

TOPIC MAPS を利用したマンガメタデータの提案

原 正一郎 内藤 求
京都大学地域研究統合情報センター 株式会社ナレッジ・シナジー

日本のマンガは海外でも広く読まれ、週刊誌から単行本・Web・携帯電話など多様なメディアへの展開が活発であり、既に一つの文化ジャンルを確立していると言える。このような背景から、マンガに関するメタデータの構築と Web による情報提供は重要であると考えられる。マンガの特徴は作品の派生と多様なメディア展開および構造にある。本稿では、マンガのこのような特徴を記述できるメタデータの定義と、メタデータを公開・操作するツールの開発について述べる。

Design of MANGA Metadata using Topic Maps

Shoichiro Hara Motomu Naitoh
Center for Integrated Area Studies, Knowledge Synergy Inc.
Kyoto University

MANGA is said to establish itself a Japanese contemporary culture. Its market expands worldwide and it is distributed by various media such as papers, Web and mobile phones. Distinguishing features of MANGA are various derivative works, media-diversity and unique structure. Metadata for the new cultural media must be necessary to organize its structure and to give access-methods to consumers under the web environment. This paper will describe our challenge of constructing MANGA metadata to organize these distinguishing features and developing tools to manipulate metadata.

1. はじめに

日本のマンガ¹は、その確立された独自の表現形式²や、発行部数³などの点から、一つの文化ジャンルを確立していると言える。その特徴として、週刊誌や月刊誌などに掲載された後に単行本としてまとめられるという出版形態、さらにアニメ・映画・小説・演劇などの多様なメディアやパロディなどの 2 次作品への展開が、他のジャンルよりも活発である点を上げることができる。また、解題や作者などに関する多様な情報が Web 上で日々流通・更新されている点も特徴である。

海外への展開も活発で、単行本として翻訳されたりドラマ化されたりする事例も多く、海外の若者が日本や日本語に興味を持つきっかけとなっている。さらに、マンガを機会とした海外における日本研究や異文化研究の事例も増えつつある[2]。日本語教育や地域研究の観点からも興味深い動向である。

このような文化的な背景から、マンガ作品のデータベース化には意義があると考えられる。ただし上記に示したマンガの特徴を考慮すると、マ

ンガメタデータは作品の派生関係や関連情報を包含したものとして組織化する必要がある。これを MARC のようなメタデータで記述するのは容易でない。そこで必要な情報を詰め込んだメタデータを新たに定義しても、項目数が多く構造も複雑になり、使うことは容易ではないと危惧される。また、マンガデータベースの構築は遡及入力作業が中心となるが、マンガ目録の専門は皆無に等しいので、参加型のデータベース構築を目指す必要がある。そのため、メタデータは可能な限り単純であることが望ましい。

また紙メディアとしてのマンガには、絶版になりやすいという問題が収集現場から指摘されている[3]。さらに紙の劣化や災害への対応として、デジタル保存の必要性も指摘されている[1]。

このような理由から、マンガメタデータは、①多様な派生関連を記述でき、②容易かつ効率的に作成でき、③画像などのコンテンツデータへの関連付けができなければならない。そこで本研究では、単行本などの単体作品の書誌情報はダブリンコア (DCMES) で記述し、メディア間の派生・時系列関係および Web 記事などとの関連情報は Topic Maps で記述することにより、構築の容易なメタデータの定義を試みた。

2. 「花より男子」

本稿では「花より男子 (はなよりだんご)」を取り上げる。「花より男子」は、神尾葉子による少女漫画作品であり、マーガレット (集英社) に 1992 年から 2004 年まで連載された。単

¹ 日本では「まんが」、「漫画」、「コミック」などとも呼ばれ、海外では comic (米) Bande Dessinee(仏)などと呼ばれているが、本稿では、それらの歴史的背景や相違には踏み込まず、マンガと呼ぶ。
² マンガの定義や表現法などについても本稿では深くは踏み込まない。
³ 出版指標年報 (2007) によると、2006 年の出版全体の推定販売部数は 345,423 万部、推定販売金額は 21,525.4 億円のところ、漫画の推定販売部数は 126,841 万部、推定販売金額は 4,810 億円である[1]。

行本は全 37 巻、完全版全 20 巻が発売されている。2005 年 9 月 12 日には発行部数を 5800 万部まで伸ばし¹、日本一売れた少女漫画と言われている。アニメ化・テレビドラマ化・映画化もされている。海外においては米国、フランス、スペイン、台湾、タイなどにおいて翻訳版の単行本が出版されている。また台湾では「流星花園」、韓国では「꽃보다 남자」という題名でテレビドラマ化されている。

このように、週刊誌から単行本、アニメ化・映画化、海外における翻訳出版など、その多様な派生作品やメディア展開の点から、「花より男子」はマンガメタデータの研究に適した作品である。また学園恋愛作品であるため一般的な会話が中心であること、海外における翻訳も多いので、テキストデータのサンプルとしても適している。本稿では日本語教育や異本研究などを想定したテキストデータ作成も試みている。

