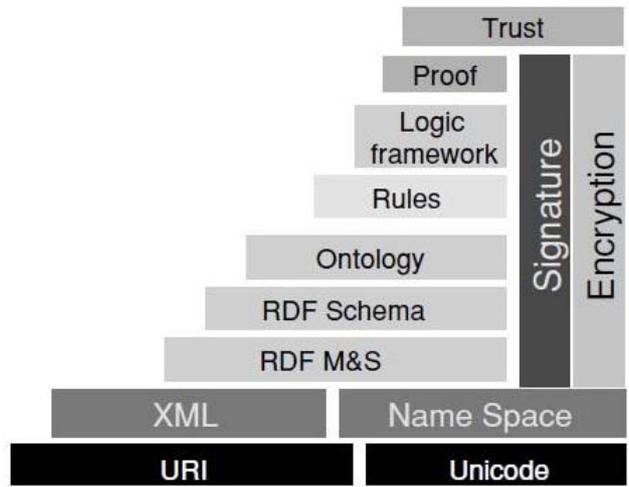


図 1.2 セマンティック Web



WWW2002,W3C trackより引用

法論が必要であるといえよう。

2. オントロジーの質

セマンティック Web では、インターネット上で分散したオントロジーを利用して、人間の複雑な要求に応えられる自動サービスを実現しようとしているが、オントロジーの質はまちまちであり、お互いに論理的に矛盾している可能性もある。分散、動的に変化する Web の世界において、どのオントロジーが信頼できるかを知る仕組み、また互いに矛盾がある場合にそれを柔軟に処理する仕組みを確立する必要がある。セマンティック Web はいまだ普及していないが、もしそのメリットが明確になれば急速に流通するものと思われる。TCP/IP がインターネットの基盤プロトコルであるように、セマンティック Web は拡張性や将来性を考慮した分散アプリケーションの基盤となるプロトコルになっていくと考える。

非文字情報を発信するのに適した提案概念

本節では、神奈川大学 21 世紀 COE プログラム「人類文化研究のための非文字資料の体系化」での

研究の一環である民族地域文化の研究資料をもとに、非文字資料から読み解くメタデータに適した提案法を論じる。

1——非文字資料のメタデータ

民俗学研究で必要な項目を Dublin Core の項目にあてはめ検証した結果、実際は何に関するメタデータかによって、グループが存在することが分かった。つまり実際の非文字資料に対応する形で、地域遺跡、郷土遺構、文化遺物に関するメタデータがそれぞれあり、他に非文字資料そのものに関するメタデータが存在する。そこで、本研究では、Dublin Core の文化人類学分野での応用として、民俗文化研究で必要とされる非文字情報のメタデータを付与する。

本稿で、付与した地域遺跡、郷土遺構、文化遺物のメタデータを以下に示す。

- 地域遺跡、郷土遺構、文化遺物の名称
- 地域遺跡、郷土遺構、文化遺物の種類
- 地域遺跡、郷土遺構、文化遺物の年代
- 地域遺跡、郷土遺構、文化遺物の存在場所（遺跡番号、所在地郵便番号、公共座標）
- 地域遺跡、郷土遺構、文化遺物の周辺関連情報

2——提案拡張子 dcterms

現在、標準化が進められている Dublin Core は、それが属する要素が一意に定まるという関係で与えられている。非文字資料を情報発信するにあたり、ある詳細化限定子に対する要素が一意に決まらなければならないという関係は、非文字資料の体系化には適切でないと判断する。

Dublin Core に対し、内容をより詳細に厳密に記述することの意義は、文化人類学の研究成果を情報発信するにあたり、かなりの比重を占めることとなる。Dublin Core の基本15要素の記述を詳細かつ正確に行うには以下の3つの観点が挙げられる。

1. 要素の意味の詳細化 (element refinement)
2. 要素の値のコード化形式の統一化 (encoding scheme)
3. 要素の値の構造定義 (value structure)

図1.3 dcterms のデータモデル例

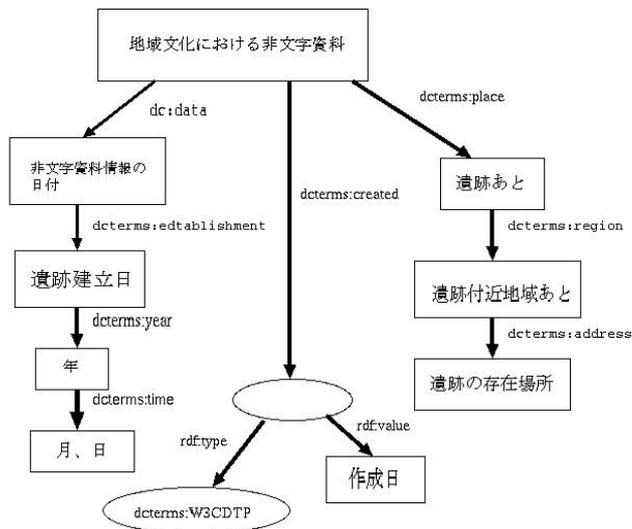
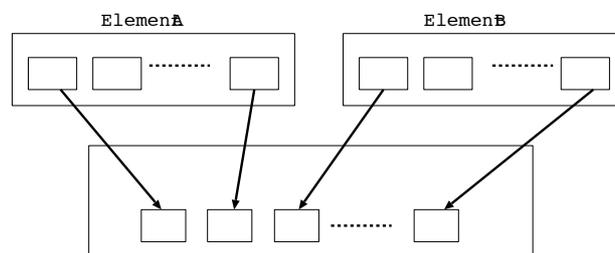


図1.4: Application dcterms



これらの記述要素を踏まえて本稿では dcterms と定義する。

3—— Dublin Core の応用分野への適用

Dublin Core にとってメタデータの相互利用性 (Interoperability) は最も重要な問題である。一方で、詳細事項は記載する限定子は研究分野によって必要性が異なるため、インターネット情報資源の記述のために共通に「必要な限定子」を決めることはそれほど容易なことではない。非文字情報が扱う研究分野毎に独自に限定子を定義することも考えられる。そうした環境におけるメタデータの相互依存性を情報資源の発見の立場から考えると、メタデータ記述の精度は落としても意味的に矛盾しないことを保証することの重要性が求められる。

Dublin Core で定義される15要素の意味定義と各々の要素を適用する際の構造を分離して捉えることで、メタデータの柔軟性を増すこととなる。以上によりメタデータの相互利用性を高めることが可能

となる。この構造を Application dcterms と定義する。

図 1.4 に Application dcterms の概略図を示す。複数のメタデータから、応用にとって必要な記述要素を選び、適切な構造的制約を与えるものである。非文字情報が持つ各々のメタデータに関して要素をばらして捉えることにより、応用毎に作成されたメタデータ間での意味的互換性が保たれる。

考察

提案概念に基づき、非文字資料を整理することで民俗学研究における、地域遺跡、郷土遺構、文化遺物それぞれは、異なるメタデータを必要とすることが分かった。これらは非文字資料に関するメタデータではあるが、体系化するにあたっては2つの方法が考えられる。

1つは、それぞれの非文字資料がすべての情報を保持する方法である。非文字資料に含まれている情報をすべて持つことで、資料としての整理がしやすくなる。反面、この方法では非文字資料に主体が置かれているため、応用が限られ地域遺跡や文化遺物などの情報に対して利用しにくいという問題がある。また各非文字資料が地域遺跡、郷土遺構のメタデータも保持するため、変更があった場合、統一的な修正が難しいほか、データベースが巨大化することとなる。

もう1つは、地域遺跡、郷土遺構、文化遺物、そして非文字資料に関する情報とに分割して、その関係を記述する方法である。この方法は、それぞれの非文字情報に属する情報を独立して管理出来る利点が挙げられる。一般に、一つの地域遺跡に対して郷土遺構、文化遺物は複数あり、メタデータもそれぞれに対応することとなる。また、地域遺跡などのメタデータはさまざまな応用が可能となる。その反面、関係を記述するため複雑になり、管理が直観的に行にくいこと、また、検索機能などが実用化技術として広く普及されてない段階にあるため効率の良い検索システムの構築が困難なことが挙げられる。

非文字情報の入力作業において、同じものを指し示す場合でも、時間軸や空間軸など、時空間推移に

よって、多少言葉が違う場合がある。これは文化人類学の研究では、当然のことで、自然言語で書く限り避けられない問題である。非文字資料メタデータ特有の語彙を洗い出し、分類を決めていく必要があるといえよう。

まとめ

本稿では、非文字資料の情報資源と情報流通の環境の現状を概観し、メタデータによる情報資源管理について述べた。その中で、Dublin Core を応用して非文字資料に適したメタデータ概念を提案した。

本研究により、非文字資料が持つ文化人類学的に意味のある情報（メタデータ）を明確化し、それを Dublin Core という国際標準に準拠した形で体系化を行った。また、文化人類学研究における非文字資料の体系化を図ったことにより、地域文化遺跡などの非文字資料が持つべきメタデータを明らかにすることが出来たと考える。本研究が、今後のメタデータを考える上での一つの参考事例になると考える。また、非文字資料に付与する情報を国際標準に準拠したメタデータとして体系化しているため、単に文化人類学の研究成果を情報発信しているに留まらず、国際的な規模での情報発信に対応していると位置付けられる。

民俗学研究における非文字資料は、これまで研究者や組織が保存、管理を行ってきた。非文字資料の多様さと柔軟性は、印刷物を中心に物理的な「もの」を扱ってきた従来の情報技術とは異なる技術を要求している。非文字資料を適切に組織化し利用者に提供することは、情報発信する際、重要な役割であり専門性が求められる点である。ネットワーク上で非文字資料を扱うには、資料を収集、蓄積、提供するという管理プロセスと資料を探し、アクセスするという利用プロセスのいずれにおいても、情報資源に関する情報、すなわちメタデータが必要とされる。また、多様な領域において色々な目的でメタデータが作られるため、情報資源を正確に管理するには、できるだけ詳細な記述のできるメタデータ規則が望

ましい。一方、情報資源の発見には、詳細記述は必ずしも求められない。ネットワーク上には多様なシステムが依存することを前提にしなければならないので、メタデータの具体的な記述方法や実現形式を認める必要があるといえよう。

大量の非文字資料を電子化し、情報発信を実用に運用開始した場合には今回定義したメタデータだけで十分であるかどうかを検証する必要がある。これまで、研究資料の多くは、当該研究者のみが管理・利用することがほとんどであったため、文化人類学的には非文字資料のメタデータとして何が必要か、適切なものの検証が必要不可欠である。

本稿は、日本の地域文化の歴史・変遷における非文字資料を中心に研究を行ってきた。日本の民俗学では、研究対象とする時代背景、地域文化の特色によって必要なメタデータが異なる。地域文化や歴史が異なった時、具体的にどのようなメタデータを追

加しなければならないのかは、今後、研究対象を拡充する際の課題となる。また、情報資源管理に関しては、さらに非文字資料を充実して民俗学資料のみに留まらず、他の文化人類学との協調を行っていく必要がある。情報流通管理に関しては、情報通信が双方向性を持った場合の機構管理を行う必要がある。

非文字資料の情報発信の実現には、さまざまな対象に関するいろいろな視点からの情報を利用できることが求められている。また、さまざまなタイプの利用者に適した利用環境を作り上げることが求められる。情報資源や情報流通に関する技術革新は激しく、これらを前提とする技術の見極めが困難である。近年、特に技術の専門化が進み、それぞれの分野のみでの研究が行われ、協調関係が欠けることがある。本研究は、これまで協調のなかった、メタデータによる情報資源の流通と管理との有効性を提案した点でも意義があると主張する。

【参考文献】

- (1) Dublin Core Metadata Initiative.
<http://www.dublincore.org/>.
- (2) Weibel,S., Kunze,J., Lagoze,C., M.Wolf: Dublin Core Metadata for Resource Discovery, RFC2413 (2002).
- (3) Dublin Core Metadata Element Set, Version1.1: Reference Description, 2003-06-02 [2004-11-11]
<http://www.dublincore.org/>.
- (4) Dublin Core Metadata Initiative, Dublin Core Qualifiers.
<http://dublincore.org/documents/dcmes-qualifiers/>.
- (5) Dublin Core Metadata Initiative, DCMI Libraries Working Group.
<http://www.dublincore.org/groups/libraries/>.
- (6) World Wide Web Consortium: Resource Description Framework (RDF) Schema and Specification 1.0. W3C Candidate Recommendation (2000).
<http://www.w3c.org/TR/rdf-schema/>.
- (7) World Wide Web Consortium: Extensible Markup Language (XML).
<http://www.w3c.org/TR/REC-xml>.
- (8) Diane Hillmann: Using Dublin Core (2001).
<http://www.dublincore.org/>.
- (9) Berners-Lee, Tim: The Semantic Web, Scientific American, Vol.284, No.5, pp.34-43 (2001).
- (10) W3C.Resource Description Framework (RDF).
<http://www.w3.org/RDF/>.
- (11) INTAP, 平成15年度 Semantic Web 技術の調査報告書(2003).
- (12) W3C, OWL Web Ontology Language Reference.
<http://www.w3.org/TR/owl-ref/>.

II Context 間の関連性を表現するメタオントロジー

はじめに

近年における情報化社会の発達は、社会全体の情報化と情報利用の変化をもたらした。従来では限られた世界・領域において情報の供給者と消費者の関係が明白であったが、現在の情報流通においては、多方向性と多領域性が特徴として挙げられる。扱われる情報と利用されるメディアは多種多様となり、それらに対応した情報技術の発達は、ネットワーク上でのさまざまな活動を可能にしている。情報の電子化技術とネットワークの発展・普及により、膨大な情報資源が電子化されている一方、利用者の目的に応じて、必要な情報を獲得・利用するためには、情報処理技術の発達だけでは解決できない問題も起きている。また、Web上に蓄積され続ける膨大な情報資源を適切に記述し、利用者にとって有効な情報を検索出来るようにすることも求められている。現在のような情報網拡大化と膨大な情報が溢れる環境において、有効な情報検索を行うためには、高度な知識処理が必要とされる。しかしながら、コンピュータはWebに存在する情報を蓄積・表示・分類したりするが、それらの情報を単にデータとして扱うだけで、情報が意味するものの理解を要するような処理をすることは出来ない。Web上の情報を理解し、利用するためには、人間の知識が必要とされる。このような背景のもと、Webコンテンツに意味情報を付与することにより、Webの有用性を飛躍的に高めようとするセマンティックWebが注目されている⁽⁹⁾。意味情報を表現するための方法として、オントロジーが用いられる。オントロジーを活用することで、ネットワーク上での情報共有・情報流通が進むことが期待されている。オントロジーはセマンティックWebを支える中核基盤の一つであり、オントロジーの概要を理解することは、知識処理の問題を扱う上で重要であると考えられる。オントロジーとは、「知識システムを構築する際の構成要素として用いられる基本概念・語彙の体系」と定義さ

れ⁽²⁾、知識ベースを構築する背景となる情報を提供する。したがって、オントロジーを参照することで、知識の理解が容易になり、対象世界をオントロジーによって構築することで、知識の共有・流通に大きく貢献すると考える。しかしながら、オントロジー構築に用いられる基本的な概念に関しても、その定義やオントロジー構築の扱いについても、明確な規範が定まっていないのが現状であり、オントロジー構築を困難にする一因となっている。そのため、対象世界の知識を取り扱う方法論の構築が望まれている。

現在、神奈川大学21世紀COEプログラム「人類文化研究のための非文字資料に体系化」では、民俗文化を知的文化遺産として研究成果を電子化し、情報発信することを目指している。本稿で扱う非文字資料とは、文字媒体として記録されることなく受け継がれてきた人間の営みである。文字に表現されることなく継承されてきた人類文化の歴史は、人間の観念・知識・知恵・行為など幅広く、文字媒体として記録され続けてきた事象とは異なる無形の文化である。文字に表現しえない人間の諸活動を資料化し、体系化することで、民俗文化の変遷を追及することが可能となり、民俗学研究的の貴重な資料となると考える⁽³⁾。しかしながら、民俗学分野の特徴として、研究資料の多数を当該研究者のみが、管理・保管している例が多い。その結果、民俗学研究における研究資料の多くが、個人のものとしてだけ存在し、広く流通されていないという問題が挙げられる。このような状況のもと、非文字資料を情報発信する際の情報共有・情報流通において以下のような問題が生じる。

1. 研究資料間の関連性が不明確である
2. 研究者間での研究資料の相互利用・相互運用が困難である

これらの問題の解決方法として、研究資料の流通を容易にする構造や項目についての標準化を図ると

ということが重要な要素となる。そこで本稿では、非文字資料に適したオントロジー構築を行うことで、非文字資料の共有と流通を図る。

本稿では非文字資料の一事例として、福島県只見町の只見方式と呼ばれる民具カードを取り上げる。民具とは、その使われ方によって意味を成し、目的や使用方法において、それぞれの役割を担っている。それゆえ、以下のような特徴を持つ。

《民具間同士の関連性が曖昧である》

上述した特徴により、民具間の関連性を把握するには、民具の用途に着目する必要があると考える。そこで、民具の用途をObjectの属性とし、用途をObjectの「目的」に関する概念と「方法」に関する概念に分け、Object同士に関連性があるContextを結び付ける。その結果、民具の用途がContextによって関連付けられると考える。また、Context間の関連性をオントロジーにより明示化することで、新たな知見・関連性の発見支援に大きく貢献すると考える。本稿では、Context間の関連性に着目し、民具の用途をオントロジーによって明示することで、民具の分類や分析を支援するための民俗学分野に特化したオントロジー構築を行う。

以下、2.2では民俗学研究のための情報発信の意義について概説し、2.3では非文字資料の情報発信における神奈川大学COEの取り組みについて述べ、2.4ではオントロジーとシソーラスについて述べる。2.5では本稿で扱う非文字資料の定義とその概略を述べ、2.6では本稿で取り上げた非文字資料の一事例である民具カードおよび、その問題点について触れる。続く2.7では民具に適したオントロジー構築を行い、2.8ではContext間に着目したオントロジー構築を行った。また、2.9では今回示したオントロジー構築における考察を述べ、最後に2.10で本稿の総括と今後の検討課題を論じる。

民俗学研究のための情報発信

従来の民俗学分野における研究手法として、聞き取り調査、フィールドワーク、文書などの記録の観察、建築や日用品から民間伝承までさまざまな事物

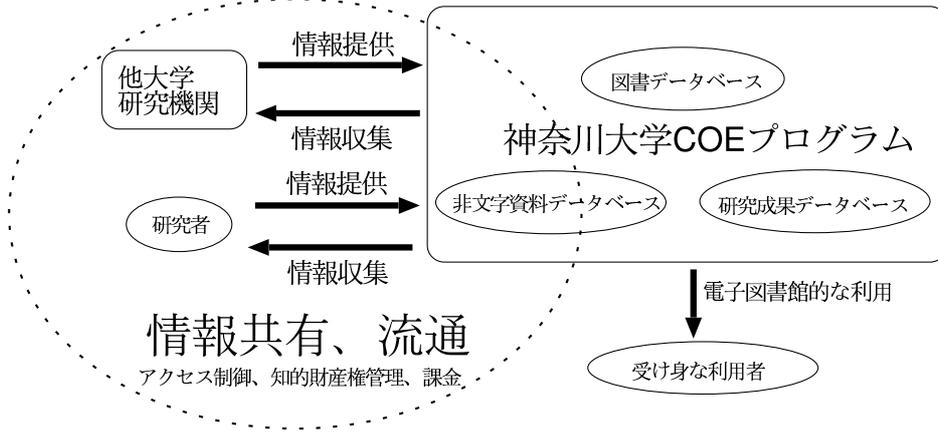
の観察などが用いられる。また歴史学、文化人類学、社会学や宗教学などと密接に関連し、ひとつの研究が民俗学とそれらの領域にまたがるものになっていることもある⁽⁴⁾。また、研究資料の多数を当該研究者のみが個人的に管理・保管している例が多い。その結果、民俗学研究における情報資源の共有ということを目指して作られるべき研究資料の多くが、個人のものとしてだけ存在し、広く流通されていないという問題が挙げられる。研究資料の流通を容易にするためには、研究資料の標準化を図ることが重要な要素となると考える。しかしながら、民俗学研究においては、経験則からの思考や論理が大きな比重を占めること、また研究手法が個別的かつ非定型であることから「標準化」ということに対しては問題がある。したがって研究資料の標準化ということが民俗学研究の「均質化」に繋がる恐れがあると考える。民俗学分野における「研究の均質化」は研究の質の低下に繋がるもので、避けるべき要素である。しかし、何らかの標準化が図られなければ、研究資料の共有と流通が図られない。特に、インターネットの普及により、このような研究分野においても情報発信や研究資料の流通、さらには研究の支援システムとしてIT技術の導入が必要不可欠となる。

以上における背景から民俗学分野に情報工学的な手法を取り込み、情報発信する研究環境の構築が望まれていると考える。また、情報工学的手法を取り込むことにより、客観的解析手段を得ることによって研究の裾野が広がると考える。

神奈川大学COEの取り組み

現在、神奈川大学21世紀COEプログラム「人類文化研究のための非文字資料の体系化」では、非文字資料の情報共有と情報流通を目指している⁽⁵⁾。COEが保有するデータベース（以下DBとする）（図書DB、非文字資料DB、研究成果DB）を、他大学や研究機関および、研究者間で相互に情報提供・情報収集することによって、情報発信が成されることが考えられる。情報発信の実現には、さまざまな対象に関する視点からの情報利用が出来ることが求めら

図2.1 情報発信概略図



れている。また、利用者に適した電子図書館的な利用環境を作り上げる必要もある。非文字資料の情報共有・情報流通には、プライバシー保護のためのアクセス制御や、著作権保護のための知的財産権管理、情報提供のための課金が含まれる。

図2.1に、非文字資料の情報発信概略図を示す。

オントロジー

本節では、本稿の中心となるオントロジーについて、その定義およびオントロジーの構成要素について概説する。また、従来研究のシソーラスおよび、その問題点について述べる。

1— オントロジーの定義

オントロジーとは本来、哲学用語であり「存在に関する体系的な理論（存在論）」という意味であるが、情報工学の立場からは「概念化の明示的な記述」と定義される⁽⁶⁾。計算機によって実世界をモデル化する時、我々は実世界に存在する（興味ある）概念の存在を認識し、概念として抽出する「概念化」を行っている⁽⁷⁾。この時、他の概念との違いや関係を同定することで、その概念を特徴付けて、その概念の意味を把握しているものと考えられる。通常は、無意識的・暗黙的であるこのような「概念化」を明示的に記述したものがオントロジーと定義されている。一方、工学的オントロジーは、計算機にも理解可能であるようにすることを目的としている。本研究では、非文字資料の共有と流通においてオントロジーの果

