A Japanese Software Keyboard Using Visualization to Promote Positive Text Input

Tomoya Yumishita and Hiroshi Hosobe[0000-0002-7975-052X]

Faculty of Computer and Information Sciences,
Hosei University, Tokyo, Japan
hosobe@acm.org

Abstract. Many people use smartphones as tools for social media and personal communication. However, the use of smartphones may cause other people's antipathy and flaming, which is often due to the contents of input texts. This paper presents a Japanese software keyboard that uses visualization to promote positive text input. By visualizing the impressions of words, it allows the user to easily input texts with positive impressions. Although its layout and operation are the same as those of normal software keyboards used by many Japanese people, its prediction conversion bar shows colored words to express plausible positive or negative impressions that the words convey. In addition, if the user presses and holds onto words, their synonyms will be displayed, which allows converting negative words to positive ones keeping meanings. This paper presents the results of the experiment conducted to evaluate its design. In addition, to prove its usefulness, this paper presents the results of the experiment on its comparison with a normal keyboard.

Keywords: Text input · Software keyboard · Visualization.

1 Introduction

With the spread of smartphones, people using social media and personal communication tools to exchange information online are increasing. For example, by using X (formerly called Twitter), people convey their opinions to many others. However, depending on how they express their opinions, bad impressions may offend recipients and cause flame wars.

Onishi et al. [10] have demonstrated that the possibility of a flame war on social media depends on how the original post is expressed. They collected and analyzed 100 past cases of posts that caused flame wars on Twitter. The results showed that the posts that had caused flame wars were more likely to contain offensive words than regular posts. Thus words with offensive meanings have the potential to offend recipients and cause flame wars.

In this paper, we propose a Japanese software keyboard that uses visualization to promote positive text input. By visualizing the impressions of words, it allows the user to easily input texts with positive impressions. Although its layout and operation are the same as those of normal software keyboards used

This is the author's version. The final authenticated version is available online at
https://doi.org/10.1007/978-3-031-93845-0_10.

by many Japanese people, its prediction conversion bar shows colored words to express plausible positive or negative impressions that the words convey. In addition, if the user presses and holds onto words, their synonyms will be displayed, which allows converting negative words to positive ones keeping meanings. Furthermore, the colors of the confirmed words are retained later. We also present the results of the experiment conducted to evaluate its design. In addition, to prove its usefulness, we present the results of the experiment on its comparison with a normal keyboard, and the results show that the proposed keyboard promoted positive text input.

The rest of this paper is organized as follows. Section 2 presents previous research related to our work. Section 3 proposes our method, and Section 4 describes its implementation. Section 5 presents a preliminary experiment, and Section 6 provides a main experiment. After Section 7 discusses our method, Section 8 gives conclusions and future work.

2 Related Work

There has been much research on text input interfaces. Masui [7] developed a Japanese predictive conversion system called POBox, which contributed to improving the speed of text input on mobile devices. Currently, such predictive conversion systems are installed and widely used in many smartphones.

There has been research on the automatic detection of flaming and flame wars on social media including Twitter. After analyzing the cases of the flame wars mentioned in Section 1, Onishi et al. [10] developed a system that judged the possibility of a flame war for an input text and pointed out the words that might cause it. Although their analysis results showed that the flame war cases contained many characteristic expressions that might offend others, the system did not obtain good evaluation results from their experiment. Iwasaki et al. [3] proposed methods for identifying potential posts for flaming on Twitter. Considering public opinions of Twitter users, they categorized flaming into criminal episodes, struggles between conflicting values, and secret exposures, and conducted experiments on the identification of posts in the first two categories. Kawakami and Iyatomi [5] proposed a system for detecting flaming on Twitter. It first reduced vocabularies by substituting similar words with representatives by using Word2Vec, and classified vector-represented sentences with a back-propagation neural network. Matsumoto et al. [8] constructed a system for judging harmful expressions on social media. They trained a distributed representation model for harmful expressions, and created a classifier based on a support vector machine.

