

1 前提：対象サービスの範囲について**【想定対象サービス】**

SNS	
検索	
動画共有	

ニュースポータル・ キュレーション	
ブログ・掲示板	はてなブログ

Q1-1 上記【想定対象サービス】欄に挙げられたもの以外に、貴社又は貴社の関係会社が事業上運営しているサービスであって、以下の①～③のいずれにも当てはまるもの（上記【想定対象サービス】欄に記載の類型以外を含む）があれば、御教示下さい。

- ① インターネットを通じ、日本国内の不特定の者によって受信されることを目的としてコンテンツ（文字、画像、映像、音声など）を媒介・送信するものであること。
- ② 上記①のコンテンツに自社以外の第三者が作成・投稿したものが含まれること。
- ③ 広告収入を主たる運営基盤としていること（※）。

※ 広告収入以外の収入（購読料、寄付金など）を主たる運営基盤とするサービスは、いわゆるアテンション・エコノミーの影響を受けにくいと考えられるため、対象外とします。

はてなブックマーク、はてな匿名ダイアリー。（カクヨム（小説投稿サービス）マンガノ（漫画投稿サービス）は他社との協業のため今回は回答対象に入れていません）

- ◆ 以下の質問に対しては、上記1の【想定対象サービス】欄に挙げられたサービス（日本国内向けのものに限ります）及びQ1-1で御回答頂いたサービス（以下、併せて「対象サービス」）について御回答下さい。
- ◆ 各質問に対しては、原則としてそれぞれの対象サービスごとに分けて御回答下さい。ただし、対象サービスの全部又は一部に共通の御回答内容となる場合は、その旨を付記して御回答下さい。
- ◆ 各質問にある取組を現時点で実施していない場合でも、実施する予定があるときは、可能な限り当該予定の内容についても御回答内容に含めて頂けますと幸いです。

2 対象サービスの規模について

Q2-1 対象サービスのうち、コンテンツの投稿に当たって利用者登録が必要なものについては、日本国内における（※1）最近の月間アクティブユーザー数（※2）として把握している数字を、いつの時点の数字かと併せて御教示下さい。

※1 後記Q2-4参照。

※2 1か月間に対象サービスに1度でもアクセスした利用者の数。

非公表

Q2-2 対象サービスのうち、コンテンツの投稿に当たって利用者登録が不要なものについては、日本国内における（※）最近の月間合計投稿数として把握している数字を、いつの時点の数字かと併せて御教示下さい。

※ 後記Q2-4参照。

非公表

Q2-3 Q2-1及び2-2いずれの数字も把握していない対象サービスについては、日本

国内における（※）事業規模を示す数字として把握しているその他の数字（登録ユーザー数など。推計値でもかまいません。）を御教示下さい。

※ 後記Q2-4参照。

登録アカウント数：約1,214万（2023年7月実績）ただし、長期休眠ユーザーや1人による重複登録、法人名義での利用も含んでおり、人数としての実数は1000万人には満たないものと考えています

Q2-4 Q2-1～2-3で御回答頂いた数字について、「日本国内における」数字であると整理された理由・基準を御教示下さい。

例：利用者登録時に入力された住所、投稿言語、投稿時の位置情報など

現在利用言語はほぼ日本語に限られており、国外利用はかなり小規模なものであると考えている

3 偽・誤情報の流通・拡散への対応方針について

Q3-1 対象サービス上での日本国内における（※1）偽・誤情報（※2）の流通・拡散への対応について、次の3点を御教示下さい（※3）。

※1 文字や音声については少なくとも日本語を含むもの、画像や映像については日本に関連する人・物・場所等の画像・映像を含むものを想定しております。それ以外の基準により「日本国内における」実態を把握している場合は、当該基準を明示いただいた上でお答えください。以下、「日本国内における」「日本国内において」「日本に関する」とある箇所につき同様。

※2 ここではひとまず大まかに、「内容や発信主体等に関して虚偽や誤解を招く事項を含む情報」とお考え下さい。

※3 既に公開している規約、ポリシー、ガイドライン等（以下「規約等」）の中で説明している場合は、当該規約等のタイトル及びURLを付記して、該当箇所を抜粋下さい。

（1）何らかの対応が必要と考える「偽・誤情報」の範囲（定義）・類型

例：災害関係、健康・医療関係、選挙・政治関係、国際紛争関係、青少年関係、自殺誘発関係、犯罪（詐欺など）誘発関係、ディープフェイクなど

（2）上記（1）の類型ごとに想定される典型事例

（3）上記（1）の類型ごとの具体的な対応方法

例：削除、アカウント停止、表示順位の降格、収益不能化、アクセス不能化、警告表示・ラベリング、ファクトチェック結果の表示、投稿時の注記の義務付けなど

（1）内容による類型定義はおこなっていない

（2）ここで問題となるフェイクニュースの多くはインプレッションや収益を目的としており、投稿される内容はその時々の大衆の興味に合わせたものになると考えられるため、内容をもとにした類型定義を行う意義があまりないと考えている

