ABSTRACT
Seeing the final score of a sports match on the Web often spoils the pleasure of a user who is waiting to watch a recording of this match on TV. This paper proposes four information clouding methods to block spoiling information, and describes implementation of a system using these methods as a browser extension. We then experimentally investigate the usefulness of the methods, taking into account their differences, differences in the variety of content, and differences in the user’s interest in sports.

Categories and Subject Descriptors
D.3.3 [Information Interfaces and Presentation (e.g., HCI)]: Hypertext/Hypermedia – user issues.

General Terms
Design, Human Factors.

Keywords
Spoilers, information clouding, anti-spoiler, Web, browser.

1. INTRODUCTION
Seeing the final score of a sports match often spoils the pleasure of a user who is waiting to watch a recorded match on TV or a delayed broadcast of the match. For example, a user who cannot watch a football match due to work commitments and who is planning to watch the match after work usually tries to avoid seeing the final score before watching the recorded match on TV. In these situations, people want to control their access to such content so as to prevent themselves from inadvertently seeing the result before they have watched the match. Unfortunately, people sometimes see or hear this information by accident while browsing the Web to check daily news, weblogs, and other information-providing services. We believe that people sometimes think, “I want to see it later, but I don’t want to see it now!”

There are various situations when spoiling information may be encountered on the Web. For example, a linked Web page might contain a spoiler such as the final score of a match, but the linked text does not mention that the linked page contains this information.

The purpose of this work is to realize an anti-spoiling system that enables users to avoid such spoilers easily. This is done through the dynamic use of clouding information on the Web, depending on what users are planning to enjoy. The system blocks content that would spoil the user’s enjoyment of a previously unwatched sports match. Specifically, we propose four information clouding methods to prevent spoilers about sports matches appearing on a Web page, and we conduct an experiment to investigate the effectiveness of each method.

2. ANTI-SPOILING METHODS
In order to detect spoiling news, an anti-spoiler system requires the utilization of certain keywords, such as players’ names, team names, and nicknames, and combines these with pattern matching about sports results. In order to detect such information, we used a sports database and a sports result dictionary [3]. We also prepared a dictionary that could be used to cloud information by inverting the result of a sports match. The dictionary groups several expressions such as {win, tie, lose, ...}, {won, tied, lost, ...}, {beat, upset, ...}, and {smile, laugh, cry, angry, feel blue, ...} and so on.

Our system automatically generates regular expressions to detect spoiling information after a user specifies a match. In addition, the system observes the user’s browsing habits and automatically generates a DOM tree of a Web page, and uses them to check whether the DOM contains spoiling information.

We proposed four information clouding methods to prevent spoilers, as follows (see Figure 1):

- **Invisible method**: This method just hides the spoiling information by setting its visibility property to hidden.
- **Censored method**: This method sets the background color of spoiling information to be the same as its text. A user can check the censored information by selecting it.
- **Increasing method**: This method substitutes incorrect information for the spoiling information by using the sports results inverting dictionary. For example, if the original
spoiling information is "Japan beats the USA," the method inserts "Japan loses to the USA" and "Japan ties with the USA" and so on.

- **Inverse method**: This method converts the spoiling information to ambiguous information randomly by using the sports results inverting dictionary. For example, the method converts "Japan 2-2 USA" into "Japan 2-4 USA" and "Japan won women’s world cup 2011" into "Japan lost women’s world cup 2011."

We developed a prototype system as a browser extension of Mozilla Firefox by using JavaScript and XUL. The system works as a sidebar of the Web browser. The browser extension also contains a sports database, the sports result dictionary and the result inverting dictionary. The database and dictionaries are updated when the system is updated.

In our system, a user specifies the target teams of a sports match from a list generated from the database. The starting date of the match is used for this because there is no news about the match before starting the match. The user also can add any other team names or player names manually on the sidebar of the browser with a starting date if they are not in the database.

Figure 2 shows a screenshot of our system. Our system appears on the left side bar of this screenshot. In this example, a user set "Mavericks vs. Nuggets" as a target match and set the censored method for information clouding, and was browsing a top page of USA Today. We can see that part of the Web page was censored by our information clouding method. Users can change the information clouding method by selecting the target radio button on the side bar.

