

トピックマップとその応用

内藤 求

motom@synergy.co.jp

株式会社シナジー・インキュベート

トピックマップは、ISO の標準であり、情報リソースとは独立した上位層（メタレイア）に位置付けられ、情報を意味的に組織化、検索、ナビゲートするための新しいパラダイムである。ISO では、関連する規格の策定についてロードマップを作成し、それに沿って検討を続けている。まず、現在、中心になって検討されている 2 種類のデータモデル、すなわち、Reference model、Standard application model の解説を試みる。次に、海外及び、国内のトピックマップの適用事例を紹介する。その後、弊社での適用事例として、「配食サービスにおける提供者と利用者の Match-Making」システムを紹介する。

Topic Maps and application of Topic Maps

Motomu Naito

Synergy Incubate Inc.

Topic Maps is a standard of ISO and is a new paradigm for organizing, retrieving, and navigating information semantically. Topic Maps make it possible through the provision of upper layer (meta layer) that is independent of information resources themselves. ISO made a roadmap to guide the activity and continue the work along with it. Two kind of data model of Topic Maps, Reference model and Standard application model are being worked on now mainly. There are many applications of Topic Maps in foreign countries. This paper describes some example of them. Then this paper describes our experience, matchmaking of user and provider in catering service.

1 . はじめに

Web の大爆発により、提供者の視点から編集された多種多様な情報が、毎日洪水となって押し寄せてくる。必要な情報をタイムリーに入手可能にするため、新しい検索エンジンや、次世代の Web といわれる Semantic Web など、様々な試みがなされている。

そのうちのひとつで、2000 年に、ISO の標準になったトピックマップについて、その概要と応用例について述べる。トピックマップは、利用者の視点から情報を分類、整理することができ、情報洪水、情報過多に対する対応方法の一つとして利用することができる。

まず、トピックマップの概要について記述する。次に、アメリカやヨーロッパでの実際の適用例を紹介する。また、国内での適用例としてシナジー・インキュベートでの使用例について記述する。最後に、まとめを行う。

2 . トピックマップ

2 . 1 トピックマップとは

トピックマップは、ISO/IEC JTC1 SC34 WG3(以降、SC34 WG3 とよぶ)で策定された標準(ISO/IEC 13250 Topic Maps)である。以降 ISO 13250 と呼ぶ。現在も、SC34 WG3 と非営利のコンソーシアムである OASIS (Organization for the Advancement of Structured Information Standards)において、ISO 13250 の改定と、多数の関連規格の策定について検討が続けられている。

トピックマップの考え方それ自体の起源は、1993 年の Davenport グループでの作業だとされている。その後、HyTime での活動を経て、1996 年から、ISO で検討が継続され、2000 年に ISO 13250 として国際標準になった。

トピックマップは、情報リソースとは独立した上位層（メタレイア）に位置付けられ、情報を意味的に組織化し、管理、検索、ナビゲートを可能にするための新しいパラダイムである。情報リソースが持つサブジェクト（概念）とサブジェクト間の関係を、情報リソースとは独立にコンピュータ上でモデル化（reify:具体化）する。サブジェクトに関連する情報リソースに対しては、リンクを張ることにより関係を明示する。

トピックマップの主要な構成要素として、以下の三大要素および3種類の Subject indicator（ポインタ）がある。

<三大要素>

- ・ topic：トピック（人間が認識する具体的または抽象的なサブジェクト / 概念を代理する要素）
- ・ association：アソシエーション（トピック間の関連を表す要素）
- ・ occurrence：オカレンス（実際の情報リソース）

<3種類の Subject indicator（ポインタ）>

- ・ topicRef（トピックへのポインタ）
- ・ subjectIndicatorRef（サブジェクトの所在を示す情報リソースへのポインタ）
- ・ resourceRef（サブジェクトを構成する情報リソースへのポインタ）