このような理由から、「花より男子」を研究対象として選択した。資料としては、日本語単行本（マーガレットコミックス：集英社）、およびその英語版（Viz Media）とタイ語版（Siam Inter Comics）を利用している。

3. マンガの構造とデータ構築

視覚メディアとしてのマンガの特徴は、内容の大部分を絵で伝達し、セリフや音などは文字で表現し、コマや吹き出しなどの独特な仕掛けにより物語の連続性を表現している点にある。読者はコマを追うことにより、ストーリーの時間的な発展を理解する。したがってマンガの内容をデータベース化するには、このようなマンガの構造を反映させる必要がある。

3. 1 画像データ

画像データのデジタル化はページを単位とした。単行本の背表紙を浅く裁断し、ページごとに 300dpi、8bit グレースケールで撮影し、TIFF および JPEG フォーマットで保存した。撮影後、製本し冊子体を復元した。ただしカバーについては見開き、300dpi、24bitRGB で撮影し、TIFF および JPEG フォーマットで保存した。各画像は、冊子の巻号（次節の Volume）とページ（次節の Page）の組み合わせで識別している。

デジタル化の過程で資料を裁断した点については批判もありうるであろう。しかし、冊子体マンガの多くは紙の劣化という問題を抱えており、最終的には紙媒体として保存するか、デジタル化して紙媒体を廃棄するかの二者択一を迫られる問題である。

実際、大量の冊子体マンガを長期間保存しようと試みても、書庫などの設備や環境維持などのコストの面から不可能であろうと思われる。また、災害などから資料を保全するために、資

料を複数用意して分散管理することも困難であろう。このような理由により、本研究ではデジタル化を優先する観点から研究を行っている。その代わりに、適切なレベルにおいて、精度の良い画像の作成を試みた。冊子体を裁断した理由は、歪みの少ない画像を得るためである。

カバーのみをカラーとしたのは予算上の都合もあったが、日本マンガのほとんどはモノクロームのページであるためカラー撮影をする必要はないという判断を行った結果でもある。しかし、精度の良い画像の作成を試みる観点からは、カラー画像とした方が良かったのかもしれない。一考する必要がある。

3. 2 構造メタデータ

テキストのデジタル化に際して、作品を単行本→巻→ページ→コマに階層化した。コマの内部は画像部分とテキスト部分から構成される。画像についてはページを基本とし、コマ単位へ分割などの処理は行わなかった。テキスト部分は、会話、オノマトペ、ナレーション・モノログ、手書き、記号から構成されるものとした（図 1）。