たす役割に着目し、「従来、知識や情報の背後に暗黙的に存在していた基盤概念を分散化し、組織化したもの」と捉える。オントロジーを活用することで、広い範囲の知識と、それらの関係性を規定する役割を果たすと考える。

2— オントロジーの構成要素

オントロジーは対象世界を説明するのに必要な概念と、それらの概念間の関係から構成される。以下に本稿で用いるオントロジーの構成要素を示す⁽⁸⁾⁽⁹⁾。

1. is-a 関係

概念の体系を記述する際の基本関係として用いられる。上位・下位関係を表す is-a 関係は、概念の一般化—詳細化の関係を表している。

2. part-of 関係（以下 p/o とする）

ある概念と、その概念を構成している部分にあたる概念（部分概念）を表す。

3. attribute-of 関係（以下 a/o とする）

ある概念を構成している属性情報。

4. instance-of 関係（以下 i/o とする）

概念とその具体例（instance）との間の関係を表す。

3— シソーラス

シソーラスとは、単語から概念を階層的に分類し構築することをいい、意味処理を行う唯一の言語資源である。単語から概念空間を構築するには概念の上位・下位関係を詳細に記述するシソーラスが必要となり、一般的に概念全体は木構造の形に体系化さ

れる。⁽¹⁰⁾言語に依存しない意味的な情報を扱うシソーラスの代表的な概念体系辞書として、「EDR電子化辞書」と呼ばれる大規模な機械処理用辞書が挙げられる。⁽¹¹⁾⁽¹²⁾EDR電子化辞書とは、コンピュータが自然言語を理解・生成するために必要な情報を処理しやすくする辞書であり、情報検索に用いることが出来る。

4——シソーラスの問題点

シソーラスにおいて、概念は単語との対応および他の概念との関係によって規定される。しかしながら、概念構造は複雑であり、単語同士の対応関係からではすべての単語の概念的関連性を見地することは出来ない。以下にシソーラスの問題点を示す。

1. 構築されたシソーラスは単語の類似性の一面のみ示す。
2. 連想性を表現出来ない

上述に示した問題点から、シソーラスは一般的に単語の類似性の判定と、単語の凡化に使われる。しかしながら、本稿で扱う非文字資料は、同一単語においても異なる概念が複数存在する。したがって、民俗学分野に特化したシソーラスが必要となる。

非文字資料

本節では、本稿で扱う非文字資料の定義を述べ、非文字資料に適したメタデータ生成および、オントロジーとシソーラスの重要性について概説する。

1——非文字資料の定義と情報発信

本稿で扱う非文字資料とは、神奈川大学21世紀COEプログラムが保持する民俗学研究資料である。非文字資料によって読み解かれる民俗文化は、個々の知識が無意識的に関係し合うことで、相互に変化していくものとする。すなわち、非文字資料とは、文字媒体として記録されることなく受け継がれてきた人間の営みに関する民俗学特有の研究資料である。非文字資料は、メタデータも包括して一つのデータとして重要な情報を含んでいる。たとえば民具の場合、民具の寸法や材質、使用法などの研究対象

に付随されるデータとが混在し合うことで、民具が再構成される。したがって、本稿で扱う非文字資料とは、メタデータからのみ、真のデータを推測することで、再構成されるものであると考える。

本稿で扱う非文字資料の具体例として本を取り上げる。書誌情報に価値がある場合、既存メディアで固定可能であるため非文字資料ではないと判断する。一方、装丁や作成された時代背景・技法に価値がある場合、既存メディアでは固定不可能であると考え、本稿で扱う非文字資料となる。非文字資料の情報発信においては、既存メディアで表現することが前提となるが、上述に述べたように既存メディアで固定するだけでは不十分である。よって、以下の2点を行う必要がある。

1. 既存メディアに写像・変換するためのモデル化・定式化を行う
2. もとの非文字資料を出来るだけ忠実に再現するためのメタデータ生成を行う

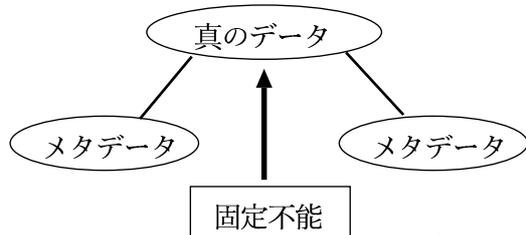
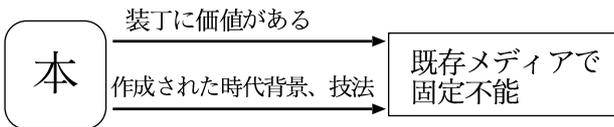
上述の2点を行うことで、非文字資料の情報発信が行われると考える。

図2.2に非文字資料の定義と一事例を示す。

2——非文字資料のメタデータ生成

メタデータを記述するための代表的なメタデータ規則にDublin Coreが挙げられる。⁽¹³⁾Dublin Coreは主に、図書館に所蔵されている図書目録のように、情報資源の対象物が既存メディアで固定可能であることを前提としている。また、情報資源の発見のメタデータであり、意味的な相互利用・運用を図るものである。しかしながら、本稿で扱う非文字資料は、民俗学分野特有のメタデータから再構成されるものである。よって、既存メディアでは固定不可能であるため、Dublin Coreでは記述不能であると考え。非文字資料は、メタデータと真のデータの関係性から推論することで、再構成されるものである。したがって、非文字資料に適したメタデータ生成を行うことで、非文字資料の再構成が成されると考える。

図2.2 非文字資料の定義と一事例



3—非文字資料に適したオントロジーとシソーラス

シソーラスに記述される概念階層は、単語の類似性から判断するため、一般的な概念のみを扱う。したがって、暗黙的な概念によって構成される非文字資料を表現するためには、十分な概念階層ではないと考える。そこで、暗黙的な概念や恣意的な意味が混在する非文字資料を適切に表現するためには、非文字資料に適したオントロジーを導入する必要がある。従来のシソーラスが扱う概念に対応する単語の類似性から、オントロジーによって関連性を導出することで、非文字資料に適したオントロジーが構築されると考える。

民具カード

本節では、本稿で扱う非文字資料の一事例である民具と民具の問題点について述べる。

1—非文字資料と民具

本稿では、非文字資料の一事例として民具を取り上げる。民具とは、先人が工夫し編み出してきた生活用具であり、古くからの生活の歩みを伝える民衆の文化財である。先人の創意工夫が民具の一点一点から読み取ることが出来、衣食住から信仰・儀礼に至る生産活動・生活ぶりが明らかになる。したがって、民具を基礎判断材料にすることで、今後の生活や地域作りに役立つと考える。民具を知ることにより、人間の営みや生活を追求することが可能となる。また、民俗文化伝承として受け継がれてきた民具を

後世に継承していく義務があると考え、非文字資料を電子化し、情報発信していく有用性が多分にあると判断する。

2—民具カード

本稿で取り上げる民具カードとは、福島県只見町に残されている民具を実測した記録カードである。民具を実際に使用した人が直接、カードに記録するという点で学術的な研究対象としても価値が高く、只見方式と呼ばれ国の有形民俗文化財に指定されている。民具カードは、客観的に実測された記録であると同時に、使用者による主観的な情報も含んでおり、只見地方の民俗資料として詳細に記述された貴重なデータである。また、経験や知恵を伝承していく上でも、資料価値の高い文化財的価値を持つ。民具カードは、表裏両面に記載されており、民具の用途は主に、その他の項目に書かれている。

図2.3にバコウグワの民具カードの一例を示す。

3—民具の事例

本節では、民具の事例として「バコウグワ」、「マグワ」、「クワ」および「タゴシラエノミヅヨケ」を取り上げる。以下に、本稿で取り上げる民具の用途例を示す。

4—民具の問題点

民具は、その使われ方によって用途がさまざまである。用途は、Objectの属性として表現される。し

—— バコウグワ ——

牛や馬に引かせて、田の土をおこす農具として用いられる。
牛や馬の口元に、はな竿を付けて子供（はなどりっ子：小学3、4年生）が、かじ取りをし、大人がバコウグワを押した。
慣れて上手になると、はなどりっ子なしでも作業出来るようになる。

—— マグワ ——

牛や馬に引かせて、田の土を砕く農具として用いられる。
牛や馬の口元に、はな竿を付けて子供（はなどりっ子）が、かじ取りをし大人（マグワオシ）がマグワを押した。
牛や馬が上手く歩いてくれないと、はなどりっ子はマグワオシに大声で叱られた。

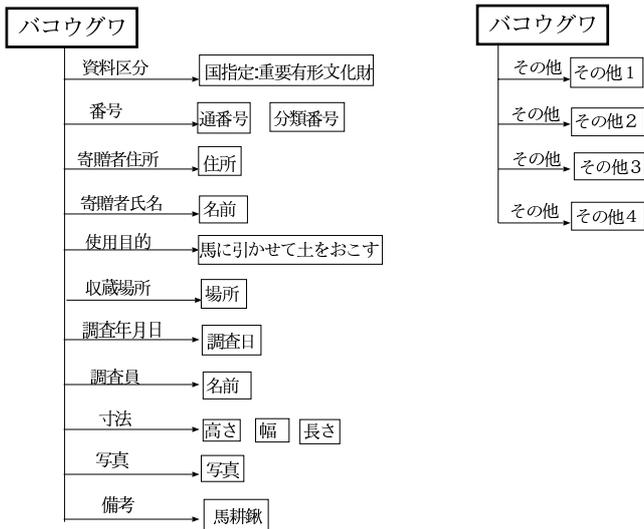
—— クワ ——

福島県只見町では、埋葬の時、墓場を掘る道具として用いられる。

—— タゴシライノミヅヨケ ——

田ごしらいの時の水よけとして用いられる。
クワの柄に付けて使った。

図2.3 民具カード



かしながら、ObjectやObjectの属性値が多岐にわたり、かつモデル化が成されていないため以下のような問題点が生じる。

《民具間同士の関係性が曖昧である》

民具間の関係には目的は違うが使われ方が似てるもの、またカード情報として共通する項目など、一見だけでは民具間の関係性が曖昧で分かりにくいものが多い。

そこで本稿では、用途をObjectの「目的」に関する概念と、Objectの「方法」に関する概念から、民具に適したシソーラスとオントロジーによって民具間の関係性を関連付けるオントロジー構築を行う。

民具に適したオントロジー構築

本節では、民具カードの基本となる項目をオントロジーで記述す。また、民具の用途を「目的」と「方法」に分離して、オントロジーの構築を行う。

1——基本項目のオントロジー

民具カードは、以下に示す3つの基本的なContext情報から成立している。

1. 民具の性質に関するもの（寸法）
2. 分類・整理に関するもの（番号）
3. 民具の用途に関するもの（目的・方法）

図2.4 基本項目のオントロジー

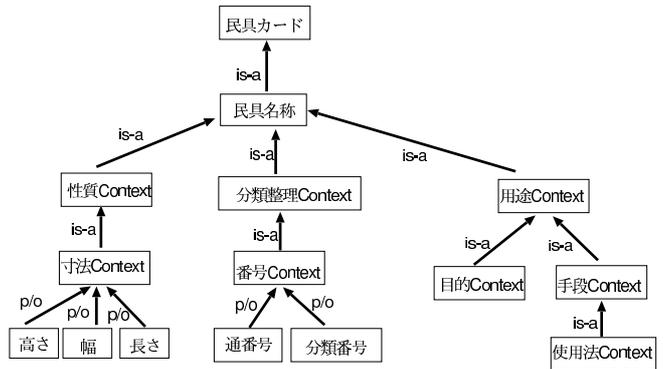


図2.5 目的のオントロジー

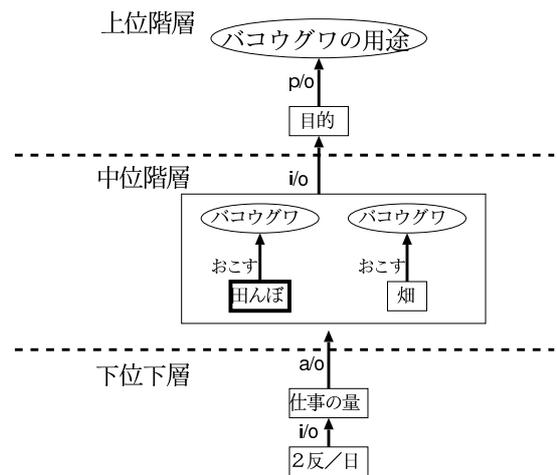


図2.4に民具に関する基本項目オントロジーを示す。さが高い。

2——用途におけるオントロジー

本節では、民具の用途を目的と方法の2つの観点からオントロジー構築を行う。

目的のオントロジー

図2.5にバコウグワの目的オントロジーを示す。

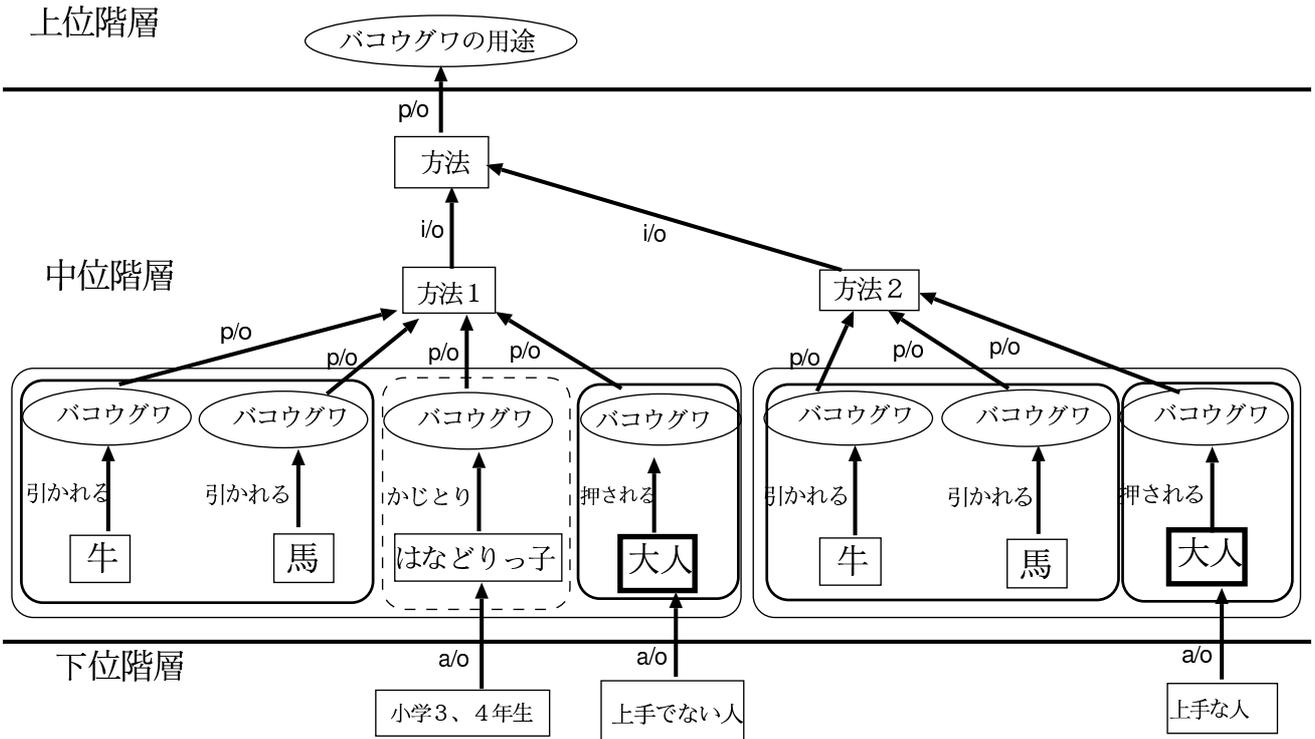
バコウグワは、田んぼや畑をおこす農具として使われ、一日の仕事量は2反とされていたことを表す階層構造を持った目的オントロジーである。階層構造を持たせたことで、一日の仕事量という属性情報が明示化される。

方法のオントロジー

図2.6にバコウグワの方法オントロジーを示す。

牛や馬がバコウグワを引く、はなどりっ子がかじを取り、大人がバコウグワを押すという行為から構

図2.6 方法のオントロジー



成されていることを表す階層構造を持った方法オントロジーである。階層構造を持たせたことで、はなどりっ子の存在有無が明示化される。

1. 田んぼを「おこす」という概念において関係がある。

以上の関係を結びつけたことにより、バコウグワとマグワの目的オントロジーに結びつきが出来た。

Context間に着目したオントロジー構築

本節では、民具の用途における Context の相違に着目し、民具間の関連性を明示化するオントロジー構築を行う。

1—— Context間のオントロジー

本節では、目的 Context と方法 Context 間に着目し、民具間の関連性を明示化する。

目的 Context によるオントロジー

図2.7にバコウグワとマグワにおける目的 Context 間による関係の一例を示す。

目的 Context のバコウグワとマグワにおける用途目的の Context に着目すると以下の関係が結び付けられる。

2—— 方法 Context によるオントロジー

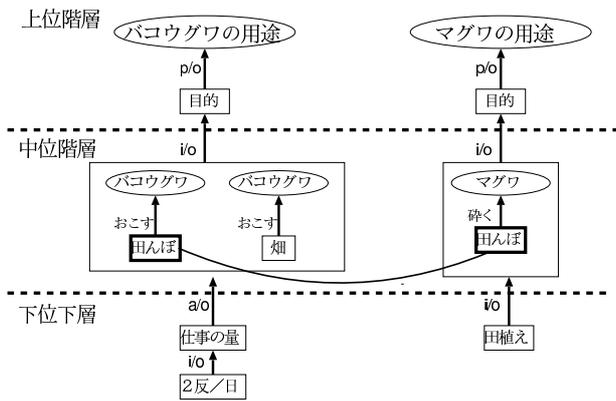
図2.8にバコウグワとマグワにおける方法 Context 間による関係の一例を示す。

方法1と方法2によるオントロジーのバコウグワとマグワにおける用途方法の Context に着目すると以下の3点の関係が結び付けられる。

1. 馬と牛に「引かれる」という概念において関係性がある。
2. はなどりっ子が「かじとりをする」という概念において関係性がある。
3. マグワオシが「押す」という概念において関係性がある。

以上の関係を結びつけたことにより、バコウグワとマグワの方法オントロジーに結びつきが出来た。

図2.7 目的Context間によるオントロジー



3— Context間に着目したオントロジー構築

本稿で取り上げた民具をシソーラスの観点より分類すると、バコウグワ・マグワ・クワは以下のような関係になる。

1. 広義語であるクワは、狭義語であるバコウグワ・マグワと上位・下位関係の概念体系で構成されている。

しかし、実際のクワの用途は、埋葬の時に墓を掘る道具として使われる。一方、田んぼで使われるタゴシラエノミツヨケは、用途方法としてバコウグワやマグワの柄に付けて使われるものである。上述のような事例からも、民具間の関係性をシソーラスによって関連付けることは困難である。したがって、民具の用途を目的と方法により、それぞれのContext間に着目しオントロジー構築を行った。その結果、タゴシラエノミツヨケの用途方法が、田んぼで使われるバコウグワとマグワの用途に深く関係していることが関連付けられた。

図2.9にContext間に着目したオントロジー構築を示す。

考察

本稿では、神奈川大学21世紀COEプログラム「人類文化研究のための非文字資料の体系化」が保有する非文字資料の情報発信の分析を支援するための民俗学分野に特化したオントロジー構築を行った。

本稿で対象とした非文字資料は、メタデータと真のデータの関係性から推論することで再構成される民俗学研究特有の性質を持つ。この非文字資料における特異性が、知識の共有と流通を図り難くする原因の一つであると考えられる。知識の共有と流通に必要なことは、その対象世界となるものの概念を明示しておくことである。そこで、知識の共有・流通に有効であるオントロジーを活用することで、非文字資料に適したオントロジー構築を行った。オントロジーは、対象世界を記述するのに必要な概念と、それらの概念間の関係記述から構成される。本稿では、民具におけるContext間の関連性に着目し、オントロジーとシソーラスを用いた民具に適したオントロジー構築を行うことで、民具間の関連性を導き出した。まず始めに、民具カードに記載されている項目を3つのContextからオントロジー構築を行った。次に、民具の構成要素である用途を目的Contextと方法Contextに分け、objectとobject同士に関連性があるContextを結び付けた。その結果、シソーラスでは関連付けることが困難であった民具の用途がContext間のオントロジーによって結び付けられた。民具は、その使われ方によって意味を成す。民具の用途がContextによって結び付けられた時、民具と民具の位相関係の記述が関連付けられると考える。さらに、Contextに階層構造を持たせたことで、民具における重要な要素である属性情報を抽出することが出来た。本稿で取り上げた民具は、さまざまな暗黙的概念や、恣意的な意味が混在するが、対象世界やContextが異なれば必要とされる概念は自ずと異なる。したがって、一般的な事象を対象とするシソーラスのみで表現することは困難である。よって、Context間の関連性に着目し、民具に適したオントロジー構築を行うことで、民具間同士の用途がさまざまな視点で切り出され、関連性を持ったオントロジー構築が成されたと考える。非文字資料に適したオントロジー構築を行ったことで、従来における各研究者間での研究対象の視点や興味の相違を吸収することが出来たと考える。また、一見関係なさそうな研究資料間の関連性の発見支援にも貢献出来たと考える。