Topic、Association and Occurrence

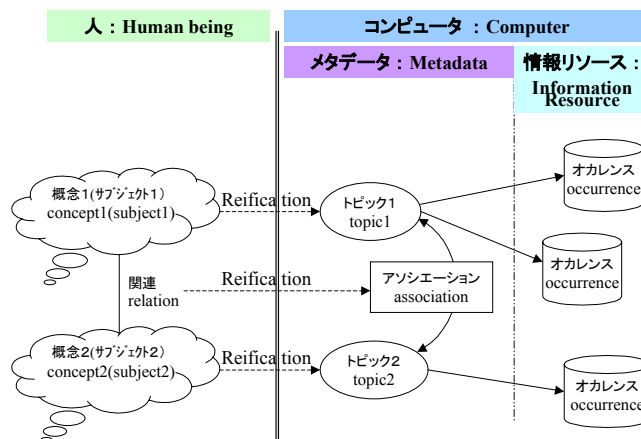


図1. サブジェクト、トピック、アソシエーション及びオカレンス

図1は、サブジェクト、トピック、アソシエーション及びオカレンス（情報リソース）の関係を表現している。トピックマップでは、認識できるすべてのものをトピックとすることができる。アソシエーションについては、XTM（XML シンタクスのトピックマップ）では、以下の2種類が標準の中で定義されている。

- ・ class-instance 関係（トピック間のクラスとインスタンス関係を表現するアソシエーションのクラス）
- ・ superclass-subclass 関係（トピック間の上位クラスと下位クラスとの関係を表現するアソシエーションのクラス）

それ以外、アソシエーションは、例えば以下のような、概念（サブジェクト）間の任意の関係を自由に定義することができる。

- ・ 等価関係（同義、類似・対比）
- ・ 階層関係（上位 - 下位、全体 - 部分）
- ・ 連想関係（並列、用途・環境、因果）
- ・ その他任意の関係

2.2 トピックマップ関連標準策定のロードマップ

今後のトピックマップの標準活動については、策定または検討中の標準および標準の位置付け、関連等、2001年12月に米国フロリダ州オーランドで開かれた SC34 WG3 の会議で作成されたロードマップで明確に示され、それに沿って作業が進められている。

Topic Maps, roadmap for further work

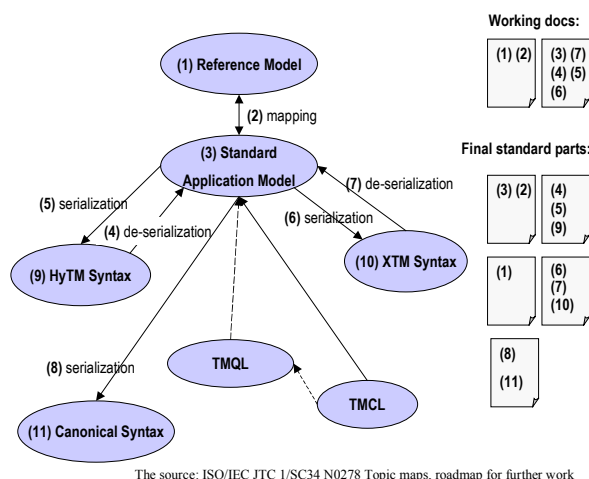


図2. トピックマップ標準策定ロードマップ

2002年5月のスペインのバルセロナの会議では、参照モデル(Reference model)、標準応用モデル(Standard application model)のドラフトについての検討が中心に行われた。以下に主な標準を列挙する。将来、ISO 13250は、マルチパート標準として、これらを束ねたものになる予定である。

日本では、ISO 13250の翻訳版、及び、XTM 1.0の翻訳版が、それぞれJIS X 4157、及び、TR X 0057として経済産業省の技術委員会承認された。現在は、出版を待っている状態である。また、現在ISOで策定中の標準も、適宜、JIS化を検討していく予定になっている。

- ・ HyTM (既にISO標準として成立): HyTimeベースのシンタックス
- ・ XTM (既にISO標準として承認済み): XMLベースのシンタックス
- ・ Reference model (参照モデル): トピックマップの2種類のデータモデルのうちの上位のもの
- ・ Standard Application Model: SAM (標準応用モデル): もう一方のデータモデル
- ・ TMQL (Topic Map Query Language): トピックマップ用の検索言語
- ・ TMCL (Topic Map Constraint Language): トピックマップ用の schema 言語

このうち、二つのデータモデル、Reference model (参照モデル)と、Standard Application Model (標準応用モデル)について、もう少し詳細に検討してみる。Reference modelは、抽象的、汎用的なデータモデルで、Standard Application Modelの上位層に位置付けられる。これに対して、Standard Application Modelは、実装のためのデータモデルであり、Reference modelの下位層に位置付けられる。

(1) Reference model (参照モデル)

