

Yahoo!ニュース、コメント欄の健全化に向けた取り組みを強化

～ AIなどのテクノロジーを最大活用し、
「投稿停止措置」の厳格化や記事単位でコメント欄を自動的に
非表示とする機能の導入など
これまで以上に誹謗中傷への厳正な対策を講じる ～

ヤフー株式会社（以下、Yahoo! JAPAN）は、日本最大級のインターネットニュース配信サービス「Yahoo!ニュース」が提供する「Yahoo!ニュース コメント」（以下、コメント欄）における言論の場の健全化を目指すため、取り組みを強化します。

「Yahoo!ニュース」は、インターネットが持つ双方向性という特性を生かし、媒体各社の記事による良質な情報発信に加えて、ユーザー個人にも発信の場を提供することで、さらなる情報の価値を創ることを目指しています。そして、その情報に触れた多くのユーザーの行動をサポートし、より良い社会の実現に貢献したいという理念を掲げています。

こうした理念のもとで「Yahoo!ニュース」では2007年より、コメント欄を提供しています。これまで多くのユーザーが、コメント欄の提供目的に即して、多様な意見を共有しあっています。



「Yahoo!ニュース」では、コメント欄の提供にあたって、ユーザーが安心して利用できることがなにより重要であると考えています。人権侵害や差別に当たりうる投稿は一切許容しておらず、ユーザー一人ひとりの多様な意見を尊重しながら、健全な言論空間を創出するためにさまざまな取り組みを行ってきました（※1）。

しかしながら、一部のコメント欄においては、「Yahoo!ニュース」が目指している目的を実現できていないという課題がありました。そこでこのたび、「投稿停止措置」の厳格化や記事単位でコメント欄を自動的に非表示とする機能の導入など、これまで以上に誹謗中傷への厳正な対策を講じます。

「Yahoo!ニュース」では、以前より誹謗中傷などの内容を含む投稿を禁止し、コメントポリシー（※2）に違反投稿の具体例をわかりやすく明示したうえで、コメント投稿フォームにも記載しています。

また、24時間体制の専門チームによる人的なパトロールに加えて、自社で開発した「深層学習を用いた自然言語処理モデル（AI）」（※3）や、独自に開発した深層学習特化型スーパーコンピュータ「kukai」を活用するなど（※4）、テクノロジーを駆使しながら、誹謗中傷をはじめとする違反コメントを1日平均約2万件削除するなどの対策を行っています。加えて、誹謗中傷などの違反コメントを複数にわたって投稿したユーザーに対して、2018年6月より「投稿停止措置」を行っています（※5）。

一方で情報の流通基盤であるプラットフォーム事業を営む企業としては、投稿者の表現の自由とのバランスを慎重に見極めながら適正に対応することも重要であると考えています。Yahoo! JAPANでは、人権救済の観点より誹謗中傷投稿に対して厳正に対処していくとともに、表現の自由にも配慮したインターネット上の言論空間全体の健全化に向けて、誹謗中傷対策の強化や透明化を進めるために取り組んでいます（※6）。

このたび強化するおもな取り組みは以下の通りです。

■AIを活用した投稿時注意メッセージの掲出内容変更と「投稿停止措置」の厳格化

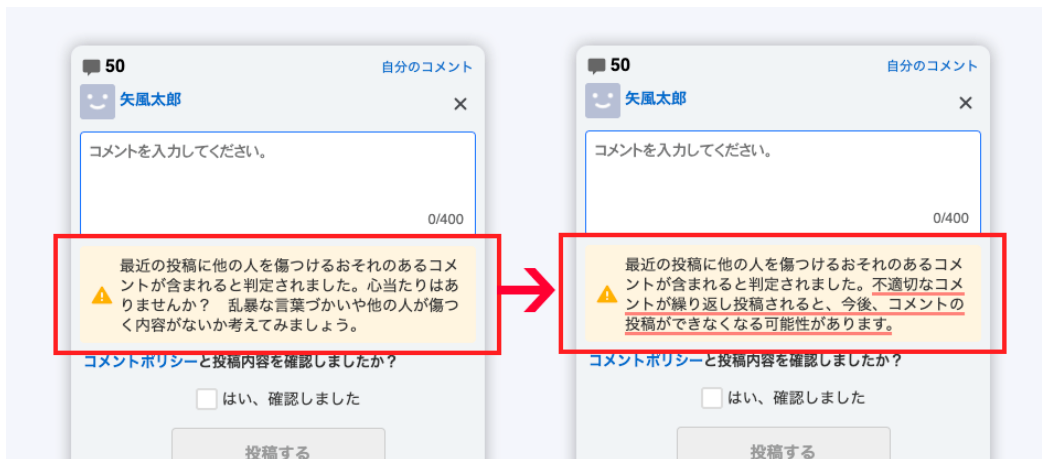
「Yahoo!ニュース」では2020年7月より、AIによって違反と判定されたコメントを複数回投稿しているユーザーに対して注意喚起し、投稿するコメント内容の再考を促すためのメッセージを掲出しています。本取り組み開始直後の同年8月と比較して、4カ月後の同年12月には、コメントを投稿したユーザー数が増加したのに対して、注意メッセージが掲出されたユーザー数が13.5%減少するなど、一定の効果が出ていることがうかがえます（※7）。

この注意メッセージの表現をより強めて、「コメントの投稿ができなくなる可能性がある」と警告する内容を掲出しています（後述の比較画像参照）。

また、「Yahoo!ニュース」は、違反コメントを複数にわたって投稿したユーザーに対して、2018年6月より「投稿停止措置」を行っています。このたび、「投稿停止措置」の対象を拡大し、誹謗中傷などの違反コメント投稿を繰り返す悪質なユーザーへの措置を、これまで以上に厳格化しています。