図2.8 方法Context 間によるオントロジー

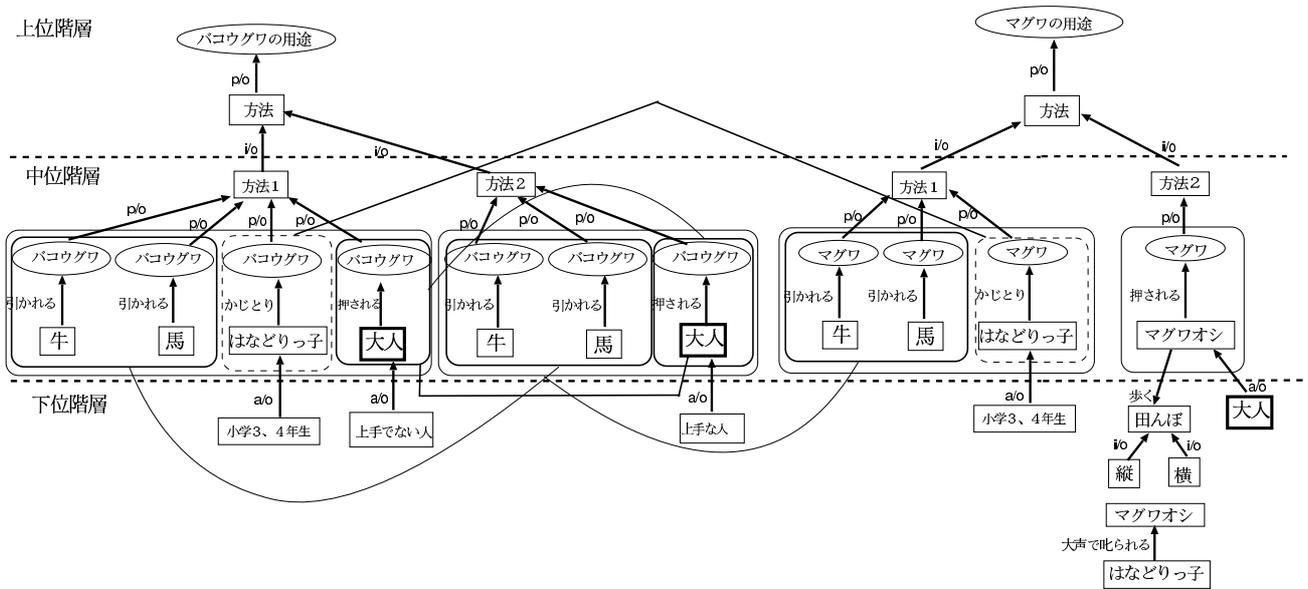
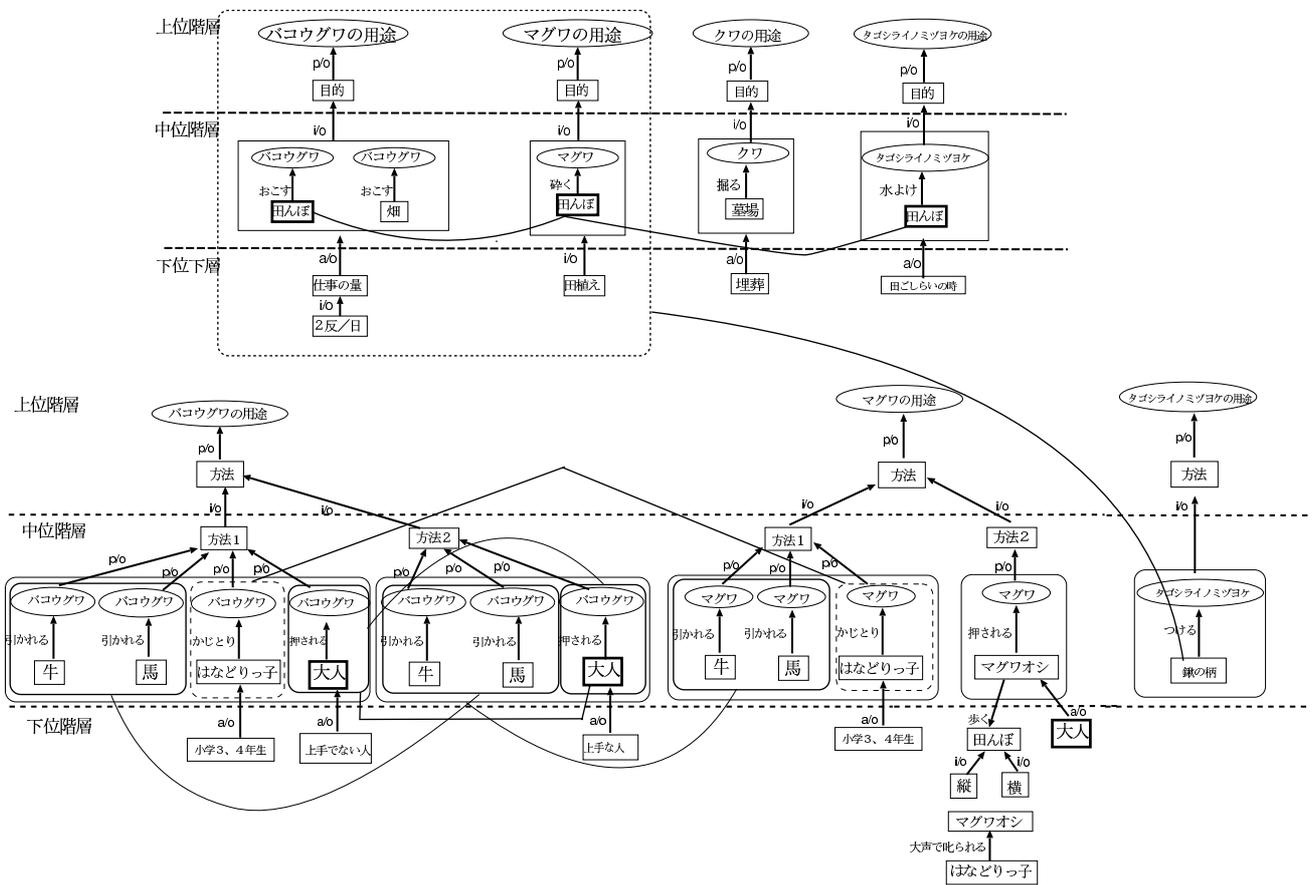


図2.9 Context 間に着目したオントロジー構築



まとめ

本稿では、非文字資料の定義とメタデータ生成を行った。また、非文字資料に適したオントロジー構築を行うことで、新たな知見・関連性の発見支援を

行った。民俗学研究における研究資料は、これまで当該研究者が管理・保管を行ってきた。民俗学分野に研究資料である非文字資料を情報発信するには、高い専門性が求められると考える。ネットワーク上で非文字資料を扱うには、資料を収集、蓄積、提供

するという管理プロセスと、資料を探しアクセスするという利用プロセスのいずれにおいても、情報資源に関する情報、すなわちメタデータが必要とされる。また、多様な領域において色々な目的でメタデータが作られるため、情報資源を正確に管理するには、できるだけ詳細な記述のできるメタデータ生成が必要であると考え。これまで、研究資料の多くは、当該研究者のみが管理・利用することがほとんどであったため、非文字資料のメタデータとして何が必要か、適切なものの検証が今後、必要不可欠であると考え。非文字資料の情報発信の実現には、さまざまな対象に関するいろいろな視点からの情報を利用できることが求められている。また、利用者に適した利用環境を作り上げることが求められる。情報資源や情報流通に関する技術革新は激しく、これらを前提とする技術の見極めが困難である。近年、特に技術の専門化が進み、それぞれの分野のみでの研究が行われ、協調関係が欠けることがある。本研究は、これまで協調のなかった、民俗学分野に情報工学的手法を取り込むことの有効性を提案した点でも意義があると主張する。

一方、知識の体系化を行うためには、基礎的概念を明確化し他の概念との関係を同定することで、概念構造において適切な位置づけをすることが必要である。オントロジーは対象世界を捉える際の基礎的概念と、対象世界との関係を明確化することで、さまざまな知識を関係付ける基盤としての役割を果たす。しかし、本稿で対象とする非文字資料には、同一対象においても異なる概念化が複数存在した。これは、対象世界での概念を情報化する社会的・歴史的・文化的背景などにより世界観が異なるためである。それぞれの立場が必要とする概念は膨大であり、立場を超えて統一して扱うことは困難である。非文

字資料におけるオントロジーの構築は、「対象世界の理解」および「対象世界を構成する概念の分析・組織化」の二つの過程を通して行われる。「対象世界の理解」では、構築するオントロジーの対象や利用目的に応じて必要な知識を収集し、オントロジーで記述すべき概念を明らかにする。一方、「対象世界を構成する概念の分析・組織化」では、抽出された概念を分析し体系的に組織化する。

本稿で取り上げた民具には、その用途により Context の相違や視点の依存性を持っていた。よって、民具の Context を峻別し、その用途を反映する形で民具間の関係性をモデル化するならば、視点や Context の変化を扱うための有効で一貫したオントロジー構築が出来る。そのため、Context 間に着目してオントロジーを構築することは、民具間の関連性の発見支援に繋がると考える。オントロジーにおいて、役割を概念化し、表現するためには、シソーラスに代表される単語の類似性から概念階層を構築する枠組みだけでなく、より高度な意味論まで考慮した枠組みが必要である。その意味においても、本稿で構築したオントロジーに含まれる Context 間の関係から、民具間同士の関連性を明示化することが可能になると考える。

本稿で示した民具に適したオントロジー構築は、民俗学分野に特化したオントロジー構築の一例である。民具における概念体系の構造化や民具間同士の関連性を明らかにすることは、従来暗黙的で明示されなかった非文字資料の情報共有・流通に重要な役割を果たすと考える。

今後の課題としては、Context 間の関連をモデル化し、大規模データベースにおける探索効率の向上を目指すことである。

【参考文献】

- (1) INTAP、平成15年度 Semantic Web 技術の調査報告書(2003)
- (2) Gruber, Thomas R.: A Translation Approach to Portable Ontology Specifications, in Knowledge Acquisition, Vol.5, pp199-220 (1999)
- (3) 神奈川大学21世紀COEプログラム研究推進会議-年報 人類文化研究のための非文字資料の体系化
- (4) 民俗学 Wikipedia.
<http://ja.wikipedia.org/wiki/>
- (5) 神奈川大学21世紀COEプログラム-非文字資料研究報告書

- (6) John, F. Sowa: Knowledge Representation: Logical, Philosophical, and Computational Foundations, Brooks/Cole (2000)
- (7) W3C Web-Ontology Working Group
<http://www.w3.org/2001/sw/WebOnt/>
- (8) 溝口理一郎、池田満：オントロジー工学序説：内容指向研究の基盤技術と理論の確立を目指して
- (9) 溝口理一郎、池田満、来村徳信：オントロジー工学基礎論-意味リンク、クラス、関係、ロールのオントロジーの意味論一、人工知能学会誌、Vol.14、No.6、 pp.8-100 (1999)
- (10) NTTコミュニケーション科学技術研究所：日本語語彙体系
- (11) 日本電子化辞書研究所：EDR 電子化辞書技術ガイド、EDR
<http://www.iinet.or.jp/edr>
- (12) EDR Home Page
http://www2.nict.go.jp/kk/e416/EDR/J_index.html
- (13) Dublin Core Metadata Initiative
<http://www.dublincore.org/>
- (14) 只見町民俗文化伝承事業報告書、2003年。—只見町の民具保全活用運動—

Ⅲ Covert Channel分析メカニズムとEJBによる情報フィルタの構築

はじめに

近年、オープンソースJavaプロダクトの充実により、Webサイトの情報閲覧や、インターネットの向う側にあるデータベースにWebを介してアクセスするシステムが、インフラとしての位置づけを着実に固めつつある。これにより新しく情報を創発し続ける社会システムを容易に構築可能であるが、反面、こうしたネットワークシステムからの情報漏えい対策が益々重要になっている。情報漏えい防止のためにはアクセス制御技術が導入されるが、導入に当たり、ポリシーなしで顧客ごとにパミッションを割り振るだけでは情報漏えいが生じる。したがってアクセス制御のためのフレームワークが必要である。本研究では、セキュリティモデルとして提案しているCommunity Based Access Control Model⁽⁴⁾に焦点を合わせた。我々の提案しているセキュリティモデルはコミュニティ・主体・客体の属性を定義し、その属性に基づいた主体、客体関係を記述する。記述にはRDF意味モデルを導入する。本稿では、EJBを用いたAccess Control Agent System⁽²⁾⁽³⁾⁽⁶⁾の構成について、設計コンセプトを示す。属性はRDFで表し、それをもとにアクセス行列に展開する。情報フィルタはこの情報、すなわち、主体と客体の競合、役割、階層、所有、プライベートの属性関係からCovert Channelを自動的に判定する。そして、属性を記したRDFデータベースにより情報フィルタルールを定め、パミッションを制御する。また、Access Control Agent SystemはSession Beanで実現する。Session BeanとEntity Beanを含むEJBにより、システムの構築を行う。Entity Beanはデータベース(RDFデータベース、アクセスルールデータベース、情報フィルタルールデータベース、ACLデータベース、XMLデータベース、リレーショナルデータベース)へのアクセスを行う。情報フィルタのCovert Channel判定のための記述はこれらのデータベース

に格納する。提案する機能を、データベースシステムのアクセス制御機能として実現することによって、システムの脅威となる情報漏えいを防止する。以下、3.2ではWebアプリケーションシステムとCovert Channelについて述べ、3.3ではAccess Control Agent Systemのアクセスルール・情報フィルタルールと情報フィルタの構成を示す。3.4ではEJBによるシステム構築について、Access Control Agent Systemの構造の観点から述べる。

WebアプリケーションシステムとCovert Channel

1—EJBの構造

JBOSSによるWebアプリケーションサーバの構築は、EJBを用いる。これにより、リモートメソッドの呼び出し、並行制御、セキュリティ面、データの永続性、などのシステムロジックがEJBサーバ側で自動化できる。また、ビジネスロジックの開発に重点をおくことができる。EJBに必要なホーム・ローカルインタフェースとディプロイメントディスクリプタの生成はXdocletを使うことによって短縮される。

2—アクセストリプルとCovert Channel

主体、客体、パミッションからなるアクセストリプルの集合、「アクセス行列」において、Covert Channelを定義する。

[定義] Covert Channel

アクセス行列において、アクセス禁止のパミッションに矛盾する情報フローをCovert Channelと呼ぶ。すなわち、

主体 S_i ($i = 1, 2, \dots$), S_j ($j = 1, 2, \dots$), ただし $i = j$

客体 O_n ($n = 1, 2, \dots$), O_m ($m = 1, 2, \dots$), ただし $n = m$

パミッション $P\{RW, \neg RW, \neg WR, \neg (RW)\}$, た

図3.1 EJB の構造

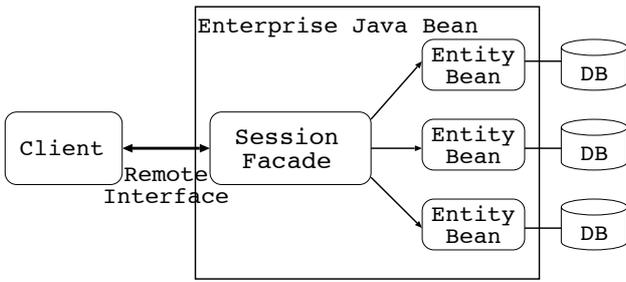
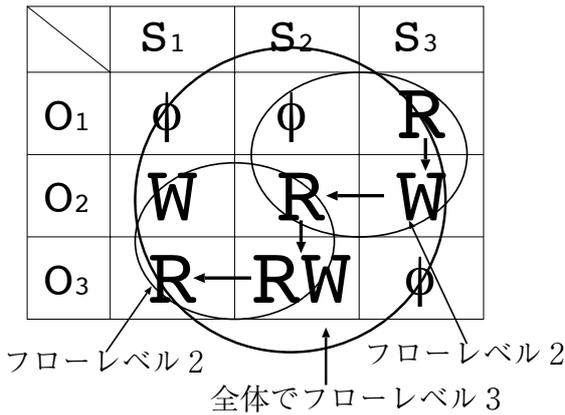


図3.2 Covert Channel とフローレベル



だし R(READ), W(WRITE) とする時、アクセストリプル

S_i, O_n, P について、

If S_i, O_n, ¬ R, And if S_j, O_n, R ∧ S_j, O_m, W ∧ S_i, O_m, R, Then Covert Channel (S_i, S_j, O_m(O_n)).

と定義する。この定義では、主体 S_j は客体 O_n と
言う名前とその内容を READ 禁止 (S_i, O_n, ¬ R)
であるにも拘わらず、主体 S_j が客体 O_n の内容を
READ し、客体 O_m にそれを WRITE し、主体 S_i が
O_m の内容から客体 O_n の内容を READ すること
(S_i, O_n, R) と表現される) によって、アクセ
ストリプル S_i, O_n, ¬ R と矛盾する結果が生じること
が示される。また、Covert Channel が引き起こさ
れる時、それに関わる主体の数によって Covert
Channel の程度 (フローレベル) を定義する。図
3.2 に Covert Channel を示す。

[定義] フローレベル

Covert Channel に関わる主体の数をフローレベル
と呼ぶ。

Access Control Agent System

1 — Access Control Agent System

システムコンセプト 属性を持った主体の集合を
コミュニティと定義する。各コミュニティではそれ
ぞれ内部で Covert Channel 分析を行い、Covert
Channel が起きないように制御する。外部のコミュ
ニティとの通信時における Covert Channel 分析は、
自コミュニティと、アクセス要求者およびその要求
について分析を行う。分析は、主体や客体の関係を
アクセス行列によって表現し、Covert Channel の検
出をする。このとき、アクセス行列の中から 2×2
の行列のパターンをすべて取り出し、フローレベル
2 の Covert Channel を分析する。

2 — Community、Subject、Object の定義

コミュニティの定義を次に示す。

[定義] コミュニティの定義

コミュニティとは、コミュニティの属性、コミュ
ニティに属する主体、および、コミュニティが管理
する客体とその属性の集まりから成る社会システム
である。

[定義] コミュニティの属性の定義

- (1) 競合：利害関係のあるコミュニティ間の関係
- (2) 連携：協調して情報を産出するコミュニティ
間の関係
- (3) 独立：競合、連携に属さない独立した無関係
のコミュニティの関係
- (4) 名前：コミュニティの名前

本稿では(1)、(2)、(3)を総じて競合と呼ぶことと
する。

[定義] コミュニティ内部の属性の定義

コミュニティ内で主体に割り当てる属性として、
役割、階層 (セキュリティレベル) と言う属性を持
つ。役割とは、コミュニティ内で組織的に決められ
るものであり職位などが相当する。主体はデータベ
ースに格納されている客体にアクセスする行為者で
ある。主体はアクセス行列では主体の名前や、その

主体のコミュニティにおける役割（部長、課長など）やセキュリティレベルとして表現される。主体の属性の定義を以下に示す。

[定義] 主体の属性

- 名前属性：主体名
- 競合属性：主体が属するコミュニティから刻印される競合属性
- 役割属性：主体の役割
- 階層属性：主体のセキュリティレベル
- プライベート属性：プライベートな情報に関する属性

客体はデータベースに格納する情報である。客体の属性の定義を以下に示す。

[定義] 客体の属性

- 名前属性：データベースに格納するための名前
- 所有属性：客体を所有する主体
- 競合属性：主体から刻印される競合属性
- 役割属性：客体にアクセスする主体の役割
- 階層属性：客体のセキュリティレベル
- プライベート属性：主体の個人情報と言う属性
- 記号属性：メタ言語、レトリックの区分

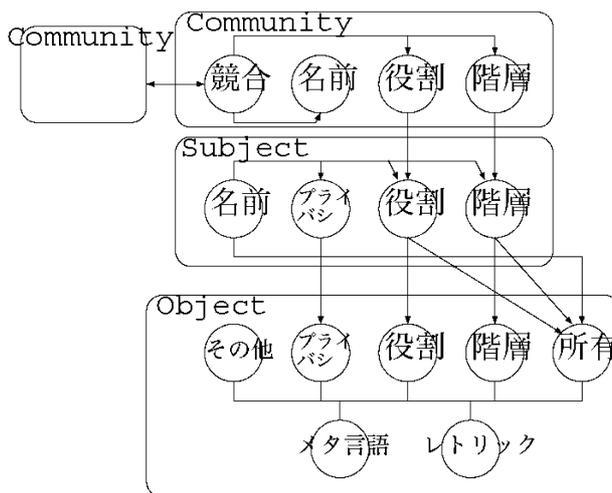
図3.3にコミュニティと主体、客体との属性関係を示す。

3— RDFによるセキュリティ構造の記述と Covert Channel分析アルゴリズム

RDFはSubject、Object、Predicateで一つのStatementを構成している。RDFの具体的な記述方法を以下に示す。ここでは主体S1の名前、所属しているコミュニティ、役割と、客体paperの内容、所属しているコミュニティ、所有している主体、どのような役割があるかを示している。

```
<?xml version="1.0" encoding="Shift_JIS" ?>
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
kxmlns:dc="http://purl.org/dc/elements/1.1/"
```

図3.3 Community, Subject, Objectの属性関係



```
xmlns:dc="http://purl.org/dc/elements/1.1/"
xml:lang="ja">
```

```
<rdf:description
rdf:about="http://ee.cs.kanagawa-u.ac.jp/cbac/">
<rdf:subject>
<rdf:attribute rdf:value="S1">
<rdf:name>Kino</rdf:name>
<rdf:community rdf:name="A_Company">
<rdf:federation>B_Company</rdf:federation>
<rdf:conflict>C_Company</rdf:conflict>
</rdf:community>
<role>Professor</role>
</rdf:attribute>
</rdf:subject>
<rdf:object>
<rdf:attribute value="paper">
<rdf:title>Academic_a-paper</rdf:title>
<rdf:community>University</rdf:community>
<rdf:ownership>Saka</rdf:ownership>
<rdf:role>Master</rdf:role>
</rdf:attribute>
</object>
</description>
</rdf:RDF>
```

RDF からアクセス行列（二次元配列）を作る方法

図3.4 アクセス行列

		役割			階層		所有		プライベート
		S1	S2	S3	S4	S5	S6	S7	S8
役割	O1	RW	R				R		
	O2		W			R	R		
	O3		RW		R				
階層	O4			R	RW			R	
	O5								
所有	O6								
	O7								
プライベート	O8								R

主体の属性、役割・階層・所有・プライベートをアクセス行列の列として書き下す。各列要素の属性間の関係はRDFで記述する。このように記述していくと、同一主体が重複する属性を持つこともあるが、それらを含めてすべてアクセス行列で表す。つまり、『S1という名前、学生という役割、セキュリティレベルという階層、S1が持っているObjectはO1』という属性を与えられている。

また、客体はメタ言語の各要素をアクセス行列の行として書き下す。XMLにおいて、型付ノード要素の構文を利用すると、本来RDFを念頭においていないXML文書も、RDFのデータモデルを適用して解釈することが可能になる。この構文を使うことによって、通常のXML文書をRDFとして処理をしたり、XMLベースのシステムをRDFベースに移行することが可能になり、データの自動連携や、再利用に適用できる。

各行要素の属性間の関係はRDFで記述する。主体と客体の属性の関係はRDFで表す。たとえば、所有については所有者と所有が刻印された客体という関係をRDFで記述する。このような状態のとき、アクセス行列は図3.4のように表わされる。

次にCovert Channel分析アルゴリズムを示す。

- (1) まずアクセス行列から ϕ を探す
- (2) その ϕ を固定し、同じ列（左右の横列）からR

かRWを探す

- (3) そのRかRWを固定し、同じ行（上下の縦列）からWかRWを探す
- (4) そのWかRWを固定し、同じ列（左右の横列）に最初固定した ϕ があつたら Covert Channel とする
- (5) 出力として Covert Channel 経路の4つの点を抜き出す