Reference Modelは、抽象的なグラフモデルである。すべてのトピックマップは、グラフであり、Reference Modelの中では、トピックマップは、"assertion"の集合と見なされる。各々の"assertion"は、特定のサブジェクト間に、強く型付けされた関係が存在することを表明する。サブジェクトは、"assertion"の中では、"role player"になる。つまり、サブジェクトは、その関係の中で特定の役割を果たす。

ドラフト Reference Modelは、以下の2つの assertion 型を定義している。

- ・ topic-subjectIndicator

(サブジェクトは、subject indicator を持つという宣言するために利用する assertion 型)

・ assertionPattern-role-playerConstraints

(assertion 型、 role、及び role player の関係を宣言するために利用する assertion 型)

assertion (すなわちトピックマップ) は、以下の 4 種類のアークと 5 種類のノードからなるグラフとしてモデル化されている。

・ 4 種類のアーク

- AP (Assertion - assertion Pattern)
- CR (Casting - Role)
- AC (Assertion - Casting)
- C (Casting - role player)

・ 5 種類のノード(endpoint)

- Assertion pattern
- Assertion
- Role
- Casting
- Role player

role、role player、assertion、assertion pattern は、すべてサブジェクトと見なされ、他の assertion の role player になりうる。

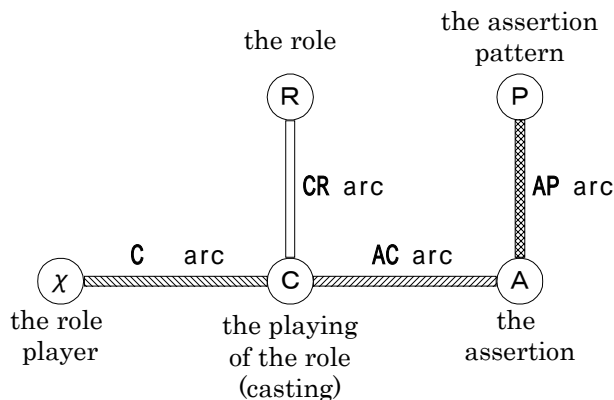


図 3 . Reference Model のグラフモデル

Reference Model での、ノード、アークについての規則を以下に示す。

- ・すべてのノードは、正確に一つのサブジェクトを表す。
- ・すべてのアークは、正確に一つの assertion の構成要素である。
- ・すべてのノードは、複数のアークの複数の endpoint になりうる。

また、別々に作成されたトピックマップを、トピックの重複を排除し、しかも、トピックを失うことなしにマージするために、subject indicator を結合点とするマージについて記述されている。これは、計算機がトピックの意味を理解してマージするのでなく、ネットワーク上のアドレス (すなわち、アドレスが同じなら、同一のトピックと判断する) を基準にマージすること示している。

(2) Standard Application Model (標準応用モデル)

Standard Application Model は、トピックマップの実装が持つべきトピックマップ情報の計算機内での内部表現、及び、構造上の制約、操作を定義している。

このドラフトで使用されているメタモデルは、XML Information Set で使用されているものと同種のものである。すなわち、トピックマップの構造を、情報セットとしてモデル化するものである。トピックマップ情報セットは、各々トピックマップのある部分を表現する情報項目から構成される。情報項目のプロパティは、トピックマップのある一部分についての情報を含む。

情報項目の型としては、以下のものがある。

- ・ Locator items
- ・ Source locators
- ・ The topic map item
- ・ Topic items
- ・ Base name items
- ・ Variant items
- ・ Occurrence items
- ・ Association items
- ・ Association role items

また、マージについて、以下の3種類のマージルールを定義している。

- ・ トピックマップのマージ
- ・ トピックのマージ
- ・ basename によるマージ

この標準に準拠してトピックマップ情報を処理できるソフトウェアモジュールをトピックマッププロセサと呼び、トピックマップ情報を内部表現から、トピックマップ構文に export することを de-serialization、逆に、トピックマップ構文から、内部表現に import することを serialization という。

2.3 Published Subject

トピックマップの重要な構成要素として、Published Subject がある。Published Subject は、多くの人々が共通に使用できると思われるトピックで、ネットワーク上で公開されているものことで、それにより、トピックマップの共有 / 交換を容易にしようとするものである。

Published Subject の標準策定、利用促進については、現在、OASIS の3つの TC(Technical Committee)が積極的に活動している。

(1) Topic Maps Published Subjects Technical Committee

この TC の目的は、Published Subject についての定義、管理、使用方法について、要求、勧告、及び最優良事例を指定することにより、その利用を促進することである。

