27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

A Method to Construct a Comic Spoiler Dataset and Analysis of Comic Spoilers

Takumi Takaku[a], Yoshiki Maki[a], Satoshi Nakamura[a]

[a]Meiji University, Nakano4-21-1, Nakano-ku, Tokyo, Japan

Abstract

People differ in their assessment of whether they consider spoilers to be problematic or not. Recent studies have investigated the effects of spoilers. However, the definition of spoilers in those studies was ambiguous, so it is difficult to say that the impact of spoilers has been investigated. In this work, we propose a definition of comic spoilers and a dataset construction method to avoid ambiguity in the definition of spoilers. Furthermore, we construct a dataset and compare our dataset with the spoilers used in previous studies to clarify the characteristics of spoilers. Then, we found that the newly constructed dataset not only covered the previous dataset but also detected important spoiler scenes that were not included in the previous dataset, such as goal scenes and scenes that reveal the killer's motive and so on, as spoilers.

© 2023 The Authors. Published by ELSEVIER B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer review under responsibility of the scientific committee of KES International

Keywords: Spoiler, Comic, Dataset

1. Introduction

Spoilers are frowned upon by those who enjoy the content, such as movies, comics, stories, etc., as they detract from the enjoyment of the content itself. However, some people do not mind spoilers and want to enjoy content after seeing them, and it is often debated whether spoilers are a problem.

Previous studies investigating the effects of spoilers have shown that they do not change the enjoyment of work and even increase it, depending on the content. In a representative study, Leavitt et al. [1] compared the degree of pleasure perceived by readers who were allowed to read the punchline of a short story before reading the story with those who were not allowed to read the punchline. The experiment results showed that readers who read the punchline, i.e. spoiled readers, rated the work as more attractive than readers who were not allowed to read it, and it became clear that spoilers make short stories more interesting. Levine et al. [2] also prepared two timing schemes for presenting spoilers in short stories before and during reading. They investigated whether there was a change in the interest perceived by the spoiled and non-spoiled readers on each occasion. As a result, they clarified that the level of interest increases when the contents of books are spoiled during reading. Furthermore, Hassoun [3] has shown that the appeal

of a comic book increases when the last panel of a page is read at the beginning. In our previous studies [4][5], we also conducted an experiment in which we set three spoiler timings: before reading the comic, after reading the first part of the comic, and after reading the middle part of the comic, and divided the participants' groups into those with spoilers and those without spoilers, and found no significant difference in enjoyment between them. On the other hand, a study by Tsang et al. [6] tested whether spoilers in films affect consumer behavior and found that spoilers make people more reluctant to watch movies.

Based on the above, spoilers may or may not influence the interest in a work. Here, our studies' results only show no change in interest after reading to the end. Still, the number of participants who lost interest in the rest of the comic after being spoiled was double compared to non-spoiled participants, suggesting that spoilers would have increased the withdrawal rate. The study by Leavitt et al. also dealt with famous works, and it can be said that it merely suggests that favorite content with a high number of sales is interesting whether it is spoiled or not, if it is read to the end following the constraints of the experiment.

In addition, the authors created spoilers in Leavitt et al. [1] and Levine et al. [2], so they are likely to have an author bias and present information that only some readers would want to know. In addition, the spoiler panels presented in Hassoun's study [3] only show the last panel of the facing page, which is hardly a spoiler. Our past studies [4][5] also have unclear points of selecting spoilers because we asked participants to choose spoiler pages from the whole of a comic, and some participants did not care about spoilers in the first place. There were also problems with the information presented as spoilers: the spoilers used by Levine et al. and Hassoun were short sentences or were only part of a page, and the amount of data needed to be more significant. Moreover, the spoilers previously used by Leavitt and ourselves only contained the final punchline, so the process up to the end of the contents was designed to be enjoyed without being spoiled. As a result, the influence of spoilers was thought to be less likely to appear.

In other words, the previous studies on spoilers' effects have been ambiguous regarding what spoilers are and in what places and media they should be presented. Therefore, spoilers have not sufficiently reduced the interest in the work, and the impact of spoilers has been less likely to be detected. Thus, in this study, based on the results of the previous experiments [4][5], we define the definition of spoilers in comics so that it is possible to clarify what kind of spoilers should be presented in the experiment. The reason for focusing on comics is that comics are published in more extensive serialization periods than other content, such as novels, dramas, animations, and movies. In Japan, comics are serialized in magazines and published as a single book combining several stories. As a result, the existence of magazine and book groups means that readers often enjoy a story at different times, and spoilers are more likely to occur than with forms of other story content.

In this study, we first clarify the definition of spoilers and ask experimental collaborators to select comic book pages that match the defined spoilers. Next, by analyzing the spoiler data obtained from the collaborators and observing the characteristics of the data, we will consider what kinds of places are likely to be avoided as spoilers. We will also consider the method to prevent spoiler effects in the future.