Recently, research on text input interfaces equipped with special systems also has progressed. Go et al. [2] proposed a software keyboard for smartphones that incorporated a reframing technique. In addition to the commonly used flick input, it had two predictive conversion bars, one for normal predicted words displayed in the lower row, and the other for converted positive words displayed in the upper row. Their experiment showed that it had improved the users' happiness. Continuing from this work, Fukasawa et al. [1] developed a positive

This is the author's version. The final authenticated version is available online at
https://doi.org/10.1007/978-3-031-93845-0_10.

reframing dictionary more automatically. Using a natural language processing model GPT-3, they created reframing pairs by generating positive expressions for negative words and phrases. Mizuno and Nakamura [9] verified that changing the fonts of the texts entered into input forms could suppress flame wars. The results showed a correlation between fonts and input contents, suggesting that conveying visual information at the time of a user's input of a text might cause the user to change the content.

There has been research on reframing also in the field of natural language processing. Ziems et al. [12] introduced a task of positive reframing, and constructed a large-scale benchmark called the Positive Psychology Frames to explain the task in terms of six theoretically motivated reframing strategies. They also used it to evaluate state-of-the-art text style transfer models. Xu et al. [11] proposed a positive reframing method for learning disentangled meaning and style representations. They adopted a simple multi-task learning-based model to fuse generation capabilities from pseudo-positive reframing datasets. Jia et al. [4] enhanced the performance of pre-trained language models on positive reframing. They introduced a multi-strategy optimization framework consisting of reinforcement training, decoding improvement, and multi-dimensional re-ranking.

3 Proposed Method

We propose a Japanese software keyboard that allows users to easily recognize the impression of the words that they use while typing texts. As shown in Figure 1(a), the layout of the text input interface is the same as that of currently popular Japanese flick input keyboards, and the key input operation is also the same as that of flick input keyboards. However, our keyboard is different in that it colors each word in its predictive conversion bar. The frame color of each word is expressed in a red-to-blue gradation, and represents the impression of the word. The closer the color is to red, the more positive the impression is; conversely, the closer it is to blue, the more negative the impression is. In addition, as shown in Figure 1(b), by the user's long pressing a predictive conversion candidate, synonyms of the target word are displayed and can be used for conversion. For the word that has been confirmed as the input, the characters themselves are colored in a red-to-blue gradation based on the color previously shown during the predictive conversion.

We expect that the proposed keyboard will reduce the input of negative texts and increase the input of positive texts. This will prevent potential damage caused by the use of social media and personal communication tools. For example, suppose that a user is about to input the sentence "He is sick." When the user types "byo" (which are the initial characters of the Japanese word "byouki" meaning "sick"), conversion predictions will be shown as in Figure 1(a). At this time, the word meaning "sick" is colored light blue, allowing the user to recognize the negative impression of the word "sick." After recognizing it, the user may reconsider and use a more colorless word "hospital" instead of the explicit word "sick," and change the sentence to a roundabout one like "He is

https://doi.org/10.1007/978-3-031-93845-0_10.

(a) (b)

