

## 災害と情報の混乱をどう防ぐか：

### SNS・生成 AI・フェイク情報の交錯と“情報の耐性”の構築

#### How to Prevent Information Failure During Disasters:

#### Building Information Resilience for Local Governments in the Age of Social Media, Generative AI, and False Information

山口 真一

YAMAGUCHI, Shinichi

(国際大学グローバル・コミュニケーション・センター 教授)

#### 【要旨】

近年、気候変動の影響で災害は増え、被害も深刻化している。それに伴い、私たちが災害情報を受け取る環境も大きく変わった。行政や報道が情報の入口を押さえていた時代は過ぎ、今は現場の映像や市民の投稿、憶測まじりの書き込み、さらには生成 AI で作られた偽の画像や動画までが、同じ勢いで SNS に流れ込む。問題は、こうした情報が命に関わる判断と隣り合わせであるにもかかわらず、「何を信じるべきか」が一瞬で揺らぐ点だ。本稿では、様々な実証研究結果から、災害時に誤情報が広がる構造や、人が誤りを信じてしまう心理、SNS と生成 AI が生む新たな混乱を整理する。あわせて、自治体が備えるべき“情報の耐性”——検証と訂正の手順、説明の積み重ね、市民参加型の公式プラットフォーム整備——について考えたい。

キーワード：『災害情報』『フェイク情報』『生成 AI』『自治体広報』『情報の耐性』

#### 1. 序：“人類総メディア時代”の災害情報

近年、災害時の情報の流れは根本的に変わった。気候変動の影響で豪雨や台風がこれまでにない規模で発生し、地域の被害が急激に進行するなか、その状況を把握する手段として SNS の存在感は一段と増している。かつては行政や報道機関が一次的な情報を握り、住民はそれを受け取る側だった。ところが、SNS の普及によって、市民一人ひとりが「発信者」となり、現場の映像や声が数分のうちに全国へ届くようになった。河川の氾濫、停電の範囲、避難所の混雑状況といった細やかな事実が、職員の足では到底追いつけない速度で可視化される。

東日本大震災を契機に、SNS は「つながる命のインフラ」として定着した。電話がつかない状況で、短い投稿が安否確認や支援要請の手段となった経験は、多くの人々の記憶に刻まれている。災害時の情報は市民の命を守る重要な資源であり、その多様化は注目・歓迎された。

しかし、当時の SNS は、まだ個人の善意が支えた素朴なツールだった。現在は大きく状況が異なる。注意を引いた投稿が広告収益につながる「アテンション・エコノミー」の時代である。人々の承認欲求や経済的インセンティブが情報流通に介入し、真偽よりも「目を引くかどうか」が優先される仕組みが定着している。この構造の中で、災害時の SNS 利用をどう位置づけるかが問われている。

加えて、生成 AI の登場が新たな混乱を生んでいる。誰でも数秒で「もっともらしい」画像や動画を作れる時代になり、現実と虚構の境界は曖昧になった。実際に、2022 年の静岡県の水害では、AI が生成した偽画像が「現地の惨状」として拡散された。情報の見た目が本物であるほど、冷静な判断は難しくなる。今後、自治体が直面する最大の課題は、発信力の強化だけでなく、「検証」「訂正」「説明責任」を組み合わせた運用の確立である。これは、単なる技術論ではなく、信頼を守る統治の問題でもある。

本稿では、災害時におけるテクノロジーと人間の関係を再考する。SNS による情報収集とその限界、フェイク情報の拡散構造、そして生成 AI がもたらす新たな課題を整理し、自治体が備えるべき「情報の耐性」を具体的に描き出す。

ここでいう耐性とは、フェイク情報を完全に排除する力ではない。むしろ、混乱の中でもフェイク情報を受け止め、検証し、正確な情報へと置き換えるプロセスの強度である。災害時の情報は、もはや一方通行ではなく、行政と市民が相互に検証しながら共有するものへと変わっている。その新しい現実には、どのような制度と技術、教育を結びつけていくか—それが本稿の主題である。

## 2. 実務の最前線：首長・自治体の SNS 活用の功罪

災害が起きたとき、情報の最前線は行政庁舎の中ではなく、SNS のタイムラインの上にある。行政やマスメディアでは把握や発表が困難な、詳細な現場の映像や声が、数分のうちに行政にも全国にも届く。誰かのスマートフォンが、報道よりも早く被害を伝える時代になった。