（3）類型に応じて対応方法を変えらるということは行なっておらず、その表現や投稿行為の悪質性、公益性の有無、投稿者のサービス利用状況に応じて対応を決定している

Q3-2 Q3-1で御回答頂いた対応方針の背景となる基本的な考え方（重視している価値や哲学）があれば、御教示下さい。

提供するサービスは、ユーザーの思いを自由に投稿するプラットフォームであり、投稿の自由は尊重したいと考えている。その一方、収益等を期待して目的と乖離した投稿利用が行われることがあり、偽情報はそれに相当するものが多い。そのような投稿は収益に対する最適化が行われる結果として投稿行動に一定の偏りが生じることが多いため、特徴的な利用状況を検知して利用制限を行うことで多くの不適切な投稿は抑制できるものと考えている。

また、ユーザーが自由に投稿を行うことにより、誤った情報に対する批判や訂正情報も寄せられることになり、結果的に中立性が保てることも期待している。

その一方、投稿者の思い込みや誤った知識により結果的にフェイク情報となってしまうような事例はある。そのような場合、運営者が利用停止等の強硬な対応を取ることは、投稿者の思い込みを強めてしまい逆効果となることがあるため、正しい情報を提供するバナーを表示するなど、より慎重な対応を要する。

Q3-3 Q3-1で御回答頂いた対応方針について、定期的な見直しや外部有識者・第三者機関等によるレビューを行っている場合は、(1)当該見直し・レビューの頻度・タイミング、及び(2)レビューの実施主体を御教示下さい。

定期的な見直しは行なっていない。ユーザーに対するルールやガイドラインを制定する場合には、法曹関係者や有識者にヒアリングを行なっている(収益目的利用に対するルール制定の際には、日本アフィリエイト協議会理事にヒアリングを行うなど)

4 偽・誤情報の発信者(投稿者)の表現の自由等への配慮について

Q4-1 Q3-1で御回答頂いた対応方針の対象となる偽・誤情報(日本国内において流通するものに限り、以下「対象偽・誤情報」)の流通・拡散に対し、Q3-1(3)で御回答頂いた対応(以下「モデレーション等」)を実施した場合、実施の事実や理由を当該対象偽・誤情報の発信者(投稿者)に通知し、又は発信者(投稿者)の知り得る状態に置いている場合、その方法(※1)、言語及び内容(※2)を御教示下さい。

※1 例えば、書面、電子メール、対象サービス上での表示など。

※2 例えば、モデレーション等を実施した事実及び理由のほか、後記5の第三者通報によるものか、第三者通報を待たず自社で検知・対応したものかといった実施の端緒や、Q4-3の苦情・不服申立て受付窓口に関する情報など。

投稿者の登録メールアドレスあてに日本語で通知を行なっている。一部海外ユーザーと思われる者に対しては英語で通知することもある。

通知内容については透明性を大切にしており、利用規約に定めなどの項目に合致するために利用停止としたか等の情報提供は行なっている。通報によるものであるか否かについては、通報者に対する逆恨みや、通報自体が不当な妨害行為であるといった主張をするなど、望ましくない結果になることがあるため、特に必要がない限りは通知をしていない。

Q4-2 対象偽・誤情報の流通・拡散に対して実施したモデレーション等について、当該対象偽・誤情報の発信者(投稿者)からの苦情や不服申立てを受け付ける窓口を設置している場合、当該窓口に関し、次の3点を御教示下さい。

(1) 苦情・不服申立ての受付方法(※)

※ 例えば、オンラインフォーム、書面、電子メール、電話など。なお、予め定められている申請フォーマットがある場合は、当該フォーマットを添付(又はURLを付記)下さい。

(2) 日本語による苦情・不服申立てを受け付けているか否か

(3) 窓口情報を日本語で公開している場合はその方法(※)

※ オンライン上で公開している場合はURLを付記下さい。

(1) 問い合わせフォーム、メールアドレス

(2) 日本語による申立を受け付けている

(3) 窓口情報は明示的には公開しておらず、汎用の問い合わせ窓口

(<https://hatena.zendesk.com/hc/ja>)を利用することを想定している。通知メールには返信先を記載している。

Q 4 - 3 Q 4 - 2で窓口を設置しているか否かにかかわらず、対象偽・誤情報の流通・拡散に対して実施したモデレーション等について、当該対象偽・誤情報の発信者（投稿者）から日本語で苦情や不服申立てを受けた場合における対応（以下「日本語苦情等対応」）（※1）に関し、次の8点を御教示下さい（※2）。

※1 苦情・不服申立て内容の合理性や問題となったモデレーション等の適否を調査・判断し、場合によってはモデレーション等を撤回するなどの対応を指します。

※2 問題となったモデレーション等が後記5の第三者通報によるものか、第三者通報を待たず自社で検知・対応したものかといった実施の端緒によって御回答内容が異なる場合は、当該実施の端緒別に御教示下さい。