4—アクセスルール

主体、客体、パミッションから構成されるアクセス行列が与えられ、その間の間接的フローを分析した時、間接的フローがCovert Channelであるかどうかを判断する基準が必要である。それは実際、間接情報フローがあつたとしても、その情報フローが脅威である場合とそうでない場合がある。たとえば、間接的フローが競合関係の間で生じる場合はCovert Channelであるが、間接的フローが連携関係（競合でない）の間で生じる場合はCovert Channelではない。そこで、3.2で述べた属性の定義を用いてCovert Channelを判定するルールを定義した。

[Covert Channel判定ルール]

競合属性ルール 2属性がそれぞれコミュニティ名属性と役割属性の組合せを含み、かつそれらが競合関係であるならば、その2属性間において競合属性間のアクセスが禁止される。

階層属性ルール 一方が階層属性（セキュリティレベル）を含み、もう一方が階層属性を含まない時、かつセキュリティレベルが機密性を目的とするならば、階層属性からの客体流出が禁止される。また、一方が階層属性（セキュリティレベル）を含み、もう一方が階層属性を含まない時、かつセキュリティレベルがインテグリティを目的とするならば、階層属性からの客体流入が禁止される。

所有属性ルール 一方が所有属性を含む時、所有属性を含まない方からのWRITEが禁止される。また、

一方が所有属性を含み、所有関係に矛盾がない時、WRITE可能である。このとき、所有属性に対するREADのアクセスは、競合属性、階層属性に従う。

プライベート属性ルール 一方が所有属性を含む時、かつプライベート属性にアクセス許可された属性の値をもう一方が持つ場合、2属性間のアクセスが許可される。また、2属性間において、同一のプライベート属性の値を持つ時、2属性間のアクセスが許可される。

情報のやり取りをする関係をアクセス制御に関する属性別に分け、この定義に従ったアクセスの制御を行う。

情報のやり取りで矛盾があった場合、Covert Channelと確定する。

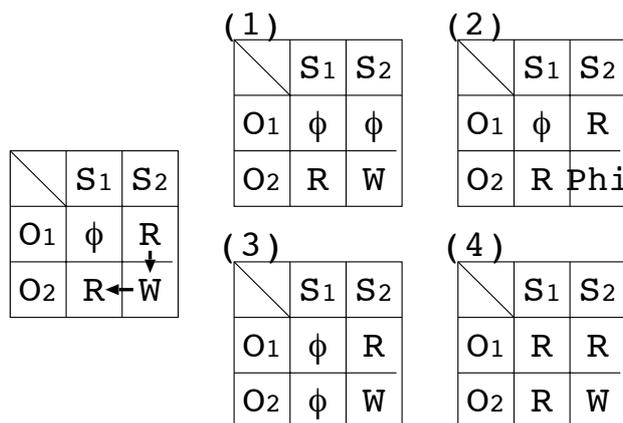
5——情報フィルタルール

主体、客体、パミッションから構成されるアクセス行列が与えられ、その中のフローレベル2のCovert Channelを分析した時、それをどう防止するかと言うことを判断する「情報フィルタルール」を必要とする。以下にそのルールを示す。また、情報フィルタのパターンについては図3.5に示す。

[情報フィルタルール]

- S1とS2が競合の場合(3)を選択する。
- S1、S2の役割によるアクセス制限のある場合、
 - 役割関係からO1、O2がS2によって産出される場合、(3)を選択する。
 - 役割関係からO1、O2がS2によって産出され、かつ、S1がS2をReadする必要がある場合、(1)を選択し、更に、S2を仮想的に追加設定する。S2はO1をRead、S2はO1を ϕ とする。ただし、S2とS2は同一主体であるとする。
- O1がS2のプライベート情報の場合、(3)を選択する。
- O1がS2の所有の刻印がしてあるならば、他のObjectに書き込んではいならない、かつそのObjectに書き込んではいならない(2)。
- セキュリティレベルが設定されている場合、

図3.5 情報フィルタルール



- S1 < S2のとき(3)を選択する。
- S1 > S2のとき(4)を選択する。
- O1 < O2のとき(3)を選択する。もしくは(2)を選択
- O1 > O2のとき(2)を選択する。もしくは(3)を選択
- S1 < S2かつO1 < O2のとき(3)を選択もしくは(4)を選択する。(O2にアクセス出来るならO1にもアクセス出来る)
- S1 < S2かつO1 > O2のとき(3)を選択する。
- S1 > S2かつO1 < O2のとき(4)を選択する。(S2がアクセス出来るならS1もアクセス可能、O2にアクセス出来るならO1にもアクセス出来る)
- S1 > S2かつO1 > O2のとき(4)を選択する。(S2がアクセス出来るならS1もアクセス可能)
- いずれにも該当しなければ(4)を選択する。

6——情報フィルタのアルゴリズム

初期の状態からアクセスがあり、情報フィルタが適応されるまでのアルゴリズムを以下に示す。

- (1) RDFで書かれたアクセス構造をアクセス行列に展開する。
- (2) Covert Channelレベル2候補をすべて検出する。
- (3) 候補に関してアクセスルールを適用し、Covert Channelかどうかを判定する。
- (4) もしCovert Channelならば情報フィルタルールを適用し、パミッションを変更する。

- (5)もし Covert Channel でなければ、パミッションはそのままとする。
- (6)アクセス要求があった場合はアクセス構造を追加し、アクセス行列を追加し、Covert Channel を分析する。

以下(1)~(5)と同じように処理をする。

7——情報フィルタアルゴリズムの考察

上記情報フィルタアルゴリズム(2)において、Covert Channelをすべて分析し、そのすべての候補に対してパミッション変更を試みるが、情報フィルタにより、パミッションを変更された場所から新たにCovert Channelが起きないようにしなければならない。そこで、記号学におけるシニフィアン・シニフィエの考え方に着目し、情報フィルタルールによって新たにCovert Channelが生じないことを示す。言葉は表現する対象に関する記号作用部(シニフィアン)と記号意味部(シニフィエ)に分離される。つまり、情報のタイトル(シニフィアン)とその内容(シニフィエ)に分け整理する。すなわち、言語を表現と言う面と内容と言う面に区別し、それぞれの面でアクセスについて考える。ただし、Covert Channel判定前のアクセス行列に対するアクセス要求を尊重した上で、Covert Channelがないようにしたいということを前提とする。

[証明]

- (1)S2はO1のシニフィアン、およびシニフィエに READしてはならない
- (2)S2はO2のシニフィアン、およびシニフィエに READしてよい

このときアクセストリプル{S2、O2、READ}がCovert Channelと判断された場合、『S2はO2のシニフィアンにREADしてはならない』が、『もともとS2はO2のシニフィエにREADすることが許可されていた』のでアクセストリプル{S2、O2、READ}を変更し、アクセストリプル{S2、O2、 ϕ }としたとき、アクセストリプル{S2、O2、 ϕ }はCovert Channelと判定される。

EJBによる Access Control Agent Systemの構築

1—— Access Control Agent Systemの構造

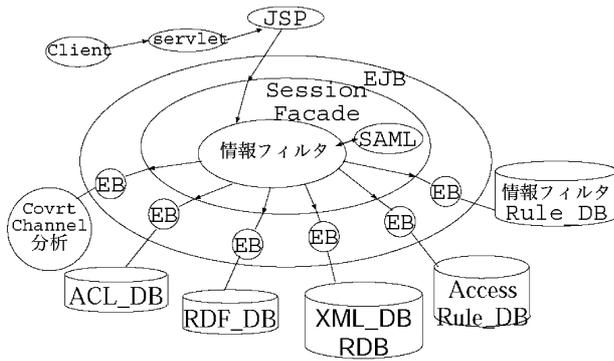
Access Control Agent Systemの構成を図3.6に示す。Access Control Agent SystemはSession Facadeパターンで構成されていて、これによりEJBの構成であるEntityの処理のパフォーマンスを向上させる。このシステムの動作は、まずクライアントから要求があり、JSP、サーブレットを介してEJBへアクセスがあるとする。EJB内ではSession BeanからEntity Beanへアクセスがある。そして、データベースへのアクセスはEntity Beanからあり、これによってデータベースの情報にアクセス出来るようになっている。データベースへのアクセスはRDFデータベース、アクセスルールデータベース、情報フィルタルールデータベース、ACLデータベースとなっている。

また、図3.7に情報フィルタの構成を示す。情報フィルタはSession Facadeパターンを構成する。情報フィルタBeanはまず、RDFデータベースから属性のセキュリティ情報を読み込む。そして、それをCovert Channel分析にアクセス行列として渡す。その結果、Covert Channelが検出された場合、そのCovert Channelの原因となっている4つの点を抜き出す。そして、その点をアクセスルールデータベースから情報を読み取り、Covert Channelかどうか判定する。Covert Channelだった場合、情報フィルタルールDBから情報を読み取り、パミッションの変更を行う。その結果を反映するために、RDFデータベースとACLデータベースに書き込む。

2—— EJBのシステム要求 (ACID)

アクセスルール、フィルタルール、Access Control Listを格納するデータベースはACID(原子性<ATOMICITY>、一貫性<CONSISTENCY>、隔離性<ISOLATION>、持続性<DURABILITY>)を満たさなければならない。これは、まとめて1つとして扱う必要のある一連の処理(トランザクション)

図3.6 Access Control Agent System の構造

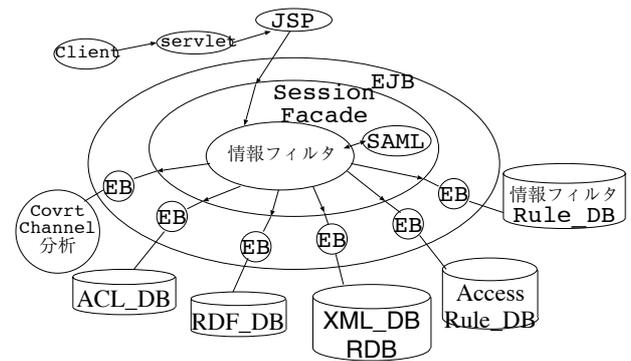


があるために、ACIDが必要になる。たとえば、あるユーザからアクセスがあり、データベースの更新を行っているとき、別のユーザからのアクセスがあった場合、データベースの更新や変更は互いに干渉しないように制御されないとならない。ACIDを満たすことで、多くのユーザが並行してデータにアクセスする場合や、システムに障害があった場合でもデータの一貫性を保ち続ける。

3— Session Facade による Session Bean

Session Beanが複数のEntity Beanのインタフェースに対して統一的なインタフェースを提供する。メリットはEntityの処理のパフォーマンスを向上させることである。特に、クライアントがアプリケーションサーバ外部に存在するリモートクライアントの場合に顕著に表れる。それは、Session Beanによってクライアント、サーバ間のネットワーク通信の数を減らすことができる。つまり、図書館システムのような大量のデータベースを扱うシステムにおいては、アクセス実行時間が増えていくので、Session Facadeパターンを適応することによってアクセス時間の増加を抑制する。Access Control Agent Systemは、情報フィルタルールDB、アクセスルールデータベース、属性間の構造データが格納されるRDFデータベース、アクセストリプルが格納されるACLデータベースと言った複数の、比較的大規模になるデータベースを統合してCovert Channel分析・制御するSession Beanである。この

図3.7 情報フィルタの構成



ような大量かつ複数のデータベースを扱うSession BeanではSession Facadeのアーキテクチャが必要になる。

4— アクセストリプルと DTO Factory

アクセストリプルの一部、たとえば客体とパーミッションはクライアントからEntity Beanにアクセスする。そしてアクセストリプルはSession BeanとしてのAccess Control Agent Systemの中の複数のスレッドにおいて使用される。したがって、アクセストリプルの修正変更を1箇所管理可能なアーキテクチャが必要となる。そこで、Entityで使用するデータオブジェクト（アクセストリプル）をクライアントとサーバ間で転送する単位とするDTO Factoryを採用し、アクセストリプルをDTOとして定義する。

むすび

競合、役割、階層、所有、プライベートと言う属性の構造からアクセスルール、また、情報フィルタルールを導いた。属性はRDF意味モデルに記述しておき、情報フィルタはEntity Beanを介してRDFから得られる情報をもとにアクセス行列を生成し、Covert Channelを自動的に判定する。また、パーミッションの制御をするように作動することを示した。セキュリティモデルはEJBの構造のSession Beanによって実現する。今後、EJBによる実装実験を行う。

【参考文献】

- (1) Role Based Access Control: ANSI, INCITS359-2004 (approved 19 Feb04)
- (2) 牛頭靖幸、森住哲也、稲積泰宏、木下宏揚：Covert Channel 分析評価のためのアクセス制御エージェントシステムの提案、コンピュータセキュリティシンポジウム2004 (2004) (3) 森住哲也、牛頭靖幸、畔上昭司、酒井剛典、稲積泰宏、木下宏揚、小柳和子：アクセス制御エージェントシステムによる安心・安全なWebアプリケーションシステム、学際的情報セキュリティ総合科学シンポジウム(2004)
- (4) 森住哲也、牛頭靖幸、畔上昭司、酒井剛典、稲積泰宏、木下宏揚：セマンティックWebシステムに於ける“Community Based Access Control Model”の適用に関する一考察、SCIS2005、3B1-4 (2005)
- (5) 酒井剛典、森住哲也、牛頭靖幸、畔上昭司、稲積泰宏、木下宏揚：Webアプリケーションシステムに適したセキュリティモデル“Community Based Access Control Model”の提案、SCIS2005、3B1-5 (2005)
- (6) 森住哲也、牛頭靖幸、稲積泰宏、木下宏揚：Covert Channel 分析評価のための Access Control Agent System の提案、JSSM 論文誌(2005)
- (7) 森住哲也、木下宏揚：セマンティックWebシステムのセキュリティモデル、技術と社会・倫理研究会(2005)
- (8) 森住哲也、木下宏揚：社会システムの中のCovert Channelについて、技術と社会・倫理研究会(2005)
- (9) 森住哲也、酒井剛典、畔上昭司、稲積泰宏、木下宏揚：機能分化的社会システムの属性に基づくセキュリティモデル、技術と社会・倫理研究会(2005)
- (10) 酒井剛典、森住哲也、牛頭靖幸、畔上昭司、稲積泰宏、木下宏揚：Covert Channel の性質に基づいた情報フィルタについて、IEICE 総合大会、A-7-9 (2005)
- (11) 酒井剛典、森住哲也、牛頭靖幸、畔上昭司、稲積泰宏、木下宏揚：Access Control Agent System における情報フィルタの機能、IEICE ソサイエティ大会、A-7-21 (2005)

Ⅳ トピックマップを用いた非文字資料における Context の表現

はじめに

民俗学における民俗資料とは、文字媒体として記録されている、いないにかかわらず、脈々と継承されてきた人間の営みであるといえる。こうした資料を研究の主題とする民俗学分野の研究手法の特徴として、研究資料の多くを当該研究者のみで管理している場合が多いことがあげられる。民俗資料には、文字資料として残せないものも非常に多く、そのことが資料の共有・相互利用を妨げている。これは、資料が多いほど手がかりが増えることを考えれば好ましくない状況であるといえる。こうした問題点により、研究資料の関係性が不明瞭となり、さらに研究者間での相互利用・相互運用が困難であるという状況が生じる。そこで、民俗学研究資料の情報共有・情報流通におけるこのような問題を解消するため、著者らは、21世紀COEプログラム「人類文化研究のための非文字資料の体系化⁽¹⁾」の一環として、民俗文化を知的文化遺産とし、研究成果を電子化、情報発信することを目指している。

これは、非文字資料そのものが持つ真のデータを文字媒体に記録・固定不能であることによると考えられる。まず最初に試みたのが、非文字情報が持つ情報をメタデータとして整理する方法を確立し、メタデータを用いた情報の流通、利用法を確立することであった。具体的には、非文字資料が持つメタデータを文字媒体に記録し、そのメタデータを用いて真のデータを再構成できると考えた。したがって、真のデータからメタデータを生成することが必要となる⁽²⁾。著者らは、非文字資料のひとつである民具資料について、Dublin Coreを用いたメタデータ生成を行い問題点を検証した。ここでは、電子図書館などで主に利用されているDublin Core⁽²⁾を応用して、非文字資料に適したメタデータを、文化人類学的に意味のある情報であるとして明確化し体系化を行っ

た。また、文化人類学研究における非文字資料の体系化を図ったことにより、非文字資料が持つべきメタデータを明らかにしたことにより、民俗資料を記述する表現力の観点から、メタデータを拡張しなければならないことがわかった。しかし、メタデータを拡張すると、相互運用性に問題が現れることがわかった。

すなわち、メタデータの拡張時の問題点からオントロジーを用いて非文字資料の共有と流通の発展を考えると、Dublin Coreは、メタデータの種類を限定することで、相互利用性を高めているので、意味の記述をわずかなメタデータで行うことは不可能であり、メタデータの意味を定義する必要性がでてくる。それを定義するものがオントロジーである。概念および概念間の関係を定義することで、メタデータが何を意味しているかをマシンで厳密に扱うことができるようになる。その結果、意味に基づいた知識処理、たとえば推論などが可能となる。非文字資料に適したオントロジーを構築し、知識共有・流通におけるオントロジーの有効性を示した⁽⁴⁾。オントロジー構築に関しては、基礎理論から民俗学にどう適用できるかを考えた。

非文字資料の情報資源と情報流通の環境を概観し、メタデータによる情報資源管理について述べた。

その結果、民具にはその用途によりContextの相違や視点の依存性を持っていることがわかった。Contextとは、一般にそのContextの相違や視点の依存性は、民具の持つContextの関係性をモデル化すれば扱うことが可能となり大規模な情報リソースにおける探索効率の向上が可能であると考えられる。

本稿では、情報の共有化を行うために「民具に適したオントロジー」をトピックマップにマッピングし、ContextおよびContext間の関係をモデル化することで、Contextの相違や視点の依存性を扱う知識処理が可能となると考えた。そこで、Contextを

表現する手法について考えられるものについて検討し、大規模な情報リソースにおける探索効率の向上に対する有効性を考慮した表現手法の提案を行っている。

以下、4.2では非文字資料の特徴および問題点について触れる。続く4.3では、本稿で用いる手段であるトピックマップについての説明およびその有効性について説明する、さらに、4.4ではトピックマップにおける非文字資料のContextの表現手法について述べる。また、4.5では今回の提案したContextの表現手法に対する考察を行い、最後に、4.6で本稿の総括および今後の課題について論じる。

非文字資料

本節では、本稿で取り扱う非文字資料の特長および問題点について述べる。また、従来研究との比較について概説する。

非文字資料とは、文字媒体として記録することのできない情報のことである。資料そのものだけでなくメタデータにも重要な価値のある資料があり、メタデータを含めた資料の全体像が真のデータであるとすれば、真の情報を表現することを目的としている。これらの情報は単純に文字媒体では表すことはできないが、情報発信のためには、資料情報をマシン上に表現しなければならない。そこで、真のデータを再構成可能であるような既存の媒体への写像・変換のモデル化を行う必要が生じる。対象の真の情報を包括的にモデル化する際に必要となるのが、オントロジーである。こうすることで、今まで表現できなかった情報の多くが固定可能となり、情報の次元が表面的な語彙のみによるものから厳密な意味に基づいたものへと飛躍的に高まる。

1——非文字資料とオントロジー

非文字資料の写像・変換の際に重要となるのがオントロジーである。オントロジーとは、哲学用語で存在論を意味し、存在とはどういったものなのかを述べた理論である。そこから派生して工学的には、オントロジーは「概念と概念間の明示」であるとさ

れている。つまり、文字そのもので存在を記述するのではなく、語彙による揺らぎのない概念を基底とした記述体系である。したがって、非文字資料をオントロジーに写像・変換することで意味に則した知識処理が可能となり、知識処理を用いて非文字資料の真のデータを間接的に保持することができる。

2——一般的なContext

ここでは、非文字資料に依存したContextを述べる前に一般的なContextの概念について説明する。一般に、Contextとは「文脈」あるいは「前後関係」、「背景」と直訳される。文字通りに、文中における論理関係の脈絡を表す。すなわち、コミュニケーションの場で使用される言葉や表現をて意義付ける背景や状況そのものを指している。相対的に定義が異なる言葉の場合は、コミュニケーションをとる2者の中でその関係性、背景や状況に対する認識が共有・同意されていることがコミュニケーションの成立の条件となる。このように、コミュニケーションを成立させる共有情報をContextというが、以上の共有情報をマシン上にモデル化する上で、知識処理におけるContextは言語的なコミュニケーションのみならず、たとえば、同じ時速10kmでも、徒歩Contextならば速い、自転車Contextならば普通、自動車Contextならば遅いというように評価が異なる。これは、速度という思考対象が、移手段というContextに依存していることを表している。

3——非文字資料におけるContext

ここで、本稿の重要な主題である非文字資料におけるContextについて、具体例を挙げつつ説明する。Contextとは、知識表現を行う対象世界において、普段我々が暗黙的に考慮しているものごとと考えられる。その暗黙的に考慮されているContextの違いによって、概念化を行う際に視点依存性などとして現れる。ここで、Contextの違いから以下の問題が生じる。

- 計算機的な意味の一貫性を保つことが困難である。

●異なる対象世界（領域）での相互理解・利用が困難である。

このように Context は、非常に曖昧な概念であるが、本稿において、Context とはあるトピックに接続されている出現、あるいは関連およびその先にあるトピックのこととする。このように定義することで、トピックマップにおいて、Context が非常に一般的なものとなる。

その定義を非文字資料にあてはめた例を以下に示す。

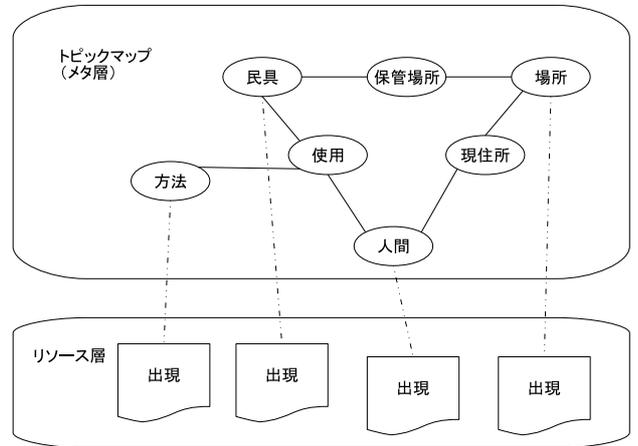
Context の具体例 使用者という概念を考えたとき、「使用者」は、「使用」という動作に関する Context と人間であるという存在に関する Context に属していると考えられる。同様に、使用者と使用目的との間には、「使用」という Context に属しているもの同士であるという関係がいえる。たとえば、民具の使用 Context 中について知りたいときに、製造 Context 中の情報を探索しても無意味である。したがって、適切な Context を設定してあげれば、関係性の高い結果が返ってくる上に、探索範囲も狭めることができるという利点が含まれる。しかし、不適切な Context を設定してしまうと逆にまともな結果が得られなくなる危険もあり、Context の表現が重要であることがわかる。