ISO 13250 の中では、Public Subject として定義されていて、XML トピックマップ (XTM) 1.0 規格の中では、Published Subject として定義されている。

(2) Vocabulary for XML Standards and Technologies TC (XMLvoc TC)

この TC の目的は、XML 標準及び技術の領域についての語彙を定義することである。それらの語彙は、トピックの参照セット、トピックの型、アソシエーションの型などになる。それにより、XML、関連する標準、及び XML コミュニティに関する情報の発見を容易にする。

また、それらの語彙を、OASIS Topic Maps Published Subjects Technical Committee の勧告に則って、Published Subjects として定義する。

(3) Topic Maps Published Subjects for Geography and Languages (GeoLang) TC

この TC の目的は、Published Subjects TC によって作成される Published Subjects ガイドに従って、言語、国、地域の Published Subjects セットを定義することである。

言語、国、地域は、広い範囲のトピックマップでしばしば使用されるサブジェクトである。最大限の再利用性、交換可能性、マージ可能性を促進するために、これらの領域をカバーする Published Subjects の標準セットが要求されている。

3. トピックマップの適用分野と事例

3.1 海外

ヨーロッパやアメリカにおいては、既に多くの事例が報告されている。特にヨーロッパでは、ヨーロッパ統合による情報の氾濫、例えば多言語の問題などもあり、その対応策として、注目されているようである。2002年1月の METAGROUP の市場調査レポートによると、ドイツの200社に対してアンケートを行い、「以下のナレッ

ジマネジメント関連の標準を知っていますか？」という質問に対して、知っているという回答があったのは、XTM:19%、RDF:13%、DAML:8%、OIL:4%で、どれも知らないという回答が 79% ということである。この調査では、RDFより、トピックマップのほうが、知名度が高いという結果が出ている。

適用分野も非常に広い範囲に渡っている。例えば、出版、Web アプリケーション開発、アプリケーション開発、アプリケーションインテグレーション、Web サービス、コンテンツ管理などの分野で利用されている。

以下に適用分野と事例を示す。

表 1 . 海外の適用事例

項番	適用分野	適用事例
1	出版のソリューション	XML カンファレンスのProceedings
		米国内国税歳入局 (IRS) の税関係の出版物
		Cogitech ホームページの作成 (http://www.cogx.com/)
		Quid : フランスの百科事典 (http://www.quid.fr/)
2	Webアプリケーション開発	ITU (IT Research and Competence in Education) - ノルウェー政府出資のオスロ大学の学部 - プロジェクト、関係者、関連ドキュメント等の共有のためのWebアプリケーション
		CSIR iWorks Ideabank (The Council for Scientific and Industrial Research) - 南アフリカ最大の調査機関 - 知識やアイデアを共有するためのWebアプリケーション
		Patrimoine - 金融ドキュメンテーション - 専門的、法的ドキュメントの提供、共有
3	アプリケーション開発	Bravo - トピックマップによるKnowledge Management - GlobalWisdom社のKnowledge Management ツール - 情報の組織化のために、トピックマップエンジンを使用 (http://www.globalwisdom.org/products/)
4	アプリケーション インテグレーション	Starbase - Starbase社のアプリケーションインテグレーション用のソフトウェア製品 - Ontopia社のトピックマップ技術を使用 (http://www.starbase.com/)
5	Web サービス	KnowledgeTaxi - ドイツの情報公開のためのWebソリューション
6	コンテンツ管理	技術情報管理システム - ドイツのHewlett-Packard Consulting社によって開発 - 技術情報の管理にトピックマップを使用

3.2 国内

国内では、まだ、あまり事例がみられない。国内の事例として、弊社での事例を以下に示す。

表 2 . 国内の適用事例

項番	適用分野	適用事例
1	文化資源情報の組織化	源氏物語トピックマップ
		古今和歌集トピックマップ
2	Thesaurus 辞書	医学用語Thesaurus
3	Web アプリケーション	配食サービスにおける提供者と利用者のMatchMaking

ここでは、「配食サービスにおける提供者と利用者の Match-Making」の事例について紹介する。

著者らは、2001年8月～2002年3月にかけて行われた、ある自治体における配食サービスの実証実験において、トピックマップを利用した料理の選択システムのプロトタイプを開発した。配食サービスの利用者が行う料理の選択を IT で支援するために、食事の提供者と利用者の仲介をするためのシステムである。

