

トピックマップ (Topic Maps) の最新標準化動向

内藤 求 (motom@synergy.co.jp)

株式会社シナジー・インキュベート

〒151-0063 東京都渋谷区富ヶ谷 1-30-22 折田ビル 2F

<http://web.synergy.co.jp/> <http://tm1.synergy.co.jp/>

あらまし

2003年3月の投票で、ISO 13250 Topic Maps は、6つのパートからなる規格に更新される案が可決され、現在、改訂のための議論が活発に行われている。また、ISO 18048 Topic Maps -- Query Language、ISO 19756 Topic Maps -- Constraint Language についても、RDF グループとの議論も含めて活動が活発化している。さらに、Published Subjects についても、OASIS において作業が進められている。これらの動向について報告する。また、適用事例、予想される利用方法についても述べる。

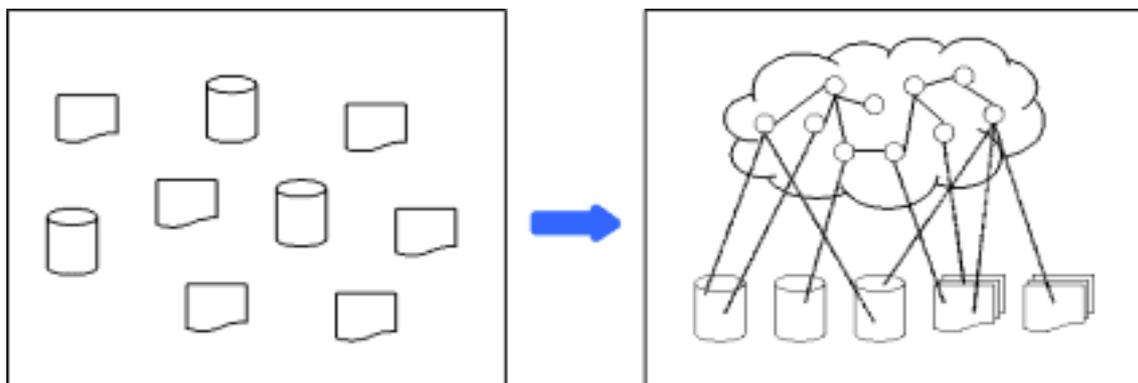
1. トピックマップとは

Topic Maps は、膨大な量の情報に対して ” 地図 (MAP) ” を与えるもので、” 個々に孤立した情報を、体系化、組織化された知識へと ” 高めて、その共有、流通、有効活用を促進する。

Steve Pepper 氏は、Topic Maps について以下のように言っている。

- ・ 知識構造の記述と情報実体への関連付けのための ISO 標準
- ・ 情報宇宙の GPS (Global Positioning System)
- ・ 大容量で相互接続された情報群のナビゲーションツール

Topic Maps は、孤立して存在する情報群を、情報実体から独立に、体系化、組織化する。また、それ自体が共有、交換、流通可能である。



Topic Maps 利用のメリットのいくつかを以下に示す。

- ・ 概念、意味レベルでの情報操作が可能
- ・ 情報実体とメタデータ/スキーマの分離
 - 情報実体の変更なしに、情報の体系化、組織化が可能
 - 概念体系のメタレイアでのモデル化と、体系に沿った情報実体へのリンクが可能
 - 情報群のカテゴリ分類、管理に有効
- ・ 主題の厳密な識別が可能
- ・ 直観的な馴染みやすさ
- ・ デジタルコンテンツの流通、共有の促進
- ・ ISO 標準
 - 情報資源の永続性の確保
 - 多様なツール選択、利用が可能
- ・ 情報資源に対する利用者視点の設定が可能
- ・ 利用者視点での情報の組織化が可能 (Personalization)

2 トピックマップの標準化動向

2.1 ISO の標準

トピックマップは、ISO/IEC JTC1 SC34 WG3 (以降、SC34 WG3 とよぶ。Convenor は、ノルウェーの Steve Pepper 氏である) で策定された標準である。正式には、ISO/IEC 13250 Topic Maps (以降 ISO 13250 と呼ぶ) といい、第 1 版は、2000 年に IS (国際標準) になっている。2002 年には、第 2 版が出版されている。現在も、SC34 WG3 において、ISO 13250 の改定と、関連規格の策定作業が続けられている。

SC34 (文書の処理と記述の言語) の WG3 (情報の関連付けを担当) の国際会議は、アメリカやヨーロッパでの XML カンファレンスなどと併設で、大体年 4 回開催されている。ISO の国際会議は、W3C、IETF や OASIS と異なり参加できるのは、National Body のメンバだけである。投票権のある参加国は、アメリカ、フランス、イギリス、ノルウェー、日本である。その他に、ドイツはこの WG では投票権を持っていないが、積極的に参加している。