2. Related Work

2.1. Studies on the effects of spoilers

As mentioned in the previous section, Leavitt et al. [1] and Levine et al. [2] have identified the impact of spoilers in short stories. Hassoun [3] also clarified the influence of spoilers in comic book panel layouts. On the other hand, Tsang et al. [6] tested whether spoilers in movies affect consumption behavior. The results revealed that spoilers make people more reluctant to watch films. They also identified the psychological factors of spoilers' influence and found that spoilers' effect can be reduced by using items other than storyline development in advertisements. Yan et al. [7] also investigated the psychological factors of the impact of spoilers. As a result, they found that the effects of spoilers varied according to the film's interpretation level. Furthermore, they discovered that spoilers concerning the story's ending and spoilers concerning the process leading up to the story's end changed the reader's anticipation before reading and the interest felt after reading the story and that spoilers concerning the story's process reduced the enjoyment more than spoilers concerning the story's outcome. Rosenbaum et al. [8] have investigated the factors behind the impact of spoilers on novels by approaching the issue from a cognitive science perspective. As a result, they found that the amount of reading experience is a factor that determines the positive and negative effects of spoilers, with those with more reading experience experiencing a decrease in enjoyment due to spoilers and those with less

reading experience experiencing an increase in enjoyment due to spoilers. However, even in these studies, the definition of a spoiler could be more precise, and it is difficult to say that the true impact of spoilers has been clarified. Johnson et al. [9][10][11] described the effect of spoilers in the audiovisual media of television and film. The results showed that there was little impact on the influence of spoilers on the level of popularity. However, they also investigated the predicted popularity of the characters and found that the reactance of this item increased due to being spoiled. We also investigated the effect of spoilers in horror films and found that while they did not affect the overall emotion or fear, they did reduce the fear at localized scenes.

2.2. Studies on preventing spoilers

Spoilers are a problem, and many people are trying to prevent spoilers systematically. Various studies have attempted to avoid spoilers through a systematic approach. The results of these studies have been used to define spoilers in comic books. Shiratori [12] clarified the characteristics of sports spoilers on Twitter that could not be determined in a related study by creating a dataset based on the degree of spoiling. They also proposed a method for judging spoilers according to the sports genre and the elapsed time in the game. Golbeck [13] proposed a method for detecting spoiler comments on Twitter by using the difference in TV program broadcast timing in separated areas. Jeon et al. [14] proposed increasing the conformance rate using SVM. Guo et al. [15] used a movie review dataset and extracted spoiler words by performing a morphological analysis. Boyd-Graber et al. [16] proposed a method for determining spoiler and non-spoiler sentences in a movie review text. Maeda et al. [17] and Ikeda et al. [18] also studied methods to prevent spoilers in review texts. In this case, spoilers in the story texts were collected from user surveys and analyzed texts that are likely to be perceived as spoilers. Furthermore, they propose a method to remove spoilers from review texts and to create spoiler-free review texts. Pang et al. [19] proposed a method for deciding spoiler synopsis sentences and non-spoiler opinion sentences by using SVM. However, the definition of a spoiler still needs to be clarified in these studies, and it can be said that research on spoilers is complex.

2.3. Studies on comic

Various studies have also been conducted on comics. Matsui et al. [20] created the Manga109 dataset for use in research and used it to generate metadata to develop comics research. In addition, Sakurai et al. [21] extend the Manga109 dataset by mapping a text line to a speaker. Ogawa et al. [22] note that the current comics dataset lacks annotations for the comics' text and panels and proposes a system that can automatically annotate them. It also reveals that the time required for annotation varies depending on the work. Tsubota et al. [23] paid attention to the faces of characters in comics and proposed a new method for classifying the faces of characters between authors. As a result, they have made it possible to perform clustering with high accuracy compared to previous methods. Daiku et al. [24] clustered genres using the stories in comic books, making it possible to separate genres cheaply. As a result, it was possible to improve the clustering accuracy.

These researches study comics from various perspectives, but none have addressed spoilers' relevance. However, they took approaches such as metadata extraction and tagging, considered effective methods for estimating spoilers.

3. Definition of a Spoiler

As mentioned in the introduction, previous spoiler studies have been problematic regarding whether the selected spoilers are indeed spoilers, as the definition of a spoiler needs to be clarified. In other words, problems with the chosen spoilers may have led to them having less impact. Therefore, based on our previous studies, we define spoilers in comics by considering what should be treated as spoilers in comics.

3.1. Discussion on previous studies

As mentioned above, in our past work [4][5], we investigated whether spoiling the final part of a comic book would change the level of interest and interest in the rest of the story. The spoilers presented to the participants were selected by ranking three pages from several volumes of a comic work prepared in advance. Multiple participants gave the

Table 1. Title of comics used in this study.