- (1) 日本語苦情等対応が可能な人員として常用している人数
 - (2) 日本語苦情等対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順（どのようなケースで用いるのかなど）
 - (3) 日本語苦情等対応にファクトチェック機関やマスメディア（新聞・放送）を関与させている場合は、当該機関等の名称及び関与手順（どのようなケースで関与させており、その意見等を日本語苦情等対応へどのように反映させているのかなど）
 - (4) 上記（3）以外の外部有識者・団体を日本語苦情等対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順（どのようなケースで関与させており、その意見等を日本語苦情等対応へどのように反映させているのかなど）
 - (5) 日本語苦情等対応の開始から完了までの目標期間を定めている場合は、当該目標期間（「開始」及び「完了」の定義を含む）
 - (6) 日本語苦情等対応の結果を発信者（投稿者）に通知し、又は発信者（投稿者）の知り得る状態に置いている場合は、その方法、言語及び内容
 - (7) 日本語苦情等対応の結果に対する再審査制度を設けている場合は、その概要（日本語による再審査要求を受け付けているか否か、及び外部有識者や第三者機関を関与させている場合はその詳細を含む）
 - (8) 特定の発信者（投稿者）（※）からの日本語による苦情・不服申立てを優先的に取り扱うこととしている場合は、当該発信者（投稿者）の類型及び具体的な優先的取扱いの内容
- ※ 例えば、国会、裁判所、行政機関、地方公共団体、その他公的機関（独立行政法人等）、国際機関、専門機関、ファクトチェック機関など

(1) 5名

(2) AIの導入はなし

(3) ファクトチェック機関やマスメディアの利用はなし

(4) 外部機関の関与はなし

(5) 2営業日以内

なお「開始」は通報フォーム送信時刻、「完了」は通報者に対する返信、通報者が返信不要と意思表示しているものについては対応システム（zendesk）でステータスを「完了」に変更した時点です。

(6) メールでの返信（日本語）

(7) 再審査制度は設けていない

(8) 優先対象はない

Q 4 - 4 Q 4 - 3で御回答頂いた内容を公開している場合は、その方法（※）を御教示下さい。

※ オンライン上で公開している場合はURLを付記下さい。

非公開です

5 偽・誤情報の流通・拡散に対するモデレーション等の手続・体制について

Q5-1 対象偽・誤情報の流通・拡散について、それにより権利を侵害されている者及び発信者（投稿者）以外の第三者からの通報（以下「第三者通報」）を受け付ける窓口を設置している場合、当該窓口に関し、次の4点を御教示下さい。

(1) 受け付けている通報手段（※）

※ 例えば、オンラインフォーム、書面、電子メール、電話など。なお、予め定められている申請フォーマットがある場合は、当該フォーマットを添付（又はURLを付記）下さい。

(2) 通報できる主体に限定を設けている場合は、当該主体

例：利用者登録済みユーザー、ログイン済みユーザー、国会、裁判所、行政機関、地方公共団体、国際機関、専門機関、ファクトチェック機関など

(3) 日本語による通報を受け付けているか否か

(4) 窓口情報を日本語で公開している場合はその方法（※）

※ オンライン上で公開している場合はURLを付記下さい。

(1) オンラインフォーム、電子メール、書面の郵送、FAXにて受け付けています。

権利侵害情報については <https://policies.hatena.ne.jp/deletion-flow>

その他の通報については <https://hatena.zendesk.com/hc/ja> の

<https://hatena.zendesk.com/hc/ja/articles/900001845703> この項目から受け付けます

(2) 制約はありません

(3) 受け付けています

(4) (1) で記載したオンラインフォームにて公開しています

Q5-2 Q5-1の窓口を設置している場合、日本語による第三者通報を受け付けた後の対応（以下「日本語通報対応」）（※1）に関し、次の8点を御教示下さい（※2）。

※1 通報内容の合理性や問題となったコンテンツの対象偽・誤情報該当性を調査・判断し、場合によってはモデレーション等を実施するなどの対応を指します。

※2 通報主体や対象偽・誤情報の類型（Q3-1（1）参照）毎に異なる場合は、当該主体別又は類型別に御教示下さい。

(1) 日本語通報対応が可能な人員として常用している人数

(2) 日本語通報対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順（どのようなケースで用いるのかなど）

(3) 日本語通報対応にファクトチェック機関やマスメディア（新聞・放送）を関与させている場合は、当該機関等の名称及び関与手順（どのようなケースで関与させており、その意見等を日本語通報対応へどのように反映させているのかなど）

(4) 上記（3）以外の外部有識者・団体を日本語通報対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順（どのようなケースで関与させており、その意見等を日本語通報対応へどのように反映させているのかなど）

(5) 日本語通報対応の開始から完了までの目標期間を定めている場合は、当該目標期間（「開始」及び「完了」の定義を含む）

(6) 日本語通報対応の結果を通報者に通知し、又は通報者の知り得る状態に置いている場合は、当該通報の方法、言語及び内容

(7) 日本語通報対応の有無・内容に関し、通報者からの不服申立てや問合せがあったときの専用窓口や処理手順を定めている場合は、その概要（日本語による不服申立て・問合せを受け付けているか否か、及び外部有識者や第三者機関を関与させている場合はその詳細を含む）

(8) 特定の通報主体（※）からの日本語による通報を優先的に取り扱うこととしている場合は、当該通報主体の類型及び具体的な優先的取扱いの内容

※ 例えば、国会、裁判所、行政機関、地方公共団体、国際機関、専門機関、ファクトチェック機関など

- (1) 5名
- (2) AIの導入はなし
- (3) ファクトチェック機関やマスメディアの利用はなし
- (4) 外部機関の関与はなし
- (5) 2営業日以内

