

Web リソースを用いたマンガ作品群における代表作の発見

塩原 百花^{†1} 三原 鉄也^{†2} 永森 光晴^{†3†4} 杉本 重雄^{†3}^{†1}筑波大学情報学群情報メディア創成学類 ^{†2}筑波大学図書館情報メディア研究科^{†3}筑波大学図書館情報メディア系 ^{†4}筑波大学知的コミュニティ基盤研究センター

1. はじめに

現在までにマンガ作品は 10 万タイトル以上出版されており^[1]、マンガに関する情報も膨大な数が Web 上に散在している。Web 上のマンガに関する情報は断片的なものが多いが、あるマンガ作品のメディアミックス状況や漫画賞の受賞状況といったデータを一箇所に集約することで、その作品の人気度、つまり面白さを評価することができるのではないかと考えた。同一のトピックを持つマンガ作品群に対してこの評価を行えば、その中で活躍している作品を発見することができる。その作品とはつまり、トピック内における「代表作」として考える。

本研究では、マンガ作品に関する Web リソースやマンガメタデータを使用し、任意の作品群から「代表作」を発見する手法を提案する。

2. マンガ作品群における代表作

本研究では、マンガ作品群における代表作を「あるジャンルやある時代・ある作者の作品群の中で、特に世間的に評価されており人気のある作品」と定義する。例えばメディア展開が活発に行われている作品は、一般のファンが多く、企業からはメディアミックスの企業利益が高いと見込まれていると推測できる。他にも、漫画賞の受賞作や連載が長期にわたり続いている作品は、その内容の面白さや作品の実力を世間から認められているのだと推測できる。

本手法ではマンガの派生(メディアミックス)作品の規模、漫画賞等の受賞歴、作品の連載期間や巻数、そして作品自体の規模や読者の注目度の目安として Wikipedia ページの記述量に基づいてマンガ作品を順位付けする(表 1)。順位付けのうち、上位の作品をその作品群の代表作とする。

“Finding the Masterpiece from a Comics Group using Web resources”

^{†1}Momoka Shiohara. School of Informatics. University of Tsukuba.

^{†2}Tetsuya Mihara. Graduate School of Library, information and Media Studies. University of Tsukuba.

^{†3}Mitsuharu Nagamori. Shigeo Sugimoto. Faculty of Library, Information and Media Studies. University of Tsukuba.

^{†4}Research Center for Knowledge Communities. University of Tsukuba.

3. 代表作発見における Web リソースの利用

本研究では、代表作の評価にメディア芸術データベース^[2]、DBpedia^[3]、Wikipedia^[4]を利用する。メディア芸術データベースは、文化庁メディア芸術デジタルアーカイブ事業によって作成されたマンガ・アニメ・ゲーム・メディアアートに関するデータベースで、マンガの書誌情報やアニメ・ゲームの作品情報を取得することができる。DBpedia、Wikipedia では、メディア芸術データベースからは手に入らないデータを収集する。例として、作品の受賞歴や実写映像化の状況などが挙げられる。

4. 代表作発見手法

本手法では代表作の発見に同一のトピックを持つマンガ作品群を用いる。例えば音楽漫画の代表作を知りたい場合は音楽をトピックに持つ漫画を収集する。収集した各作品について、表 1 に示した評価項目を取得する。例えば作品 A の巻数が全 5 巻であれば巻数の評価項目値 x_A は 5、作品 B が全 10 巻であれば評価項目値 x_B は 10 とする。次に評価項目別に偏差値 t_x を求める。作品 A・B・C において各巻数の項目値が $x_A = 5$ 、 $x_B = 10$ 、 $x_C = 6$ である場合、巻数に関する作品 A の偏差値 t_{x_A} は 41、B の偏差値 t_{x_B} は 64、C の偏差値 t_{x_C} は 45 となる(小数点以下は四捨五入した)。このようにして全評価項目において値の偏差値を求め、一つの作品に関する n 個の評価項目の偏差値全てを足し合わせた $\sum_{i=1}^n t_i$ をその作品の総合スコアとし、総合スコアから作品 A の総合偏差値 T_A を算出する。総合偏差値 T の高さで作品を順位付けし、上位の作品ほど代表作に値するものであると推定する。

5. 評価実験

実験において、同一トピックを持つ作品の収集には Wikipedia のカテゴリを用いた。カテゴリとは分類やキーワードを示すために Wikipedia の記事ごとに付与されているものである。例えば漫画「ドラゴンボール」には、「漫画作品」と「週刊少年ジャンプの漫画作品」「鳥山明」「格闘技漫画」「冒険漫画」といった複数のカ