Genre	Title of Comics	Number of Volumes
Battle	All You Need Is Kill	1 - 2
Sport	GIANT KILLING	1 - 5
Suspense	Prophecy	1 - 3
Romance Story	ReRe Hello	1 - 6

page the highest ranking. In addition, by preparing three stages of spoiler timing (before reading, after reading up to the beginning, and after reading up to the middle), we investigated and analyzed the effect of spoilers at each stage of the reading process.

The results showed that the influence of spoilers on the level of interest was not observed at any of the spoiler timings. However, the level of interest in the continuation of the story significantly decreased when a spoiler was given after the middle part of the story. In this experiment, readers who were spoiled at any time were satisfied with the contents of the scenes after the middle of the book. These results suggest that what is abhorred as a spoiler is content that comes immediately after where the reader is reading.

Here, spoilers are considered to be intensely dependent on the context of how far readers have read initially. For instance, if a spoiler of volume 20 is presented to a person who has read volume 1 of a series, they are unlikely to recognize it as a spoiler. In other words, to construct a dataset of spoilers, it is necessary to ask the reader to select spoilers for each scene, depending on the appropriate situation. Therefore, the chosen spoilers in previous studies may be inappropriate. Also, in the dataset constructed in the previous study, it was directed that readers could select only three pages from a comic, which were subjectively perceived as spoilers, and were asked to assign a ranking. Since the number of pages was limited to three, only a few could be selected as spoilers, despite initially having many more spoiler pages. Pages that should have originally been presented as spoilers could not be given. On the other hand, another problem in constructing the spoiler dataset has been including users who do not care about spoilers. It is thought that the judgment of whether a page is a spoiler for them increases the blurriness regarding whether the page was a spoiler.

From the above points, the spoiler datasets constructed in previous studies had problems regarding context, the number of selectable pages, and attitude toward spoilers. It is essential to solve these problems to build a spoiler dataset.

3.2. How to define spoilers in comics

In this study, to define spoilers clearly, we describe the criterion for judging whether a page is a spoiler page or not as "a page p in episode $N+1$ that would be objectionable to others who have read up to episode N of the comic when it is presented." By using others as the criterion, even if there are people in the dataset constructer who do not mind being spoiled or who want to enjoy the content after being spoiled, it is thought that by having them think of others who do not wish to be spoiled, dataset constructer who do not mind being spoiled could choose the spoiler scene. Also, we limited the target audience to readers who had read the previous episode. This was because, from the earlier studies, we realized that spoilers affecting readers were likely to be on a page immediately after the reader had read it.

Next, when the readers were asked to select spoilers, they were asked to tag the spoilers according to the degree of spoilers. There are three types of tags, namely blue, yellow, and red. The blue tag is set as "not a spoiler." The yellow tag is set as "a page that would make people angry if shown to others," which means pages that annotators themselves consider as spoilers while other readers may not. The red tag is set as "a page that definitely makes people angry if shown to others," which means pages that both annotators and other readers consider spoilers. Readers were asked to tag the page based on this information. The experiment was conducted by having the participants tag pages that were considered spoiler pages according to the three criteria mentioned above so that they could recall what the intended readers were thinking by experiencing the work themselves. This makes remembering what the intended readers think through their experiences easier.

4. Spoiler Dataset Construction

Based on the criteria for spoilers in comics described in section 3, we constructed a dataset with a clear definition of spoilers (hereafter referred to as the "spoiler dataset") so that no ambiguity remained when it was used in experiments. We will also use the constructed spoiler dataset to determine how it differs from spoiler datasets constructed in previous studies, which pages are likely to be spoilers for the reader, and what kind of pages are likely to be spoilers for the reader.

4.1. How to construct a spoiler dataset

The names and volume numbers of the comics from which the spoiler pages were collected are shown in Table 1. The four comics used in this study were selected from the four genres used in previous studies [4][5]. This was done to find the differences between the spoiler dataset used in the earlier studies and the one created in this study. Participants who constructed this spoiler dataset were five university students aged between 21 and 25, and we did not consider whether they had read the comic. The flows of the dataset construction were as follows.

(1) Our system asked participants to read the Nth story. N starts from 1.
(2) Our system listed the N-1th thumbnail pages and asked participants to browse them briefly.
(3) Our system asked participants to select a spoiler for each page of the Nth story based on the criteria for selecting spoilers defined in the previous section. They were also asked to choose one of three levels (blue, yellow, red) for the degree of spoilers (see Fig 1).

The participants were asked to repeat these steps (1) to (3) until the final episodes of each work so that they could select spoilers for all episodes.

