

SS研座談会 第1回

「マンガをビッグデータとして扱う

～来たれ！SS研のマンガ好きたちよ！～」討議骨子

1. 検索について

①マンガの版数による違い

- ・マンガは版を重ねると修正が入り、例えば雑誌で掲載されたものと、単行本では違うものになっていることがある。これは異本として解釈できる。
- ・どのように変わっていったかは興味がある。
- ・N次創作はどのように位置づけたいのか検討が必要である。

②キャラクターによる検索

- ・非常に長大な作品の一部を読みたい場合に、登場人物（キャラクター）をキーとして検索したくなる可能性がある。
- ・ネット上の画像共有サービス等でキャラクターデータベースがあり、タグクラウド系と階層系（分類表）に大別できる。
- ・同名のキャラクターの区別は問題である。
- ・名前データベースの構築と維持管理更新も問題である。

③ストーリーによる検索

- ・ストーリーの検索をしたい1つの場面は「この話の続きが読みたい」である。

④元ネタの検索

- ・パロディ的な作品の元ネタを検索したくなる場合がある。

⑤その他の検索

- ・コマ割りによる検索も考えられる。この場合、条件入力方法は「このようなページに似た作品」であろう。

⑥推薦システム

- ・推薦システムは利用者として欲しい機能である。ネタバレをせずにお勧めだけしてくれるのが望ましい。
- ・利用者による作品の評価に際し、匿名化は良い面と悪い面がある。

2. デジタル化について

①キャラクターの判別

- ・画像処理の観点からキャラクターの表情分析は難しい。
- ・顔抽出は写真の顔抽出のアルゴリズムでは全くうまくいかない。
- ・連載の開始時と終了時で、人間でも同一人物か判断できない場合もある。文脈を把握した上で追従しないと難しい。
- ・コマとキャラクターのリンク付けもまた難しい。
- ・更にはキャラクターの描かれ方（デフォルメされているか）などの情報も抽出したい。

開催情報

■日時：2013年7月31日(水) 13:30～17:30

■場所：富士通(株)汐留本社6階 ユーザコミュニティサロン

■参加者：6名

※工学研究者、データベース研究者、理学研究者、元漫画家、等の方々

②文字抽出

- ・吹き出しの文字自体の抽出は可能であるが、それとキャラクターのリンク付けは難しい。
- ・地の文やオノマトペの抽出も難しい。

③マンガ理解システム

- ・コマの進行を理解するシステムが必要である。
- ・抽出した文字列が作品中でどのような意味を持つか、別途判断させなければならない。



3. 保存について

①デジタルデータ

- ・ネット上には大量にスキャンされたマンガが現状存在する。著作権の問題に加え、このようなデータは完全性と網羅性が保証されていないのが問題である。
- ・網羅性に関しては、著者が作品一覧を出してくれるのが一番望ましい。
- ・将来的には、スキャンした画像と、その内部の情報を別レイヤーで持つようなデータ構造で保存されるのが利用者（読者）にとっては有用であろう。
- ・長期連載のマンガの場合、将来的にはこのようなデジタル化と検索システムが必要なのではないと思われる。

②図書館

- ・図書館での保存は著作権の問題はクリアする。一方、現状マンガは図書館においても特殊分野でほとんど扱われておらず、新規にどう図書館での保存を開始することが可能かは必ずしも明らかではない。
- ・出版社側がそもそも図書館で扱ってほしくないという事情もある。
- ・作家としては作品を残したいと考えている。タダでもいいから誰かに読んでもらえる状態でないと散逸消滅する。
- ・著作権は、個人の作家の意思とは別に出版社が持っている場合がある。

③保存方法

- ・紙媒体としては出版社も自社の雑誌が完全に揃っていない場合があり、過去のものを残そうという活動には至っていない。
- ・国内数ヶ所だけ紙媒体での保存も行っている図書館がある。
- ・デジタルデータとして残す場合にも記録媒体自体の寿命と媒体の規格の寿命の問題はある。
- ・天災の事を考えると、分散クラウドは一つの解である。
- ・作品自体とは別にマンガの制作過程に関わった資料等をアーカイブしたいという希望もある。この場合、研究者・専門家が必要である。

以上