3. EXPERIMENT

3.1 Overview
We conducted an experiment to investigate the effectiveness and the participants’ subjective impressions of the four proposed information clouding methods described above. Fifty-eight undergraduate and graduate students participated (40 men and 18 women; 21 - 27 years old).

1. The system gave the following instructions to the participants. "Imagine there is a big sports game that will be held during your working hours, so you are trying to avoid seeing spoiler information about this game. Suddenly, your boss asks you to check the breaking news on the Web. Therefore, you have to check this news while still avoiding the spoiler information."

2. Then, the system presented a screenshot, applying one of the four information clouding methods, and the participants were asked to say what the breaking news was. They were then asked to answer seven questions using a 7-point Likert scale about the presented information clouding method (Table 1).

3. After experiencing all four methods, the participants were asked to fill in a questionnaire about their degree of interest in sports using a 7-point Likert scale (No interest: 1 point, Very interested: 7 points).

3.2 Evaluation
In this experiment, we prepared four different sports news items about the Japanese men’s national football team, the Japanese women’s national football team, the Japanese baseball team, and Kei Nishikori who is a famous Japanese male tennis player.

There were three independent variables in this experiment; differences in information clouding methods (method factor), differences in the types of sports (sports factor), and the participants’ degree of interest in watching sports matches (preference factor). Regarding the preference factor, the participants were divided into two groups (sporty/non-sporty) according to the number of points they assigned when answering the questions about the degree of interest in sports; that is, the 27 participants who assigned fewer points (1 to 3 points) were the non-sporty group, and the 31 participants who assigned more points (4 to 7 points) were the sporty group. Also, there were seven dependent variables; seven questions in the questionnaire about the participants’ subjective impressions.

Specifically, the method and sport factors were analyzed by combining the preference factor. In addition, all the participants correctly identified in all sessions what the breaking news was, so no participants were excluded from the evaluation process.
3.3 Results

3.3.1 Method factor × Preference factor

The seven dependent variables in the questionnaire were analyzed by a 4 (method factor) x 2 (preference factor) mixed ANOVA (Figure 4), and the results showed that there were significant differences in the method factor in all seven questions. We then confirmed that the evaluations of our proposed information clouding methods ranged widely, and that the evaluations of invisible and censored methods received higher evaluations than those of the increasing and inverse methods.

In questions 1 and 2, there were significant and marginally significant differences in the interaction effects between method and preference factors. This means that the participants’ degree of interest in sports clearly affected their evaluations of the information clouding methods. In particular, the participants in the non-sporty group preferred the censored method while those in the sporty group preferred the invisible method. In question 6, there was a significant difference in regard to the preference factor, which means that the participants in the sporty group assigned a higher evaluation to the anti-spoiler system.

Next, to confirm the effect of the types of the sports matches on the above analysis, we carried out the same analysis on all types of sports, and we then observed the following phenomena:

- Men’s football: There was no significant difference regarding interaction effects in questions 1 and 2.
- Women’s football: Similar tendency to overall trend.
- Baseball: There was a significant difference regarding the preference factor in question 3. In particular, the participants in the sporty group gave higher evaluations.
- Tennis: There were significant differences regarding the method factor in questions 3 and 7.

From these results, we confirmed that the evaluations of men’s football were different to evaluations of other sports, especially in questions 1 and 2, and those for tennis showed smaller effects regarding the method factor compared to the other methods.

3.3.2 Sports factor × Preference factor

Seven dependent variables in the questionnaire were also analyzed by a 4 (sports factor) x 2 (preference factor) mixed ANOVA (Figure 5). The results showed that there was a significant difference in regard to the preference factor in question 6, so we confirmed that that the participants in the sporty group
assigned a higher evaluation to the anti-spoiler system than those in the non-sporty group. In questions 1, 2, 5, and 7, there were marginally significant differences in regard to the sports factor, and in particular, the evaluations of men’s football were lower than the evaluation for the other sports. Moreover, in question 7, there was a marginally significant difference in the interaction effect, so the participants in the sporty group assigned higher evaluations for baseball and lower ones for tennis. Next, to confirm the effect of the types of information clouding method on this analysis, we carried out the same analysis using all the information clouding methods, and we then confirmed the following phenomena:

- **Invisible method:** There were significant differences regarding the preference factor in questions 1, 2, and 6.
- **Censored method:** There were significant differences regarding the sports factor in questions 1 and 5.
- **Increasing method:** There was a significant difference regarding the interaction effect in question 1.
- **Inverse method:** There was a significant difference regarding the sports factor in question 3.