4.2. Results

We scored spoilers based on the degree of spoilers tagged (blue tag: 0 points, yellow tag: 1 point, red tag: 2 points). Fig 2-5 shows the results of the average score of the five participants for each work, page by page. Also, Fig 6 shows the results of assigning scores according to the ranking of the spoiler pages (1st: 5 points, 2nd: 3 points, 3rd: 1 point) for the works used in this study from the spoiler dataset constructed in the previous survey.

Comparing the spoiler dataset created this time and the spoiler dataset made last time shows that the former acquired many spoilers. The pages that scored highly in the previous dataset also scored high in the current dataset, but we needed help selecting pages with similarly high scores.

A comic-by-comic analysis shows that in the comic title "All You Need Is Kill," the pages with high spoilers were
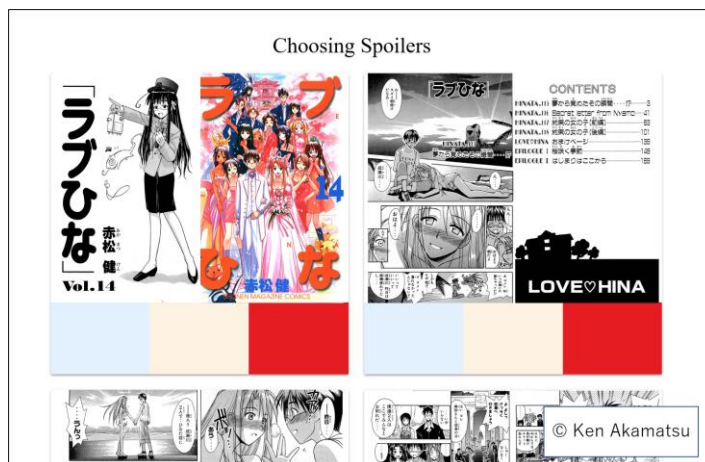

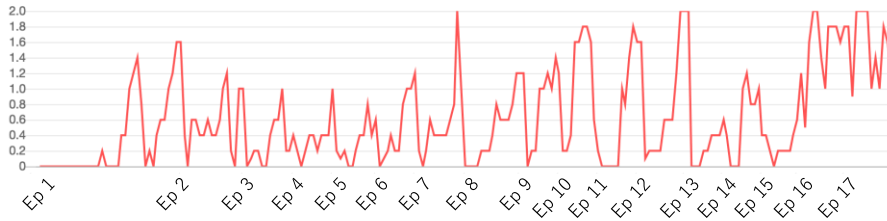Fig. 1. A system screenshot to select degree of spoilers.

Fig.2. Transition of degree of spoilers in "All You Need Is Kill."

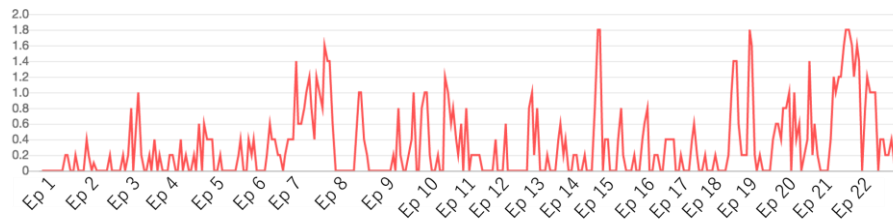
Fig.3. Transition of degree of spoilers in "GIANT KILLING."

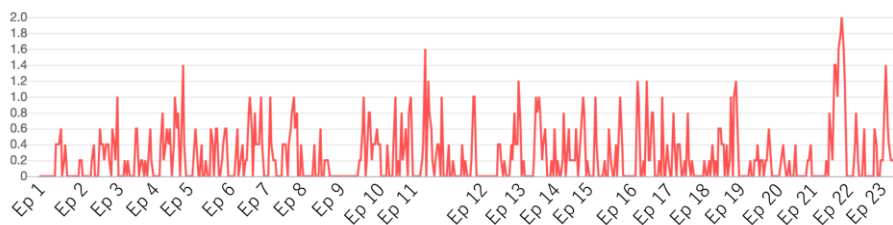Fig.4. Transition of degree of spoilers in "Prophecy."

Fig.5. Transition of degree of spoilers in "ReRe Hello."

found in episodes 7, 10 to 13, and 16 to 17. Episodes 7, 10, and 11 depict scenes in which the heroine's secret is revealed. Episodes 12 and 13 depicted scenes in which the core secret of the story is told, and episodes 16 and 17 depicted scenes that would be the story's punchline. Compared to the previous spoiler dataset, the scenes in episodes 16 to 17 were mainly selected as the end of the story in previous studies. The scenes in episodes 12 and 13 were the core secret of the story. They were not chosen as spoiler pages in previous studies, which means these episodes were not considered sufficient for a spoiler page in previous studies.