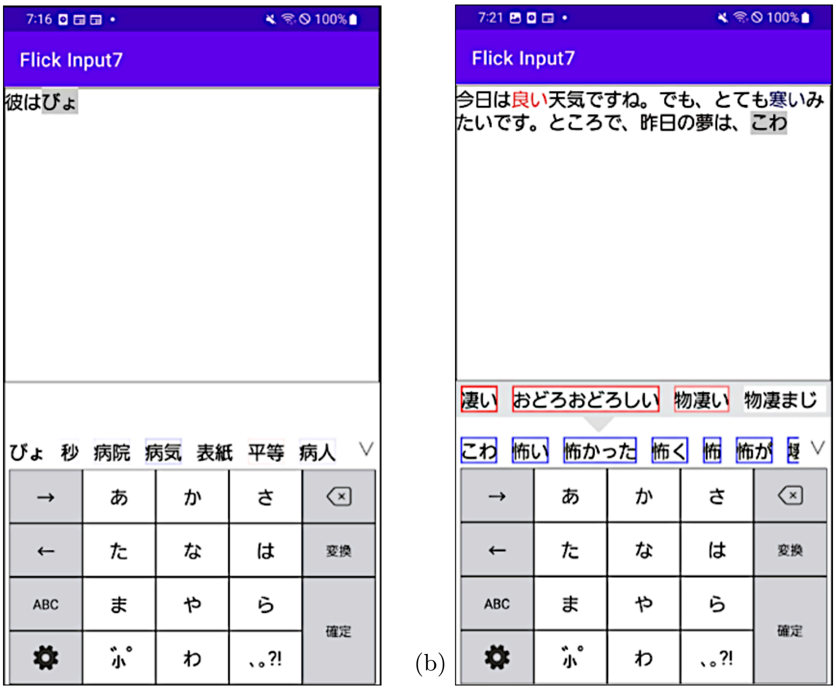


Fig. 1. Proposed Japanese software keyboard.

being treated at the hospital." In addition, the confirmed words are colored as in Figure 1(b), where the word meaning "good" is colored red, and the word meaning "cold" is colored blue. Therefore, the user can recognize the impression of the sentence at a glance even after the user finished the input of the sentence. Thus the user can revise the entire sentence to a more positive one.

An important advantage of the proposed keyboard is that the way of typing is almost the same as that of commonly used Japanese software keyboards. The only difference is that the words in the predictive conversion bar and the confirmed words are colored, which will not significantly change the cognitive load on the user. Therefore, we think that our keyboard can provide users with information about the impressions of words in a semi-unconscious manner while keeping their acceptance.

4 Implementation

Based on the method proposed in the previous section, we implemented a software keyboard in Java and Python. We adopted an Android smartphone Samsung Galaxy M23 5G with a 6.6-inch screen. We implemented a keyboard of ten-key type, which is common for Japanese software keyboards. Also, we implemented a flick input method, in which characters are entered by flicking, and a toggle input method, in which characters are entered by repeatedly tapping the same keys. To display conversion candidates in the predictive conversion bar, we used the kana-kanji conversion service of the Yahoo! Developer Network.

To color the impressions of words, we created an emotional polarity dictionary that quantified the impression of each word. This dictionary assigns a real number ranging from -1 to 1 to each word to express its impression such as the positive and the negative. To create the dictionary, we used Google Cloud's Natural Language API and the Balanced Corpus of Contemporary Written Japanese (BCCWJ) [6]. We used the created dictionary to color words in the predictive conversion bar. It operates as follows. First, each text is divided into words by applying the morphological analysis tool Kuromoji. Then the polarity dictionary is used to assign to each word a numerical value between -1 and 1, and its color is determined according to the numerical value. If there exist multiple words, the word with the largest absolute polarity value is colored.

We also implemented the synonym conversion bar by using Japanese WordNet. It operates as follows. First, when a word displayed in the predictive conversion bar is touched with a finger, Japanese WordNet is used to obtain all synonyms for the word. Then the polarity values of all synonyms are compared, and the top five are displayed. When a word is confirmed in the predictive conversion bar or the synonym conversion bar, the color of the word is determined, and the confirmed word in the text box is colored.

5 Preliminary Experiment

We devised several methods of text coloring. In this section, we present a preliminary experiment on the user evaluation of the text coloring methods. In conducting this experiment, we followed the regulations of Research Ethics Involving Human Subjects enacted by our university.

5.1 Method

We recruited six participants (four males and two females) for this experiment. All of them were university students who were 20.4 years old on average. We asked them to enter texts with different settings of the proposed keyboard and then rate them on a five-point scale.

There were two evaluation items, i.e., coloring locations and color intensities. We asked the participants to evaluate three types of coloring locations, i.e., words themselves, the backgrounds of words, and the frames of words. Figure 2 shows the examples of the coloring locations. We asked the participants to evaluate three types of color intensities, i.e., light, normal, and dark. The participants first evaluated the coloring locations, for which the color intensity was set to normal. Next, they evaluated the color intensities, for which the coloring location was set to the one that they had rated the best.

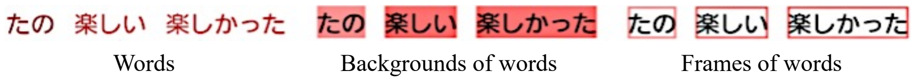


Fig. 2. Locations of coloring.

