

[P14] マンガ・アニメ・ゲームを対象とした作品と出版物の構造的関係を記述した LOD データセット開発

○内海祐希¹⁾, 三原鉄也²⁾, 永森光晴²⁾, 杉本重雄²⁾

¹⁾ 筑波大学情報学群情報メディア創成学類, 〒305-8550 茨城県つくば市春日 1-2

²⁾ 筑波大学図書館情報メディア系

E-mail: s1511426@u.tsukuba.ac.jp

A LOD Dataset for Describing Structural Relationships of Work and Publication in Manga, Animation and Game

UTSUMI Yuki¹⁾, MIHARA Tetsuya²⁾, NAGAMORI Mitsuharu²⁾, and SUGIMOTO Shigeo²⁾

¹⁾ College of Media Arts, Science and Technology, the University of Tsukuba, 1-2 Kasuga, Tsukuba, Ibaraki, 305-8550 Japan

²⁾ Faculty of Library, Information and Media Science, University of Tsukuba

【発表概要】

近年、マンガ・アニメ・ゲーム (MAG) を保存・活用するためにデータベース整備の取り組みが始まっている。しかし、既存のデータセットにはマンガ・アニメ・ゲームの作品間や作品と出版物間の関連が明示的に記述されておらず、情報の利活用が困難である。本研究では、Web 上での MAG 作品と出版物の関連に基づく検索のための基盤の構築を目的として、それらの構造的関係を記述するデータセットの開発を行った。大量の MAG 作品についてのデータを作成するために、複数のデータベースが提供する MAG に関するデータを機械的に変換・統合する手法によりデータセットを作成した。

1. はじめに

近年、マンガ・アニメ・ゲーム (以下、MAG と呼ぶ) を日本のポップカルチャを代表する文化資源として保護することを目的に、これらに関する出版物や資料をアーカイブする取り組みが産学官を挙げて推進されている。この取り組みのひとつとして文化庁では MAG に関する出版物や資料を統合的に検索できるデータベースの整備を行っている (メディア芸術データベース (開発版) [1])。しかし本来異なるメディアである MAG 作品と出版物の関連に基づくデータの整備は困難であり、それぞれの情報の横断検索に留まっている。

そこで本研究では、Web 上での MAG 作品と出版物の関連に基づく情報検索のための基盤の構築を目的として、MAG 作品と出版物の構造を記述する Linked Open Data (LOD) データセットの開発を行った。

2. マンガ・アニメ・ゲーム作品とその出版物の構造的関係

MAG 作品はメディアミックスやシリーズ展開によって個別作品の枠を超えて世界観や設定を共有した作品が制作されることが多い。こうした作品間、作品と出版物間の関連は MAG 作品の探索の際の重要な手がかりとなる。

しかし、例えばマンガ単行本やアニメのテレビ放送、ゲームパッケージのシリーズといった、作品と出版物を繋ぐ構造的な関係のあり方はメディアによって異なる。こうした MAG の作品間や作品と出版物間の関連は一般的な書誌情報を記述する目録規則では網羅されていない。また、複数のマンガ単行本やアニメのテレビ放送、ゲームパッケージで構成されるシリーズに即した出版物との間の関連も十分に記述されていない。そのため、既存のデータベースでは複数の作品や出版物をメディアとシリーズを超えて探索することが難しい。

3. 関連研究

大石らは、Wikipedia[2]の記事と構造の情報から MAG 作品を同定し構造化を行った[3]。この研究では作品となりうる実体を情報源から発見して記事の上下関係から構造を推定しデータセットを作成している。Jacob らは、作品と出版物に関する実体を記述するデータモデルを提案している[4]。この研究ではゲーム作品とパッケージ、その間をつなぐ中間的な実体をプラットフォームの違いなどでデータモデルとして定義している。

4. Linked Open Data 技術に基づいた MAG 作品と出版物の構造的関係につ いてのデータセットの開発

Linked Open Data (LOD) とはウェブ上で機械処理に適したデータを公開・共有するための技術の総称である。機械処理に優れており、データを実体の関連として記述するのが MAG のデータ作成に適しているため、本研究ではこれに基づいてデータセットの開発を行う。

膨大な数に上る MAG の作品および出版物のデータをなるべく網羅的に作成するため、複数の既存のデータセットを変換・統合してデータセットを作成した。その際、作品と出版物のほかに、それらを繋ぐ「コミックス版」のような出版物を集約する中間的実体と構造的関係を表現する必要がある。

5. データセット作成手法

本研究では Wikipedia、メディア芸術データベース(MADB)、しよばいカレンダー[5]のデータを利用してデータセットを作成した。これらを利用することで、MAG すべてで作品と出版物の構造的関係の記述が可能である。

MAG それぞれに作品と出版物、中間的実体、さらに MAG を横断する上位の作品を持つデータモデルを構造的関係に基づいて定義した。

Wikipedia からは MAG 作品の実体とその構造の情報、ゲームの中間的な実体を取得した。MADB からはマンガ出版物と作品、その中間的な実体、ゲームパッケージの実体の情報を取得した。しよばいカレンダーからはアニメの作品とテレビ放送、中間的な実体の情報を取得した。それぞれが持つ作品実体の情報から、タイトルの編集距離が最も近く、発表年が同じものでマッチングを行い、作品実体の統合を行った。こうして作成したデータを RDF のトリプルで表現し LOD データセットを作成した。

表 1. 作成できた実体の数

MAGを横断する作品		26,382
マンガ	作品	106,319
	中間的実体	398,761
	出版物	556,579
アニメ	作品	5,712
	中間的実体	105,244
	出版物	360,046
ゲーム	作品	21,437
	中間的実体	21,437
	出版物	40,655

6. データセット作成結果

提案手法によるデータセットの作成結果を 表 1 に示す。ランダムサンプリングして目視で検証を行い、すべての実体で約 9 割の精度があることを確認した。

7. おわりに

データセット作成ルールの改善、利用するデータセットの増加によって更なるデータセットの拡充が期待される。

謝辞

本研究は JSPS 科研費 JP 16H01754 の助成を受けたものです。

参考文献

- [1]メディア芸術データベース.
<https://mediaarts-db.bunka.go.jp/>
(参照日 2019-1-25)
- [2] Wikipedia. <https://ja.wikipedia.org>
(参照日 2019-1-25)
- [3] 大石康介.“Wikipedia を用いたマンガ・アニメ・ゲームの作品実体の同定とカテゴリ情報に基づく構造化”. 筑波大学, 2018.
- [4] Jacob Jett, Simone Sacchi, Jin Ha Lee, Rachel Ivy Clarke. A Co0nceptual Model for Video Games and Interactive Media. Journal of the Association for Information Science and Technology Volume 67 Issue 3,2016,p505-517.
- [5] しよばいカレンダー. <http://cal.syoboi.jp/>
(参照日 2019-1-25)



この記事の著作権は著者に属します。この記事は Creative Commons 4.0 に基づきライセンスされます(<http://creativecommons.org/licenses/by/4.0/>)。出典を表示することを主な条件とし、複製、改変はもちろん、営利目的での二次利用も許可されています。