Analyzing User Requests for Anime Recommendations

Jin Ha Lee, Yuna Shim
University of Washington
Mary Gates Hall, Suite 370
Seattle, WA 98195
+1 206.685.0153
jinhalee@uw.edu, yunas@uw.edu

Jacob Jett
Center for Informatics Research in Science & Scholarship
Graduate School of Library & Information Science
501 E. Daniel St., Champaign, IL 61820
+1 217.244.2164
jjett2@illinois.edu

ABSTRACT
Anime is increasingly becoming recognized as an important commercial product and cultural artifact. However, little is known regarding users' information needs and behavior related to anime. This study specifically attempts to improve our understanding of how people seek anime recommendations. We analyzed 546 user questions in natural language, collected from a Korean Q&A website Naver Knowledge-iN, where users are asking for anime recommendations. The findings suggest the importance of establishing robust metadata for the seven commonly used features for anime recommenders (i.e., title, genre, artistic style, story, character description, series title, and mood) in digital libraries, as well as allowing users to specify known anime and series titles as examples for seeking similar items, or examples of the kinds of items to be excluded.

Categories and Subject Descriptors
H.3.7 Digital Libraries: User Issues; H.5.1 Multimedia Information Systems: Animations

General Terms
Human Factors.

Keywords
Anime; User; Recommendation; Metadata.

1. INTRODUCTION
Anime (i.e., Japanese Animation) emerged from Japan's strong heritage of comics in the 1960s and has been part of mainstream Japanese culture since the 1970s. Its global popularity has been gradually increasing [10], owing to successful filmmakers such as Miyazaki Hayao and steady penetration into broadcast marketplaces (e.g., Cartoon Network's Toonami). As anime's user base is continuing to grow in the global market, it is important to provide users from different regions and cultures with efficient ways to seek anime information to facilitate the discovery of new titles. However, little research has been done on cross-cultural users' information needs or seeking behavior around anime. Therefore we do not have enough empirical user data to derive design requirements for an effective anime recommender with cross-cultural considerations, resulting in reduced discoverability.

This study attempts to address this gap and aims to improve our understanding on cross-cultural anime users' information needs and behavior, in particular, when they seek recommendations. We specifically seek to answer the following research question: Which information features do people describe and how do they describe them when seeking anime recommendations?

JCDL'15, June 21–25, 2015, Knoxville, TN, USA.
ACM 978-1-4503-3594-2/15/06.
http://dx.doi.org/10.1145/2756406.2756969

2. BACKGROUND AND PRIOR WORK
Anime is a rich and complex form of multimedia, consisting of various components such as images, sound, speech and text [7]. Although there are a number of studies focusing on organization of and access to movies (e.g., [2],[8]), anime, as a cross-cultural information resource, is different from conventional movies in terms of genres, characters, and content, and thus poses unique search/discovery challenges for users. Among these challenges are:

- exact title (*Shingeki no Kyojin* vs. *Attack on Titan*),
- names of cast members,
- unique genres (e.g., スーパーロボット *(Super Robot)*), and
- cross-medium visual styles (e.g., shōnen, shōjo)

Previous studies on multimedia information retrieval systems and their users found that user queries tend to be shorter and less formulated than queries in traditional information retrieval systems because of the lack of web searching mechanisms and representational congruity [4]. This may also be the case for anime related queries. For this reason, rather than attempting to analyze search queries in a particular anime website(s), we chose to collect users' natural language questions from relevant Q&A websites. Several prior studies have successfully analyzed and utilized user questions for automatic classification of queries [3], to understand users' motivations for participation [9], and to investigate how they seek music information [6]. We also chose to analyze online user questions rather than observe how they conduct an artificial search task in an experimental setting in order to understand how they behave in a real-life search setting.

3. STUDY DESIGN AND METHOD
Content analysis of user questions was adopted as the main method for this research [5]. The user questions were collected from Naver Knowledge-iN, the largest online Q&A community in South Korea. Naver limits the number of questions that are viewable, and thus we mined the maximum number of questions retrievable using a combination of keywords such as 애니, 애니메이션, 추천 (anime, animation, recommend(ation)). A total of 1,326 questions were collected in 2013, and we manually filtered out the irrelevant questions that were not seeking anime recommendations (e.g., known-item searches). After examination, 546 were deemed relevant. A typical question looks like: "I am looking for fantasy action anime, over 50 episodes with a bright tone (unlike *D.Gray-man* or *Black Butler*)." All the examples of user questions presented here are translated by the first author who is fluent in both English and Korean.

The questions in the dataset were coded over 11 months through an iterative process of coding sub-samples and refining the codebook. In order to test the intercoder reliability of the features, two coders independently coded 100 queries based on the revised codebook. The coded data were transformed into a binary representation based on the absence (0) and presence (1) of each feature, and Krippendorff's alpha-coefficient was calculated. The global alpha-coefficient value of all features was .81 exceeding the minimum value of .8 to be considered reliable [5].

