Metadata for manga: Metadata vocabulary, schema and tools

Shigeo Sugimoto
Faculty of Library, Information and Media Science
University of Tsukuba
Japan
sugimoto @ slis.tsukuba.ac.jp
Outline

• Background of this talk
• Introducing manga
• Why metadata for manga?
  • Pop-culture resource description
  • Digital and network-oriented production and publishing environment
• Manga metadata projects at Tsukuba
  • Manga Metadata Framework
  • Identification of \textit{Works} using bibliographic records
  • Software tools to help create and access
• Issues for the future
Some Background Issues

• Manga – Comics, Graphic Novels – is a major genre of Japanese pop-culture resources
  • Many spin-off products of manga, e.g., anime and games
  • Important genre since early stage of digital publishing on mobile phones

• Manga was not collected by libraries because they were regarded as a sub-culture resource
  • Lack of well-developed resources for manga, e.g. authority files
  • Efforts to describe manga
    • Recent manga collection development at libraries and museums, e.g. Kyoto International Manga Museum, Manga collection at JSC of Monash, etc.
    • Media Arts Database by Agency for Cultural Affairs of the Japanese Government

• This talk covers mainly from technological aspects of metadata
  • Intellectual Property Rights (IPR) issues are crucial but not mentioned
Introducing Manga

• Manga - a Japanese term meaning comics, graphic novels
  • what is manga?, how it looks like?, and who creates?

• Manga and their related resources
  • Primary manga entities
  • Primary entities created from manga entities – spin-off products
  • Secondary entities about manga

• Production process and components of manga
Manga

Various publishing styles – magazines (serials), monographs, Web pages.

Multi-modal content – images, texts, symbols, layouts and grammar
Different Styles of Digital Manga

- **Manga**
- **Social stream**
- **Book style**
- **Scroll style**
Manga Fans Community

• Comic Market – Big Market and Forum for Amateur Creators and Fans
  • 3 days event, twice a year
  • 35,000 individuals and groups, mostly non-professionals
  • 500K+ people at a single event
Web Resources created by Manga Fans

Wikipedia Article “Naruto”

A review about manga in Amazon.co.jp
Various Resources related to a “Work” of manga
- NARUTO, a popular manga, as an example -

User-generated Article
- Story
- Character

Bibliographic Records
- Libraries
- Commercial providers (publishers, bookstores)

Entities of NARUTO
Items/Manifestations, Works

Transmedia products
- Animation
- Music
- Game
- etc.

Reviews and Chatting on SNS

2015/8/12
Components of manga – another view

• Manga has many components which are connected each other
  → Manga creation is to manage them
Production Process – Storyboard (*Name*)

- Manga creator makes many storyboards (called “*Name*” in Japanese publishers and creators community) to explain their plan of the story and their ideas about the graphic expressions to his/her collaborator, e.g., publishers and editors.
- Edit a *Name* where graphic objects and their descriptions are given.

Instances created in a production process
Structure a Manga page

Page

Scene

Frame

Dialog

Character

Symbol

Components

2015/8/12
Why metadata for Manga?

• Fundamental changes in the whole lifecycle
• Creators’ environment
  • From “paper and pencil” to drawing tools for production
• Publishing environment changes
  • from print-only publishing to print & network publishing
• Readers’ environment
  • from print-based reading to network-assisted reading, e.g., SNS
• Information environment --- Resources born in Digital
  • What is the difference between digital manga and digital movie?
  • Various resources used in a unified environment – controlled by metadata
What metadata for Manga?

• Two viewpoints – bibliographic description and structural description
  • Bibliographic description – primarily as finding aids
    • Libraries – use standards for and conventions at libraries, e.g., MARC and Item-level description
    • Non-Library – project-based databases, publishers/booksellers’ sites, Internet resources, e.g., Wikipedia
  • Structural description
    • Structure of published materials – relationships between contents in serials and monographs, monographs and series
    • Structural description of constructs of manga – Manga specific features, e.g., serial number, volume, page, frame, texts, onomatopoeia
    • Relationship description of a manga and related resources – spin-off products, leading characters, etc.
Entities in a typical commercial publishing model

1. Publication in a magazine as a series
   (if it gets good feedbacks)
2. Publication as a monograph
   (if it gets good feedbacks)
3. Publication of the monograph as a series
4. Spin-off products – animations, games, character goods, etc.
Entities in a typical commercial publishing model and their descriptions

1. Publication in a magazine as a series → Each manga in a single volume is not described in library bibliographic data
2. Publication as a monograph → Item level description
3. Publication of the monograph as a series → Work?
4. Spin-off products – animations, games, character goods, etc. → Relationships among the entities
   → Fans create rich information resources for these entities
Entities in metadata description

• Manga Entities (FRBR Group 1)
  • Stories in a magazine, Monographs, Series
  • Item/Manifestation, Work

• Related Entities (FRBR Group 2 & 3)
  • Creators, Publishers, etc.
  • Leading Characters, Subjects, Genres, etc.
  • Spin-off Products

• Secondary resources
  • Wikipedia articles, Fan created sites
  • Review sites, etc.
Some studies on Manga Metadata at Tsukuba

• Manga Metadata Framework (MMF)
  • Three aspects of MMF and its base schema
    • bibliographic description – based on FRBR
    • structural description – based on TV-Anytime (a standard for digital video broadcasting)
    • intellectual entities (contained in the manga)
  • First presented at DC-2009, Seoul, Korea
  • Manga Path Language – Xpath-based scheme for structure description

• Identification of FRBR Works using bibliographic data

• Software tools for production of and access to manga
Structural Description

• Bibliographic Entities
  • An Item of a monograph, a title of a magazine, a Work of manga, etc.