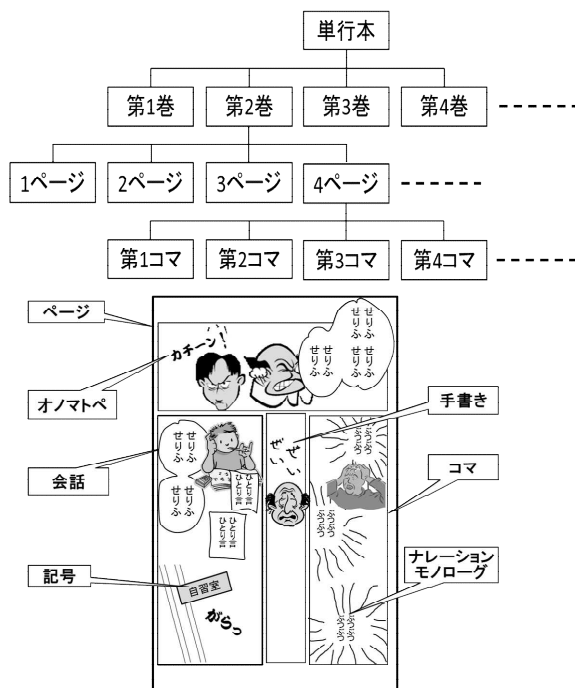


図 1. マンガの物理的構造

テキストデータは、各会話、オノマトペ、ナレーション・モノログ、記号、手書きを 1 レコードとした。レコードの構造は以下のように定義した。

- ID : テキスト識別子
- Volume : 冊子体の巻号

¹ 発行部数の根拠が曖昧であるため、あくまでも参考値である[3]。

ID	Volume	Page	Coma	Attribute	Speaker	String	Overlap	Comment
27	1	P008	C01	f1	makiko	いーなあ/あたしたち/なんて/ビニールの/バッグよおお//		
28	1	P008	C01	f1	makiko	ねっ/つくしちゃん//		
29	1	P008	C01	o1		カチーン//		
30	1	P008	C02	n1		ビニールの/バッグの/どこが悪いッ		
31	1	P008	C02	n2		安いし丈夫だし/おまけに/雨の日でも/持ち歩けるのよっ//		
32	1	P008	C02	n3		どーでも/いいけど/なんで/そんなに/煽りへつらうわけえ! ?//		
33	1	P008	C03	f2	makiko	どうしたのオ/つくしちゃん// 鼠が/あらいいけど//		
34	1	P008	C03	h1		はあはあ//		
35	1	P008	C03	f3	tsukushi	な...っ/なんでも/ないのよ/真木子ちゃん// 階段/のぼったら/ほら動悸が//ーねっ//		
36	1	P008	C04	f4	makiko	つくしちゃんて/ホントに/大人しいよね		
37	1	P008	C04	f4	makiko	髪だっ/て/おさげほだけば/セッター/かわいいのに//		
38	1	P008	C04	n4		あぶない/あぶない//		
39	1	P008	C04	n4		つい興奮して/自分が/出てしまった//		
40	1	P008	C04	i1		2-C//		
41	1	P008	C04	o2		がらっ//		
42	1	P009	C01	o1		しーん//		
43	1	P009	C01	f1	makiko	おは...//		
44	1	P009	C02	f2	makiko	な.../なに/この雰囲気//		
45	1	P009	C03	f3	classmate	しっ// 樹本が/来てんだよ//		
46	1	P009	C03	f4	makiko	...// 樹本くんで/ずっ/と/登校拒否/してた...?//		
47	1	P009	C04	f5	tsukushi	すごい/やつれた//		
48	1	P009	C04	h1		奮気ただよってる//		
49	1	P009	C05	f6	classmate	退学する/らしいよ//		
50	1	P009	C05	f6	classmate	今日/荷物/とりに来た/みたい//		
51	1	P009	C05	h2		ほそ/ほそ//		

図2 テキストデータ例 (日本語) マーガレットコミックス花より男子①, 集英社, p.8-9

ID	Volume	Page	Coma	Attribute	Speaker	String	Overlap	Comment
27	1	P008	C01	f1	makiko	HAIYOU PUT OUR CHEESY VINYL BAGS TO SHAME		
28	1	P008	C01	f1	makiko	DOESNT SHETSUKUSHI?		
29	1	P008	C01	o1		RRK		
30	1	P008	C02	n1		WHAT'S WRONG WITH VINYL?		
31	1	P008	C02	n2		IT'S CHEAP IT'S DURABLE IT'S RAIN-PROOF		
32	1	P008	C02	n3		I'M NOT PUTTING DOWN VINYL TO SUCK UP TO HER		
33	1	P008	C03	f2	makiko	YOU OKAY,TSUKUSHI?YOU'RE BREATHING AWFUL HARD.		
34	1	P008	C03	h1		HUFHUF		
35	1	P008	C03	f3	tsukushi	IT'S N-N-NOTHING,MAKIKO... JUST WINDED... FROM CLIMBING THE STAIRS...		
36	1	P008	C04	f4	makiko	YOU'RE SO QUIET,TSUKUSHI!		
37	1	P008	C04	f4	makiko	EVE YOUR HAIRYOU'D LOOK REALLY CUTE IF YOU UNDDID YOUR BRAIDS.		
38	1	P008	C04	n4		EASY,EASY.		
39	1	P008	C04	n4		DONT LOSE YOUR COOL		
40	1	P008	C04	il		2-C		
41	1	P008	C04	o2		KREEEK		
42	1	P009	C01	o1		SHHHHHHH		
43	1	P009	C01	f1	makiko	GOOD..MORNING?		
44	1	P009	C02	f2	makiko	W-WHY IT SO QUIET?		
45	1	P009	C03	f3	classmate	SHHKJKOMO'S HERE		
46	1	P009	C03	f4	makiko	KIMOT O?THE ONE WHO'S BEEN OUT OF SHOOL..?		
47	1	P009	C04	f5	tsukushi	HE LOOKS SO GAUNT.		
48	1	P009	C04	h1		Like a ghost.		
49	1	P009	C05	f6	classmate	HE'S PICKING UP HIS STUFF.HE MUST BE LEAVING EITOKU.		
50	1	P009	C05	f6	classmate			
51	1	P009	C05	h2		TNK TNK		