なお「開始」は申立者がフォーム送信した時刻、「完了」は申立者に対する初回の返信、申立者が返信不要と意思表示しているものについては対応システム（zendesk）でステータスを「完了」に変更した時点です。

特に権利侵害情報については申立者に事情聴取などを行う必要があり、情報削除手続自体が完了するまでの時間の基準は設けていません。

また、文書の場合、送受信のタイムラグが生じるため時間の基準は設定していません。

- (6) メールでの返信（日本語）文書でメールアドレスの記載がない場合は、FAX、郵送での返信
- (7) 再審査制度は設けていない
- (8) 優先対象はない

Q5-3 対象偽・誤情報の流通・拡散について、第三者通報を待たず自社で検知・対応（以下「検知・対応」）（※1）することがある場合、次の5点を御教示下さい（※2）。

※1 検知内容の合理性やモデレーション等の必要性を調査・判断し、場合によってはモデレーション等を実施するなどの対応を含みます。

※2 対象偽・誤情報の類型（Q3-1（1）参照）ごとに異なる場合は当該類型別に御教示下さい。

- (1) 検知・対応に向けた調査を実施する頻度・ケース
- (2) 検知・対応にAIその他の機械的手段を利用している場合は、当該手段の概要及び利用手順（どのようなケース・頻度で用いるのかなど）
- (3) 検知・対応にファクトチェック機関やマスメディア（新聞・放送）を関与させている場合は、当該機関等の名称及び関与手順（どのようなケースで関与させており、その意見等を検知・対応へどのように反映させているのかなど）
- (4) 上記（3）以外の外部有識者・団体を検知・対応に関与させている場合は、当該外部有識者・団体の名称及び関与手順（どのようなケースで関与させており、その意見等を検知・対応へどのように反映させているのかなど）
- (5) 検知してからモデレーション等を実施するまで（モデレーション等を実施しない場合は実施しない旨決定するまで）の目標期間を定めている場合は、当該目標期間

(1) 頻度を設けて網羅的に調査を行うことはない。トップページ、ランキング等のメディアページについては随時目視で確認をしている

- (2) AIの導入はなし
- (3) ファクトチェック機関やマスメディアの利用はなし
- (4) 外部機関の関与はなし
- (5) 検知後すみやかに対応することとしており目標期限は設けていない

Q5-4 Q5-2及び5-3で御回答頂いた内容を公開している場合は、その方法（※）を御教示下さい。

※ オンライン上で公開している場合はURLを付記下さい。

非公開です

6 偽・誤情報の流通・拡散への対応状況について

Q6-1 2022年中及び2023年中それぞれに関する（※1）次の20点）として把握し

ている数字があれば、御教示下さい（※2）。

※1 別の期間の数字のみ回答可能な場合は、期間を明示した上で当該数字を御教示下さい。

※2 可能であれば、対象偽・誤情報の類型別（Q3-1（1）参照）、モデレーション等の方法別（同（3）参照）の数字も併せて御教示下さい。また、可能であれば、（9）～（12）については通報主体別（Q5-1（2）参照）、（18）・（19）については発信者（投稿者）の類型別の数字も併せて御教示下さい。

（1）対象偽・誤情報（※）についてモデレーション等を実施した件数

※ Q4-1のとおり、対象偽・誤情報は「日本国内において流通するもの」に限りますが、対象偽・誤情報に関する数字を把握していない場合で、何らかの推計値や、対象偽・誤情報に関する数字を含むものと思われる別の数字を把握しているときは、どのような数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。

（2）上記（1）のうち、A I その他の機械的手段のみによって検知・対応した件数

（3）上記（1）のうち、A I その他の機械的手段と人間の手を組み合わせてモデレーション等を実施した（第三者通報への対応過程でA I その他の機械的手段を利用した場合を含む）件数

（4）上記（2）（3）のうち、A I その他の機械的手段による判断結果に誤りがあることが事後的に判明した（上記（3）において人間の手によって補正された場合を含む）件数

（5）上記（1）のうち、ファクトチェック機関の意見等を反映してモデレーション等を実施した件数

（6）上記（1）のうち、マスメディアの意見等を反映してモデレーション等を実施した件数

（7）上記（1）のうち、上記（5）（6）以外の外部有識者・団体の意見等を反映してモデレーション等を実施した件数

（8）上記（5）～（7）のうち、ファクトチェック機関、マスメディアその他の外部有識者・団体の意見等に誤りがあることが事後的に判明した件数

（9）対象偽・誤情報に関する日本語による（※）第三者通報の総数

※ 日本語による第三者通報の件数を把握していない場合で、言語以外の基準により日本国内における同様の数字として把握している数字があるときは、当該基準を明示頂いた上で当該数字を御回答下さい（推計でもかまいません）。また、そのような数字も把握していない場合で、日本国内における数字を含むものと思われる別の数字を把握しているときは、どの範囲の数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。