トピックマップ

本節では、本稿の手段となるトピックマップについて説明する。

トピックマップは、知識を記号化し、この記号化された知識を関係がある情報リソースに結び付ける技術である。図4.1に示すように、トピックマップは、論議の主題（Subjects）を表現するトピック（Topics）、主題間の関係を表現する関連（Associations）および主題と主題に関連する情報リソースを結び付ける出現（Occurrences）によって組織化される。ISO/IEC JTC1 SC34 WG3 で策定された標準（ISO/IEC 13250 Topic Maps）であり、現在では広く活用されている。タクソノミーやシソー

図4.1: トピックマップの概観



ラス、オントロジーなどと同じく知識表現の手段の一つである。主な特徴を挙げると、情報リソースと明確に分離されており、情報リソースに依存しないという利点や、自由に関連性を定義可能であり、自由度が高いという利点がある。また、トピックマップには標準で主題の有効範囲を設定できるスコープという機能や、同一の主題と同定できるものについては併合できるという機能が盛り込まれているのでネットワーク上で、セマンティックネットワークを共有する際に優位であるという利点もある。

1——トピックマップの主要な構成要素

図4.1に示すように、トピックマップはトピック、関連、出現の3つの主要な構成要素からなる。

●トピック (Topics)

人間が認識する具体的または抽象的な主題/概念を表す要素。

●関連 (Associations)

一つ以上の主題間の関係の表現。関連は役割を持っている。

●出現 (Occurrences)

主題および情報資源間の関係の表現する。出現は型（主題）を持ち、出現によってリンクされた主題および情報資源間の関係の性質を記述することが可能。

●公開主題識別子 (Published Subjects Identifier)

トピックや関連といったメタ層に属するノードを Subjects と呼んでいるが、それらはweb上に公開することができ、サブジェクトを共有できる。さらに、トピックマップ自体が持つサブジェクトのマージ機能によりトピックマップ同士が接続され、シームレスなセマンティクスネットワークが形成される。

●有効範囲 (Scope)

トピックマップのトピック名、関連、出現についての有効範囲を設定できるものである。この有効範囲が同じ集団のトピックマップを集めることで、任意の Context についてのトピックマップを抽出できると考えた。以下に、スコープの典型的な利用例を挙げる。

- 国を表すトピック名：English (“country”)、Japanese (“国”)、Norwegian (“land”)
- 異形名を持つトピック： (“ニナワ”、“ニナ”);
- 関連が成立するドメイン：subject area (history, biography, culture, politics); historical period (ancient, classical, medieval, modern, past, present)

2——トピックマップ利用のメリット

トピックマップは、意味ネットワークを形成するメタ層とデータを保持するリソース層が明確に分離されている。したがって、データモデルの変更なしにオントロジーの拡張が可能となる。RDFなどと異なり、その構造を知らなくても機械的に処理すればよいので、情報の多様性の表現力や変化に対する頑強性が高い。トピックマップのメリットは、柔軟性が高いと言うことに尽きる。マージ機能やスコープを用いたフィルタリングを利用できるトピックマップを利用した方が、メリットが多い。述語論理の表現からはやや逸脱するので推論可能性、知識表現の柔軟性が高い。

3——トピックマップにおけるオントロジー表現

オントロジーを「概念と概念間の定義と体系化」とするならば、オントロジーにおける概念を「トピック」に、概念間の関係を「関係」に写像すること

で、オントロジーをトピックマップで表現可能といえる。したがって、オントロジーで記述可能な知識表現は、トピックマップでも同様に可能であるといえる。具体的には表4.1のような対応関係になっている。

このように、トピックマップはオントロジーを表現することが可能となっている。

トピックマップを用いた Context の表現手法

Contextとは、暗黙的に考慮している主題であると考えられる。本稿では、トピックマップの特長を生かし、Contextを表現する手法について述べる。

1—— Context間の関係性の表現手法

Contextを表現するには、いくつかの手法が考えられるが、本稿ではトピックマップの主題に Scope で直接有効範囲を与える Context 表現およびトピックマップの隣接したノードを有効範囲とする Context 表現、Contextのトピックによる Context 表現について説明する。

● Scope の利用して Context を判断する方法

(図4.2)

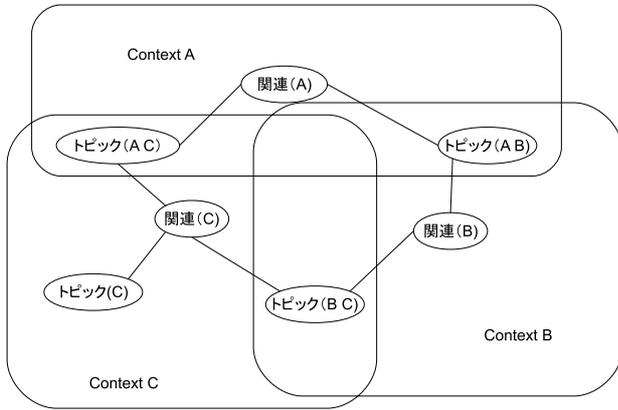
この方法は、共通の Scope を持っている主題を Context とする方法である。図中の主題に注目するとわかるが、主題は Scope を持っている。図ではカッコ内の英字で表記されている。Scope は空白文字で連結することで、任意の個数保持することができる。Scope A に着目した場合の一般的な Context は、図中の枠で囲われた Context A 内の Subject 集合である。

この方法は、最も自由度が高い方法であり、Scope の集合と Context の集合の関係の定義もいく

図4.1: オントロジーとトピックマップの要素間の対応

OWL/RDF	トピックマップ
主体	トピック
述語	関連
客体	トピック/出現

図4.2: スコープを用いたContext 表現



つか考えられる⁽¹⁰⁾。(集合が等しくなければならない、Scopeの集合がContextの集合サブセットあるいはスーパーセットでなければならないなど) トピックマップのつながりに依存せずに、任意の主題の集合をContextとすることができ、表現力が非常に高いという利点がある。一方で、Scopeはあらゆる視点や背景などが対象となりうるので、Scopeの記述量はすぐに現実的な量を超えてしまうという欠点がある。

したがって、Scopeの使用例ででてきた言語を指定するといったようなグローバルなContextについては、Scopeによる記述が望ましく、ドメインに依存しているContextなどは、他の記述方式を用いた方がよいといえる。

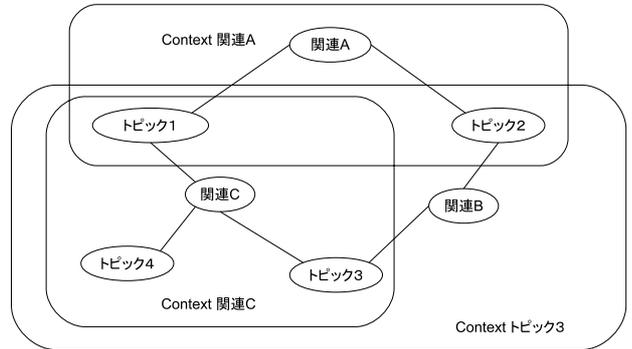
● 隣接している主題からContextを判断する方法

(図4.3)

あるトピックに接続されている出現、あるいは関連およびその先にあるトピックをContextとする方法である。図中のSubjectに着目すると関連に注目したときは、その隣接するSubjectであるトピックの集合トピックに注目したときは、その隣接する関連およびその関連の隣接サブジェクトの集合全体をContextとするものである。

トピックマップの構造のみを利用している方法なので、新たに特別な操作が必要ないという利点がある。しかし、これはデータモデルとしてはトピックマップそのものであり、クエリや推論エンジンが隣接するSubjectをContextとして扱うように考慮し

図4.3: 隣接する主題を用いたContext 表現



て設計されたものでなければならないという欠点がある。また、Contextの表現力は、今回提示する手法の中で一番低い。

● Context用トピックからContextを判断する方法 (図4.4)

Contextを表すトピックに接続されている関連およびその先にあるトピックをContextとする方法である。図中のトピック、関連からContextを示すトピックへとリンクが張られているが、これは抽象的なもので、関連を含みそれを経由してリンクされている。また、Contextのトピックから関連につながっているが、これも抽象的なものであり、関連を具体化し、その具体化したトピックとContextを示す関連で接続されている。

この方法は、Context自身をトピックとすることで、text間の関係性を記述できるという点が他のモデルと異なる特徴である。Contextを表すトピックはContextの関連で接続され、Contextが、どのトピックとどのような関係にあるのかわかるようになっている。

2—— Context 間の関係性の表現手法

これまで、Contextの表現手法について述べてきたが、Context間関係性についても表現しなければ、トピックの単純なタクソノミーを生成したに過ぎない。Context用トピックを用いて、Context間関係性を記述したものを図に示す。これは、データモデルとしては、トピックマップ自体の関係性を記述する上位の層に位置する。

このようにして表現した Context 間の関係によって、Context そのものを推論によって導くことを可能にすれば、資料に暗黙に隠されていた意図などが発見できると考えられる(図4.4)。

考察

本稿では、民具資料のオントロジーをマッピングしたトピックマップを作成し、トピックマップの特徴を活用した Context の表現についての提案を行った。Context の表現は、知識表現における重要な問題のひとつである。

Context とは、普段我々が暗黙的に考慮しているもののことである。民俗資料をもとに人間の営みの歴史の変遷を明らかにし、生活文化の意味を説明すること目的としている民俗学などにおいては、民俗資料の背景に隠された常民の意図を読み解く際に Context が重要となる。Context の表現にトピックマップを用いたが、その理由は、メタ層とリソース層が分離されているためオントロジーの変更が柔軟に可能だからである。また、トピックマップは、すでに稼動しているシステムが多く、それらのシステムとの共有・相互利用することが考えられる。また、トピックマップの機能であるマージ機能、スコープ機能などを活用することで知識の共有・フィルタリングが易しくなる。

Context の表現方法にもいろいろな方法が考えられた。本稿で取り上げただけでも、トピックマップの標準機能である Scope を用いる方法、トピックマップの隣接ノードを Context と判断する方法、Context 用のトピックマップにより、トピックマップを判定する方法などが挙げられた。それぞれの手法は、一長一短であり、それぞれ適した場所に適用する必要があることがわかった。特に、Scope はグローバルな Context を記述するためだけに用いるほうがよいと考えられる。また、簡易的には Context を特別に記述しなくても隣接ノードを Context とする方法があるので、クエリなどの記述を工夫すれば、Context を考慮した検索・処理が行えるといえる。しかし、Context 間の関係を表すには、Context 自身

図4.4: Context 用トピックを用いた Context 表現

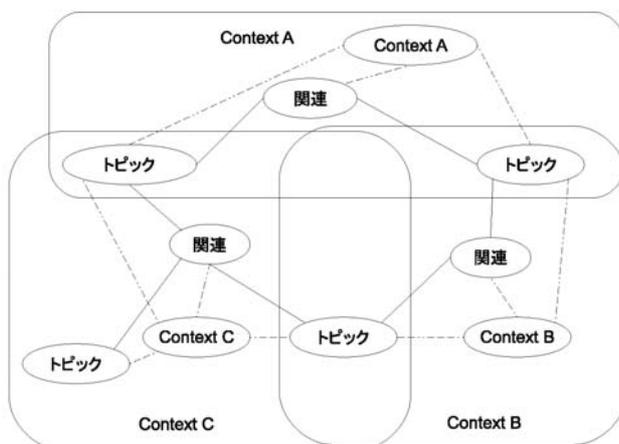
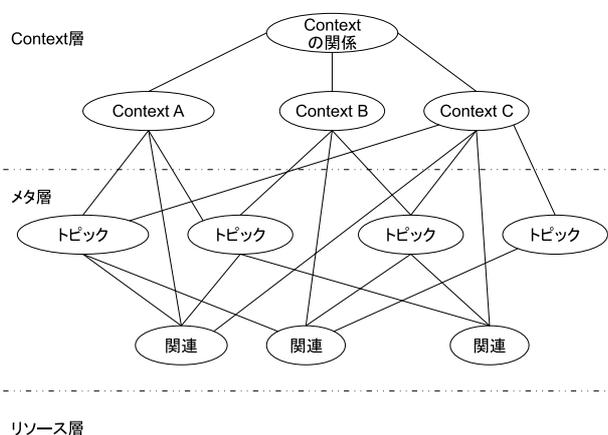


図4.4: Context 用トピックを用いた Context 表現



をトピックとする方法しかない。Context をトピックとすることで Context 間の関係を記述でき、Context がどのような働きをしているかを定義可能になる。そうすることで、一目見ただけではわからない Context を推論によって導き、隠れた関係性の候補を導出することが可能となり、民俗学研究に大いに役に立つと考えられる。

さらに柔軟性を高めるには、トピックマップの具体化という機能を用いて関連を具体化し、関連についても Context のトピックで接続できるように発展させるのが望ましいと考えられる。Context の表現を試みたことで、常民の意図などといった Context を記述することが可能となる。

まとめ

本稿では、非文字資料における Context の表現手法について述べた。その中で、トピックマップを

用いた Context の表現手法を提案した。

Context には、常民の隠れた意図や、研究者の無意識な視点の違いなどさまざまなものが考えられる。これらの Context を表現することは、研究者間の民俗資料の共有・相互利用を進めた際に起こりうる問題を解消することになる。社会的背景・歴史的背景などによって世界観が異なることによる視点の違いである。したがって、Context を表現することが非常に重要となる。ISO により標準化されたトピックマップというメタモデルを表現する規格を活用し、Context の表現手法を明らかにした本研究によ

り、非文字資料のオントロジーにおいて Context の表現が可能となり、よりいっそう民俗学の研究に適した情報の共有・相互利用の実現性が高まったといえる。

本稿の Context の表現のモデルを生かすには、トピックマップのクエリ言語を活用する必要がある。

今後の課題としては、Context 同士の対応関係にどのような種類があり、どのような活用方法があるのかを検討することが必要である。また、トピックマップのクエリ言語を活用し、暗黙的な Context を導き出せることの証明を行う必要がある。

【参考文献】

- (1) 神奈川大学 21 世紀 COE プログラム 人類文化研究のための非文字資料の体系化
<http://www.himoji.jp/>
- (2) Dublin Core Metadata Initiative
<http://www.dublincore.org/>
- (3) 木下慶子、稲積泰宏、木下宏揚、森住哲也：COE における非文字資料の共有と流通
- (4) 木下慶子、稲積泰宏、木下宏揚、森住哲也：非文字資料に適した Ontology 構築
- (5) 木下慶子、村上敦志、稲積泰宏、木下宏揚、森住哲也：デジタルアーカイブにおける Ontology の活用
- (6) The TAO of Topic Maps
<http://www.ontopia.net/topicmaps/materials/tao.html>
- (7) ISO/IEC 13250-2 Topic Maps -Data Mode
- (8) トピックマップ入門
[http://www.knowledge-synergy.com/topicmaps/document/TopicMaps Introduction.pdf](http://www.knowledge-synergy.com/topicmaps/document/TopicMaps%20Introduction.pdf)
- (9) XTM 1.0
http://www.y-adagio.com/public/standards/tr_xtm/xtm-main.htm
- (10) Towards a General Theory of Scope
<http://www.ontopia.net/topicmaps/materials/scope.htm>

V 情報カプセルを用いた著作権管理

はじめに

現在、音楽や映像などのデジタルコンテンツがコンピュータ上で広く利用されているが、インターネット普及が急速に促進され、デジタルコンテンツを配信、利用するにあたって、著作権の保護や利用の権限をどのように取り扱い、どう制御するのが問題となっている。この問題に対し、デジタルコンテンツを従来からある著作物と同等に扱い、著作権法を適用するアプローチが考えられる。また、一方で、独自のライセンスを基に配布や改変を実施する権限を広く認めるオープンソース的アプローチも考えられる。オープンソース的アプローチは利用・配布の利便性の向上を目指すことになる。このように、デジタルコンテンツがインターネットを介して産出、利用されるコミュニティからの要求はさまざまである。したがって、著作権（所有権）を主張する権利者とそれを利用するユーザがインターネットで接続された世界においては、さまざまな権限に対する要求に柔軟に対応可能なシステムが求められる。

データがデジタル化され、デジタルコンテンツとなる場合、複製しても劣化しないという利点があるため、著作者の権利が侵害されやすいという問題点がある。そこで一般にこれらを流通させる場合 DRM（デジタル著作権管理）が利用される⁽¹⁾。また、異なった複数のコンテンツの1部分を引用した2次的コンテンツ（著作物）を作る際に編集方法について提案されている。しかし、権利管理の調停への対応が不足している。すなわち、従来技術はたとえば著作権法で認められている範囲内であってもコンテンツの一部の参照が制限されざるを得なくなる場合や、他人への譲渡の制限を強くせざるを得ない場合が生じる。このため、ユーザ側の利用の権利まで損なわれてしまう。ライセンスを用いた権利流通の基盤の研究として、DRMを基に権利とライセン

スをマッチングする方式では⁽⁶⁾、利便性の向上が図られている。しかし、ライセンスの利用において制限が多く、途中の変更の考慮がされていない。

そこで、本稿ではまず著作物をデータベース側で管理するツールとして、Dublin Core Metadata Initiative⁽²⁾（以下、Dublin Core）に着目する。Dublin Coreは、簡潔なメタデータを定義してネットワークを介した情報源へのアクセスを促進することを目的に作られたものである。メタデータはデータに関する構造化されたデータであり、それを使うことによってデータベースへのアクセスの相互操作性、柔軟性、拡張性を向上させる効果がある。また、Dublin Coreはメタデータのコア要素の1つとして「権利管理」が定義されている。そこでこの要素を著作権、利用の権限のために拡張する。本稿が着目するシステム構成は、「拡張Dublin Coreとアクセス制御リストによりデータベースを管理するエージェント」、「著作権や利用の権限を制御するエージェントを伴うコンテンツのカプセル」、「ユーザ側のエージェント」である。データベースには著作者が作成した著作物が格納される。データベースの管理エージェントは、データベース内部のデータの読み書きに関しては従来のアクセス制御技術が適用される⁽³⁾。しかし、一度データベース内部のコンテンツがネット上に流通すると、一般には制御不能になってしまう。本稿では既存の情報カプセルにTake-Grantと情報フィルタ、エージェントを導入し、コンテンツ利用の利便性を向上する方法を提案する。このシステムにより、コンテンツの利用の利便性の向上が期待される。また、エージェントを導入することで、カプセル側のカプセルエージェントとユーザ側のアクセス制御エージェントが交渉し、通信のトラフィックの低減、レスポンスの向上、権利の充足性の検査、充足させるための推論メカニズム、n次利用の複雑な権利関係の調停ができるという利点がある。

1——デジタル著作物の問題点

デジタルコンテンツの著作物に関する特徴を下記に示す。

- 完全な複製が容易に生成可能。
- 流通が容易。
- 改変・加工が容易。

このような特徴を持つデジタルコンテンツの流通管理として必要な技術的要件を下記に示す。一般的に著作権が保護されている状態とは、広く一般に配布するコンテンツに対して、以下の条件を満たしていることである。

- 適正なn次利用が促進されること。
- 不正なコピー行為に対する防止策が講じられていること。

インターネットを介してコンテンツを配布するためには、そのコンテンツがどこにあるのかを効率よく知るシステムが必要である。そのための一つの手段として、コンテンツのオントロジーとしてのメタデータを情報の分類、整理、検索に利用する。このようなシステムにおいて、デジタルデータの特徴を生かし、コンテンツ利用の利便性を向上するために、本稿では下記に着目する。

1. 流通後に利用条件やコンテンツの変更が可能であること
2. 適正な2次利用が可能であること

2——コンテンツの保護

コンテンツの保護という点では、いわゆる著作権をはじめとする権利関係とデータの暗号化などの技術要素に分けられる。

電子透かし

世間一般に知られているコンテンツの保護技術の一つとして、電子透かしがある。電子透かしとは、コンテンツに情報を追加し、入れた情報を読み出す技術であり、

- (1) 使用者には分からない様に情報を入れることで、音声・図画・動画を劣化させずに情報を入れることができる
- (2) 入れた情報は改竄できない様になっていて、情報の切り出しや追加を行っても入れた情報を正しく取り出せる

の2点が満たされる必要がある。この特性を利用して、権利情報や流通経路情報をコンテンツに追加して、権利の管理・保護に使用する。実際に使用している例も見られるようになったが、課題も残っている。第一に、時間がかかり過ぎる。最低でも実視聴時間は必要であり、実際には強度にもよるが相当の時間がかかってしまう。これではライブでは使えないし、VOD (Video On Demand) でも実用に達しているとは言えない。図・画などの静止画か音楽への適用が現在の実用範囲である。第二に、コンテンツの劣化が発生する。透かし技術は進歩しているが、並べて比べれば違いは明らかであり、製作者の理解を得るレベルまでは達していないのが実情である。更に抑止効果は期待できるものの、結局視聴は可能であるという点にも大きな課題が残る。追跡ロボットとの組み合わせで不正利用の追跡・調査あるいはタイムリーな発見のマーク付けに利用するなどの応用も考えられていて、色々使い道はありそうだが、コンテンツの保護という点においては不足している。

DRM (暗号化システム)

コンテンツ保護の本命は現在のところ暗号化である。権利取得者以外は利用ができないというのがコンテンツ保護のあるべき姿である。ただし、DRMの方式は一本化⁽⁶⁾されているわけではなく、適用する場合にはストリーミングを提供しているベンダの

製品を個々に使用するか、要素技術の提供を受けて独自に開発するかである。ダウンロード型のコンテンツの場合は確かに単純であり現在でも一本化の可能性もあるが、ストリーミング方式の場合はそれ自体それなりのレベルには到達してきたものの発展途上であり、スタンダードとして一本化されるまでには至っていない。今後の拡張への対応を考えた場合、当面ベンダの製品を使うのが自然な選択であるといえる。

DRM 技術

DRMの主な技術として、Rights DescriptionとRights Enforcementがある。

Rights Descriptionは、コンテンツに関するさまざまな権利を記述・表現するための技術である。著作権者（Holder）と流通業者（Holderからコンテンツの使用許諾を受ける役割、Distributor）、販売業者（Distributorから販売許諾を受ける役割、リテイラ）、エンドユーザの間に関する権利は、以下の4つである。

1. コンテンツに対してHolderが所有する権利
2. DistributorがHolderから使用許諾される権利
3. リテイラがDistributorから販売許諾される権利
4. エンドユーザがリテイラから利用許諾される権利