図3 テキストデータ例 (英語) Boys over Flowers 1, Vis Media, p.8-9

ID	Volume	Page	Coma	Attribute	Speaker	String	Overlap	Comment
27	1	P008	C01	f1	makiko	ตั้งเลย อย่างเรามีเซ่ ก็แค่ กระเป๋า พลาสติค		
28	1	P008	C01	f1	makiko	นะอะ ซึ่คืเซ่ ตั้ง		
29	1	P008	C01	o1		เปรี๊ยะ		
30	1	P008	C02	n1		กระเป๋า พลาสติคมัน เสียหาย ตรงไหน		
31	1	P008	C02	n2		หิ้งถูก ทั้งทันทาน กันน้ำ ได้ด้วย		
32	1	P008	C02	n3		แล้วทำไม ต้องทำเป็น ประจบประแจง ขนาดนั้นด้วย		
33	1	P008	C03	f2	makiko	เป็นไป๋ไป๋ หายใจแรง จัง		
34	1	P008	C03	h1		หะระ หะระ		
35	1	P008	C03	f3	tsukushi	ปะ...เปล่าหรอก มาคิโกะ จัง ขึ้นบันได รู้สึกเหนื่อย นะ		
36	1	P008	C04	f4	makiko	สิ่คืเซ่ตั้งเนีย เpsilon ร้อย จริงๆ นะ		
37	1	P008	C04	f4	makiko	อีก หางเบี่ย น่ารัก เขียว		
38	1	P008	C04	n4		ต้อง ระวังๆ		
39	1	P008	C04	n4		อย่า มีอารมณ์ เสียจะ มีเชื้อง		
40	1	P008	C04	il		2-C		
41	1	P008	C04	o2		แกกรัก		
42	1	P009	C01	o1		เร่-ยบ		
43	1	P009	C01	f1	makiko	อรุณ...		
44	1	P009	C02	f2	makiko	ปะ... เป็นอะไรกัน ไปหมด		
45	1	P009	C03	f3	classmate	ฮู้ว คิโมโตะ มา นะ		
46	1	P009	C03	f4	makiko	คิโมโตะ หายไป ตั้งนาน		
47	1	P009	C04	f5	tsukushi	อู โทรม เขียว		
48	1	P009	C04	h1		เห็นว่า จะ ลาออก		
49	1	P009	C05	f6	classmate	วันนี้ เลยก มาเก็บ ของ		
50	1	P009	C05	f6	classmate	ซุบซิบ ซุบซิบ		
51	1	P009	C05	h2		ซุบซิบ ซุบซิบ		

図4 テキストデータ例 (タイ語) สาวแกร่งแรงเกินร้อย 1, Siam Inter Comics, p.8-9

Page : 各冊子体におけるページ
Coma : 各ページにおけるコマ番号
Attribute : テキストの種類。会話（吹き出し）(f)、手書き(h)、ナレーション・モノログ(n)、オノマトペ(o)、記号(i)、テキストのないページ(z)を定義している
Speaker : 話者
String : テキスト本体
Overlap : オノマトペなどが複数のコマに跨がっている場合、それらのコマ番号を記載する
Comment : データ作成上のコメント

本研究では、これをマンガの構造メタデータと呼ぶ。ID はテキスト識別子であるが、Volume、Page、Coma の組み合わせにより決定できる。

「花より男子」のような少女マンガでは、不定型なコマ割が多く、コマ順の同定を困難にしている。このような場合、本研究ではデータ作成者がストーリー展開を追いやすい順序を採用した。

1 コマ中に会話、オノマトペ、ナレーション・モノログ、記号、手書きが混在している場合、レコードの順序は「読み」順に従うが、その決定もデータ作成者がストーリー展開を追いやすい順序を採用した。

String はテキストデータの本体である。日本語テキストについては、文節ごとに「/」を区切りとして付与した。また会話が複数の文から成っている場合、文ごとに「//」を区切りとして付与した。これは多言語パラレルテキストなどの研究を想定しているためである¹。図 2 から図 4 にテキストデータの一部を示す。

3. 3 書誌メタデータ

マンガメタデータには、①ページやコマの順序などの物理的情報と、テキストの種類や発話者などの論理的な情報を記述したもの、②週刊誌に掲載された 1 作品、1 冊の単行本、1 枚の DVD、画像データなど発行メディアに依存した作品情報を記述したもの、③あるマンガに関する創作物としての包括的な情報（主題、タイトル、作者など）を記述したものなど多様である。

前述の構造メタデータは①に対応し、会話や発話者などから画像や関連メタデータへのアクセスポイントとなる。

②がいわゆる書誌メタデータに相当する。本研究では、容易かつ効率的なメタデータ作成を実現するために、データ項目を DCMES の 15 項目に限定した。また著者 (contributor)、出版社 (publisher) や内容 (description) などに関する詳細な情報の多くは既に Web 上に登録されているので、自ら記述するのではなく Web などへの

参照を基本とした。例えば単行本の第 1 巻の DCMES データは以下ようになる。なお PSI は関連概念への参照（後述の Topic Maps で説明する）、URI は Wikipedia などへの参照である。

Title : 花より男子 (だんご) 1
Creator: 神尾葉子 [PSI]
Subject : 恋愛[PSI]
Description : 日本の少女漫画作品。全 37 巻の単行本の第 1 巻 [URI]
Publisher : 株式会社集英社[PSI]
Contributor : 株式会社創美社 [PSI]
Date : 199210
Type : Still Image [PSI]
Format : 印刷物,205p;18cm
Identifier : ISBN4088480287
Source : 花より男子 [PSI]
Language : Ja [PSI]
Relation : 花より男子[PSI]
Coverage : 平成 [PSI]
Rights : Yoko Kamio [PSI]

しかし、これらのメタデータは独立しているために、マンガ情報を集中的に収集したければ、データベースや Web などの多くの情報を自分で探索し整理しなければならない。例えば「花より男子」を国会図書館の NDL サーチ[5]で検索（タイトル＝「花より男子」、著者・編者＝「神尾葉子」）すると、「花より男子 第 10 巻」はマーガレットコミックス（単行本）のシリーズであることが分かる。つまり図 1 の単行本と各巻の関係についての情報は収集できる。しかし、これが「マーガレット」という隔週漫画雑誌に掲載されたものの派生作品であることや²、米国では Boys Over Flowers という題名で翻訳出版されていることは、Web などを調べないと分からない。また試し読みができるサイトも分からない（例えば、http://www.s-manga.net/comics/cm_20050525_mg_sgc_4088551133_hanayoridango-kanzenban-1k.html）。