In the comic title "GIANT KILLING," the pages with many spoilers were 16-17, 20-21, 27, and 46-47. The scenes depicted in the pages with high spoiler counts were all in which the goal scene or the results of crucial matches were known. The target contents were related to several games. Still, in the previous spoiler dataset, only the goal scenes of the last game and the results of the previous match were included. However, the pages depicting the goal scenes and the effects of the most critical games were collected for each match in the current dataset. From this, it can be said that the previous dataset was insufficient regarding situational spoilers.

In the case of "Prophecy," episodes with pages with a high degree of spoilers were episodes 7, 15, 18, and 21. These three episodes were selected because, in episode 7, the scene depicting the background of the criminals; in episode 15, the location where one of the characters betrays the plot; in episode 18, the intro leading to the punchline; and in
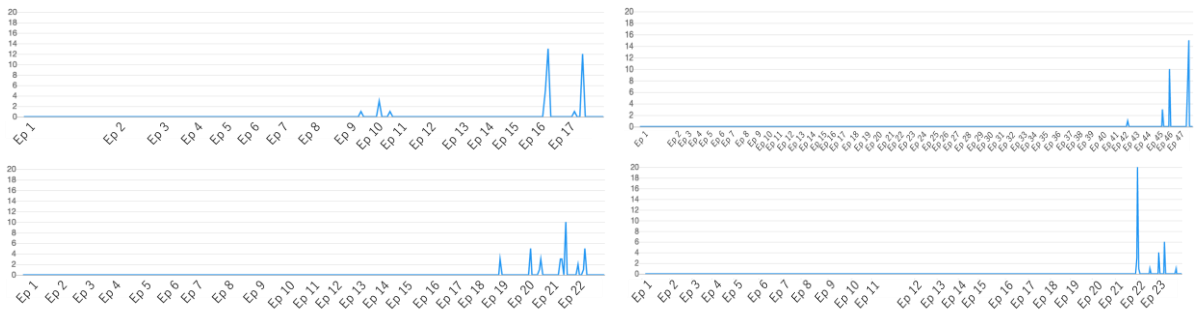
Fig. 6. Transition of degree of spoilers, constructed in previous studies.
(Left-above: "All You Need Is Kill", Right-above: "GIANT KILLING", Left-bottom: "Prophecy",
Right-bottom: "ReRe Hello")

episode 21, the scene depicting the punchline of the story and the motives of the criminal. Compared with the previous spoiler dataset, only the set that would be the story's punchline was chosen for the last dataset spoiler, whereas other scenes were also obtained this time. Since these scenes were essential in the story, it can be said that the previous dataset was inadequate.

In "ReRe Hello," the episodes with a high number of spoiler pages were episodes 11, 21, and 23. The scenes depicted in episode 11 were about the romance between the protagonist's friends, and those in episodes 21 and 23 were about the liaison between the protagonist and the hero character. In the previous spoiler dataset, only the pages describing the romance between the protagonist and hero characters were selected and retrieved. In contrast, however, we could also include friends' affairs in this dataset. Still, the evaluation value was low for these pages, and the scenes with high importance were also selected well enough in the previous dataset.

4.3. Discussion

In all comics, a comparison between the previous spoiler dataset and the current spoiler dataset showed that the latter was able to collect more spoiler pages than the former because the number of pages that could be specified was different. However, pages not contained in the previous dataset were also selected as spoilers with high scores, indicating that the previous spoiler dataset needed to be revised. The reasons for the increase in the number of pages selected as spoilers could be due to the rise in the number of selectable spoilers or because the spoilers' criteria were changed, which allowed readers to choose spoilers for those who had read up to the first episode, rather than spoilers for the story as a whole.

The results of each comic showed that in "All You Need Is Kill," spoilers were selected for pages that revealed the core content of the story, which had not been obtained in the previous spoiler dataset. This suggests that the spoiler influence experiment in the previous study did not sufficiently spoil the story and may not have been able to measure the influence of spoilers. Also, in "GIANT KILLING," only the result of the last match was selected as a spoiler in the previous spoiler dataset. In contrast, the results of this study suggest that the earlier matches and goal scenes were also chosen as spoilers, so it is difficult to say that the influence of spoilers was measured adequately, which means the spoiler impacts were not measured sufficiently. In the case of "Prophecy," spoilers for the punchline were also selected. The results of this study suggest that it is difficult to say that the effect of spoilers on the results was measured. However, in the case of "ReRe Hello," only a few changes in the pages were spoiled compared to the previous spoiler dataset. Even the pages that were rated highly were pages that had little to do with the main character's story, so it can be said that the effect of spoilers was sufficiently measured.