情報 / 知識も含めたある物について、取引 / 共有が成り立つには、提供者と利用者が同一の物にたどり着き、

それを認識しなければならない。利用者は、利用者の概念空間を様々なナビゲートしながら、ときには、概念の木構造に沿って根元から枝葉に向かってたどり着く場合もあるだろうし、ときには、枝から枝へジャンプしながら辿り着くかもしれないし、直接、目的とする葉にたどり着くかもしれない。提供者は、提供者の概念空間の中で、提供者の分類に則って提供品を位置付けているかもしれない。

具体的な物として料理を考えた場合、利用者がある料理を選ぶ道のりには、いろいろなものが考えられる。例えば、肉、魚、野菜などの材料の視点から選択する場合、日本料理、中国料理、スペイン料理などの国、地域の視点から選択する場合、その他にも、健康に注目している人は、栄養素の視点から選択したり、健康状態がよくない人は、症状にあった料理を選択したりする場合などが考えられる。

他方、提供者（例えば、レストラン）が提供する料理は、立脚するコンテキストの中で、一般的と考えられる分類に則ってメニューが作成されており、メニューの中で位置付けられている場合が一般的と考えられる。レストランの種類によっては、メニューがある特定の料理に偏っていたり（いわゆる専門店）幅広い分類の料理をそろえていたりもする。

料理を選択した状態とは、利用者と提供者、この場合、実体としての物（料理）について共通の認識を持ち、それが特定/指定された状態と考えられる。

利用者側の視点からの料理の分類（オントロジ）と、提供者側の視点からの料理の分類（オントロジ）を、それぞれ、個別にトピックマップで表現し、利用者、提供者が共通に認識出来る物（料理）を定義し、それぞれのトピックマップの中で共通に利用した。別々に作成したトピックマップを、トピックマップが持つ mergeMap 機能でマージすることにより、料理を結合点(binding point)として、専用のマッチング機能の必要無しに、利用者視点からの料理の選択及び提供者の選択（利用者と提供者のマッチメイキング）が可能になった。