Rights Descriptionとして、OASIS（Organization for the Advancement of Structured Information Standards）において標準化の検討が進められているXrML（Extensible rights Markup Language⁽⁷⁾）やW3C（World Wide Web Consortium）において標準化の検討が進められているODRL（Open Digital Rights Language⁽⁸⁾）などがある。

Rights Enforcementは、Rights Descriptionによって記述・表現された権利をコンテンツの利用者（主にエンドユーザ）が侵害しないように保護・監視するための技術である。保護・監視する対象として、コンテンツ自体とそのコンテンツのメタデ

ータに記述された権利情報がある。

現状の主なDRMシステム（WORM〈Windows Media Rights Management⁽⁹⁾〉やRDMCS〈RealSystem Media Commerce Suite⁽¹⁰⁾〉など）では、両者を明確に区別しておらず、むしろシステム固有のRights DescriptionとRights Enforcementが不可分な形式で提供されている。具体的には、コンテンツを施錠可能で安全なコンテナに格納したカプセルをエンドユーザに配布し、そのカプセルを解く鍵情報とRights Description（利用端末の限定や利用回数など）を含むライセンスを別途発行することで、コンテンツの利用制御を行っている。このとき、ライセンスのRights DescriptionはDRMシステムに固有の情報であり、DRMシステムに固有のRights Enforcement（カプセル化や利用制御など）を実現するために最適な方法で記述されている。このため、Rights DescriptionとRights Enforcementの依存関係が強く、たとえばRights Descriptionを別のRights Enforcementを持つDRMシステムへ適用することができない。

しかしながら、本来Rights DescriptionとRights Enforcementは独立な技術であり、両者を不可分に扱うことは、DRMシステムの柔軟性と相互運用性を大きく損なうことになる。Rights DescriptionとRights Enforcementを独立なものとして、必要に応じて相互に選択可能なDRM方式が望まれている。

3—— RDF

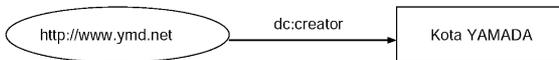
RDF の仕様

RDFの仕様では、資源間の構造的な情報であるRDFグラフとその記述と移植のための文法、およびその意味論を定義している。RDFを用いることで資源に対して構造的なメタデータを付加することができる。資源とは、URI（Uniform Resource Identifier⁽¹¹⁾）を用いて表現できるものすべてを指し、本稿では主にコンピュータ上に存在する資源を対象としている。また、RDF Schema⁽¹²⁾は、RDFデータのスキーマ情報を記述するための仕様である。RDF Schemaを用いることで、RDFグラフを表現する際に用いるクラスやプロパティを定義できる。

図5.1 RDF

http://www.ymd.netにはKota YAMADAという著者がいる

主語 (Subject)	リソース	http://www.ymd.net
述語(Predicate)	プロパティ	著作者
目的語(Object)	プロパティの値	Kota YAMADA



RDFグラフは、資源間の2項関係を表現することができる文 (rdf : Statement) と呼ばれる基本単位によって構成されている。たとえば、文を用いることによって、「This paper is authored by Kota YAMADA.」という情報を表現することができる。文は主語 (rdf:subject, 「This paper」) と述語 (rdf: predicate, 「is authored by」)、目的語 (rdf: object, 「Kota YAMADA」) という三つの部分によって構成されている。文の集合を用いることで、主語と目的語を頂点とし、述語が有向辺に対応したラベルを持つ有向グラフ構造の情報を表現することができる。

RDF Schemaは、RDFデータで用いる語彙を定義するための仕様で、RDF Schemaを用いることで、クラス (rdfs : Class) やプロパティ (rdfs : Property) を定義することができる。さらに、プロパティの定義域 (rdfs : domain) や値域 (rdfs : range)、クラスやプロパティを継承 (rdfs : subclassOf, rdfs : subPropertyOf) を定義することができる。これらのクラスやプロパティは、他のRDFデータ内で表現される資源のタイプ (rdfs : type) として利用する。すなわち、RDFデータは、RDFスキーマデータのインスタンス情報である。RDFスキーマデータも通常のRDFデータと同様に、文によって構成されている。そのため、RDFスキーマデータで用いる語彙は、RDF Schemaの仕様で暗黙的に定義されている。すなわち、RDF Schemaを用いて定義するための暗黙的な文は、メタスキーマ情報であるといえる。RDFでは、URIで記述できるものはすべて資源であるため、クラスやプロパティはもちろん、RDF Schemaの仕様で定義さ

れた暗黙的な語彙もすべて資源である。

RDF データの特徴

RDFの特徴は、データサイズに関する面と、RDFグラフの構造に関する面の二つの側面から、現在普及されている代表的なRDFデータRSS⁽¹³⁾、FOAF⁽¹⁴⁾、Dublin Core、Wordnet⁽¹⁵⁾、ODP (Open Directory Project)⁽¹⁶⁾、Gene Ontology⁽¹⁷⁾の概要を説明する。まず、現在のRDFデータはデータサイズの面から大きく二つに分類できる。一つは、WordnetやODP、Gene Ontologyで、語彙やゲノムなどの資源を体系化するためのRDFデータである。このため、一般に大きなサイズを持つ。他方では、RSSやFOAF、Dublin Coreで、Web上に実際に存在する画像やページなどの資源に対してメタデータを付加することを目的としたRDFである。このため、一般にそのサイズは小さい。留意すべき点として、後者のRDFデータは単体で用いることはあまりなく、複数のRDFデータを用いて初めてメタデータとして知的に扱われることができる。

RDFグラフの構造に関して説明する前に、RDFグラフを形成する述語 (有向辺) の種類について説明する。RDFデータは、述語によって構造化されている。述語は大まかに2種類に分類することができる。ここでは仮に、ある文を経由すると有向グラフの構造が発散するものを発散述語、ある文で終端する述語、あるいはその文を経由したあとは収束に向かうものを終端述語と呼ぶことにする。終端述語は、目的語がリテラルであるものが一般的で、主語を直接的に修飾するために用いられる。たとえば、「名前」や「所有」などがある。発散述語は、主語と目的語が同一レベルのクラスに属するもので、主語と目的語の関連を示すために用いられる。たとえば、「恋人関係」や「関連研究」などがある。これらの分類は、RDFデータの構造に深く関わっており、発散述語の割合が増加するにつれて複雑さが加速度的に増加することが多い。

次にRDFグラフの構造について説明する。現在普及しているRDFデータのグラフの構造は、単体ではほとんど閉路を含まないような単純な構造であることが一般的で、単体のRDFデータで閉路を多

く含んだものはほとんど無いと言ってもよい。実際、前述した六つのメタデータ規格では、いずれも単一文書では、木あるいは非巡回有向グラフの構造で、ほとんど閉路を含まない。その理由として、RDFデータ内で用いられる述語は終端述語が多く使われており、発散述語はあまり用いられていないためである。これは、一つのRDFデータで大量の発散述語を用いたRDFグラフを構成することは困難であると考えられる。

本稿では、その中でDublin Coreに着目して、権限の調停に用いる。

Dublin Core

メタデータはデータについてのデータであり、インターネットに分散配置されたデータベースの中のようなデータがあるかを記述し、検索効率を上げる目的がある。Dublin Core Metadata Element Setは、この様に多種多様なメタデータを効率的に参照、交換する必要最小限のメタデータの組み合わせ（メタデータセット）として開発された。⁽⁴⁾

Dublin Coreは15個の要素である⁽⁵⁾タイトル（情報資源に与えられた名前）、作成者（情報資源の内容の作成に主たる責任を持つ者）、主題およびキーワード（情報資源の内容の主題）、内容記述（情報資源の内容の説明）、公開者（情報資源を提供している主体）、寄与者（情報資源の内容に貢献している者）、日付（情報資源のライフサイクルにおけるイベントに関連する日）、資源タイプ（情報資源の内容の種類またはジャンル）、形式（情報資源の物理的または電子的形式）、資源識別子（与えられた環境において、情報資源の一意に定まる参照）、出处（現在の情報資源が作り出される源となった情報資源の参照）、言語（情報資源の内容を記述する言語）、関係（関連情報資源への参照）、時間的、空間的対象範囲（情報資源の範囲または領域）、権利管理（情報資源に含まれる、ないしは関わる権利に関する情報）で構成される。その中で、著作権に関する要素にrightsがある。これは、リソースが保持する、あるいはリソースに適用される権利に関する情報で、通常、リソースの権利管理宣言やそのような情

報を提供するサービスについて言明されて、知的所有権、著作権などのさまざまな財産権などの情報を含むことが多い。この要素がない場合は、リソースの権利に関していかなる憶測もたてないことになっている。また、rightの拡張要素として、accessRightsがある。これは、誰がリソースにアクセスできるかについての情報もしくはセキュリティステータスの提示で、アクセス権は、プライバシー、セキュリティまたはその他の規則に基づいたアクセスあるいは制限に関する情報を含んでいる場合がある。

しかしながら、権利管理における要素における拡張がなされておらず、また、利用目的が明白でないなど、Dublin Coreの表現が不足している。

提案システム

前述において、DRMでは、2次利用におけるユーザがコンテンツ利用の柔軟性における問題がある。また、著者が流通しているコンテンツの権限の設定を変更することができない。そこで、それらを解決するシステムを提案する。そのためにまず、Dublin Coreを著作権に関する部分を拡張する。利用目的を明白にし、それを権限と見なすことによって、ルールを設定する役割を担う。ルールを設定し、カプセル化を行う役割として、アクセス制御エージェントを用いる。カプセル化したコンテンツを管理する役割として、アクセス制御エージェントを用いる。権限を調停する方法として、Take-Grantを用い、権限の調停をカプセルエージェントとアクセス制御エージェント間で行う。それらを用いて柔軟な2次利用を行うシステムを構築する。そうすることにより、ユーザが意図する2次利用が可能になり、また、著者が流通中のコンテンツの権限の変更が可能になり、著者とユーザの双方におけるコンテンツ利用の利便性が向上する。

1——配布、利用の定義

デジタル著作物を利用する権限のひとつである配布と利用について説明する。

図5.2 提案システム

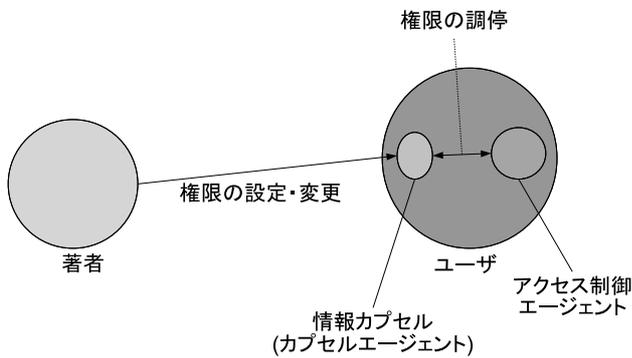
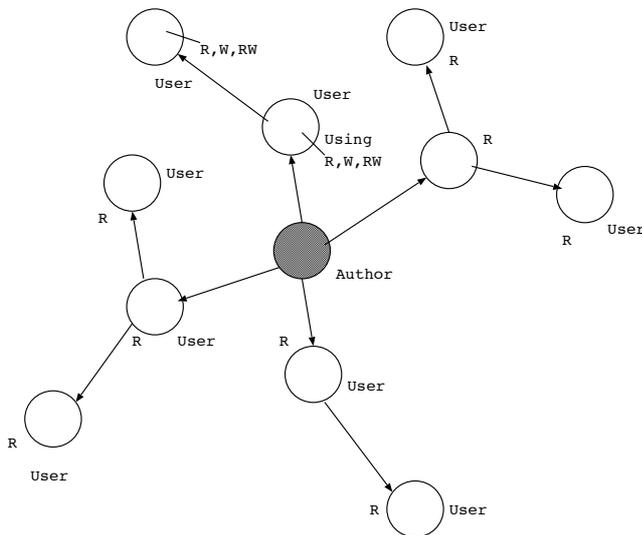


図5.3 コンテンツの利用・配布



配布の定義：エンティティからエンティティへ情報を移動させること。

利用の定義：コンテンツの複製、あるいはR、あるいはRWを伴う操作。

次配布の定義：起点エンティティから終点エンティティまでの枝の数をnとする配布。2次配布とは、著者から配布されたコンテンツを他のユーザに再配布することである。そのとき、そのコンテンツは複製可能とする。それをn回繰り返すことをn次配布という。

1次利用の定義：Rをする操作、もしくはコンテンツの変更を伴わない利用をする操作。

n次利用の定義：少なくとも1つのコンテンツの変更を伴う利用をする操作。すなわち、コンテンツをRWすることをいう。ただし、複製におけるRWの操作を除く。たとえば、2次利用とは、配布されたコンテンツを再編集して、2次

著作物として他のユーザに配布することである。⁽²¹⁾それをn回繰り返すことをn次利用という。

再配布の定義：コンテンツの元となる作者ではない作者からコンテンツを入手することである。

デジタルコンテンツにおいて、オリジナルの著作物をコピーすることによる問題として、次のものがある。

- コンテンツ作成者の権利の正当性
- バージョンの管理

たとえば、コピーした著作物をコピーした人があたかも著作者を主張することがある。また、コンテンツのバージョンが更新されて、ユーザが勝手にコンテンツの仕様を変更し、最新でないバージョンとの互換性が損なわれる可能性がある。上述により、違法コピーなどの不正利用が行われるので、再配布をできないようにする、またはルールの強化によって再配布を制御するシステムが必要になってくる。⁽¹⁹⁾⁽²⁰⁾

また、カプセルを2次的、3次的に利用するにあたって、下記前提条件を想定する。

- カプセルの権限は、著作者のカプセルエージェントが管理・制御していることである

2——拡張Dublin Core モデル

著作権を保護するために必要な機能は以下のとおりである。

1. 利用目的が明白である
2. 著者の名前とコンテンツが対応していること
3. 著者が流通しているコンテンツの権限を更新、削除することが可能

著作権に拡張したDublin Coreモデルを図5.4に示す。Dublin Coreの表現を、著作物に関する権利関係や利用目的などについてさらに拡張したもので、著作権情報の詳細が既存のものよりも明白になって

著作権情報の詳細が既存のものよりも明白になっている。図5.4において、著作物の権利が何なのかを確認する。著作権とは狭義の著作権である複製権、上演・演奏権、上映権、公衆送信権、口述権などに対応する。利用条件とは、その中の一つの権利の有無あるいは金銭との関連の組合せである。RDFではこのような各権利に対応して書き下す。

権利ごとに3つに分類される。利用目的1では、著作者の権利がないときで、たとえば、著作権法での著作物の著作権の期間が終了したときである。利用目的2では、著作者の権利がある状態で、ユーザの利用における課金が発生しない場合である。利用目的3では、著作者の権利があり、かつユーザが利用する際に課金が発生する場合である。それぞれの権利において利用目的が何なのかをコンテンツのダウンロードと複製、改変の3つに分類できる。利用目的2と3では複製において、配布する際に配布後のコンテンツが残って使える状態なのか残ってても使えないまたは再配布後にコンテンツ自体削除されるかの2つに分類される。利用目的2において、それぞれの目的において著作権の使用における許可するかどうかに承認が必要なのか、承認なしでも利用できるのかに分類される。利用目的3において、許可するかどうかについては課金によって制御される。また、2次的、3次的に利用または配布においてこのモデルを適用すれば、著作権情報の権限が保持され、安全に流通できるようになる。

提案法においては、権限の設定に利用する。

権限変更のルール設定

図5.4における拡張Dublin Coreのルートからリーフまでの1本の枝を $ak(k=1,2,\dots,n)$ とするとき、コンテンツに対する処理の制限の定義を以下のように設定する。

$$Pa1 < Pa2 < Pa3 < \dots < Pan$$

このとき、拡張Dublin Coreの枝 ak の集合 $\{ak\}$ の最大要素を $\max\{ak\}$ とすると、複数のカプセルの情報処理する拡張Dublin Coreの権限は、 $\max\{Pak\}$

である。

3— Take-Grant による考え方

コンテンツのカプセルがどの場所に存在しても、著者などの権限が維持されるというモデルとしてTake-Grantを考える。権限のGrantをカプセル側のエージェントが持ち、Takeをユーザ側のエージェントが持つことによって権限を分割、照合する。Take-Grantの照合方法として拡張Dublin Coreモデルを適用する。そこでまず、Take-Grantの定義について説明する。

Take-Grantの定義：著者とコンテンツの関係、著者名、パミッション、利用権、利用回数、配布回数を与える権利（Grant権）および受ける権利（Take権）

特権ユーザに対して、著者が利用権を決め、コンテンツを配布あるいは利用させる場合のTake-Grant権の制御を下記に示す。

1. Take 権をコンテンツ所有者から送ってもらう（配布するにあたって、コンテンツのR、W、RW権を決めるのは著作者である）
2. 著作者に関する情報（著者とコンテンツの関係、著者名、パミッション、利用権、利用回数、配布回数を与える権利）をコンテンツのデータとして記述し、著作者はそれをGrant権に入れる
3. コンテンツのGrant権とユーザのTake権を照合する
4. コンテンツが更新された場合、Take権が更新されるので、照合後、カプセル側のエージェントが著者の元に戻り、更新情報を得て、再照合する

4— エージェントの動作

提案法の核となるカプセルエージェントとアクセス制御エージェントのそれぞれの役割について説明する。

カプセルエージェント

図5.4 拡張Dublin Core モデル

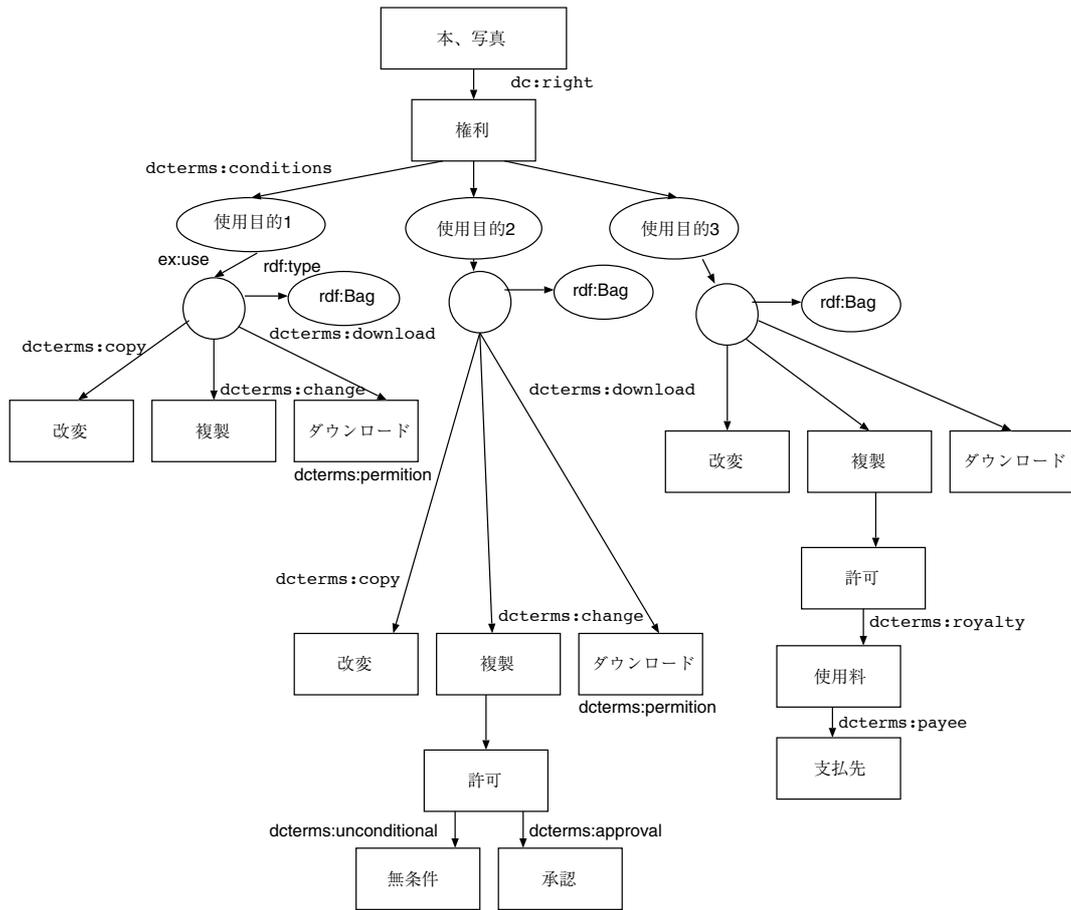
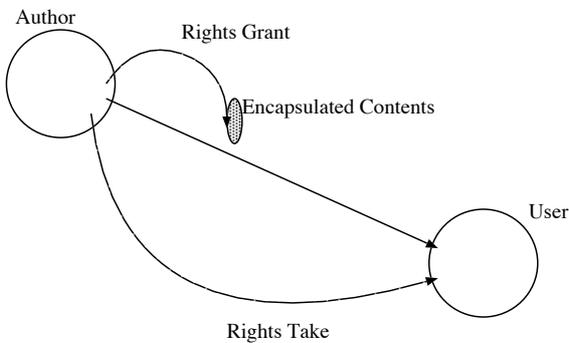


図5.5 Take 権と Grant 権



本システムはエージェントによって利用条件やルールを制御している。カプセルエージェントは、著者のアクセス制御エージェントによって生成される。カプセルエージェントの機能として、配布先のユーザの権限を確かめていく。その際に確認する事柄は、ユーザのアクセス制御エージェントのTake権とカプセルが持っているGrant権が一致しているかどうかである。配布後にはユーザがカプセルを利用したと履歴としてカプセルのデータに書き込む。

もし、そのユーザがこのデータを元に新たな著作物を作る（二次利用）権限を与えられているならば、そのユーザのアクセス制御エージェントから作られたカプセルエージェントは元のカプセルエージェント権限の継承をする。また、カプセルの権限を変更するときには、権限変更のルールにしたがって設定する。

アクセス制御エージェント（情報フィルタ）

アクセス制御エージェントは、データベースとユーザの間に位置する。アクセス制御エージェントは、著作権を管理する拡張Dublin Coreデータベース、および情報（コンテンツ）のパミッションに基づいてデータベースのアクセス制御を実行する情報フィルタの機能を持つ。データベースのアクセス制御はデータがデータベースの中にある間は制御可能であるが、一度データベース管理領域から外に出たデータ（情報）に関しては追跡制御が難しくなる。この問題に対して、アクセス制御エージェント