この問題を解決するには、各派生作品の書誌メタデータ（図 1 では単行本）をまとめる上位階層あるいはデータ項目を定義し、そこに③の包括的な情報を記述する必要がある。作品の多様な派生情報などを記述する枠組みとして FRBR (Functional Requirements for Bibliographic Records) の適用が考えられる[6,7,8]。

FRBR では書誌の記述を、著作 (Work)、表現形 (Expression)、体現形 (Manifestation)、個別資料 (Item) の 4 段階でとらえる。「著作」は内容のまとまりに対応する部分であり、冊子体など特定のメディアにはとられない。

¹ データ作成上の試みとして行ったもので、利用に関する詳細な検討は今後の課題である。

² 例えば JAPAN/MARC MARC21 には 780 (先行記入) や 785 (後継記入) などがあるので、これらを利用すれば派生情報を記述することは可能であると思われる。

「花より男子」の包括的な情報が該当する。本研究では、これをマンガの著作メタデータと呼ぶ。「表現形」は著作を何らかの形で表現したものであり、単行本などに関する包括的な情報が該当する。「体现形」は表現形を具体的なメディアとして実現したものである。単行本1巻ごとの包括的な情報が該当する。本研究における書誌メタデータは、ここに該当すると考えられる。これらの具体的な一点一点が「個別資料」である。

本研究では FRBR モデルを参照し、著作メタデータ、書誌メタデータおよび構造メタデータを図5のように構造化し、これをマンガメタデータと定義した(説明は4.1)。前述のように、体现形に対応する書誌メタデータは DCMES を基本としている。同様に、著作および表現形に対応するメタデータについても、データ項目の意味は DCMES を基本としている。

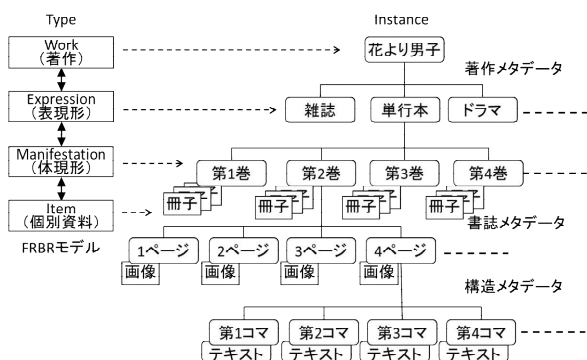


図5 マンガメタデータの全体構造

4. Topic Maps によるメタデータの統合

本研究では、メタデータの実装において Topic Maps を利用する[9]。Topic Maps は、問題領域に関する知識を、概念 (topic)、概念間の関係である関連 (association)、概念と関係する情報リソースへのリンクである出現 (occurrence) で表現する。トピック、関連、出現は型を持ち、それらを Topic Maps Ontology と呼んでいる。

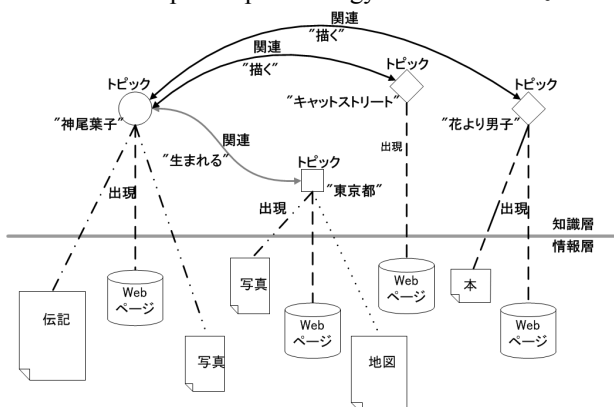


図6 Topic Maps の基本モデル

例えば「神尾葉子は花より男子を描く」を Topic Maps の基本モデルで表現した例を図6に示す。この文を問題領域とすると、『神尾葉子』と『花より男子』が Topic Maps の「概念(主題)」となる。また、これら2つの「概念」は「関連」により結びつけられている。また『神尾葉子』にはファンページや Wiki 記事などの情報リソースが存在しており、これらの情報リソースと『神尾葉子』は「出現」により結びつけられている。『キャットストリート』や『東京』も同様である。Topic Maps では、このようにして問題領域を表現する。

Topic Maps の「概念」や「関連」には型があり、『神尾葉子』のような具体的な対象は、該当する型のインスタンスとして記述する。図6には型を表示していないが、例えば『神尾葉子』は『人』型のインスタンス、『花より男子』は『著作』型のインスタンスである。「関連」についても、例えば『描く』は(ここでは同じ名前であるが)『描く』型のインスタンスである。ところで『神尾葉子』には同名異人が存在するように、一般的に名前は識別子として不適切である。そこで Topic Maps では IRI (Internationalized Resource Identifier) というユニークな識別子を利用する。この識別子を主題識別子 (Subject Identifier) といい、公開された主題識別子を PSI (Published Subject Identifier) という。