（10）上記（9）の第三者通報を契機として対象偽・誤情報についてモデレーション等を実施した件数

（11）上記（9）の第三者通報を受け付けてからモデレーション等を実施するまで（モデレーション等を実施しない場合は実施しない旨決定するまで）に要した期間の平均値

（12）上記（9）のうち、モデレーション等の有無・内容に関し、通報者からの不服申立てや問合せがあった件数

（13）Q4-2の窓口への対象偽・誤情報に関する発信者（投稿者）からの日本語による（※）苦情・不服申立て件数

※ 日本語による苦情・不服申立て件数を把握していない場合で、言語以外の基準により日本国内における同様の数字として把握している数字があるときは、当該基準を明示頂いた上で当該数字を御回答下さい（推計でもかまいません）。また、そのような数字も把握していない場合で、日本国内における数字を含むものと思われる別の数字を把握しているときは、どの範囲の数字であるか明示頂いた上で当該数字を御回答下さい。以下同じ。

（14）Q4-2の窓口以外を通じた対象偽・誤情報に関する発信者（投稿者）からの日本語による苦情・不服申立ての件数

（15）上記（13）（14）のうち、当該苦情・不服申立てに基づき、問題となったモデレーション等を撤回した件数

（16）上記（13）（14）の苦情・不服申立てを受け付けてから、問題となったモデレーション等を撤回するまで（撤回しない場合は撤回しない旨決定するまで）に要した期間の平均値

（17）上記（13）（14）のうち、Q4-3（7）の再審査要求があった件数

（18）上記（13）（14）のうち、Q4-3（8）の類型に該当する特定の発信者（投稿者）による苦情・不服申立ての件数

（19）上記（18）の苦情・不服申立てを受け付けてから、問題となったモデレーション等を撤回するまで（撤回しない場合は撤回しない旨決定するまで）に要した期間の平均値

(20) モデレーション等の有無・内容に関し、発信者（投稿者）、通報者との間で、訴訟・仲裁・調停など裁判所等の外部紛争解決機関が関与する手続に発展した件数

偽・誤情報とスパム行為、迷惑行為について対応時に区別をしていないため、対応実態の数値は取得していません。

Q6-2 Q6-1で御回答頂いた数字について、いわゆる透明性レポート等の形で公開している場合は、その方法（※）と頻度（例えば、年1回など）を御教示下さい。

※ オンライン上で公開している場合はURLを付記下さい。

はてなブックマークについて透明性レポートを3ヶ月に1度公開していますが、偽・誤情報のみを独立した形で集計することは今後も難しいと考えています。

7 令和6年能登半島地震関連の偽・誤情報の流通・拡散への対応状況について

Q7-1 2024年1月1日に発生した令和6年能登半島地震に関連する偽・誤情報の流通・拡散に対する取組状況を御教示下さい。特に、現時点までにモデレーション等を実施した主なコンテンツとモデレーション等の方法別（Q3-1（3）参照）件数を御教示下さい。

災害発生後、目視にて確認をしましたが、偽・誤情報の拡散が確認できなかったため特に対応を行っていません。

Q7-2 令和6年能登半島地震に関連するコンテンツのうち、その内容が誤りであるとするファクトチェック機関のファクトチェック記事が存在するコンテンツについて、Q7-1で御回答頂いた件数に含まれる場合は、当該件数を御教示下さい。

同上

Q7-3 令和6年能登半島地震に関連する偽・誤情報の流通・拡散への対応として、上記4及び5にて御回答頂いた回答と比べて体制を強化した場合は、強化した内容を御教示下さい。

※ 例えば、対行政機関窓口の設置、モデレーション等を担当する人員の増強など。

既存人員で対応が可能でした

Q7-4 令和6年能登半島地震に関連する偽・誤情報の流通・拡散への対応に関し、他の事業者、ファクトチェック機関その他のファクトチェック関連団体又は日本の行政機関（※）との間で情報共有等の連携を実施している場合は、その詳細を御教示下さい。

※ 例えば、内閣府（防災担当）、警察庁など。

特にありません

Q7-5 地震その他の災害に関連する偽・誤情報の流通・拡散への対応に関し、今後（さらに）体制を強化する予定・余地があれば、強化する内容を御教示下さい。

現状では不要であると考えています

Q7-6 地震その他の災害に関連する偽・誤情報の流通・拡散への対応に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

今回は偽・誤情報が拡散されるということがありませんでしたが、政府としての公式情報をワンストップで確認できるページの URL があれば、災害に関する言及があるページに自動的にその URL への誘導枠を設置するなど、正しい情報への誘導がしやすくなると思います

8 レコメンドやモデレーション等に関する透明性・アカウントビリティ確保に向けた取組について

Q8-1 対象偽・誤情報を含む対象サービス上で表示される日本語のコンテンツについて、アルゴリズムを用いてモデレーション等の要否・内容を自動的に決定している場合、当該決定に当たって用いられる主なパラメータ及び当該決定の過程（当該パラメータが当該決定にどのように寄与しているのかなど）を御教示下さい。

はてなブックマークの注目コメント選定において、LINE ヤフー株式会社の提供する建設的コメント順位付けモデル API を社内独自のアルゴリズムと併用して利用していますが、詳しいパラメータや決定過程については逆用されることを防ぐため非公開としています。