なお、投稿停止措置を受けたユーザーについては、異議申し立てのための問い合わせ窓口を用意し、措置の内容についてもあらためて確認を行っています。

<注意メッセージの比較>



■記事単位でコメント欄を自動的に非表示とする機能の導入

本日より、一定以上の投稿数のある記事のコメント欄を対象に、AIが判定した違反コメント数などの基準に従い、コメント欄を自動的に非表示とします。

なお、対象となるコメント欄の非表示においては、客観的かつ恣意性のない基準を設けるとともに、個別の違反コメントの削除および違反コメント投稿者に対する厳正な対処も同時に進めていく必要があると考えています。

今後、実施した非表示対応を検証して、適切に行われているか適宜確認をするとともに、有識者やユーザーなど外部の意見を伺いながら、改善を積み重ねていく予定です。

※本機能の詳細について

<コメント欄が非表示になった場合の通知画面（イメージ）>



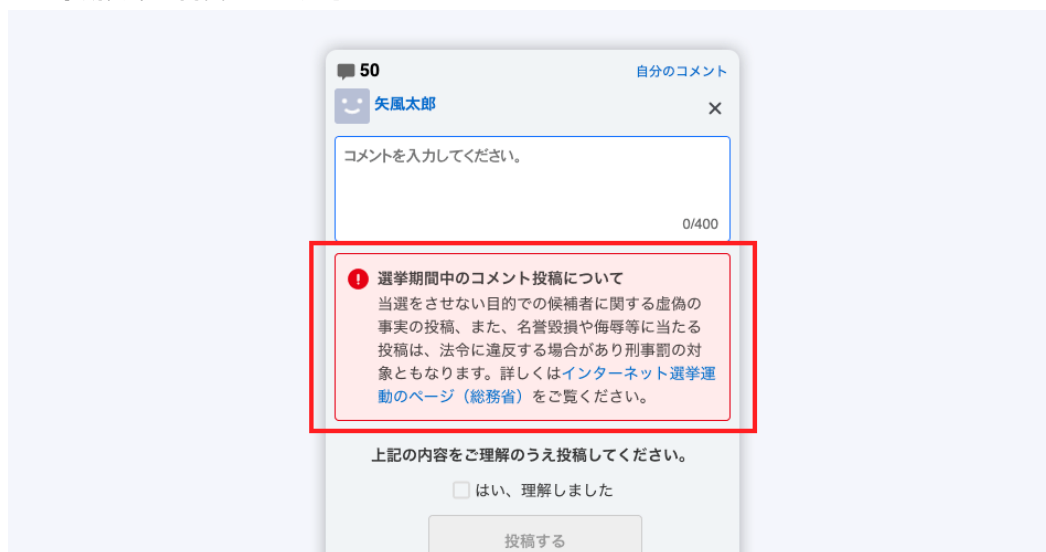
■衆議院選挙期間中における注意メッセージの掲出

第49回衆議院選挙の選挙運動期間中（2021年10月19日～10月31日）に、「Yahoo!ニュース」の政治カテゴリに掲載されるすべての記事において、コメント投稿時に特別な注意メッセージを掲出します。

公職選挙法によると、「当選を得させない目的をもって公職の候補者に関し虚偽の事実を公にし、又は事実をゆがめて公に」するなどの行為は、禁止されています。また、名誉毀損や侮辱などの行為も法律で禁止されています。（※8）

「Yahoo!ニュース」では、多様な意見の尊重と公正な選挙の実現に配慮し、コメント投稿時に注意メッセージを掲出して注意喚起を行うことで、選挙運動期間中に、選挙の公正を阻害したり、または選挙結果に影響を及ぼしうる虚偽の事項についての書き込みなどがなされたりすることを抑止します。

<選挙期間中に掲出される注意メッセージ>



「Yahoo!ニュース」は今後も、投稿者に再考を促して誹謗中傷などの投稿を防止する取り組みの強化や、多様な意見を提示するAIの開発に取り組むなど、さまざまな施策を実行していく予定です。それにより、コメント欄で投稿される多様な考えや意見によって、ユーザーがニュースに対する興味や多角的な視点を持つきっかけを提供していきます。

そして、社外のパートナー企業や有識者などとも連携しながら、日本のインターネットにおける健全な言論空間を構築し、より良い社会の実現に向けて全力を尽くします。

※1：[Yahoo!ニュース コメント 取り組みまとめ](#)

※2：[Yahoo!ニュース コメントポリシー](#)（2021年3月25改定）

※3：[大規模深層学習モデルによるYahoo!ニュース「不適切コメント」対策](#)（Yahoo! JAPAN Tech Blog）

※4：[Yahoo!ニュース、最先端の「深層学習を用いた自然言語処理による判定モデル（AI）」と独自スパコン「kukai」を活用し、コメント対策を強化](#)（2020年3月6日発表）

※5：[Yahoo!ニュースのコメント機能 アカウントの投稿停止措置について](#)（2018年6月12日発表）

※6：[誹謗中傷対策に関する有識者会議を2020年に設置](#)

※7：[Yahoo!ニュースにおけるAIを活用した投稿時注意メッセージの掲出効果について](#)（2021年2月16日発表）

※8：[インターネット選挙運動解禁（公職選挙法の一部を改正する法律）の概要](#)（総務省サイト）