このように Topic Maps の構造は比較的単純である。また出現を利用することにより、メタデータの項目 (トピック) と画像などのコンテンツとの関連付けが容易となる。IRI・PSI により、Web 上に分散しているトピックも利用できる。さらに問合せ言語である tolog (後述) を利用すれば、メタデータの意味的構造に基づく多様な検索も可能である。Topic Maps の導入により、マンガメタデータをデータの単なる羅列ではなく、概念のネットワークとして表現できると期待される。

以上の理由から、本研究ではメタデータの実装において Topic Maps を採用した。これにより、メタデータ本体およびメタデータ記述対象への精密なアクセスや発見が容易になる。さらに、対象の問題領域に新たな価値を創造・付加する可能が開けるなど、メタデータの有用性を高めることができると考える。

4.1 Topic Maps によるマンガメタデータの表現

前述のように、本研究では FRBR を参照モデルとしてマンガメタデータを構造化している。図5は、これを Topic Maps 的に整理したものである。図の左部分が FRBR を参照モデルとした Topic Maps Ontology で、右の部分がマンガメタデータのインスタンスおよびその関係を表現し

ている。図には描かれていないが、派生関係も関連として表すことができる。

図 5 の書誌メタデータの実体は DCMES の 15 項目である。これを Topic Maps Ontology で表現すると図 7 のようになる。ここでは DCMES の各項目を概念型（トピック型）、項目間の関係を関連型としている。Topic Maps Ontology では、トピック型を四角、関連型を線で表現する。また四角の中の文字列がトピック型の名前である。線の中央付近の文字列が関連型の名前、さらに四角の近くの文字列が関連役割型の名前を表す。

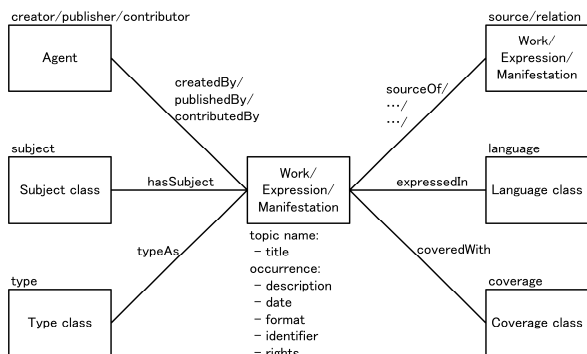


図 7 書誌メタデータの Topic Maps Ontology

図 7 に基づき定義した書誌メタデータの各項目を以下に示す（下線付きは DCMES）。

Topic Types(8): Work, Expression, Manifestation, Agent, Coverage, Language, Subject, Type

Occurrence Types (5): Description, Date, Format, Identifier, Rights

Association Types(9): createdBy, hasSubject, publishedBy, contributedBy, typeAs, sourceOf, expressedIn, relatedWith, coveredWith

Association role Types (18): creator, subject, publisher, contributor, type, source, language, relation, coverage、および各型に付随する役割(9)

Title : Work, Expression, Manifestation の Topic Name として利用する。

図 5 の構造メタデータは 3.2 で述べた。これを Topic Maps で表現した Topic Maps Ontology を図 8 に示す。

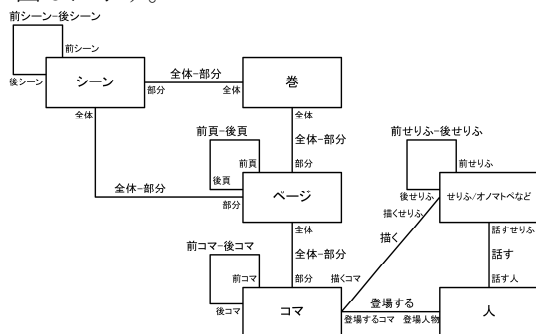


図 8 構造メタデータの Topic Maps Ontology

図 8 に基づいて定義した構造メタデータの各項目を以下に示す。

Topic Types (6) : 巻、シーン、ページ、コマ、せりふ/オノマトペなど、登場人物

Occurrence Types (3) : 属性、文字列、画像（これらは図 2、3、4 の Attribute と String およびページの画像データに対応する。）

Association Types (8) : 全体-部分、登場する、描く、話す、前シーン-後シーン、前ページ-後ページ、前コマ-後コマ、前せりふ-後せりふ

Association role Types (16) : 全体、部分、登場するコマ、登場人物、描くコマ、描くせりふ、話す人、話すせりふ、前ページ、後ページ、前コマ、後コマ、前せりふ、後せりふ、前シーン、後シーン

次に Topic Type のインスタンス例を示す。

『巻』型インスタンス

トピック型 : 巻

トピック名 : 1 巻

スコープ : ja(日本語)

主題指示子 (Subject Identifier) :

[http:// ... /psi/hanadan/volume_10001](http://.../psi/hanadan/volume_10001)