This result indicates that spoilers do not only occur at the end of the story but can exist in a wide range of stories and that more spoilers may be needed to reduce the interest in the work. In other words, spoiler influence studies that present only a few parts of the story as a spoiler may not be called spoiler influence in the true sense, especially since they show scenes from the latter half of the story. In the future, we plan to conduct experiments on the influence of spoilers in comics based on the definition of spoilers and the spoiler dataset we have created.

In addition, we asked five participants to read each comic and define whether there were spoiler pages or not in this spoiler dataset construction. We consider five participants as the minimum number required to define a spoiler. On

the other hand, if the number of annotators is increased, all pages will probably be registered in the dataset as spoilers. Therefore, we plan to further study the appropriate amount of spoilers in the dataset in the future.

5. Preliminary Investigation of the Automatic Detection of Spoiler Page

Next, we will consider the metadata necessary to implement the automatic assessment system of spoilers. Here, various methods exist for OCR (Optical Character Recognition). Among these, we will consider using the Mokuro [25], which allows users to retrieve images from e-comics and obtain text information through OCR analysis.

Fig 7-10 showed a line graph plotting the number of characters in the comics for every two pages for all the comics
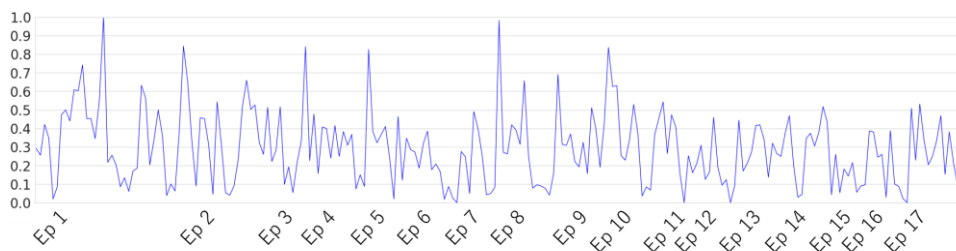

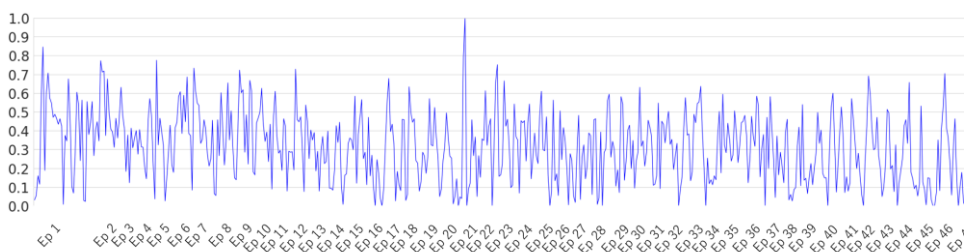Fig. 7. The results of OCR using the Mokuro in comic "All You Need Is Kill."

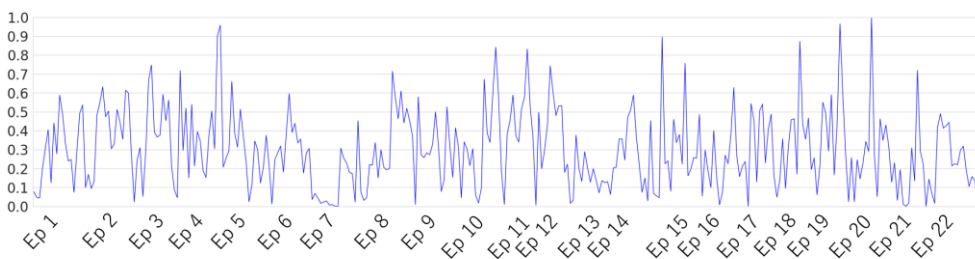Fig. 8. The results of OCR using the Mokuro in comic "GIANT KILLING."

Fig. 9. The results of OCR using the Mokuro in comic "Prophecy."
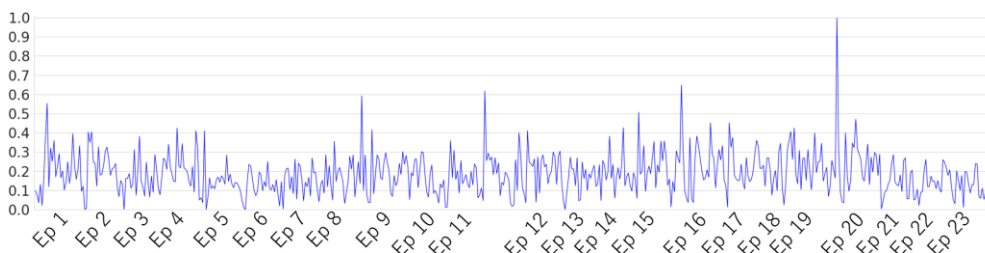

Fig. 10. The results of OCR using the Mokuro in comic "ReRe Hello."

used in the above dataset construction. Here, the graph's horizontal axis is the number of episodes in each comic, and the vertical axis is the normalized number of characters. We decided on the number of characters for every two pages because the images available in the Mokuro system were based on a single-page spread of a comic strip, so the minimum number of pages that could be analyzed was two.