4. RESULTS AND DISCUSSION

4.1 Features of Anime in Users' Questions

In total, 13 features of anime were identified in more than five questions. Table 1 shows the features with their definition, count of questions containing each feature, and percentage over the number of questions. The number of features commonly used was relatively small; users tended to heavily rely on a handful of top ranked features. Anime title was the most frequently used term, and series title was also commonly used. Several features describing the content of anime such as genre, style, story, character description, and mood were also highly ranked. The other six features were used in less than five percent of all questions. The "other" category included features like person or corporate names, availability of anime, edition information, etc.

Table 1. Features of anime mentioned in user questions

Feature	Description	#	%
title	Name of anime titles	294	53.85
genre	Genre/theme of the anime	193	35.35
style	Description of artistic style of the anime	83	15.20
story	Story, narrative, or plot of the anime	73	13.37
character description	Characters' personalities or traits portrayed in the anime	59	10.81
series	Name of anime series	52	9.52
mood	Overall tone or atmosphere of the anime	47	8.61
audience	Intended audience of the anime	26	4.76
temporal information	Specific year or a range of years when the anime is published	19	3.48
length	Number of episodes (for TV series) or running time (for movies)	18	3.30
scene	Description of a specific kind of scene wanted (e.g., hot spring scene, kiss scene)	13	2.38
character-name	Names of specific anime characters (e.g., Naruto, Lelouch)	8	1.47
format	Distribution medium of the anime	8	1.47
other	Other characteristics	10	1.83

Among the lessons derived from the results of this study, relevant to the development of an anime recommendation service are:

1) similarity-based relationships among anime titles are extremely important to users
2) narrative genre-labels, theme-labels, and mood terms vary considerably across cultures creating a complex topical landscape for users to navigate (e.g., unique terms like 먼치킨 (munchkin), 백합 (baekhap))
3) artistic style of anime is prominently featured in user questions, yet its description tends to be generic, and
4) descriptions of overall story/narrative or specific characters are also commonly used (e.g., *"protagonist's talent gets noticed by a renowned person…and everyone else gets jealous")*

4.2 Other Contextual Information

In addition to the information features of anime found in questions, context particular to each user was prominently featured in many of the questions. These questions included qualitative information such as: personal tastes and contextual information about themselves (appeared in 163 questions), reaction (appeared in 102 questions), criteria for exclusion (appeared in 140 questions), and instructions for recommenders (e.g., provide images or short summaries) (appeared in 112 questions).

5. CONCLUSION

The findings of our analysis of 546 user questions suggest that there is a need for robust metadata that better captures and describes anime content features such as genre, artistic style, story, character types, and mood. Such metadata would support a more well-rounded recommendation service that improves users' abilities to expand their queries beyond known-title searching strategies. While users have a deep and descriptive understanding of genre and mood labels, the vocabulary used to express artistic styles remains simple and their information needs were frequently more qualitative and contextually sensitive than ordinary library metadata and recommenders services typically support (e.g., their personal tastes). In our future work, we plan to establish user-centered taxonomies of genre, mood, and trope based on the description given in user questions so that it can be used to facilitate discovery of anime in digital libraries. We also plan to conduct semi-structured interviews to learn more about how users seek and evaluate anime recommendations, involving users from different cultural backgrounds to identify any cross-cultural differences or challenges that may exist.

6. REFERENCES

[1] Bloch, D. A. and Kraemer, H. C. 1989. 2x2 kappa coefficients: measures of agreement or association. *Biometrics,* 45, 269-287.

[2] Cameron, A. 2010. Approaches to cataloguing moving image: some practical experiences. *Catalogue & Index*, 159, 2-4.

[3] Harper, F. M., Moy, D. and Konstan, J. A. 2009. Facts or friends?: distinguishing informational and conversational questions in social Q&A sites. In *Proc. CHI,* ACM, 759-768.

[4] Jansen, B. J., Goodrum, A. and Spink, A. 2000. Searching for multimedia: analysis of audio, video and image Web queries. World Wide Web 3, 4, 249-254.

[5] Krippendorff, K. 2004. *Content analysis: an introduction to its methodology*. Sage, Thousand Oaks, CA.

[6] Lee, J. H. 2010. Analysis of user needs and information features in natural language queries seeking music information. *JASIST,* 61, 5, 1025-1045.

[7] Maghrebi, H. and David, A. 2006. Toward a model for the representation of multimedia information based on users' needs: economic intelligence approach. In *Proceedings of m-ICTE*, 195-200.

[8] Martins, P., Langlois, T. and Chambel, T. 2011. MovieClouds: content-based overviews and exploratory browsing of movies. In *Proc. MindTrek '11,* ACM, 133-140.

[9] Nam, K. K., Ackerman, M. S. and Adamic, L. A. 2009. Questions in, knowledge in?: a study of Naver's question answering community. In *Proc. CHI*, ACM, 779-788.

[10] Poitras, G. 2008. Contemporary anime in Japanese pop culture. In *Japanese visual culture: Explorations in the world of manga and anime*. M.E. Sharpe, Armonk, N.Y.