From these results, we confirmed that the evaluations of the invisible and censored methods were affected by the different types of sports, while the evaluations of increasing and inverse methods were not.

### 4. DISCUSSION AND CONCLUSION

Recently, Leavitt et al. [2] reported that story spoilers do not spoil stories, but instead they help the users when reading. However, our study suggests that the sports spoiler is not same as the story spoiler; that is, the experimental results showed that the sporty group had much more need of the anti-spoiling system than the non-sporty group. We can say that people who like sports feel that spoiling information is a daily problem. This result showed that there is a general need for our system. We believe "Sports Spoilers do Spoil Sports!"

Those in the non-sporty group felt websites about men’s football could be viewed more safely than tennis sites because spoiling information was prevented by the information clouding method. The reason for this result is that men’s football has become very popular in Japan over the last fourteen years since the Japan men’s national football team made their first appearance in the FIFA world cup in 1998. Therefore, participants in the non-sporty group might be interested in men’s football as a special sports event and they might feel it would be useful at critical moments.

In regard to tennis, sporty participants considered the spoilers were less successful at concealing results compared to the non-sporty participants in all information clouding methods. On the other hand, the sporty group participants misunderstood the results more, compared to the non-sporty group participants. The sporty group participants often guessed the wrong result of a sports match when reading information clouded content. This means that our anti-spoiling system succeeded in confounding sporty participants’ predictions of the final score, and our information clouding methods work well.

We found that the invisible and censored methods were evaluated more highly than the increasing and inverse methods. One reason for this evaluation was that participants found it difficult to judge whether information was trustworthy or not in the increasing and inverse methods. This feeling of insecurity might have affected the participants’ evaluations. Here, if users use the inverse method or increasing method for daily sports such as baseball, they may not feel comfortable using the anti-spoiling system because they always encounter incorrect information. On the other hand, if users use the increasing or inverse methods for critical sports matches such as the FIFA world cup final, our system may work well. In the future, we will make a comparison of information clouding methods based on the novelty of a sports match.

We used our system during the 2010 FIFA world cup and felt no stress because the system worked very quickly. In fact, our system could prevent spoiling information appearing on most Web pages within one second. Our system could detect 102 of 108 items of spoiling information published in three major Japanese Web news portals. In addition, our system detected 142 items as spoiling information even though the total number of spoiling information items was 108. We believe that an anti-spoiling system must take care of recall of detecting spoiling items because people strongly think to avoid them. In the future, we are planning to increase the precision of our system by using machine learning methods.

In daily usage of our system, by using the invisible and censored methods, we were sometimes worried about the prevented content. This means that the information prevented by the invisible and censored methods increases the cognitive load. In particular, in the invisible method, users have no means of obtaining prevented information. The increasing method sometimes destroys the structure of a Web page because this method simply inserts fake DOM objects before/after the DOM object of spoiling information. Users sometimes get a surprise to see an incorrect big upset result when using the inverse method. On the other hand, when users have an explicit intention to avoid spoiling information, the inverse method works well. In the future, we are planning to monitor user’s eye movements while using information clouding methods and to analyze the effects of information clouding.

In this work, we only applied our information clouding methods to sports matches. Ikeda et al., [1] researched how to judge whether a movie review contains a plot spoiler by using a machine learning method in order to remove the plot from the review. We can apply our methods not only to spoilers of sports matches but also to spoilers of movies, books, and TV games on review sites. In the future, we are planning to apply our system to such content.

### 5. ACKNOWLEDGMENTS

This work was supported by Grant-in-Aid for Young Scientists (A) (#23680006) from JSPS.

### 6. REFERENCES