表1 作品の評価項目とそのデータの収集元

項目名	データ収集元
受賞歴	DBpedia
連載期間	メディア芸術データベース
マンガの巻数	メディア芸術データベース、DBpedia
映像化(アニメ)	メディア芸術データベース
映像化(実写)	DBpedia
ゲーム	メディア芸術データベース
関連書籍	メディア芸術データベース
Wikipedia の記述量	Wikipedia

カテゴリが付与されている。実験では、カテゴリ「野球漫画」を持つ 170 作品と「大学を舞台とする漫画作品」を持つ 60 作品に本手法を適用した。表 2 と表 3 に実験結果の上位 10 作品のタイトルと偏差値、参考として評価項目の一部である「巻数」「メディアミックスの回数(アニメ・映画・ドラマ・ゲーム化における回数。偏差値算出時にはメディア別に評価しているが、ここでは総数を記す。)」 「受賞回数」を示した。

6. 評価と考察

実験結果では、メディアミックスの回数が多い作品や巻数が多い作品(すなわち長期連載の作品)が上位になった。DBpedia・Wikipedia の作家の項目には、その作家の代表作を示すプロパティ「prop-ja:代表作」が付いている場合がある。今回の 2 カテゴリ 20 作品の作家 20 名中、「prop-ja:代表作」を持つ作家は 16 名おり、その全てに本実験で求めた作品が含まれていた。これにより本手法で取得した作品は一般的に認知度が高いものだと考えられる。また、作品の人気の出る傾向は各々で違う。したがって着眼点が偏らないよう、異なる観点の評価項目が必要である。今回使用した評価項目の観点は大きく 3 つに分けることができる。「マンガのステータス(連載状況や Wikipedia ページを記述するファンたち)」、「メディアミックスの活発さ」、「専門家からの評価(受賞歴)」である。今回の実験結果では、バランス良く項目値が高いものもあれば、ある一点にのみ特化したものもある。項目値に則した評価を行いつつ、その作品の強みを活かした順位付けができています。

7. 今後の課題

今後の課題として以下の 3 点を挙げる。一つ目は収集データの精度の問題である。DBpedia 及び Wikipedia では構成の差異や表記ゆれにより正確なデータが取り出せない場合がある。データの抽出の精度を上げるか、何らかの整備されたデータを手にいれなければならない。二つ目は評

表2 Category:野球漫画(170 作品中上位 10 作品)

順位	タイトル	偏差値	巻数	メディアミックスの回数	受賞回数
1	巨人の星	111.87	19	21	0
2	MAJOR	102.95	78	11	1
3	タッチ	95.32	26	9	1
4	あぶさん	79.69	107	0	1
5	ドカベン	78.80	48	5	0
6	ダイヤのA	72.81	25	1	2
7	キャプテン	72.42	26	5	1
8	野球狂の詩	72.39	17	5	0
9	おおきく振りかぶって	71.22	15	4	2
10	グラゼニ	63.10	16	0	3

表3 Category:大学を舞台とする漫画作品(60 作品中上位 10 作品)

順位	タイトル	偏差値	巻数	メディアミックスの回数	受賞回数
1	ああっ女神さまっ	91.65	48	11	1
2	のだめカンタービレ	77.29	25	9	1
3	生徒諸君!	71.43	24	3	1
4	ハチミツとクローバー	69.44	10	4	2
5	げんしけん	66.33	9	6	1
6	神戸在住	65.29	10	2	1
7	もやしもん	64.22	13	4	2
8	天才柳沢教授の生活	62.52	33	1	1
9	東京大学物語	62.51	34	4	0
10	動物のお医者さん	60.82	12	1	0

価項目の検討である。現在は 8 項目を用いているが、より精度を上げるために項目数の追加が必要であると考えている。三つ目は項目への重み付けである。現在は全項目の偏差値を同様の重み付けで評価しているが、適切な重み付けによって、より精度の高い結果を得ることができるようになるだろう。

8. おわりに

本論文では Web リソースを用いて任意の作品群の中から代表作を発見する手法を提案した。この手法により Web 上に散らばるデータの新たな活用方法を提案し、データに基づいた信頼できるマンガ作品の評価を行うことが可能になった。今後は作品評価の精度を高めていくと共に、別のメディア芸術作品でも同様に代表作を発見できるよう取り組んでいきたい。

参考文献

- [1]凸版印刷株式会社編.平成 26 年度メディア芸術デジタルアーカイブ事業委託業務成果報告書.2015,144p.
- [2]メディア芸術データベース.
<https://mediaarts-db.jp/>
- [3]DBpedia. <http://ja.dbpedia.org/>
- [4]Wikipedia. <https://www.wikipedia.org/>