Q8-2 Q8-1で御回答頂いたパラメータを用いてモデレーション等の要否・内容を自動的に決定していることについて、既に公開している規約等の中で説明している場合は、当該規約等のタイトル及びURLを付記して、該当箇所を抜粋下さい。

アルゴリズムを逆用した掲載内容・順位の操作を避けるため、アルゴリズムの詳細は公開しておりません。かわりに、アルゴリズムを設計・運用する上で遵守しているポリシーを以下にご紹介します。として、評価基準を公表しています。

<https://b.hatena.ne.jp/guide/guideline#adjusting-display-order-with-algorithms>

Q8-3 対象サービス上で表示される日本語のコンテンツ（広告を含みます）について、アルゴリズムを用いて表示先や表示順位等のレコメンドの要否・内容を自動的に決定している場合、当該決定に当たって用いられる主なパラメータ及び当該決定の過程（当該パラメータが当該決定にどのように寄与しているのかなど）を御教示下さい。

はてなブックマークコメントには、注目コメントを優先的に別タブで表示する仕組みがありますが、そのタブに掲載されるコメントを選定する際に使われています。

Q8-4 Q8-3で御回答頂いたパラメータを用いてコンテンツの表示先や表示順位等のレコメンドの要否・内容を自動的に決定していることにつき、既に公開している規約等で説明している場合は、当該規約等のタイトル及びURLを付記し、該当箇所を抜粋下さい。

アルゴリズムによる表示調整については、はてなブックマークのガイドライン

<https://b.hatena.ne.jp/guide/guideline#adjusting-display-order-with-algorithms>にて明示しています。

また、利用規約第9条 (<https://policies.hatena.ne.jp/rule#rule09>) は、「当社は、本サービスを監視する義務を負いませんが、本規約に反する、あるいはそのおそれがある行為や情報開示がある場合には、当該情報の削除や掲載場所の移動、外部サイトのクローラー巡回対象からの除外、および当該行為を行った登録ユーザーの登録取り消しなどを行う場合があります。」と定めており、不適切な情報発信に対するレコメンド順

位の変化は「掲載場所の異動」に相当するものとして取り扱っています。

Q8-5 Q8-1及び8-3のアルゴリズムについて、透明性・アカウントビリティ確保に向けた取組の一環として、特定の第三者（※1）に対して特定の目的（※2）で開示している場合、当該第三者の名称又は類型その他の詳細（※3）を御教示下さい。

※1 例えば、秘密保持契約を締結した研究機関や研究者など。

※2 例えば、透明性・アカウントビリティ確保状況に関する監視・研究目的など。

※3 例えば、開示目的、開示に至った経緯など。

現在のところ開示していません。

Q8-5 Q8-4で既に関示しているか否かにかかわらず、Q8-1及び8-3のアルゴリズムについて、今後、特定の第三者に対して特定の目的で開示することの可否・継続性に関するお考えがあれば、お聞かせ下さい。

現在、一部通報状況や、監視によるスパム行為の取締について はてなブックマーク透明性レポートを公開しています。

<https://bookmark.hatenastaff.com/entry/2023/12/26/160000>

今後、このような形での情報公開はすすめていきたいと考えています。

9 広告エコシステムの透明性・アカウントビリティ確保に向けた取組について

Q9-1 対象サービスのうち、広告収入を主たる運営基盤としているものについて、日本に関するコンテンツの発信者（投稿者）へ広告収入の分配を行っている場合、分配の基準を御教示下さい。特に、コンテンツのPV数やいいね、リポスト等の拡散行為の対象となった数という定量的な基準以外の定性的な基準（※）や、災害発生時を含む有事において適用される特別な基準が存在する場合は、当該定性的な基準や特別な基準を具体的に御教示下さい。

※ 例えば、コンテンツに偽・誤情報が含まれる場合には基準となるPV数等から除外する、そのようなコンテンツには広告を掲載しないなど。

弊社と利用者でのレベニューシェアに当たる仕組みは特にありません

Q9-2 Q9-1で御回答頂いた内容を公開したり、普及啓発に向けた取組を実施したりしている場合は、その方法・内容（※）を御教示下さい。

※ オンライン上で公開・普及啓発等している場合はURLを付記下さい。

該当なし

10 AI・ディープフェイク技術への対応状況について

Q10-1 上記4で御回答頂いたもののほか、対象サービスに関連してAIシステムを開発し、又は対象サービスにAIシステムを組み込んで第三者に提供し、若しくは自ら利用している事例がある場合は、その概要を御教示下さい。

ブログ記事の内容に基づいたタイトルをAIで生成する機能を提供しています。

Q10-2 Q10-1で御回答頂いたAIシステムの開発・提供・利用に関し、総務省及び経済産業省が策定・公表した「AI事業者ガイドライン案」（※1）において、「生成AIによって、内容が真実・公平であるかのように装った情報を誰でも作ることができるようになり、AIが生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まって

いることを認識した上で、必要な対策を講じる」ことが求められていることを受けて、何らかの対策（※2）を講じている場合は、その概要を御教示下さい。

※1 AI戦略会議第7回（2023年12月21日）資料1-3

https://www8.cao.go.jp/cstp/ai/ai_senryaku/7kai/13gaidorain.pdf

※2 例えば、生成AIが事実と異なることをもつともらしく回答する（ハルシネーション）リスクへの対策としての開発上の工夫、AIシステム提供時におけるハルシネーションリスクに関する透明性・アカウントビリティ確保の取組（ユーザーへの情報開示等）、AIの学習過程を含む開発メンバーや提供先・利用者のリテラシー向上に向けた取組など。