5.2 Results

The results of the preliminary experiment are shown in Figure 3. While Figure 3(a) shows the results of the coloring locations, Figure 3(b) shows only the results of the color intensities when the frames of words were used as the coloring location, which was evaluated by four participants.

Coloring the frames of words was generally highly rated. The percentage of the participants who gave a lower rating to the backgrounds of words was higher than the other locations. The reason for this is that the coloring of the backgrounds made the words difficult to see. There were opinions that, although coloring words themselves was easy to read, it was difficult to distinguish differences among the visualized impressions of the words.

There was no large difference in the color intensities. In addition, the results show that preferences differed depending on the participants. Although four participants chose the frames of words, the ratings of the color intensities were not very distributed. Therefore, we think that the color intensities are not an important factor.

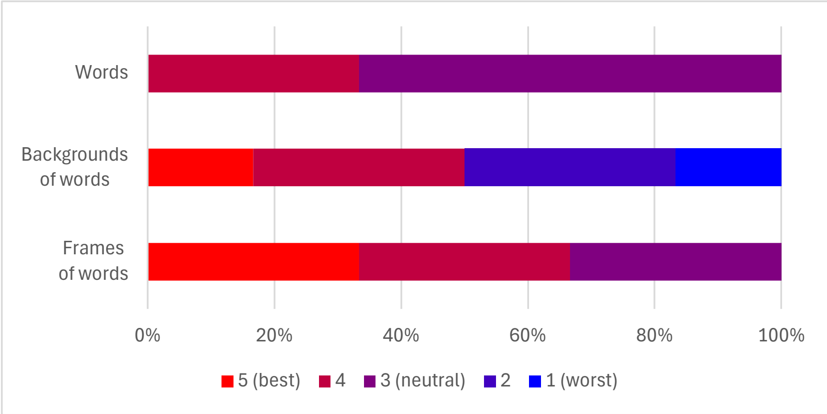
6 Main Experiment

Based on the results of the preliminary experiment in Section 5, we decided to color the frame of words with the normal intensity for our final visualization. We then conducted a main experiment to demonstrate the usefulness of the proposed keyboard. Figure 4 shows how the proposed keyboard was operated during the main experiment. In conducting this experiment, we again followed our university's regulations of Research Ethics Involving Human Subjects.

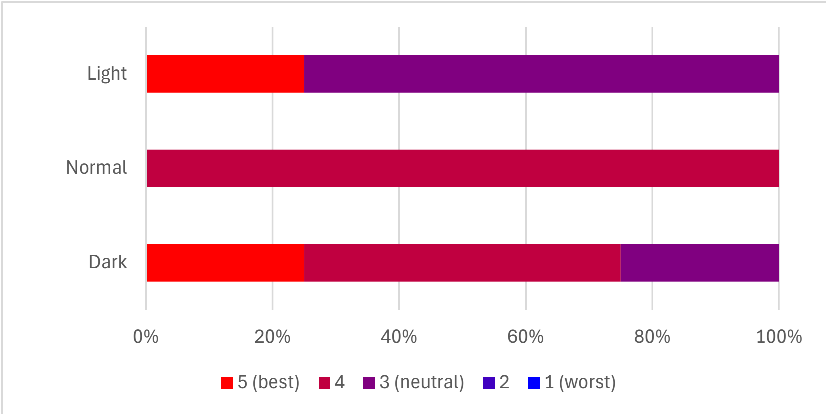
6.1 Method

We recruited six participants (two males and four females) who were university students and 20.5 years old on average. In the experiment, the participants viewed six creative works, such as paintings, comics, and novels, and wrote reviews by using the normal keyboard and the proposed keyboard. At the start of the experiment, we asked them to write sentences including at least two strong points and two weak points. In addition, we informed them that the works that they were viewing had been created by the first author or someone whom they knew, and that their reviews would be conveyed to the creators. This simulated a situation similar to that of regular interactions using communication tools.