『ページ』型インスタンス

トピック型 : ページ

トピック名 : P008

スコープ : ja(日本語)

主題指示子(Subject Identifier):

[http:// ... /psi/hanadan/page_10001_P008](http://.../psi/hanadan/page_10001_P008)

出現(画像), スコープ(en(英語)):

en/01000010.jpg

出現(画像), スコープ(ja(日本語)) :

ja/01000009.jpg

出現(画像), スコープ(th(タイ語)) :

th/01000010.jpg

『コマ』型インスタンス

トピック型 : コマ

トピック名 : P008-C01

スコープ : ja(日本語)

主題指示子(Subject Identifier) :

[http:// ... /psi/hanadan/coma_10001_P008_C01](http://.../psi/hanadan/coma_10001_P008_C01)

『せりふ/オノマトペなど』型インスタンス

トピック型 : せりふ/オノマトペなど

トピック名 : ねっ/つくしちゃん//

スコープ : ja(日本語)

主題指示子(Subject Identifier):

[http:// ... /psi/hanadan/script_10001_1000028](http://.../psi/hanadan/script_10001_1000028)

出現(属性), スコープ ja(日本語) :

せりふ (吹き出しの中)

出現(文字列), スコープ en(英語) :

DOESN'T SHE, TSUKUSHI?
 出現(文字列), スコープ ja(日本語):
 ねっ/つくしちゃん//
 出現(文字列), スコープ th(タイ語):
 เหนอะ ซึกุชิ จัง

さらにトピックが関連を持つ場合は関連役割も持つ。『せりふ/オノマトペなど』型のインスタンスが持つ関連役割の例を以下に示す

関連役割: "P008-C01"トピックとの間の『描く』関連における『描くせりふ』役割

関連役割: "makiko"トピックとの間の『話す』関連における『話すせりふ』役割

4. 2 Topic Maps による実装例

以下に Topic Maps による実装例を示す。なお実装したデータは、「花より男子」(単行本)の日本語版、英語版、タイ語版の各第1巻である。各言語版において、ページ数は158、コマ数は550、せりふ/オノマトペ数1,239、登場人物18である。

Topic Maps の統合開発・運用ツール Ontopia に含まれる DB2TM 機能を使用し、図2、3、4のExcelデータから構造メタデータ topic map¹を生成した。書誌データ topic map は、DCMESの15項目をカラムに持つExcelデータを作成し、同じくDB2TMを使用して生成した。

構造メタデータ topic map と書誌メタデータ topic map は個別に作成し、必要に応じてマージできるようにした。これら topic map の Web アプリケーションは、同じく Ontopia に含まれる Navigator Framework を利用して開発した。

Web アプリケーションが持つ主な機能は以下のとおりである。

- トピック型ごとのインスタンス一覧表示、
- インスタンスの詳細表示
- グラフィック表示
- 文字列検索
- tolog 問合せ
- 関連に基づくトピック間のナビゲート

一致	Topic
(En) Thanks. (Ja) ごくろうさま// (Th) ขอโทษ นะ	ごくろうさま//
(En) THANKS, MKINO. (Ja) あっ/ありがとう/牧野さん// (Th) ขอโทษ นะ มากัดโนะ	あっ/ありがとう/牧野さん//
(En) THANK YOU... (Ja)ありがとう// (Th) ขอขอบคุณありがとう//
(En) I WANTED TO THANK HIM AGAIN WHEN HE WAS ALONE... (Ja) F4から/はなれて/ひとりの時に/もういちど/お礼言いた/かったのに...// (Th) อยากทําใจทําส ขอขอบคุณเขา อีกครั้ง...	F4から/はなれて/ひとりの時に/もういちど/お礼言いた/かったのに...//
(En) I WANTED TO THANK FOR YESTERDAY. (Ja) あたし/~~~~/昨日のお礼/ちゃんと/言いたかったの// (Th) ฉัน ฉันทอยาก ขอขอบคุณจริง ๆ เรื่อง เมื่อวาน	あたし/~~~~/昨日のお礼/ちゃんと/言いたかったの//

図9 「thank」による文字列検索例

¹ Topic Maps はトピックマップ技術を意味する場合には使用し、topic map は個別のトピックマップを示すときに使用する。

Topic map の Web アプリケーションの例として、「thank」をキーワードとして文字列検索した結果を図9に示す。左側のカラムにはヒットした文字列を含む「せりふ/オノマトペなど」が、英語、日本語、タイ語で併記されている。右側のカラムには、該当するトピックへのリンクが表示されている

Topic Maps では、単なる文字列検索機能に加え、問合せ言語を利用した高度な検索も可能である。Topic Maps の問い合わせ言語には、述語論理に基づいた tolog[10]がある。Tologの例として、『人』の『つくし』が発話するページとコマを問合せる式を以下に示す。

```
using psil for i"http://www. ... /psi/hanadan/"
select $PAGE from
psil:part-of($PAGE : psil:whole, $COMA :
psil:part),
psil:appear($COMA : psil:coma-appearing-person,
psil:person_tsukushi : psil:person-in-coma)
order by $PAGE?
```