学習データが自身の投稿したブログ記事であるという機能の性格上、偽情報の生成に使われるリスクはほぼないと考えています

Q10-3 対象サービス上でAIが生成した日本に関するコンテンツが投稿された場合に、これを検知するツールや、投稿時のラベリングを投稿者に義務付ける規約等を導入している場合は、その詳細及び実効性担保に向けた工夫（もしあれば）を御教示下さい。

特に導入していません

Q10-4 以上御回答頂いたほか、対象サービス上でAIが生成した日本に関するコンテンツが投稿された場合に対応するため、他のAI関連事業者（開発者・提供者・利用者）との間で連携・協力に向けた取組（※）を行っている場合は、その詳細を御教示下さい。

※ 例えば、情報共有、資金援助、共同開発その他の技術提携・実装など。

特に対応していません

11 ファクトチェックの推進に向けた取組について

Q11-1 上記4及び5で御回答頂いたもののほか、日本国内に所在するファクトチェック機関その他のファクトチェック関連団体（※1）との間で連携・協力に向けた取組（※2）を行っている場合は、その詳細を御教示下さい。

※1 日本国内に向けたサービスを提供する等の活動をしているものを含みます。

※2 例えば、情報共有、資金援助、ファクトチェックを容易にするツールの開発や技術提携・実装、人材育成・交流、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

特に対応していません

Q11-2 ファクトチェック機関その他のファクトチェック関連団体との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

ファクトチェック機関との具体的な連携方法について特に情報がなく、どのように取り組めば良いのかが不明です。

体制が少人数のため、協力内容によっては担当者を手配しづらいこともあります。

Q11-3 その他、日本国内におけるファクトチェックの推進の観点から行っている具体的な取組（※）があれば、詳細御教示下さい。

※ 例えば、ファクトチェックを容易にするツールの開発・提供、ファクトチェックを実施する人材の育成など。

ユーザーによる自然発生的な対抗言論に期待しています

12 マスメディア（新聞・放送）との連携状況について

Q12-1 上記4及び5で御回答頂いたもののほか、日本国内に所在するマスメディア（新

聞・放送) (※1) との間で連携・協力に向けた取組 (※2) を行っている場合は、その詳細を御教示下さい。

※1 日本国内で活動しているものを含みます。

※2 例えば、情報・ノウハウの共有、資金援助、人材交流、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

特に連携は行なっていません

Q12-2 マスメディアとの連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

Q11-2 と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

13 利用者のICTリテラシー向上に向けた取組について

Q13-1 日本国内に所在する教育機関・普及啓発機関や消費者団体・利用者団体 (※1) との間で連携・協力に向けた取組 (※2) を行っている場合は、その詳細を御教示下さい。

※1 日本国内で活動しているものを含みます。

※2 例えば、対象サービスの利用者のICTリテラシー向上に関する定期的な意見交換、教育教材の共同開発、教育・普及啓発イベントの共催・後援など。

安心協等の団体に加盟し、ICT教育に係る取り組みに協力するなど随時行なっています

Q13-2 教育機関・普及啓発機関や消費者団体・利用者団体との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

Q11-2 と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

Q13-3 その他、日本国内における対象サービスの利用者のICTリテラシー向上の観点から行っている具体的な取組 (※) があれば、詳細御教示下さい。

※ 例えば、教材作成、普及啓発活動など。

サイト内「週間はてなブログ」での啓蒙記事の公開

(<https://blog.hatenablog.com/entry/2021/07/29/160000>) や、コミュニティガイドライン (<https://policies.hatena.ne.jp/community-guideline>) として弊社の考える健全なコミュニティのあり方を発信するなど、自社サービス内での情報発信を行なっていません。他社、他団体との協力については現時点ですすめている具体的な取り組みはありませんが、協力できることがあれば協力したいと考えています

14 研究開発の推進に向けた取組について

Q14-1 上記10で御回答頂いたもののほか、対象サービスに関連して、日本国内に所在する研究機関・研究者・アカデミア (※1) との間で、特に認知科学、心理学、情報工学、計算社会科学、法学などの分野において連携・協力に向けた取組 (※2) を行っている場合は、その詳細を御教示下さい。

※1 日本国内で活動しているものを含みます。

※2 例えば、情報共有、資金援助、共同開発その他の技術提携・実装、イベントの共催・後援など。

特に取り組みは行なっていません

Q14-2 研究機関・研究者・アカデミアとの連携・協力に向けた取組に際し、支障となって

いるもの・ことがございましたら、詳細御教示下さい。

Q11-2と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

Q14-3 研究機関や研究者等向けに、対象サービス上の情報流通の状況に関する実データをAPI等を通じて提供している場合は、提供先その他の詳細（※）を御教示下さい。