This is the author's version. The final authenticated version is available online at
https://doi.org/10.1007/978-3-031-93845-0_10.



(a)



(b)

Fig. 3. Results of the preliminary experiment on (a) the coloring locations and (b) the color intensities.

The reviews were analyzed and compared with two different methods using Google Cloud's Natural Language API. The first method was to directly request the polarities of the entire reviews from the API. The second method was to obtain the polarity of each sentence from the API after manually separating the sentences into strong and weak points.

At the end of the experiment, we asked the participants to give an overall evaluation of the proposed keyboard by answering the following questions:

- Operability: Can the proposed keyboard be operated as easily as normal keyboards?
- Suitability of coloring: Do the colors of words suit their impressions?
- Recognizability of impressions: Can the impressions of words be recognized from their colors?
- Change of typing: Do you feel that the proposed keyboard changed the way of your typing compared to the keyboard that you normally use?

https://doi.org/10.1007/978-3-031-93845-0_10.

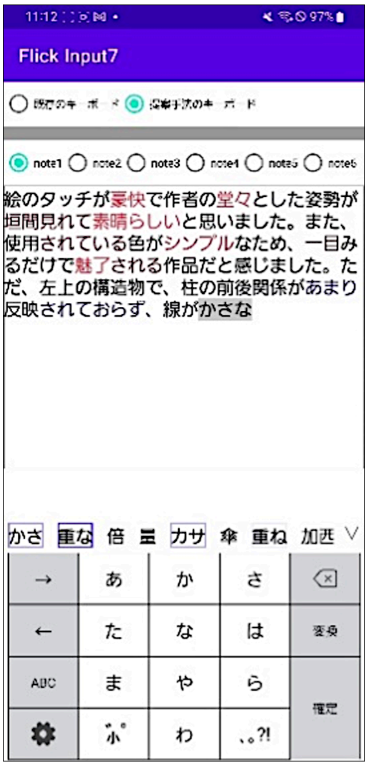


Fig. 4. Operation of the proposed Japanese software keyboard during the main experiment.

6.2 Results of the Reviews

Table 1 shows the mean polarities for the overall reviews of each participant. Four out of the six participants obtained higher polarities with the proposed keyboard than with the normal keyboard. Also, Table 2 shows the mean polarities of the sentences corresponding to the strong and the weak points. For both points, four out of the six participants obtained higher polarities with the proposed keyboard than with the normal keyboard. However, paired t-tests did not indicate significant differences to these results.

6.3 Results of the User Evaluation

The results of the user evaluation are shown in Figure 5. The mean rating for the first question on the operability of the proposed keyboard was 3.2. While some participants commented that it was comparable to the normal keyboard, participants C, D, and F said that it took a long time for predictive conversions to appear.

Table 1. Comparison of the polarities of the reviews.

Participant	Normal keyboard	Proposed keyboard
A	**0.041**	−0.001
B	0.141	**0.269**
C	0.050	**0.071**
D	0.090	**0.235**
E	**0.094**	−0.035
F	0.003	**0.140**

Table 2. Comparison of the polarities of the strong and the weak points.

	Strong points		Weak points	
Participant	Normal kbd.	Proposed kbd.	Normal kbd.	Proposed kbd.
A	0.887	**0.933**	**−0.647**	−0.796
B	**0.957**	0.930	−0.821	**−0.688**
C	**0.926**	0.894	−0.742	**−0.725**
D	0.949	**0.959**	−0.763	**−0.488**
E	0.934	**0.961**	**−0.630**	−0.741
F	0.945	**0.951**	−0.824	**−0.687**

The mean rating for the second question on the suitability of coloring words was 3.3. Participants B, E, and F commented that it was basically suited. On the other hand, participants A and D felt that the meanings of words by themselves were suited, but that their meanings in the actual contexts were not suited and could not be fully grasped.

Notably, the mean rating for the third question on the recognizability of the impression of words was 4.7. The participants answered that the simple red-to-blue gradation allowed them to instantly recognize the impressions of words when they saw them at a glance.