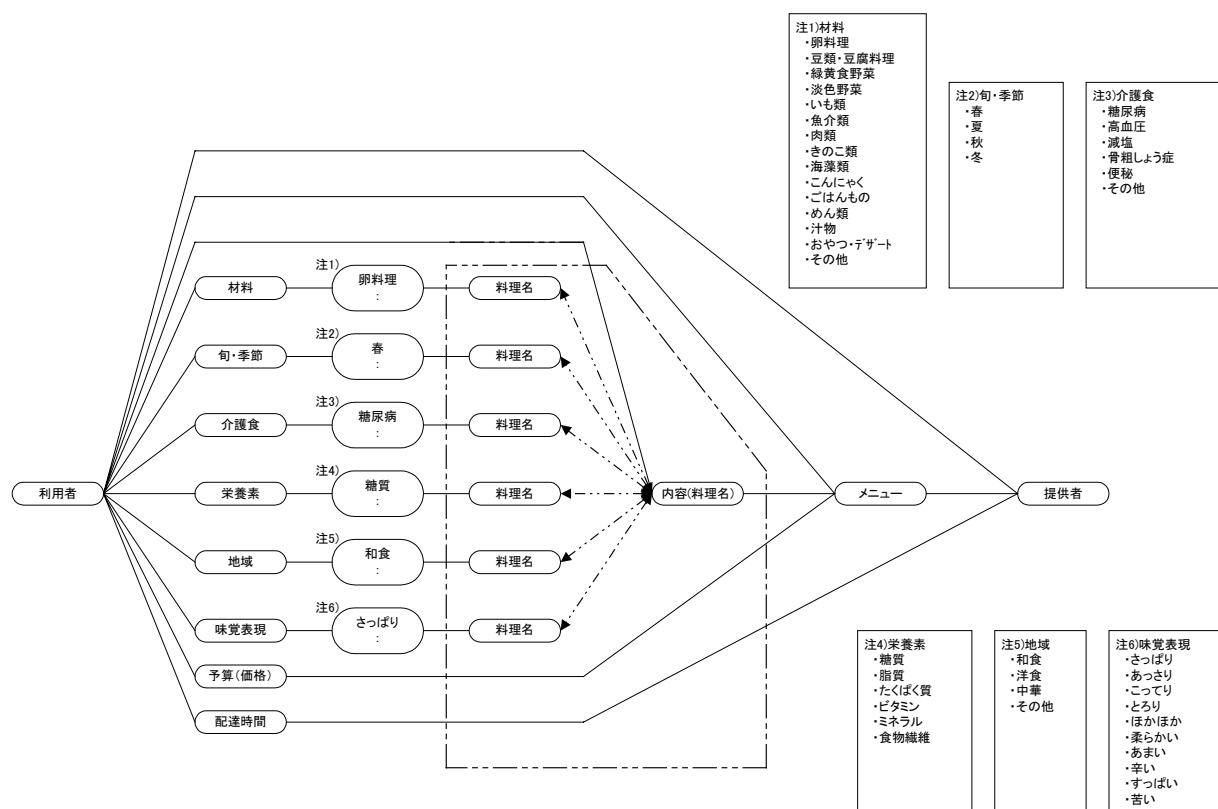


図4 . 利用者視点のオントロジと提供者視点のオントロジの結合

利用者と提供者が共通に認識できる物（料理）は、Published Subject として定義可能と考えられる。それにより、Published Subject が結合点（Binding point）になる。また、今回のプロジェクトでは、共通に認識できる

物として、料理を選んだか、これは、他の物にも適用可能である。

利用者一人一人が個別にトピックマップを作成する。提供者も自分達の分類に則って、それぞれのトピックマップを作成する。その中では、Published Subject として定義されたトピックをお互いに共有する。それぞれのトピックマップを、共通の場(マーケットプレイス)に持ち込んで mergeMap することにより、Published Subject を結合点(Binding point)として、利用者/提供者一人一人にパーソナライズされたマッチメイキング機構が出来あがる。

4. むすび

トピックマップは、非常にシンプルである。トピック、アソシエーションを自由に定義することができ、情報リソースと関連付けることができる。それらにより、任意の問題領域でのオントロジを容易に作成することができる。つまり、利用者自身による視点設定、それに基づく情報の組織化が可能になる。

いろいろなアプリケーションにトピックマップを適用することで、直観的、ユーザフレンドリなユーザインターフェースの構築が可能になり、その付加価値を高めることができると考えている。また、情報過多(Infoglut)対策にも有効である。

オントロジ及びその構成要素であるトピック、アソシエーションに対する計算機処理も容易で、データ処理から、情報処理へと進化してきた、計算機利用を、知識処理へと進化させる可能性を持った方法(道具)の一つとして多様な利用方法が考えられる。

参照情報と文献

- [1] Steve Pepper, "The TAO of Topic Maps." Conference Paper at ACM CIKM'99 .
- [2] International Standard Organization, "ISO/IEC 13250 Information Technology - SGML Applications - Topic Maps", ISO/IEC 13250:2000.
- [3] John F. Sowa, "Knowledge Representation Logical, Philosophical, and Computational Foundations", Brooks/Cole, ISBN 0-534-94965-7 .
- [4] TopicMaps.org, "XML Topic Maps (XTM) 1.0" .
- [5] Topicmaps.net, "Topicmaps.net's Processing Model for XTM 1.0, Version 1.0.1" .
- [6] Steve Pepper, "Modelling Topic Maps", Tutorial Handout, XML 2001, Orlando, Florida .
- [7] ISO/IEC JTC1 SC34, "Topic maps, roadmap for further work", ISO/IEC JTC1 SC34 N0278 .
- [8] Michel Biezunski, Steven R. Newcomb, "The Topic Maps Technical Workshop", KT 2001, Austin, Texas .
- [9] Martin S Lacher, Stefan Decker, "On the Integration of Topic Maps and RDF Data", Proceedings of SWWS '01, pp 331-344, 2001 .
- [10] Graham Moore, "RDF and TopicMaps An Exercise in Convergence", XML Europe 2001, Berlin .
- [11] Mari Nagase, Motomu Naito, "Application and Evaluation of Topic Maps for the Cultural Resource Data --- Experiment with the Graphic Data of "Genjji Monogatari (the Tale of Genji)" --- ", KT 2002, Seattle, 2002.
- [12] ISO/IEC JTC1 SC34, "A Draft Reference Model for ISO 13250 Topic Maps", ISO/IEC JTC1 SC34 N0298 Rev.1
- [13] ISO/IEC JTC1 SC34, "The Standard Applicatoin Model for Topic Maps", ISO/IEC JTC1 SC34 N0299
- [14] Kal Ahmed, "Topic Maps - A practical Introduction With Case Studies", XML Europe 2002, Barcelona, Spain .