※ 例えば、提供開始（予定）時期、対価の額等の提供条件など。

特にありません

Q14-4 Q14-3で提供していない場合において、過去に提供していたことがある場合は、提供先、提供期間及び提供を終了した理由を御教示下さい。

特にありません

Q14-5 Q14-3で提供していない場合（Q14-4の場合を含みます。）、将来的に提供を開始（再開）することの可否についてお考えをお聞かせ下さい。

使用規約第8条で、コンテンツの学術利用については手当されており、
<https://policies.hatena.ne.jp/rule#rule08> 協力に対してはポジティブです

Q14-6 その他、対象サービスに関連した研究開発の推進の観点からの具体的な取組（※）があれば、詳細御教示下さい。

※ 例えば、電子透かし技術の開発、ディープフェイク検知・対抗技術の開発、自動ファクトチェック技術の開発など。

特にありませんが、将来的に何らかの取り組みができればと考えています

15 サイバーセキュリティ関係機関等との連携について

Q15-1 サイバーセキュリティ関係機関（※1）との間で、対象サービスに関連して連携・協力に向けた取組（※2）を行っている場合は、その詳細を御教示下さい。

※1 日本国内に向けたサービスを提供する等の活動をしているものを含みます。

※2 例えば、情報共有、資金援助、共同開発その他の技術提携・実装、イベントの共催・後援など。

IPAから脆弱性情報の共有は受けていますが、偽情報対策という面において特に連携はしていません

Q15-2 サイバーセキュリティ関係機関との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

Q11-2と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

Q15-3 その他、対象サービス上の対象偽・誤情報の流通・拡散への対応として、サイバーセキュリティ対策との連携の観点から行っている具体的な取組（ISACへの参加など）があれば、詳細御教示下さい。

特にありません

16 行政機関や地方公共団体等との連携について

Q16-1 日本の行政機関や地方公共団体等との間で連携・協力に向けた取組（※）を行っている場合は、その詳細を御教示下さい。

※ 例えば、情報共有、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援、官民連絡会議（米国における Global Internet Forum to Counter Terrorism (GIFST) や選挙の integrity に関するワーキンググループ (<https://perma.cc/JC5V-MARG>) に相当するものなど）の設置など。

特にありません

Q16-2 行政機関や地方公共団体等との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

Q11-2 と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

17 国際機関等との連携について

Q17-1 国際機関や専門機関等との間で連携・協力に向けた取組（※）を行っている場合は、その詳細を御教示下さい。

※ 例えば、情報共有、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

特にありません

Q17-2 国際機関や専門機関等との連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

Q11-2 と同様、連携に関する具体的な情報不足、担当者の確保が難しいことがある

Q17-3 その他、国際連携の観点から行っている具体的な取組（国際会議への出席など）があれば、詳細御教示下さい。

特にありません

18 その他のステークホルダーとの連携状況について

Q18-1 日本国内に所在するその他のコンテンツ制作主体（※）との間で連携・協力に向けた取組を行っている場合は、その詳細を御教示下さい。

※1 ジャーナリスト、クリエイター、ネットメディアなど。日本国内で活動しているものを含みます。

※2 例えば、情報・ノウハウの共有、資金援助、人材交流、対象サービス上でのコンテンツの優先的表示、イベントの共催・後援など。

特に取組はありません

Q18-2 日本国内に所在する他の電気通信事業者（※）との間で、ゼロレーティングサービスの享受などを通じた連携を行っている場合は、その詳細を御教示下さい。

※ 日本国内で事業を運営しているものを含みます。

特に取組はありません

Q18-3 日本国内に所在する他のプラットフォーム（※1）事業者やメタバース関連事業者（※2）と連携・協力に向けた取組（※3）を行っている場合は、その詳細を御教示下さい。

※1 SNS、検索サービス、動画共有サービス、ブログ・掲示板、ニュースポータル、キュレーションなど。

※2 日本国内で事業を運営しているものを含みます。

※3 例えば、情報共有、技術提携・実装、人材育成・交流、イベントの共催・後援など。

セーファーインターネット協会、安心協、モバイルコンテンツフォーラムなど、業界団体を通じての取り組みや、有志企業担当者との勉強会は随時開催しています。

Q18-4 Q18-1～18-3のステークホルダーとの連携・協力に向けた取組に際し、支障となっているもの・ことがございましたら、詳細御教示下さい。

行われているイベントや勉強会に関する情報共有の仕組みがあればと思います

19 特に紹介したい取組について

Q19-1 以上御回答頂いたほか、偽・誤情報の流通・拡散への対策をはじめ、デジタル空間における情報流通の健全性確保に向けた貴社の取組として共有可能なものがございましたら、当該取組の詳細(※1)を御教示下さい(※2)。

※1 例えば、当該取組を実施するに至った背景・課題、当該取組による効果など。

※2 この質問への御回答内容は、前記【御回答に当たっての留意事項】(3)のとおり、本検討会において「偽・誤情報対策に係る取組事例集(仮称)」としてとりまとめ、公開させて頂く予定です。

現時点では詳細情報を共有可能な形のものはありません

20 その他

Q20-1 以上御回答頂いたほか、デジタル空間における情報流通の健全性確保に関する課題や今後の対応にあたっての基本的な考え方、具体的な方策などについて、御意見や補足等ございましたら御教示下さい。

以上