The mean rating for the fourth question on changing the way of typing was 3.5. Participants C and F commented that they were able to input more carefully than usual in such a way that the blue color would not appear. Participant D also said that the synonym conversion function enabled the selection of words with better impressions. On the other hand, participant A, who usually paid attention to phrasing, said that the change was not very noticeable.

Finally, in addition to the four questions, we obtained opinions about the synonym conversion function. Participants D and F said that they rarely used it since it required a long press and took time to show synonyms. After the experiment, we checked the operation logs of each user, and found that participant D used the synonym conversion function five times on average for the review of each work, but that the other participants hardly used it.

https://doi.org/10.1007/978-3-031-93845-0_10.

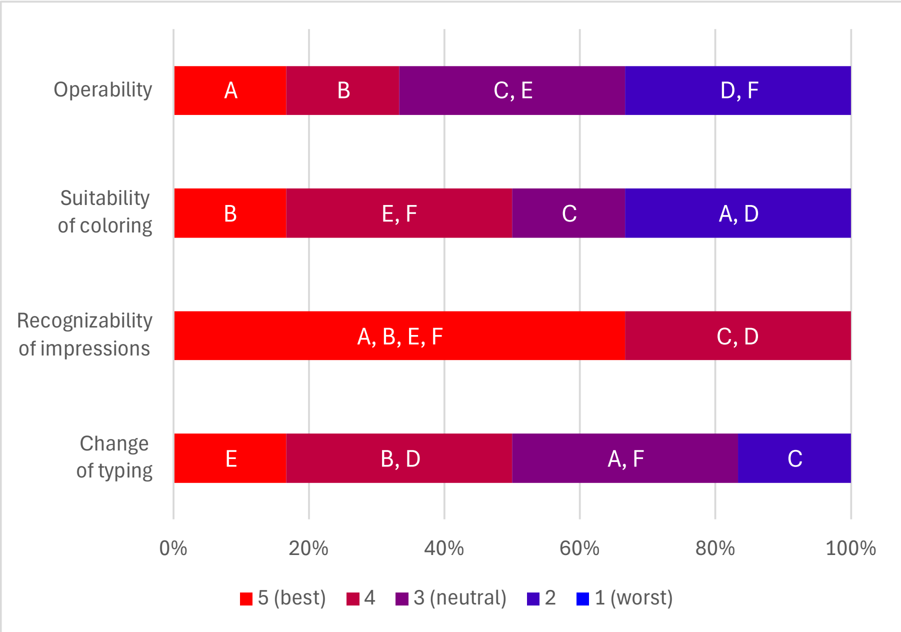


Fig. 5. Results of the user evaluation. Letters A to F indicate the participants who gave the ratings to the corresponding questions.

7 Discussion

We discuss the results of the main experiment below.

7.1 On the Results of the Reviews

In the results of comparing the reviews in Subsection 6.2, four out of the six participants obtained higher polarities when using the proposed keyboard for the overall reviews, the strong points, and the weak points. This indicates that, although the proposed keyboard has the potential to promote positive text input, it is not powerful enough to cause changes in all users.

According to Table 2, participant D increased the impressions of the weak points by 0.275 when using the proposed keyboard. Participant D used the synonym conversion function the most frequently, which we think to be the major factor. From this, we can say that the synonym conversion function can promote positive text input. However, since the other participants than D hardly used it, its implementation needs to be improved.

The maximum of the differences in the polarities of the strong points was 0.046, indicating little change. On the other hand, for the weak points, increases of more than 0.1 were observed from two participants in addition to D. Therefore, we can say that the proposed keyboard is mainly effective when typing negative texts.

https://doi.org/10.1007/978-3-031-93845-0_10.

7.2 On the Results of the User Evaluation

On the first question about the operability of the proposed keyboard, many comments were made about the slow display of the predictive conversion function. Although the delay was partly due to the coloring process, the major factor was the use of the kana-kanji conversion service of the Yahoo! Developer Network to implement predictive conversion. This problem can be solved by implementing the predictive conversion function within the application. It should be noted that no other comments were made about the first question.

On the second question about the suitability of coloring words, differences in the meanings of words in the actual contexts were pointed out. This was due to the implementation focused only on the words being converted. We think that this can be solved by considering the already entered sentences in determining the colors of the words being converted.