この検索結果を図10に示す。結果として、『つくし』が発話するページとコマの一覧が得られる。ここで一覧の中の特定のページまたはコマを選択すると、そのページまたはコマにナビゲートできる。このように、メタデータの構造に即した詳細な粒度での問合せが可能である。

Execution time 16 milliseconds
 Result rows 163

PAGE	COMA
P005	P005-C03
P008	P008-C03
P009	P009-C04
P010	P010-C02
P013	P013-C01
P013	P013-C03
P013	P013-C04
P014	P014-C05

図10 tologによる検索結果例

5. 今後の展望

現状は少数のサンプルに対してメタデータを試行的に作成した段階であり、メタデータを増やすことが当面の課題である。

メタデータの品質を保証するためには、メタデータ記述のためのガイドラインが重要である。メタデータ全体の構成にはFRBRを採用しているが、詳細な検討を進めている途上である。特に多様な派生情報を記述する上で、DCMESのSource項目およびRelation項目の役割とFRBRとの関連についての再考が必要であると考えている。またCreator項目、Publisher項目、

Contributor 項目、Subject 項目のための典拠データや PSI の整備も重要な課題である。

マンガに関するリソースは、紙媒体や Web を問わず、急速に劣化・消滅しつつある。したがってリソースのデジタル化やアーカイブ化は最優先課題であるが、これらを有効活用するためのメタデータ整備も同様に不可欠である。そのためには、参加型によるメタデータの作成・管理・維持・公開を低コストかつ効率的に行えるような仕組みが必要である。技術的には、入力・保守が容易な環境・ツールの構築、メタデータ間の動的・柔軟なリンク・マージの実現が必要であると考えている。

これらにより、マンガ関連のリソースを広く世界に発信するとともに、後世への継承もより円滑に行えるようになることを期待している。例えば、子供の頃に読んだ「あのマンガのあのシーン」をもう一度読みたい、「ある時代の文化・流行・風潮・思想」などを調べたい、といったニーズに応えられることを目指している。さらに、本研究の仕組みは、娯楽面での利用はもちろん、学習教材や複雑化し急速に進歩する機器類のユーザガイドや整備マニュアルへも応用できるなど、実用面での新しい展開が期待できる。

本稿はメタデータ構築が主要テーマであり、テキストデータの利用については言及していない。しかし、資料の異同については考察する必要がある。本稿執筆時点において、各言語による資料間の異同は装丁の都合によるページのみである。ページ内のコマや吹き出しに異同はない。しかし、週刊誌から単行本としてまとめられる過程で、コマ割りや吹き出しあるいはストーリーそのものが改変されることは、マンガでは頻繁に見られることである。このような異同も考慮するならば、現状の構造メタデータでは不十分であり、何らかの工夫が必要である。文献[11]などを参考として、構造メタデータの拡張を行う予定である。

その他にも、アニメやドラマなど単行本以外のメディアに対応していないなど、考慮すべき課題を抱えているが、地道に取り組んでいきたいと考えている。

参考文献

- [1] 内記 稔夫, 秋田 孝宏 : CA1637 - 日本における漫画の保存と利用、カレントアウェアネス No.293 2007年9月20日、<http://current.ndl.go.jp/ca1637>、国立国会図書館。
- [2] 熊野七絵、廣利正代情報処理学会 : 「アニメ・マンガ」調査研究—地域事情と日本語教材—、国際交流基金日本語教育紀要第4号、55-69、2008。
- [3] 京都大学図書館機構 - ハーバード日記 : 司書が見たアメリカ : 【34. 日本のマンガを保存する図書館 - Cartoon Research Library, Ohio State University】

<http://www.kulib.kyoto-u.ac.jp/modules/wordpress/wp-trackback.php/61>

- [4] Wikipedia:花より男子、<http://ja.wikipedia.org/wiki/%E8%8A%B1%E3%82%88%E3%82%8A%E7%94%B7%E5%AD%90>
- [5] 国立国会図書館サーチ : <http://iss.ndl.go.jp/>
- [6] IFLA: FUNCTIONAL REQUIREMENTS FOR BIBLIOGRAPHIC RECORDS Final Report, http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf
- [7] 野村聡美、両角彩子、森永光晴、杉本重雄 : マングのためのメタデータモデルを目指したマンガのアーキテクチャの分析、デジタル図書館 (36), 3-14, 2009。
- [8] 宮田洋輔 : 日本の図書目録における書誌的家系 : J-BISC における調査と先行研究との比較分析、Library and Information Science、N0.61、91-117、2009。
- [9] 内藤求 : トピックマップ入門 (セマンティック技術シリーズ)、東京電機大学出版局、2006。
- [10] tolog A topic map query language: <http://www.ontopia.net/topicmaps/materials/tolog.html>
- [11] 安岡孝一 : マングにおける異本研究、公開シンポジウム『情報の構造とメタデータ』、京都大学人文科学研究所、pp.3-20、2012。