は、その情報フィルター機能の一部を Grant 権と言うモデルでカプセルエージェントに委託し、データベース制御の外部でもパMISSIONの制御が可能な様にする。

配布者側のアクセス制御エージェントはそのときに Grant 権をカプセルに与える。コンテンツ要求するユーザに Take 権を付与する。ユーザ側のアクセス制御エージェントはカプセルエージェントからの権限が一致しているかどうかを確認し、一致しているのであれば Take 権とカプセルの Grant 権を調停し、カプセルエージェントと共に拡張 Dublin Core に基づいたコンテンツ処理を行う。カプセル内にある暗号化された復号鍵をコンテンツと共にディスプレイへ運び、ディスプレイ上で暗号化された復号鍵を復号化し、その鍵を使ってコンテンツを表示可能にする。ユーザの利用目的が2次著作物をつくる場合には、カプセルエージェントの権限をもとにして、アクセス制御エージェントが編集されたデータをカプセル化する。そのときに、著者のカプセルエージェントに加えて新たなカプセルエージェントを生成する。

また、カプセルを削除した場合、カプセルの入手先のアクセス制御エージェントに配布先を送信する。

5—直列な n 次利用および n 次配布におけるシステム動作

カプセル化コンテンツの2次配布を禁止するメカニズムにおいて、カプセル内のエージェントが起動しないまま配布され、使用しないまま他ユーザへ送られて、カプセルのエージェントが1次配布と認識されてしまう可能性がある。それが起こらないためには以下の条件が必要である。

- コンテンツの配信後にエージェントが起動している。
- カプセルが配信されたときに履歴として残す。
- 上述2つの条件のもとで2次配布がなされたときにそのユーザが閲覧可能である。上記では、カプセルの再配布においてコンテンツをコピー&ペーストしてかつ再配布後でも再配布者がコンテンツが閲覧できることを想定している。

次に、再配布後では再配布者がコンテンツが見れない状態であることを想定すると、以下の条件になる。

- コンテンツの配信後にコンテンツが利用される時にエージェントが起動。
- コンテンツが利用しているときに履歴として残す。
- 上述2つの条件のもとで2次配布がなされたときにそのユーザが閲覧できないまたはコンテンツ自体を削除する。

n 次利用

本システムにおいて、エージェントによる制御によって、n 次利用と n 次配布を安全に行う手順を示す。2次利用では、以下のような流通を行う。

1. ユーザが著作者（または配布者）に欲しい情報を要求し、利用目的を伝える。
2. 著作者側のアクセス制御エージェントが著作物を著作権モデルに適用し、カプセル化してカプセルエージェントを生成する。
3. カプセル化コンテンツをユーザに配布する。
4. カプセルエージェントがユーザ側のアクセス制御エージェントと権限が一致しているかどうか確認する。
5. ユーザがコンテンツの中の欲しいデータを抽出し、それを元に編集し、2次著作者となる。
6. アクセス制御エージェントが他のユーザの欲しい情報の要求と利用目的を受け取り、カプセル化して、著作者のカプセルエージェントに加えて新たなカプセルエージェントを生成する。
7. 新たに作られたカプセル化コンテンツを他のユーザに配布する。

上記を繰り返すことにより、n 次利用を安全に行える。

n 次配布

2次配布では以下のような流通を行う。

1. ユーザが著作者（または配布者）に欲しい情報を要求し、利用目的を伝える。
2. 著作者側のアクセス制御エージェントが著作物を著作権モデルに適用し、カプセル化してカプセルエージェントを生成する。
3. カプセル化コンテンツを利用者に配布する。
4. カプセルエージェントが利用者側のアクセス制御エージェントと権限が一致しているかどうか確認する。
5. カプセルエージェントがユーザにたどり着いたときにユーザの存在を確認し、履歴としてカプセルに書き込む。
6. ユーザのアクセス制御エージェントは、カプセルを再配布する際にカプセルを複製する。そのとき、複製されたカプセルエージェントにおいてはユーザの履歴が入っているとする。
7. 他のユーザへ再配布する。

上記を繰り返すことにより、コンテンツのコピー&ペーストを行われて再配布をしたあと、再配布者のデータベースの中のコンテンツが閲覧可能な場合のn次配布を安全に行うことができる。次に再配布後に再配布者のデータベースにコンテンツが利用できない状態になるまたはコンテンツが削除されるときの流通を以下のように行えば、n次配布を安全に行うことができる。

1. 利用者が著作者（または配布者）に欲しい情報を要求し、利用目的を伝える。
2. 著作者側のアクセス制御エージェントが著作物を著作権モデルに適用し、カプセル化してカプセルエージェントを生成する。
3. カプセル化コンテンツを利用者に配布する。
4. カプセルエージェントが利用者側のアクセス制御エージェントと権限が一致しているかどうか確認する。
5. カプセルエージェントが利用者にたどり着いたときにユーザの存在を確認し、履歴としてカプセルに書き込む。
6. 利用者のアクセス制御エージェントは、カプセル

- ルを再配布する際にカプセルを複製する。そのとき、複製されたカプセルエージェントにおいては利用者の履歴が入っているとする。
7. 他のユーザへ再配布する。

並列なn次利用におけるシステム動作

m個の著作物における並列な2次利用では、以下のような流通を行う。

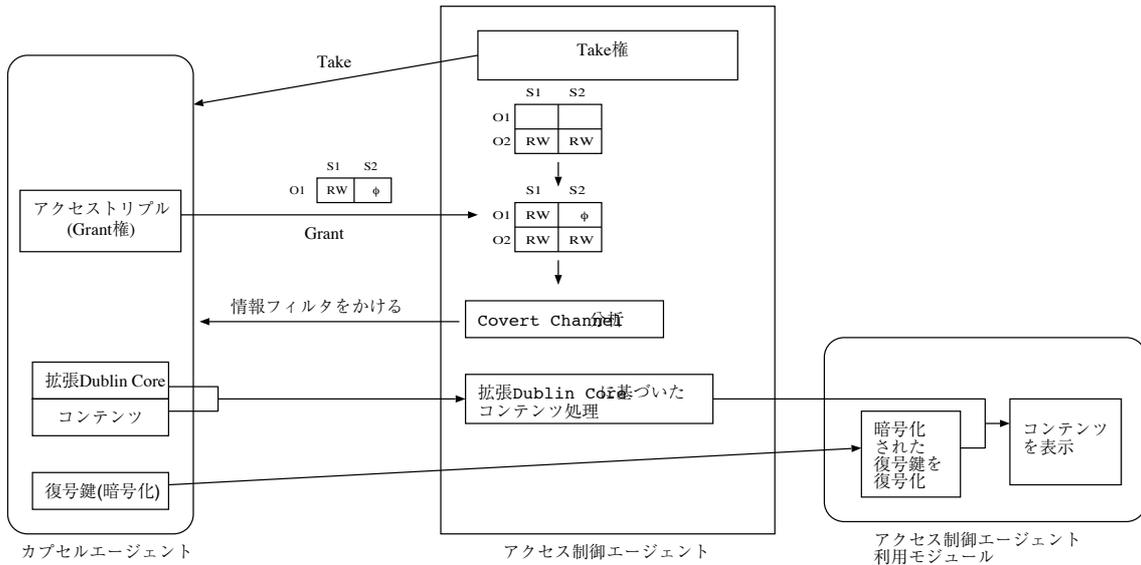
1. ユーザが複数の著作者（または配布者）に欲しい情報を要求し、利用目的を伝える。
2. 各々の著作者側のアクセス制御エージェントが著作物を著作権モデルに適用し、カプセル化してカプセルエージェントを生成する。
3. カプセル化コンテンツをユーザに配布する。
4. カプセルエージェントが利用者側のアクセス制御エージェントと権限が一致しているかどうか確認する。
5. ユーザが複数のカプセル化コンテンツの中の欲しいデータを抽出する。
6. 複数のコンテンツの抽出データからの利用権を確認し、n次までの利用回数のカウンタが少ない方に合わせて、複数のデータを元に編集し、2次著作者となる。
7. アクセス制御エージェントが他のユーザの欲しい情報の要求と利用目的を受け取り、カプセル化して、著作者のカプセルエージェントに加えて新たなカプセルエージェントを生成する。
8. 新たに作られたカプセル化コンテンツを他のユーザに配布する。

上記を繰り返すことにより、m個の著作物における並列なn次利用を安全に行える。

6——コンテンツ利用機構

著者のアクセス制御エージェントをS3、著者が作ったカプセルエージェントをC3、ユーザのアクセス制御エージェントをS1とすると、カプセルエージェントとアクセス制御エージェントとのやり取りは以下ようになる。

図5.6 提案システムのプロトコル



に処理する。

1. アクセス制御エージェント S1が「S1」のTake権をカプセルエージェント C3に伝える。
2. C3は主体と客体、パMISSIONの3つから構成されたアクセストリプル (Grant権)を持ち、それをS1へ。
3. S1はアクセス行列からCovert Channel分析を行い、対象コンテンツに情報フィルタを掛ける。
4. C3内の拡張Dublin CoreモデルをS1と照合し、それに対するコンテンツ処理を施す。
5. カプセル内に暗号化された復号鍵をS1がアクセス制御エージェントモジュールに持っていったあと、暗号化された復号鍵を復号し、モジュール上でコンテンツを復号する。

コンテンツ利用機構における2次利用

コンテンツ利用機構における2次利用について説明する。コンテンツ利用機構のプロトコル(4)の処理後、カプセル内のデータまたはその一部を抽出する。次に2次コンテンツの作成として、抽出したデータを引用・編集する。カプセル化をして、カプセルエージェントが生成され、1次コンテンツのカプセルエージェントの権限を継承する。複数のコンテンツから抽出した場合またはコンテンツ利用の権限を変更する場合は、コンテンツ利用のルールにしたがって設定する。n次利用する際も2次利用と同様

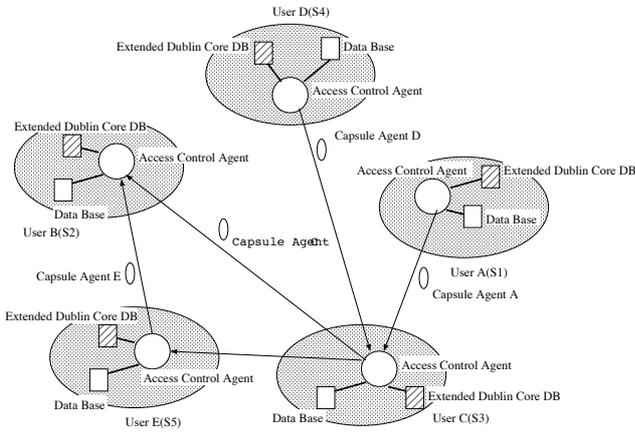
7—情報カプセルのネットワークの形成

著者は流通しているカプセルの権限を変更するためのカプセルのネットワークについて説明する。提案する情報カプセルの仕様として、情報カプセル内のカプセルエージェントは、配布元のユーザーとリンク先のユーザーを記録し、そのユーザーのアクセス制御エージェントが管理する。また、カプセル内には著者の位置が記されていて、カプセルエージェントが管理している。

著者はカプセルを更新するときのコンテンツ利用のメカニズムは、以下のとおりである。

1. 著者はカプセルの権限を更新する際、コンテンツを利用するユーザーのTake権を更新する
2. 更新されたTake権と情報カプセルのGrant権をエージェント同士でネゴシエーションをする
3. アクセス制御エージェントはカプセルエージェントに著者が権限を変更したことを紹介 (図5.8 <1>)
4. カプセルエージェントはカプセルを著者に送信 (図5.8 <2>)
5. 送信されたカプセルは著者のアクセス制御エージェントによって更新され、新たなGrant権

図5.7 2次利用



6. 更新されたカプセルを元のユーザに送信 (図5.9 <4>)
7. 更新されたカプセルのエージェントは、リンク元とリンク先のカプセルのカプセルエージェントに更新情報を更新 (図5.9 <5>)

流通中にコンテンツが削除された場合

情報カプセルの作成者または配布者をC1、C1から配布されたカプセルの利用者をC2、C2から配布されたカプセルの利用者をC3 (以下、C4, C5, ..., Cn) とし、カプセルが流通している。コンテンツの流通中、情報カプセルCj+1が配布元のカプセルCjが削除された場合、Cj+1を持つてるユーザは利用できない。Cj+1を持つてるユーザが利用できるように修復するためのプロトコルを以下に示す (図5.10)。

1. 情報カプセルCjのアクセス制御エージェントはカプセルをもらったユーザが所持しているカプセルCj。1のアクセス制御エージェントにリンク先を送信
2. Sj-1のアクセス制御エージェントはカプセルエージェントにCjからの送信データを手入
3. Sj-1のアクセス制御エージェントはカプセルのリンク先をSj+1に設定

また、前述のプロトコルが成り立たない場合においては、以下の動作を行う (図5.11)。

図5.8 カプセルの更新(1)

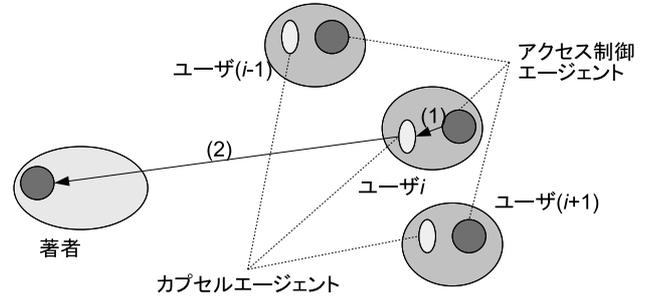
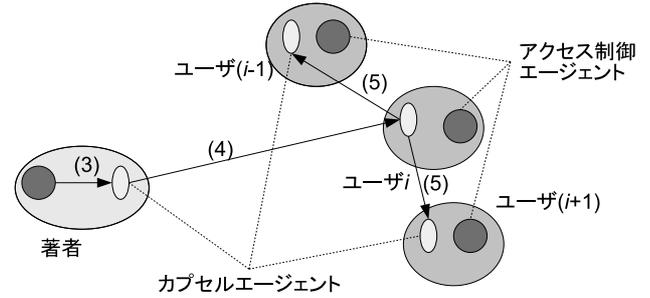


図5.9 カプセルの更新(2)



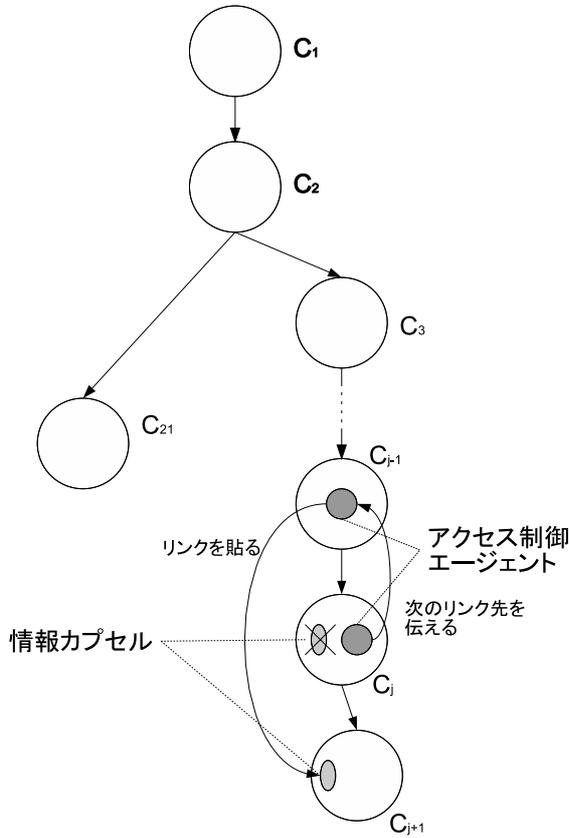
1. 無くなったカプセルからリンクされてるユーザCj+1にあるカプセルエージェントは著者に自身のカプセルにリンクを貼ってくれるユーザを紹介するよう依頼
2. 著者のアクセス制御エージェントは、
 - (1) 著者があらかじめ決めたユーザのカプセルエージェントにCj+1からの依頼を送信
 - (2) 流通中のユーザからランダムに選び、選んだユーザのカプセルエージェントにCj+1からの依頼を送信
3. 依頼を受けたカプセルエージェントは、Cj+1のカプセルにリンクを貼る

この2つの操作のいずれかを行うことにより、流通中のカプセルが削除された場合における削除されたカプセルのリンク先のカプセルが使用可能になる。

8—コンテンツ利用における例

コンテンツ利用の例として、オンデマンドで2次利用のソースとなるコンテンツを入手して権利を調停する場合を示す。

図5.10 著者を介さない修復プロトコル



1. テレビ局A(B)はテレビ局C 向けのコンテンツ A(B)を作成し、2次利用を許可する権限を拡張 Dublin Core モデルに基づき設定し、Take 権を テレビ局C に送信し、アクセス制御エージェント A(B)がコンテンツをカプセル化する
2. テレビ局A(B)のカプセルを入手し、テレビ局 C 上のアクセス制御エージェントC はテレビ局 A(B)からのTake 権とカプセルA(B)のGrant 権をカプセルエージェントA(B)と共に照合する
3. カプセルA(B)の権限を基にカプセル内のデータまたはその一部を抽出し、新たなコンテンツ C を作成し、拡張Dublin Core に特権ユーザ向けの権限を権限変更のルールにしたがって設定する
4. 特権ユーザにTake 権を送信し、コンテンツC をカプセル化する
5. 特権ユーザ1がテレビ局Cのカプセルを入手し、アクセス制御エージェントU1とカプセル エージェントCがTake-Grantにおける権限の照合し、権限にしたがって利用する

6. テレビ局はテレビ局向けの権限にしたがって (2) から (4) の行程を行い、それ以外の特権ユーザは(5)を行う

システムの効果

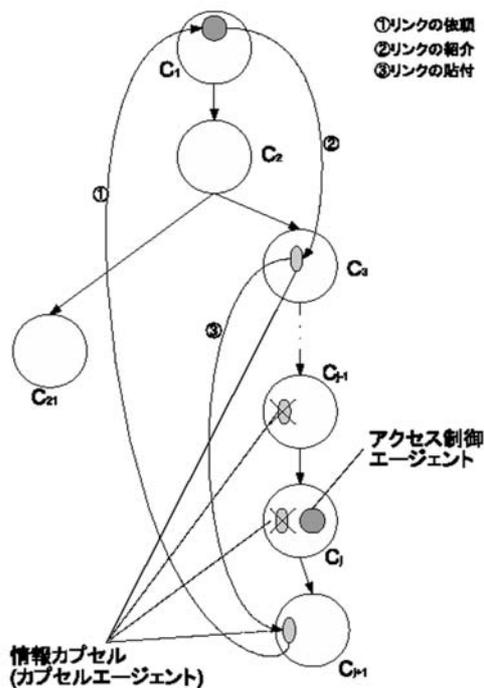
提案法の効果として、Take-Grantを用いることで、ユーザの利用権をカプセル化したコンテンツのエージェントとユーザ側のエージェントが調停していくことで著作権所有者の権利が保護される。著者がTake 権を更新し、ユーザに配布したあとにカプセル内のGrant 権を照合し、更新したことをカプセルエージェントが認識する。そのカプセルは著者の元に直接的に戻り、著者自身がコンテンツの権限の変更ができる。また、ユーザがカプセル内のコンテンツを利用する際の著者からの許諾をカプセルエージェントが代理として行われる。これらによって、著者の権限が守られ、かつコンテンツ利用における利便性が向上した。

結論

著作物をデータベース側で管理するツールとして、Dublin Core に着目し、そのコア要素の1つ、「権利管理」を著作権、利用の権限のために拡張した。著作権とその利用の権限を制御するシステム構成は、「拡張Dublin Core とアクセス制御リストによりデータベースを管理するエージェント」、「著作権や利用の権限を制御するエージェントを伴うコンテンツのカプセル」、「ユーザ側のエージェント」である。データベースの管理エージェントは、データベース内部のデータの読み書きに関しては従来のアクセス制御技術が適用される。

本稿では、ネット上に配布・利用されるコンテンツが著作物管理のポリシーをコンテンツを運ぶエージェントが継承し、ユーザ側のエージェント、およびデータベース側のエージェントと協調して権限を制御する枠組みを示し、Take-Grantを用いて、カプセル内の情報を保護された状態でコンテンツ利用の利便性の向上を示した。

図5.12 コンテンツ利用の例



の利便性の向上を示した。

提案システムの枠組みは、さらに「流通している個人カプセルを削除する」という用途への拡張が考えられる。また、このシステムは著作権の管理に関する仕組みであるが、同様のシステムによって、著作権だけでなく、個人情報をカプセル化して保護するシステムとしても拡張することが可能であると思われる。

今後、これらの拡張機能を含め、システムを実装する具体的方法について検討する。

【参考文献】

- (1) さまざまな DRM のスタンス、
<http://www.jagat.or.jp/story memo view.asp?StoryID=4242>
- (2) D. Brickley and R. V. Guha: Resource Description Framework (RDF) Model and Syntax Specification
<http://www.w3.org/TR/REC-rdf-syntax>
- (3) 牛頭靖幸、森住哲也、稲積泰宏、木下宏揚：Covert Channel 分析評価のためのアクセス制御エージェントシステムの提案、コンピュータセキュリティシンポジウム2004 (2004.10.20～22)
- (4) 杉本重雄：Dublin Core について—概要—
<http://avalon.ulis.ac.jp/sugimoto/Articles/JohoKanri0702.htm>
- (5) Dublin Core (ダブリン・コア)：書誌情報メタデータの共通語彙
<http://www.kanzaki.com/docs/sw/dublin-core.html>
- (6) 西岡秀一、高田智規、山本隆二、阿部剛仁、川村春美、大村弘之、曾根原登、有沢博：デジタルコンテンツに関する権利流通基盤の構築、情報処理学会論文誌：データベース、Vol.45、No.SIG 7 (TOD22)、pp.243-254.
- (7) Wang, X., Lao, G., DeMartini, T., Reddy, H., Nguyen, M. and Valenzuela, E.: XrML . eXtensible rights Markup Language, Proc. 2002 ACM workshop on XML security, pp.71.79, ACM Press (2002).
- (8) IPR Systems: Open Digital Rights Language (ODRL)
<http://www.odrl.net/>
- (9) Microsoft: Windows Media Rights Management (WMRM)
<http://www.microsoft.com/>
- (10) RealNetworks: Real System Media Commerce Suite (RSMCS)
<http://www.realnworks.com/>
- (11) T. Berners-Lee, R.T. Fielding, and L. Masinter: Uniform resource identifiers (URI): Generic syntax
<http://www.isi.edu/in-notes/rfc2396.txt,1998>. RFC2396
- (12) World Wide Web Consortium: RDF vocabulary description language 1.0: RDF Schema
<http://www.w3.org/TR/rdf-schema/,2004>. W3C Recommendation 10 Feb. 2004.
- (13) RSS-DEV Working Group: RDF Site Summary (RSS) 1.0
<http://web.resource.org/rss/1.0/> (2000).
- (14) D. Brickley and L. Miller: FOAF vocabulary specification
<http://xmlns.com/foaf/0.1/,2003>. RDFWeb Namespace Document 16 Aug. 2003
- (15) G.A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller: Introduction to WordNet: An on-line lexical database
<http://www.cogsci.princeton.edu/wm/> (1993)

- (16) Netscape: Open Directory Project
<http://dmoz.org/>
- (17) Gene Ontology Consortium: Gene Ontology.
<http://www.geneontology.org/>
- (18) 豊城かおり、寺西裕一、奥田剛、下條真司、宮原秀夫：コンテンツの編集を考慮した権利管理機構の提案と実現、情処学研報、2002-DPS-106、 pp.241-246 (2002).
- (19) 中江政行、細見格、市山俊治：ユーザ要求に適合したサービスを提供するカプセル化コンテンツ、電子化知的財産・社会基盤、pp.3-11 (1999).
- (20) 山田尚志、石原淳、加藤拓：マルチメディア時代のコピープロテクション、電子情報通信学会論文誌A、Vol.J87.A,No.6、 pp.734-745 (2004).
- (21) 寺西裕一、豊城かおり、奥田剛、下條真司、宮原秀夫：ASIA: 派生コンテンツの利用制約管理が可能な情報配信システム、電子情報通信学会誌B、Vol.J86.B, No.8,pp.1463-1475 (2003).