The mean rating for the third question on the recognizability of the impressions of words was as high as 4.7. This shows that adding colors to normal keyboards can convey further information to users. Although we added only positive and negative information in this paper, we can say that users may be able to recognize other kinds of information from coloring-based visualization.

From the results of the fourth question on changing the way of typing, one of the reasons why the experimental results did not differ significantly was found to be that we had told the participants that the reviews would be sent to the creators of the works. In addition, since the review targets were paintings, comics, and novels, the input texts were not very stimulating. For example, we expect that differences in input contents would be greater if the review targets were posts that had caused flame wars on social media.

8 Conclusions and Future Work

In this paper, we proposed a Japanese software keyboard using visualization to promote positive text input. The final experiment showed that the percentage of the participants who had entered texts with better impressions was higher when using the proposed keyboard than when using the normal keyboard. Also, we found that coloring texts allowed users to recognize positive and negative information.

In the future, we plan to improve the software keyboard based on the proposed method. Specifically, we will make the synonym conversion function easier to use, make the predictive conversion function display faster, and implement a coloring method that considers contexts. Also, we will reconsider experimental methods and more reliably demonstrate that the proposed keyboard promotes positive text input.

Acknowledgment

This work was partly supported by JSPS KAKENHI Grant Number JP24K14904.

References

1. Fukasawa, H., Go, K., Fukumoto, F., Li, J., Kinoshita, Y.: Creating a positive reframing dictionary using machine learning. In: Proc. HCI Int. CCIS, vol. 1836, pp. 411–417 (2023)
2. Go, K., Moriya, Y., Kinoshita, Y., Li, J., Fukumoto, F.: Happy text entering: Promoting subjective well-being using an input method for presenting positive words and phrases. In: Proc. OzCHI. pp. 153–158. ACM (2021)
3. Iwasaki, Y., Orihara, R., Sei, Y., Nakagawa, H., Tahara, Y., Ohsuga, A.: Identification of flaming and its applications in CGM: Case studies toward ultimate prevention. In: Proc. ICAART. vol. 1, pp. 639–644. SCITEPRESS (2014)
4. Jia, S., Cao, B., Gao, Q., Cao1, J., Liu, B.: Positive text reframing under multi-strategy optimization. In: Proc. COLING. pp. 429–447 (2025)
5. Kawakami, M., Iyatomi, H.: Development of a flaming detection system for Twitter. In: Proc. FSS. pp. 705–708. J-SOFT (2016), in Japanese
6. Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., Koiso, H., Yamaguchi, M., Tanaka, M., Den, Y.: Balanced corpus of contemporary written Japanese. Language Resources and Evaluation **48**, 345–371 (2014)
7. Masui, T.: An efficient text input method for pen-based computers. In: Proc. ACM CHI. pp. 328–335 (1998)
8. Matsumoto, K., Tsuchiya, S., Miyake, T., Yoshida, M., Kita, K.: Flame prediction based on harmful expression judgement using distributed representation. Int. J. Technol. Eng. Stud. **4**(1), 7–15 (2018)
9. Mizuno, H., Nakamura, S.: Proposal and verification of a method for suppressing flame wars by changing fonts in input forms. IPSJ SIG Tech. Rep. **2020-GN-110**(10), 1–7 (2020), in Japanese
10. Onishi, M., Sawai, Y., Komai, M., Sakai, K., Shindo, H.: Building a comprehensive system for preventing flaming on Twitter. In: Proc. JSAI Conf. pp. 3O1-3in:1–4 (2015), in Japanese
11. Xu, S., Fukumoto, F., Li, J., Go, K., Suzuki, Y.: Learning disentangled meaning and style representations for positive text reframing. In: Proc. INLG. pp. 424–430 (2023)
12. Ziems, C., Li, M., Zhang, A., Yang, D.: Inducing positive perspectives with text reframing. In: Proc. ACL. vol. 1, pp. 3682–3700 (2022)

https://doi.org/10.1007/978-3-031-93845-0_10.