From these figures, it can be seen that in the battle genre "All You Need Is Kill" and in the sports genre "GIANT KILLING," the episodes with a high degree of spoilers tended to have a low value on characters. Especially in "GIANT KILLING," all the episodes with a high degree of spoilers, which were episodes 16-17, 20-21, 27, and 46-47, had a character value lower than 0.25. This could be because the characteristic scenes of spoilers in sports genre comics often include descriptions that can be understood at a glance, such as depictions of scoring goals or the scores of match results. The same can be said for battle genre comics since their characteristic scenes of spoilers tend to be spoiled from comics' images rather than texts.

On the other hand, "Prophecy" in the suspense genre and "ReRe Hello" in the romance genre, the episodes with high spoilers tended to have some low value on characters, while mainly had characters values above average. Especially in "Prophecy," episodes 7 and 21 had a value lower than 0.2, while episodes 15 and 18 had a higher than 0.7. This could be because the main characteristic scene of spoilers in suspense genre comics is likely to be when the criminal or mystery is revealed, and these scenes could be depicted with both images and texts. Therefore, the episodes with a high degree of spoilers lead to a high value on characters. The same can be said for romance genre comics because their characteristic scenes of spoilers will mainly be a scene in which one's love is expressed, and this scene is likely to be depicted with both images and texts.

This suggests that many of the spoilers in "All You Need Is Kill" and "GIANT KILLING" had a high value for characters and, thus, are considered helpful to some extent in determining spoilers. Similarly, for "Prophecy" and "ReRe Hello," it is thought that spoilers can sometimes be judged by focusing on characters of both high value and low value.

In the future, we will use these results to refine the spoiler page assessment further. However, it is hardly possible to say that the method we used in the present study, which focuses on the value of characters for every two pages, can handle all spoiler patterns. Therefore, in the future, we are considering the value of characters for every episode to identify which episodes can be a high degree of spoilers, including which part of the episode can be detected as spoilers. Furthermore, we will use machine learning to validate the accuracy of the assessment. In addition, a unique characteristic in comics is onomatopoeia, which has been studied to analyze comics [26]. Thus, we believe that onomatopoeia can be used to detect spoilers scenes, and we are planning to use the COO [27] to recognize onomatopoeia automatically in comics.

6. Conclusion

In this study, based on the results of previous studies, we considered and defined which parts of spoilers for comic readers, what kind of spoilers can be used for testing the effects of spoilers, and how to collect more objective spoilers. To select spoilers with a high degree of objectivity, we defined the definition as "a page p in episode $N+1$ that would be objectionable to others who have read up to episode N of the comic when it is presented." Furthermore, we reconstructed the spoiler dataset of spoiler pages according to the definition and compared the dataset created at this time with the previous one. Then, we found that the newly constructed dataset not only covered the previous dataset but also detected important spoiler scenes that were not included in the previous dataset, such as goal scenes and scenes that reveal the killer's motive and so on, as spoilers.

In the future, we plan to verify the influence of these reconstructed spoilers through experiments. In addition, we plan to automate the spoiler page judge by using the spoiler dataset constructed this time. For that, the frame information of the parts where the spoiler scene has occurred will be acquired, and the page frame information will be used to determine the spoiler.

Acknowledgment

References

[1] Leavitt, J. D. and Christenfeld. N. J. S. (2011) "Story Spoilers Don't Spoil Stories." *Psychological Science*, **22 (9)**: 1152–1154.

[2] Levine, William. H., Betzner, Michelle. and Kevin, S. (2016) "The Effect of Spoilers on the Enjoyment of Short Stories." *Journal of Discourse Processes* **53**: 513–531.

[3] Hassoun, D. (2013) "Sequential outliers: the role of spoilers in comic book reading." *Journal of Graphic Novels and Comics*, **4**: 346–358.

[4] Maki, Y., Shiratori, Y., Sato, K. and Nakamura, S. (2018) "A Consideration to Estimate Spoiling Pages in Comics" *International Symposium on Affective Science and Engineering ISASE2018* **4:** 1–6.

[5] Maki, Y. and Nakamura, S. (2017) "Do Manga Spoilers Spoil Manga?" *The Sixth Asian Conference on Information Systems*: 258–262.

[6] Tsang. A. S. L. and Yan, D. (2009) "Reducing the Spoiler Effect in Experiential Consumption." *Advances in Consumer Research*, **36**: 708–709.