Ⅵ Webでの提示を想定した浮世絵の技法の表現法

はじめに

浮世絵とは、江戸時代に成立した、人々の日常生活や風物などを描いたものである。肉筆画も含まれるが、一般的に浮世絵というと、多色刷りの木版画（錦絵）を指すことが多い。浮世絵にはさまざまな技法が存在する。いくつかの技法を以下に示す。

から摺り（からずり） 絵の具を用いずに、版木を強く押し当てて摺ることによって、紙に凹凸をつける技法。それによって、紙に無色の線や図形が表現される。

きめ出し（きめだし） 「から摺り」と良く似ていて、凹面の版木を作り、紙を当てて裏からたたき出して紙にゆるやかな凸面を作る技法。

このような技法によって紙面に凹凸がつけられ、鑑賞する際にさまざまな表情を見せる。非文字資料として浮世絵をみると、絵画としての一面だけでなく、制作された時代背景、技法であったり、技法によって生じる表面の微妙な質感であったりと、既存メディアで固定不能な情報を多く持っている。

近年、急速なWWWの発展により、非文字資料を閲覧できる電子図書館や電子美術館が増えてきている。そこでは浮世絵に関しては二次元カラー画像としてデジタル化され保存されている⁽⁶⁾⁽⁷⁾。

浮世絵には技法により表面にわずかに凹凸がついたものが存在する。しかし、Webで閲覧できる非文字資料の画像は、主に紙面全体に光を照射し撮影されたものである。そのため、現在のWebでの提示法では浮世絵の立体的表現を判断することは困難であり、作者が意図して制作した表現を鑑賞者に伝えることが出来ない。非文字資料を情報化するには既存メディアで表現することしか出来ないため、人類遺産の蓄積と文化情報の発信のためには、元の非

文字資料を出来るだけ忠実に表現し、浮世絵の技法に適した方法で凹凸を表現した画像を生成し、Webで提示することが必要である。

三次元的な油彩画の表示手法として、非接触式三次元形状測定器を用いることで、油彩画表面の凹凸形状を表すデータを獲得し、テクスチャをマッピングしていく手法がある[8]。しかし、この手法に用いられているような特殊な機器を使用する機会に誰しも恵まれているわけではない。そこで、特殊な機材を使用せずに、浮世絵の技法による立体的表現を鑑賞者に伝える手法を考える。

上記をふまえ本研究では、神奈川大学21世紀COEプログラム「人類文化研究のための非文字資料の体系化」の一環として、浮世絵を記録精度が高く、保存性・情報発信に優れたデジタルデータの形で記憶し、浮世絵の技法である「から摺り」によって生じる浮世絵表面の凹凸を、浮世絵写真の凹凸部に影をつけることによって表現した二次元カラー画像を生成する。また、最終的にWebで情報発信を行い研究者間での情報共有を容易に行うことを目指しているため、Webで閲覧者が光源を自由に変化させられることと同等の効果を得られるインタフェースを作成することを想定し、任意の一定方向から光源を照射した際に生じる影を表現する画像を生成する手法を提案する。

本研究では、特殊な機材を使用せず、カメラとストロボを用いて3枚の写真画像を撮影し、使用する。

提案手法

浮世絵の表面には技法により、わずかな凹凸がついたものが存在する。本研究で使用する浮世絵を図6.1に示す。この浮世絵には「から摺り」という技法により、凹凸がついている。凹凸は主に、雪の粒の形を表現した細かな円形の凹凸と、木の幹に雪が積もっていることを表現した線状の凹凸がある。図

6.2に凹凸部分の拡大図を示す。

本研究では、これらの凹凸部分に任意の一定方向に影を生成し、一定方向から光を照射したものと同等の効果を得られる画像を生成する。

提案手法

浮世絵の凹凸を表現した画像を生成する提案手法の概要を以下に示す。

1. 浮世絵の撮影
2. 輝度値補正
3. 差分画像の生成
4. 二値化处理
5. ノイズ除去
6. 凹凸部のみが現れた画像の生成
7. 影の生成

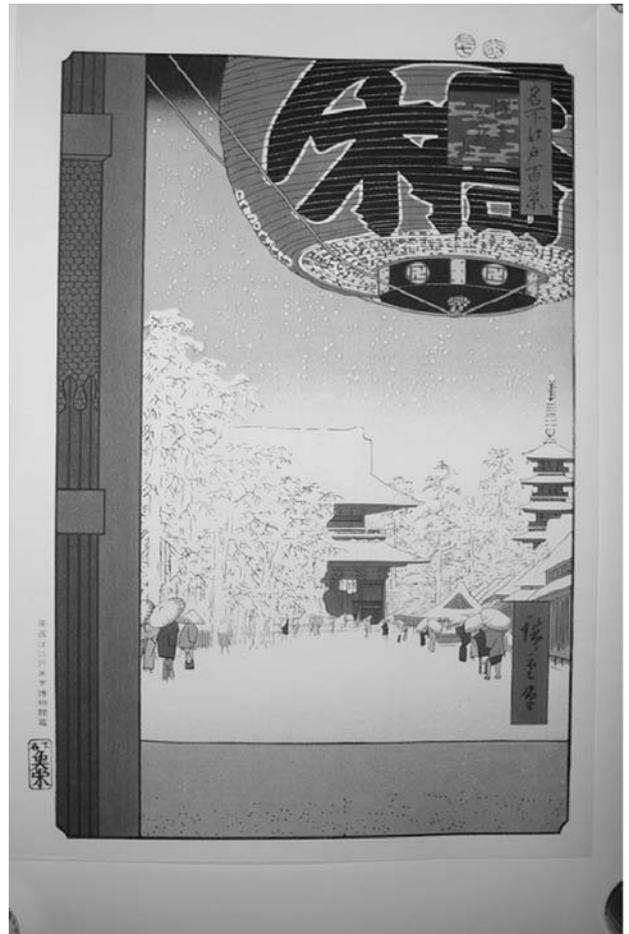
浮世絵の撮影

凹凸を表現した画像を生成するために浮世絵をカメラで撮影する。

撮影条件の概略図を図6.3に示す。

まず、暗室に浮世絵を平面に置き固定し、浮世絵の中心から垂直方向の上部55cmにカメラを固定する。さらに、浮世絵の中心から上下方向1m35cmの位置にストロボを固定し、浮世絵に対し45度の角度で設置する。[A]の位置からのみストロボを照射した画像[A]、[B]の位置からのみストロボを照射した画像[B]を撮影する。カメラとほぼ同じ位置から浮世絵に垂直にストロボを照射した画像[C]を撮影する。画像[A]、[B]、[C]を2336×3504の分解能で獲得する。画像フォーマットとして、汎用性が高くUNIXにおいて標準的な画像形式である、ppmバイナリ形式とpgmバイナリ形式を用いる。画像[A]、[B]をpgm形式、画像[C]をppm形式として保存する。画像[A]、[B]は差分画像を生成する際に用いる。画像[C]は提案法により影を生成する際のベースとなる画像として用いる。

図6.1 本研究で使用する浮世絵



輝度値の補正

提案法による画像を生成するにあたり、画像[A]、[B]には画像の輝度値にむらが生じてしまっている。撮影条件の図6.3から分かるように、浮世絵の上部と下部では光源からの距離が異なるため、照度にむらが生じてしまう。輝度値にむらが生じると、後の工程である差分画像を生成する際に、凹凸がうまく現れないなどの不具合が発生する。ここでは、画像[A]、[B]に対して、縦方向、横方向にそれぞれ輝度値の補正を行う。以下に縦方向の輝度値を補正する手法を示す。

まず、浮世絵の余白部分を用いて縦方向に20画素ごとの輝度値を取り出す。浮世絵の余白を用いる理由として、浮世絵に使用されている和紙の反射率で測定することにより、より良く補正ができるからである。

それを数箇所において取り出し、平均した値に対

して最小二乗法により近似関数を決定する。取り出した平均の値と最小二乗法による近似関数のグラフを図6.4に示す。最小二乗法により求めた近似関数にもとづき、縦方向の輝度値を補正する。

同様の処理を横方向にも行い、画像[A]、[B]の輝度値が補正される。

図6.2 凹凸部の拡大画像

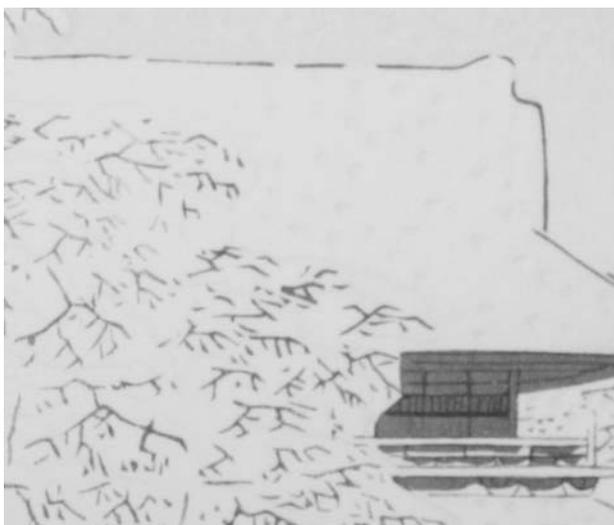


図6.3 撮影条件

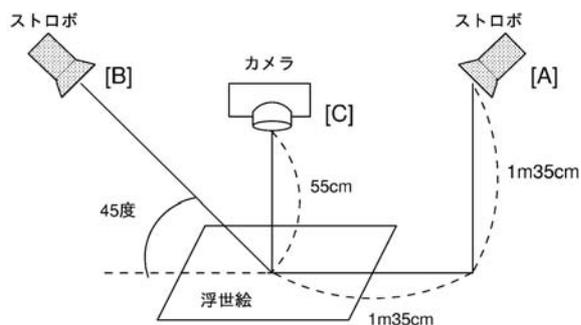
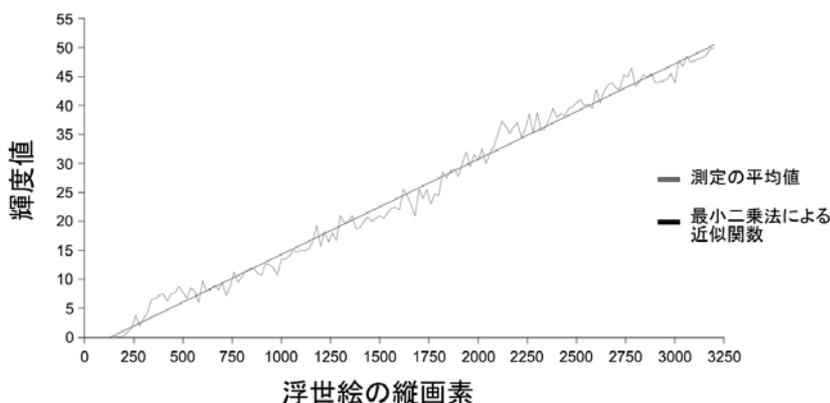


図6.4 縦方向における平均の輝度値と近似関数



差分画像の生成

浮世絵の凹凸位置検出のため、輝度値の補正をおこなった画像[A]、[B]の差分画像を生成する。画像[A]、[B]には照射角度の違いにより、紙面の凹凸部に生じる影の位置が異なる。画像の凹凸部以外の紙面は理想的には差が生じないため、画像[A]、[B]の差分画像を生成することにより、凹凸位置が差分として画像に現れる。生成した差分画像の凹凸部の拡大図を図6.5に示す。

画像[A]、[B]はグレイマップ画像のため、図6.5に示す差分画像は、黒い画素ほど差が少なく、白い画素ほど差が生じていることを示している。

二値化処理

ここでは、後の工程であるノイズ除去や影位置の生成を簡単にするため、差分画像に対して二値化処理を行う。閾値を決定する際、閾値を小さくすれば凹凸ははっきりと現れるが、比例してノイズも多くなる。ある程度のノイズは後の工程で除去できるが、閾値を小さくしすぎると凹凸部とノイズが連結してしまい、除去することができなくなってしまう。そのため、閾値は凹凸部がはっきりとわかる程度の最大の値にするのが望ましい。適当な閾値を決定し、二値化処理を行い二値画像を生成する。閾値を13として二値化処理した画像の凹凸部の拡大図を図6.6に示す。

図6.5 差分画像

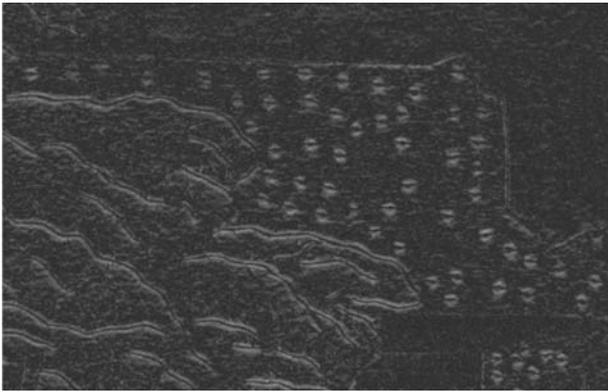


図6.6 閾値13で二値化処理を施した二値画像



図6.7 周囲長20以下の連結領域を除去した二値画像

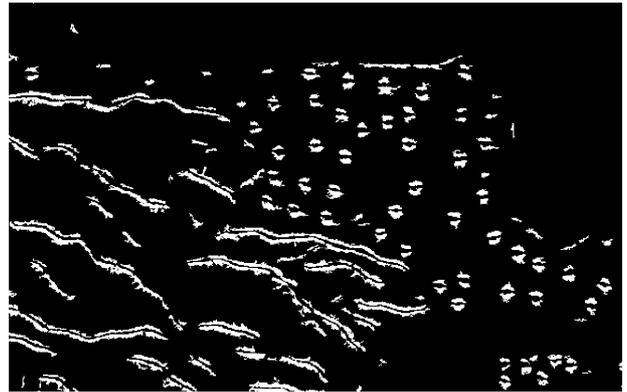


図6.8 ノイズ除去した画像の拡大図



図6.9 入射する光により生じる円形凹凸部の影

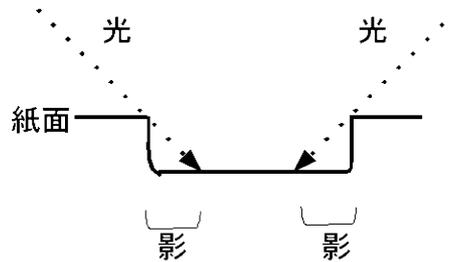


図6.10 凹凸部のみが現れた画像



ノイズ除去

図6.6に示すように、二値化処理を施した画像には、数ドットから十数ドットの大きさのノイズが多数現れる。ノイズが生じたままだと後の工程で影を生成する際、ノイズ部分にも処理を行ってしまう。そこで、ノイズ除去のため、二値画像中の連結領域に対して輪郭線追跡を行い、連結領域の周囲長 L を得る。指定した値よりも短い周囲長の連結領域を画像から除去する処理を施す。指定する値を大きくしすぎると、凹凸部の連結領域も除去されてしまうため、指定する値は凹凸部が除去されない最大の値をとることが望ましい。二値画像に対して、周囲長20以下の連結領域を除去した画像の凹凸部の拡大図を図6.7に示す。

凹凸部のみが現れた画像の生成

ノイズを除去した画像の円形凹凸を拡大した画像を図6.8に示す。図6.8を見てみると、円形凹凸部上

下に差分が現れている。これは図6.9のように、浮世絵の凹凸部に対して45度の角度で上下からストロボが照射されるためである。このように上下に別れてしまっていると、影を生成する位置を決定することができないため、この上下の差分に挟まれた部

分を、上下部分と同じ白画素にする処理を施すことにより、凹凸部のみが現れた画像が生成される。ノイズを除去した二値画像の凹凸部の上下の差分に挟まれた部分を白画素にし、凹凸部のみが現れた画像の凹凸部の拡大図を図6.10に示す。

影の生成

凹凸部のみが現れた画像を用いて、影を生成する位置を決定する。ここでは、凹凸の下部に影を生成する場合を述べる。まず、画像の下方向から走査していき、黒の背景画素から白の図形画素（凹凸部）に変化する位置を探索する。検出した位置から上方向に数ドットを影をつける位置として決定しておく。影を生成するベースとなる画像は、正面からストロボを照射し、影が現れていない画像[C]を用いる。影をつける位置を決定した画像における影をつける位置と、画像[C]において同じ位置の画素のRGBに対してHSI変換を行う。HSI変換を行った画素の明度Iを適当な値で下げることにより、画像[C]の影を生成したい位置の画素を暗くすることができる。ノイズ部分にも影を生成してしまうが、HSI空間で処理を行うことによって、元の色彩を保ったまま明度を落とし、目立たなくすることができる。RGB空間で影となる色を生成しようとする、ノイズ部分にも同じ色が乗ってしまい非常に目立ってしまう。最後にHSI逆変換を行い、RGB空間に戻し画像が完成する。画像[C]の凹凸部の下側に影を生成した画像の凹凸部の拡大図を図6.11に示す。

実験結果

正面から光を照射し、影が現れていない画像[C]の拡大図を図6.12に示す。本研究の提案法により画像[C]の凹凸部下側に影を生成した画像の拡大図を図6.13に示す。

考察

提案法による生成画像と画像[C]を比較すると、

図6.11 凹凸部下側に影を生成した画像



画像[C]には全体的に光があたっているため、凹凸部に影が現れず、凹凸を認識することは困難である。提案法による生成画像には、凹凸部下側に影があるため、浮世絵の持つ立体的表現を認識しやすくなっており、上方向45度から光を照射した「から摺り」の効果を表現できていることがわかる。また、任意の方向に影を生成することが出来るため、光源の位置を変化させたことと同等の効果のある画像を生成することができる。これは、今後Webで情報発信する際、光源位置を自由に变化させられるような操作を行いたい場合に有効であると考えられる。

今回、提案法による画像を生成するにあたり、差分画像を生成する際、浮世絵上部や余白部分に大きなノイズがのってしまい、それが最後に影位置を決定する工程まで残ってしまった。それにより、影を生成する際にノイズ部分の明度も下げてしまい、凹凸部にのみ明度を下げるという理想的な画像を得ることができなかった。ノイズが生じてしまった原因として、写真撮影の際、平面に浮世絵を固定したが、浮世絵は糊状のもので台紙に張られており、表面は全体的にわずかながら波をうっている。なるべく平らになるように余白部分におもりを置き固定して撮影したが、浮世絵が摺られている部分の波うちを固定することはできなかった。その波うちの影響で光が均一に照射されず、光源から距離の影響以外の輝度値のむらが生じてしまい、輝度値を最小二乗法で補正する際にうまく補正できなかったと考えられる。また、浮世絵上部の赤の部分で多く差分が現れてしまったことから、浮世絵に使用される顔料の色彩による反射率の違いがあると思われる。

図6.12 影が現れてない画像[C]



図6.13 提案法により生成した画像



まとめ

本研究では神奈川大学 21 世紀 COE プログラム「人類文化研究のための非文字資料の体系化」の一環として、非文字資料を忠実に表現する手法を提案した。非文字資料は、既存メディアでは表現不能な情報を数多く持っている。しかし、非文字資料を情報化するためには、既存メディアで表現することしか出来ない。今回、提案法による画像を生成することで、非文字資料が持つ記録化することが難しいものを情報化することができた。さらに今後、最終的に Web で情報発信を行うことを想定している。Web で情報発信することにより非文字資料の情報共有と流通が図られ、研究の活性化につながると考えられる。

今後の課題として、凹凸をより忠実に表現できるような影の生成法や、輝度値のむらを少なくするための撮影条件の改良、更なるノイズ除去の手法を考えなければならない。また、Web で浮世絵を提示する際に、自由にあらゆる位置から光を照射したものと同等の効果のあるような操作ができるインタフェースを作る必要がある。

【参考文献】

- (1) 神奈川大学 21 世紀 COE プログラムシンポジウム報告 2 非文字資料とはなにか ～人類文化の記憶と記録～、(2006).
- (2) 木下、稲積、木下、森住：COE における非文字資料の共有と流通、信学技法、SITE2005-24、pp.125 (2005).
- (3) 神奈川大学 21 世紀 COE プログラムシンポジウム報告 1 版画と写真－19 世紀後半 出来事とイメージの創出－、(2006)
- (4) 浮世絵技法用語説明、<http://www.tt.rim.or.jp/haruto/ukiyoe/yougo/gihou.html>
- (5) 浮世絵、<http://ja.wikipedia.org/wiki/>
- (6) 那珂川町馬頭広重美術館、<http://hiroshige.bato.tochigi.jp/>
- (7) 中山道広重美術館、<http://museum.city.ena.gifu.jp/index1.html>
- (8) 及川、亀田：鑑賞を目的とした油彩画の 3 次元デジタルアーカイブ手法、IMPS2006、I-4、18、pp.93-94.
- (9) 最小二乗法について、<http://szksrv.isc.chubu.ac.jp/>
- (10) 安居院、長尾：知的画像処理、昭晃堂
- (11) 田村：コンピュータ画像処理、オーム社
- (12) 吹抜：画像のデジタル信号処理、日刊工業新聞社
- (13) 安居院、長尾：画像の処理と認識、昭晃堂
- (14) デジタル画像処理編集委員会：デジタル画像処理、CG-ARTS 協会
- (15) 安居院、長尾：C 言語による画像処理入門、昭晃堂
- (16) HSV 色空間、<http://ja.wikipedia.org/wiki/HSV>