国内の受け皿として、(社) 情報処理学会 情報規格調査会によって、JTC1/SC34 専門委員会が組織され、日本 National Body として、国際会議、投票、議論などに参加している。トピックマップ関連の規格とその状況を以下に示す。

(1) ISO/IEC 13250: Topic Maps

トピックマップの中心となる規格で、以下のパートから構成され、それぞれのパートについて作業が進められている。

- ・ ISO 13250-1: Topic Maps -- 概観及び基本概念
2003/8 月のモントリオールの会議までに、working draft を作成する。
- ・ ISO 13250-2: Topic Maps -- データモデル
2003/5 月のロンドンの会議での決定事項を反映させた新バージョンを出来るだけ早く作成し、CD 投票用に提出する。
- ・ ISO 13250-3: Topic Maps -- XML 構文
2003/5 月のロンドンの会議での決定事項を反映させた新バージョンを出来るだけ早く作成し、CD 投票用に提出する。
- ・ Topic Maps -- HyTime 構文 (HyTM)
各 National Body がユーザがいるか報告し、もし全然いなければ、13250 の次の版では HyTM を削除する。
- ・ Topic Maps -- Canonical XML 構文
エディタに working draft の提出が求められている。
- ・ Topic Maps -- 参照モデル
2003/8 月のモントリオールの会議に向けて、利用者要求文書を作成する。

* 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 として日本規格協会から出版されている。

(2) ISO/IEC 18048 : Topic Maps Query Language(TMQL)

2003/8 月のモントリオールの会議に向けて、利用者要求文書と利用者ユースケース文書を作成する。

(3) ISO/IEC 19756 : Topic Maps Constraint Language(TMCL)

2003/8 月のモントリオールの会議に向けて、利用者要求文書とユースケース文書を作成する。

規格の策定において、最近、トピックマップ陣営の Steve Pepper 氏、及び、Lars Marius Garshol 氏と、RDF 陣営の Eric Miller 氏、Dan Brickley 氏、及び、Libby Miller 氏といったメンバで、いろいろ活発に意見交換が行われ、可能な部分は、相互補完していこうとする動きが出ている。特に制約言語や検索言語については、共通化の可能性を模索する動きが出ている。

ロンドンで、2003 年 5 月に、ISO SC34 の国際会議とほぼ同時に開かれた XML Europe

2003 においても、Steve Pepper 氏から”Curing the Web’s Identity Crisis: Subject Indicators for RDF”、Lars Marius Garshol 氏から”Living with Both Topic Maps and RDF”という発表があり、相互補完や、資源の共有化の提案がなされている。

2.2 Published Subject

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をネットワーク上で公開し、それを使用することにより、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。SGML/XML の標準化、利用促進の団体である OASIS にて標準の策定、利用促進のための検討が行われている。また、前述したように、XML Europe 2003 において、Steve Pepper 氏より、Published Subjects による RDF の弱点（主題の識別性の確保）の補完が提案されている。現在、OASIS には、Published Subject に関して、以下の三つの TC（Technical committee）がある。

(1) Topic Maps Published Subjects Technical Committee

(<http://www.oasis-open.org/committees/tm-pubsubj/>)

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を目的としている。

Published Subjects についての入門書である”Published Subjects: Introduction and Basic Requirements”が2003年5月19日の投票で、TC recommendation の最終 DRAFT になることが可決された。

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

(<http://www.oasis-open.org/committees/xmlvoc/>)

XML 標準及び技術の領域についての語彙を Published Subjects として定義することを目的としている。

XML Vocabulary に関連する定義、トピックマップなどを収集する作業が進められている。

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

(<http://www.oasis-open.org/committees/geolang/>)

言語、国、地域の Published Subjects セットを定義することを目的としている。

その一部が、<http://psi.oasis-open.org/iso/639/> で公開されている。

3 トピックマップの適用分野と事例

3.1 海外の事例

欧米では、既にかかなりの適用事例がある。XML Europe 2002、XML 2002、及び、XML Europe 2003 などのヨーロッパやアメリカの XML カンファレンスでは、それぞれ、10 前後のトピックマップ関係のセッションがあり、最も発表件数の多いテーマの一つになっている。以下、Techquila 社の Kal Ahmed 氏の分類をベースに適用分野と事例を列挙する。

以下の事例の詳細については、「平成 14 年度 将来型文書統合システム標準化調査研究委員会 成果報告書」、財団法人日本規格協会に記述してある。

(1) 出版のソリューション (Publishing Solutions)