• Manga as a published material has its structure
  • Title = a collection of stories
  • Story = a series of episodes (or sub-stories)
  • Episode = a series of scenes (logical structure) and a series of pages (graphical structure)
  • Scene/Page = a series of Frames
  • Frame = composed of graphic objects and texts, and linked to other frames

• Manga Path Language – XPath-based scheme to express structure of manga
Overall structure of MMF

Structural Description
- Title
- Story
- Episode
- Scene
- Page
- Frame
- Non-Frame
- Visual Object

Bibliographic Description based on FRBR

Intellectual Entities
- Work
- Expression
- Manifestation

contains
has subject

character
genre
title
creator
theme
place
time

2015/8/12
Overall structure of MMF

Title
Story
Episode
Frame
Visual
Object

character
genre
title
creator
theme
place
time

contains
has subject

TV Anytime (MPEG 7)
Title
Story
Episode
Scene
Page
Frame
Non-Frame
Visual Object

character
onomatopoeia
texts

FRBR Group 1 Entities
FRBR Group 3 Entities

2015/8/12
Structure a Manga page
- an Xpath-based description -
Identification of Works from bibliographic data[1]

• Work of FRBR is a natural entity for us to identify manga
  • Bibliographic records at libraries are created based on items
  • Authority records are not well developed for sub-culture resources
• Identification of Works of FRBR from Bibliographic Records provided by Kyoto International Manga Museum
• Use DBpedia in the identification process as a semi-authority
• Use identified Work entities to find translations in other languages
• Another approach using BIBFRAME (not yet published)

[1] Identification of works of manga using LOD resources: an experimental FRBRization of bibliographic data of comic books, WL. He, T.Mihara, M.Nagamori and S.Sugimoto, JCDL 2013
Software tools to support production and access

• Digital Manga Storyboard (Name) Editor: Metadata Creation with Manga Authoring
  • Improve the usability of resources of manga to provide authoring tool for storyboard of manga
  • Store them with its metadata created automatically

• Manga Collection Viewer
  • Visual presentation of collections of manga
  • Need Ontology for Manga – vocabularies for genres, characters, etc.
    • Lack of Authority records
    • Challenges to use Web resources, e.g. DBpedia, as a quasi-authority
Some Instances in Manga Ontology

FRBR Work-Manifestation
- mo: Volume
- mo: Publication
- mo: Theme
- mo: Painting Style
- mo: Character
- mo: Format

FRBR Person
- mo: Editor
- mo: Creator
- mo: Reviewer
- mo: Storywriter

FRBR Corporate Body
- mo: Organization
- mo: Publisher

mo: Manga Work

mo: Manga Medium

mo: Product
- mo: Brand
- mo: Information Magazine
- mo: Magazine
- mo: Institution
- mo: Shop
- mo: Production
- mo: Software
- mo: TV Program
- mo: Event

mo: History
- mo: Award
- mo: Dojin
- mo: Term
- mo: Incident
- mo: Movement
Linking manga resources via Manga Ontology

User-generated Article
- Story
- Character

Bibliographic Information
- Title
- Publisher

Manga Ontology

Transmedia
- Animation
- Music

Review on SNS

2015/8/12
Manga Collection Viewer – a facet-based tool

Initial view

- Each circle shows a facet (type/genre)
- Hierarchical relations between upper concepts and their children
Manga Collection Viewer – a facet-based tool

Zoom into a circle
This is a view when user select “mo:Theme” as a facet
Works in a type/genre
This is the collection of circles when their theme “mo:Robot” is selected.
Create manga in digital environment

• Tools to help creation are crucial
  • Tools to draw manga
  • Tools to describe and maintain descriptions of manga, i.e., Manga metadata

• Drawing and authoring tools are well used

• Tools for Manga metadata are not well developed

• Some lessons learned in our projects
  • Describing metadata after production is too expensive
  • Not only the final manga products but also those created in the production process are useful for re-editing, re-creation and re-use of the manga and the ideas used in the production
Digital and network-oriented production and publishing environment

• Current situation: Born in digital and published in page-oriented format, i.e., print-publishing and page-oriented digital format
  • Web comics – a new stream

• “Born in digital” from a perspective for the future
  • Use structural metadata for functions based on fine granule components
  • Promote usage of creator-created and reader-created metadata
  • Bridge to other visual media resources, e.g. digital video
  • Link to other resources on the Internet

• Some fundamental issues
  • Identification scheme of “content”, e.g., a story, a page, a frame, a character, etc.
  • Need standards and authority files for physical, digital, and conceptual entities

2015/8/12
Digital Resources on the Internet

Resources

- Concepts, etc.
- Content
- File
  - Binary data
- Format

User Environment

- Print
- Mobile
- PC

Medium + Content

Unbundle

Access

Users
A Metadata Centric View of Digital Resource
A Metadata Centric View of Digital Resource
Issues for the Future: Global vs. Local – Sharing Vocabularies is a challenge

• Metadata vocabularies – the basis of interoperable metadata

• Vocabulary sharing – different requirements
  • Personal (individual, small group) vs. Community
  • Subject-wise Local community vs. Global community
  • Geographically Local community vs. Global community
  • Temporally Local community vs. Global community

• Vocabulary sharing for “Interoperability across Communities”
  • DCMI’s goal
  • Metadata registries as an infrastructure for sharing vocabularies across communities and over time
Thank you very much

sugimoto @ slis. tsukuba. ac. jp