

Yahoo! JAPAN News Strengthens Measures to Improve the Integrity of the Comment Section

- Maximizing the use of AI and other technologies, Yahoo! JAPAN will take severer measures against slander, such as stricter "suspension of postings," and will introduce a function that automatically hides the comment section for some articles.

Yahoo! JAPAN News, one of the largest Internet news distribution services in Japan, will strengthen its efforts to improve the soundness of the forum for speech in "Yahoo! JAPAN News Comment" (hereinafter "Comment Section").

Taking advantage of the interactive nature of the Internet, Yahoo! JAPAN News aims to create even more value for information by providing a place for individual users to contribute comments in addition to delivering the high quality information published by media companies. Yahoo! JAPAN's philosophy is to support the actions of the many users who come into contact with these pieces of information, and to contribute to the realization of a better society.

Yahoo! JAPAN News has provided the Comment Section since 2007 based on this philosophy. Many users have shared their diverse opinions with each other in line with the purpose of the Comment Section.



(For illustrative purpose only)

In providing the Comment Section, Yahoo! JAPAN News believes that it is of utmost importance that users can use the section with peace of mind. Yahoo! JAPAN News has a zero-tolerance policy for postings that could be construed as human rights violations or discrimination, and has taken a variety of measures to create a healthy discourse space while respecting the diverse opinions of individual users^{*1}.

However, some of the Comment Sections have not been able to achieve the objectives that are intended by Yahoo! JAPAN News. In response, Yahoo! JAPAN News will be taking stricter measures against slander and defamation, including severer measures such as suspension of postings and the introduction of a function to automatically hide the Comments Section for individual articles.

From before, Yahoo! JAPAN News has prohibited postings that contain slanderous or defamatory content, and has clearly stated specific examples of offensive postings in its comment policy^{*2}, as well as in the comment submission form.

In addition to human patrols by a dedicated team working around the clock, Yahoo! JAPAN News also uses its deep learning-based natural language processing model (AI)^{*3} and an independently developed supercomputer "kukai" that specializes in deep learning processing^{*4}. By making full use of technology, an average of approximately 20,000 slanderous and defamatory rule-violating comments are deleted per day. Additionally, since June 2018, users who have posted multiple slanderous and other offensive comments have been subject to measures that suspend them from posting comments^{*5}.

Meanwhile, as a company engaged in the platform business, which serves as the foundation for information distribution, Yahoo! JAPAN believes it is important to take appropriate measures while carefully maintaining a balance with the freedom of expression of those who post their comments. In addition to strictly addressing defamatory postings from the perspective of human rights protection, Yahoo! JAPAN is working to strengthen its anti-

defamation measures and promote transparency in order to ensure the integrity of the forum for speech on the Internet, with due consideration for freedom of expression^{*6}.

The main initiatives to be strengthened are as follows.

■Change in Wording of AI-Based Warning Message, and Implementation of Stricter “Suspension of Posting Comments”

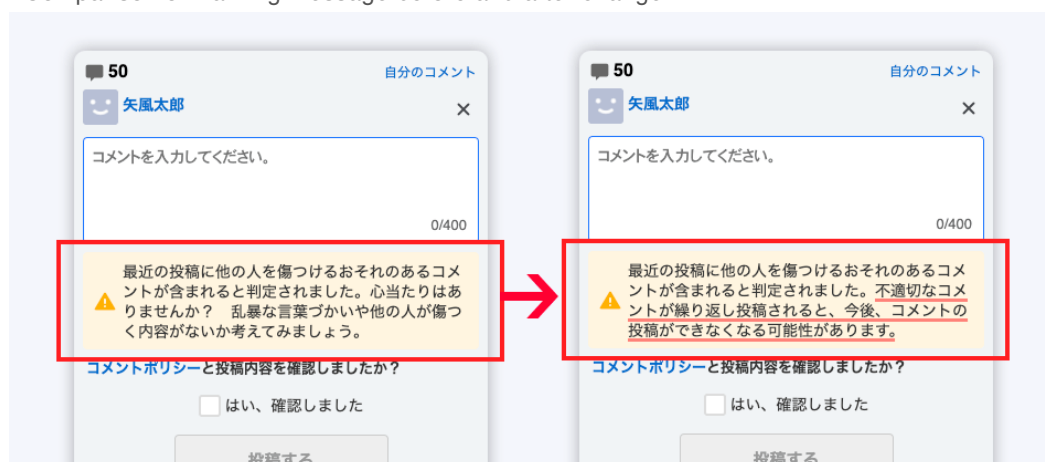
Since July 2020, Yahoo! JAPAN News has been alerting users who have more than once posted comments that are judged by AI to be rule-violating, by showing them a message prompting them to rethink their comments. The number of users receiving this warning decreased by 13.5% in four months from August 2020, immediately after the start of this measure, to December 2020 while the number of users posting comments increased, indicating a certain degree of effectiveness^{*7}.

Yahoo! JAPAN News has strengthened the wording of this cautionary message to include a warning that users may be barred from posting comments (see comparison image below).

Aside from this, since June 2018, Yahoo! JAPAN News has been taking a measure of suspending posting of comments by users who have posted multiple rule-violating comments. The scope of the "suspension of posting" has been expanded to include more stringent measures against malicious users who repeatedly post rule-violating comments such as slander.

For users who have been suspended from posting, a contact point has been set up to file an objection, and the details of the suspension are reviewed.