ここでは、印刷物、電子出版の両方を含めている。

- ・ XML カンファレンスの Proceedings
- ・ アメリカの内国税歳入局 (IRS) の税関係の出版物
- ・ トピックマップベースの ホームページの作成 (Cogitech)
- ・ フランスの百科事典 (Quid)

(2) Web アプリケーション開発

- ・ ITU (Network for IT-Research and Competence in Education) の Web サイト
<http://www.itu.no/>
- ・ CSIR iWorks Ideabank (The Council for Scientific and Industrial Research)
- ・ Patrimoine - 金融ドキュメンテーション

(3) アプリケーション開発

- ・ Bravo - トピックマップによる Knowledge Management

(4) アプリケーションインテグレーション

- ・ Starbase 社のトピックマップ応用製品
- (注) Starbase 社は、2003 年 1 月に、Borland 社に買収された。

(5) コンテンツ管理

- ・ アメリカエネルギー省 (DOE: US Department of Energy) での適用事例

(6) e-Learning, telelearning

- ・ 教材のナビゲーションアプリケーション

<http://www.LmTM.de/>

(7) Web ポータル

- ・ [forskning.no](http://www.forskning.no/) 科学技術情報のポータルサイト
- ・ <http://www.forskning.no/>
- ・ [forbrukerportalen.no](http://www.forbrukerportalen.no/) 消費者情報のポータルサイト
- ・ [matportalen.no](http://www.matportalen.no/) 食品情報のポータルサイト
- ・ [avhending.no](http://www.avhending.no/) 不動産販売のポータルサイト
- ・ [hoyre.no](http://www.hoyre.no/) 政党のポータルサイト

3.2 日本での事例

日本でも、徐々に普及が進んでいる。既に、いくつかの企業やアカデミック機関で利用の検討や、プロトタイプ的な適用が試みられている。ただ、企業秘密にかかわり、名前を出せないケースがある。弊社は、複数のプロジェクトにおいてプロトタイプシステムを構築している。また、弊社のサイト (<http://tm1.synergy.co.jp/>) から、いくつかのトピックマップの事例を見ることができる。

(1) 標準及び標準活動の分類

- ・ XML Vocabulary (XML 関連規格、標準化組織) についてのトピックマップ

(2) Web 上での MatchMaking

- ・ 配食サービスにおける提供者と利用者の MatchMaking

(3) 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

(4) Thesaurus

- ・ Thesaurus 辞書のトピックマップ化

4. 今後の利用予想

前述したトピックマップの適用事例から、今後の以下のような方向で、広く利用されていくことが予想される。

(1) 文書横断的な統合インデックス

ドキュメントとは独立(ドキュメントの外部)に、一つのドキュメントに止まらず、存在するすべてのドキュメント横断的、統合的なインデックスが作成される。それらは、柔軟性にとんでいて、必要な主題を見つけやすくする。

(2) オントロジの表現と操作及びそのアプリケーションでの利用

オントロジ、シソーラス等を基に、対象領域の主題が概念的に体系化され、それに則って情報が編成され、利用者の概念体系を基にした視点の設定を可能にする。

(3) ダイナミックなコンテンツ生成

要求した時点で、いろいろな情報源からコンテンツがダイナミックに生成される。コンテンツの実体がなく情報源が更新されるのに伴い内容が変化する。場合によっては、再現性の確保が困難になる。

(4) コンテンツのインテグレーション

コンテンツが地理的、システムの分散されたいろいろな情報源(リポジトリ、DB、アプリケーション、Web サービス)から、統合され作成される。

(5) コミュニティによる作成、共有

特定多数、不特定多数のコミュニティの中での協業を通してコンテンツが作成され、共有、再利用、改善される。

これらの特性を通して、コンテンツに含まれる知識をより有効に活用できるようになると考えられる。特に(2)のアプリケーションでのオントロジ利用については、様々な可能性が考えられ、また、今のアプリケーションのブレークスルーも期待できる。

5 . まとめ

トピックマップの標準化と利用は、着実に進んでいるように思われる。RDF との相互補完、制約言語、検索言語の共通化のための試みも行われている。トピックマップは、既に実用化されている技術であり、オントロジ及びその構成要素である語彙(トピック)、語彙間の関連(アソシエーション)に対する計算機処理も容易である。ネットワーク上に無秩序に存在する大量のコンテンツを、概念体系に沿って組織化、体系化することを可能し、その有効活用を促進する。そのため、今後、いろいろな適用が試みられていくと思われる。