[7] Yan, D. and Tsang, A. S. L. (2016) "The misforecasted spoiler effect: Underlying mechanism and boundary conditions." *Journal of Consumer Psychology* **26(1)**: 81–90.

[8] Rosenbaum, J. E. and Johnson, B. K, (2016) "Who's afraid of spoilers? Need for cognition, need for affect, and narrative selection and enjoyment." *Psychology of Popular Media Culture* **5(3)**: 273–289.

[9] Johnson, B. K. and Rosenbaum, J. E, (2018) "(Don't) Tell Me How It Ends: Spoilers, Enjoyment, and Involvement in Television and Film." *Media Psychology* **21(4)**: 582–612.

[10] Johnson, B. K. and Rosenbaum, J. E, (2015) "Spoiler alert: Consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation." *Communication Research* **42(8)**: 1068–1088.

[11] Johnson, B. K., Udvardi, A., Eden, A. and Rosenbaum, J. E, (2020) "Spoilers go bump in the night: Impacts of minor and major reveals on horror film enjoyment." *Journal of Media Psychology* **32(1)**: 14-25.

[12] Yuji, S. Maki, Y. Satoshi, N. and Takanori, K. (2018) "Detection of Football Spoilers on Twitter" *The 10th International Conference on Collaboration Technologies (CollabTech 2018)*: 129–141.

[13] Golbeck, J. (2012) "The Twitter Mute Button, A Web Filtering Challenge." *Proc.2012 ACM Annual Conference on Human Factors in Computing Systems*.

[14] Jeon, S., Kim, S. and Yu, H. (2013) "Don't Be Spoiled by Your Friends: Spoiler Detection in TV Program Tweets." *Proceedings of the 7th International Conference on Weblogs and Social Media (ICWSM)*.

[15] Guo, S. and Ramakrishnan, N. (2010) "Finding the storyteller: automatic spoiler tagging using linguistic cues." *Proceedings of the 23rd International Conference on Computational Linguistics*.

[16] Boyd-Graber, J., Glasgow, K. and Zajac, J. S. (2013) "Spoiler Alert: Machine Learning Approaches to Detect Social Media Posts with Revelatory Information." *Proceedings of the 76th Annual Meeting of the American Society for Information Science and Technology (ASIST)* **45**.

[17] Maeda, K., Hijikata, Y. and Nakamura, S. (2016) "A Basic Study on Spoiler Detection from Review Comments Using Story Documents" *Web Intelligence (WI)*: 272–277.

[18] Ikeda, K., Hijikata, Y. and Nishida, S. (2010) "Proposal of deleting plots from the reviews to the items with stories" *In: Proceedings of SNSMW* **6193**: 346-352.

[19] Pang, B. and Lee, L. (2004) "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts." *Proc. of ACL'04*: 271–278.

[20] Matsui, Y., Ito, K., Aramaki, Yuji., Fujimoto, A., Ogawa, T., Yamasaki, T. and Aizawa, K. (2017) "Sketch-based manga retrieval using manga109 dataset" *Multimedia Tools and Applications* **76(20)**: 21811–21838.

[21] Sakurai, T., Ito, R., Abe, K. and Nakamura, S. (2022) "A Method to Annotate Who Speaks a Text Line in Manga and Speaker-Line Dataset for Manga109" *The 5th International Workshop on coMics Analysis, Processing and Understanding*.

[22] Ogawa, T., Otsubo, A., Narita, R., Matsui, Y., Yamasaki, T. and Aizawa, K. (2018) "Object Detection for Comics using Manga109 Annotations" *Computer Vision and Pattern Recognition* (arXiv preprint arXiv: 1803.08670).

[23] Tsubota, K., Ogawa, T., Yamasaki, T. and Aizawa, K. (2018) "Adaptation of manga face representation for accurate clustering" *SA '18 SIGGRAPH Asia 2018* (**15**).

[24] Daiku, Y., Iwara, M., Augereau, O. and Kise, K. (2018) "Comics Story Representation System Based on Genre, Document Analysis Systems (DAS)" *13th IAPR International Workshop*: 257–262.

[25] kha-white. Mokuro, URL : https://github.com/kha-white/mokuro, 2022.

[26] Sugita, R., Okajima, M., Komatsu, T. and Nakamura, S (2022) "Considering the Meanings and Effects of Frames without Onomatopoeias in Japanese Comics" *The 5th International Workshop on coMics Analysis, Processing and Understanding*.

[27] Baek, J.H., Matsui, Y. and Aizawa, K. (2022) "COO: Comic Onomatopoeia Dataset for Recognizing Arbitrary or Truncated Texts" *European Conference on Computer Vision (ECCV)* **arXiv: 2207.04675**.