<Comparison of warning message before and after change>



(For illustrative purpose only)

■Introduction of a Function That Automatically Hides Comment Section for Individual Articles

Starting today, for articles with more than a certain number of posts, the Comment Section will

be automatically hidden based on the number of offending comments and other criteria determined by AI.

In addition to setting objective and non-arbitrary standards for the hiding of applicable Comment Sections, Yahoo! JAPAN News believes that it is necessary to delete individual offending comments and take strict action against offending commenters at the same time.

Going forward, Yahoo! JAPAN News plans to analyze the effectiveness of this measure, review as needed whether it is applied appropriately, and make improvements seeking advice and opinions of experts, users, and other external parties.

*For more details of this function, please refer to:

<https://news.yahoo.co.jp/newshack/information/20211019.html> (Japanese only)

<Image of screen when Comment Section is hidden>



(For illustrative purpose only)

■Display of Warning Message during Lower House Election

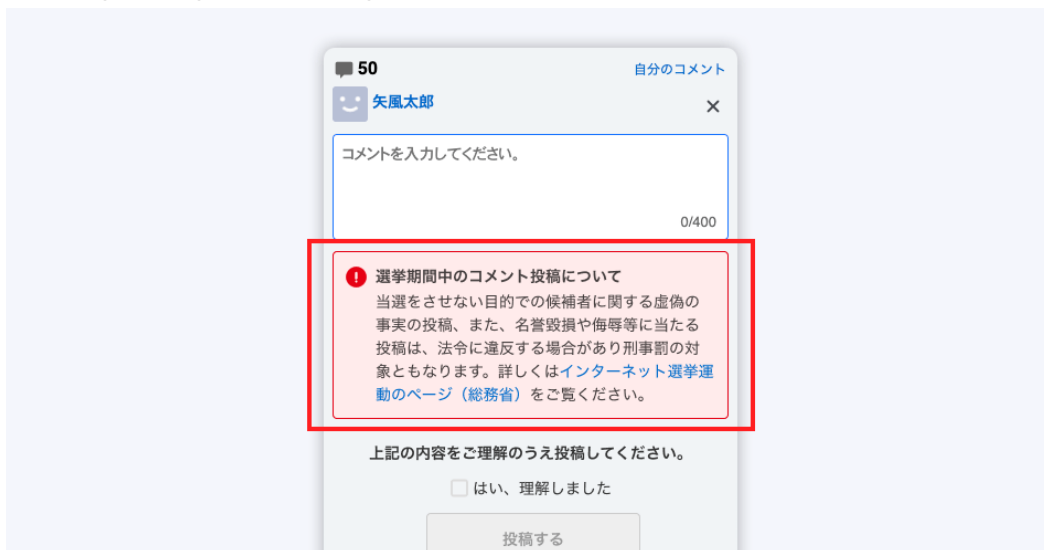
During the campaign period for the 49th Lower House election (October 19–October 31, 2021) in Japan, when users submit comments, a special cautionary message will be displayed in all articles published in the political category of Yahoo! JAPAN News.

The Public Offices Election Act of Japan prohibits acts of publicly announcing false facts or distorted facts concerning a candidate for public office for the purpose of preventing them from being elected. In addition, acts such as defamation and insult are also prohibited by laws^{*8}.

With consideration for respect for diverse opinions and the realization of fair elections, Yahoo! JAPAN News will post cautionary messages when users submit comments. This will

discourage the posting of false information that may interfere with the fairness of the election during the election campaign, or affect the election results.

<Warning message shown during election period>



(For illustrative purpose only)

Yahoo! JAPAN News will continue to implement various other measures in the future, including strengthening of its efforts to prevent defamatory posts through a mechanism to encourage users to reconsider before posting, and the development of AI that can present a variety of opinions. With these measures, Yahoo! JAPAN News will provide users with the opportunity to develop interest in news and multiple perspectives through the diverse views and opinions posted in the Comment Section.

Working with external partner companies and experts, Yahoo! JAPAN News will make its utmost effort to build a sound forum for speech and contribute to the realization of a better society.

*1 [Initiatives Undertaken by Yahoo! JAPAN News Comment \(Summary\)](#) (Japanese only)

*2 [Yahoo! JAPAN News Comment Policy](#) (Revised on March 25, 2021) (Japanese only)

*3 [Countermeasures Taken by Yahoo! JAPAN News against Inappropriate Comments Utilizing Its Large-Scale Deep Learning Model](#) (Yahoo! JAPAN Tech Blog) (Japanese only)

*4 [Yahoo! JAPAN News Strengthens Countermeasures against Inappropriate Comments Using Its Cutting-Edge “Judgment Model \(AI\) Based on Natural Language Processing Using Deep Learning” and Its Proprietary Supercomputer, “kukai”](#) (March 6, 2020) (Japanese only)

*5 [Yahoo! JAPAN News Comment Function: Suspension of Postings for Accounts](#) (June 12, 2018) (Japanese only)

*6 [Expert Panel to Discuss Countermeasures against Defamatory Postings Established in 2020](#) (Japanese only)

*7 [Effects of the AI-Based Warning Message Displayed in Yahoo! JAPAN News](#) (February 16, 2021) (Japanese only)

*8 [Overview of Lifting of the Ban on Internet Election Campaigns \(Act for Partially Revising the Public Offices Election Act\)](#) (Ministry of Internal Affairs and Communications) (Japanese only)

Unless otherwise specified, English-language documents are prepared solely for the convenience of non-Japanese speakers. If there is any inconsistency between the English-language documents and the Japanese-language documents, the Japanese-language documents will prevail.