そうした状況のなかで、市長や知事が自ら SNS で発信し、また、情報収集する姿も当たり前のものになりつつある。行政が市民に直接語りかけ、共感を得ながら情報を届ける手段として、SNS は確かに力を発揮している。だが同時に、その使い方を誤れば、ほんの一言が大きな混乱を生むこともある。

2025 年 8 月、熊本市で大雨による被害が広がった際、大西一史市長は自身の X(旧 Twitter) で「浸水の状況が分かる写真や動画を投稿してください」と市民に呼びかけた。すると数時間のうちに 300 件を超える投稿が集まり、市が進めていた現地調査とあわせて活用され、浸水エリアの特定を早めることにつながったという。投稿は 1000 万回以上閲覧され、災害時に SNS を通じて行政と住民が協力し合う姿を示す好例となった。

一方で、同じ時期の福岡市では、別の形で情報の難しさが表れた。高島宗一郎市長は、SNS

上に一般市民が投稿した「河川氾濫の動画」について「偽情報だ」と指摘したが、その後、実際に氾濫が起きていたことが確認され、市長は誤りを認めて謝罪した。高島市長は平時から SNS を積極的に活用し、個人発信を重視してきたが、そうした“SNS に強い”首長であっても、災害時の緊迫した場面では、限られた情報と時間の中で判断を誤ることがあるといえる。

また、2025 年 11 月には、宮城県女川町が公式アカウントで発信した「クマの出没情報」が、住民からの目撃写真として投稿した画像を添付して注意を呼びかけたものの、のちに画像が生成 AI によって作られた偽のものであったことが判明し、町が投稿を削除して訂正と謝罪を行う事態も発生した。深夜の住宅街にクマが立つ衝撃的な一枚は、多くの住民に不安を与え、SNS 上でも広く拡散されたが、実際には存在しない“被害”だった。行政側は住民の安全確保のため迅速に対応したつもりでも、提供された画像が虚偽であれば判断の基盤そのものが揺らぐ。

これらの事例は、自治体が直面する真偽不明の情報が溢れる環境、つまり、限られた時間の中で真実と虚偽を識別しなければならない検証の最前線の困難さを物語っている。

### 3. 災害時フェイク情報の三重構造：心理・経済・AI

#### 3.1 災害時にフェイクが広がる背景

災害が発生すると、正確な情報と同じ速さで、誤った情報も拡散する。能登半島地震のとき、X（旧 Twitter）では「#助けて」「#SOS」と書かれた投稿が次々と広がった。ところが、架空の住所を含む偽の救助要請が少なくない数含まれていた。

他にも様々なフェイク情報が拡散した。過去の津波映像を「今回の被害」として投稿した動画もあり、中には数百万回以上再生されたものもあった。さらに、「地震は人工的に起こされた」「原発から放射能が漏れている」など、根拠のない主張も拡散した。被災地が混乱しているときほど、人々は情報を求め、信じたいものを信じてしまう。そこに、災害時のフェイク情報の危うさがある。

フェイク情報が広がる背景には、人の心理が深く関わっている。災害時、人々は強い不安や恐怖を抱く。その不安を少しでも和らげようとして、誰かに知らせたくなる。「自分の投稿が誰かの助けになるかもしれない」という気持ちは善意そのものだ。しかし、その善意が裏目に出ることがある。筆者の調査では、「不安だったから」「人の役に立ちたかったから」という理由で誤情報を拡散した人が少なくなかった（山口ほか, 2024）。悪意よりも善意の拡散のほうが、実は広がりやすい。これは災害時のフェイク情報の根深い特徴である。

もうひとつの要因は、SNS そのものの仕組みだ。今の SNS は、人々の関心を集めることに価値が置かれるアテンション・エコノミーの中で動いている。つまり、多くの人に見られること自体が利益になる構造だ。特定の条件を満たす投稿者は広告収入を得ることができるため、「注目される投稿」が経済的な価値を持つ。災害時のように人々の関心が集中する

場面では、刺激的な言葉やショッキングな画像ほど拡散されやすくなる。「事実かどうか」よりも「どれだけ話題になるか」が優先される構造が、結果的に誤情報の増幅を招いている。

### 3.2 災害時デマの「定番」パターン

災害時に広まるデマには、いくつかの典型的なパターンがある。多いのは、被害の規模や原因をめぐるものだ。過去の津波や火災の映像を「今回の災害の映像だ」と偽って投稿する事例は繰り返し見られる。能登半島地震でも東日本大震災の映像が流用され、数百万回も視聴された。原発の状況についての根拠のない噂も後を絶たない。

また、犯罪行為に関するデマも目立つ。避難所での略奪や、外国人窃盗団が押し寄せているといった虚偽の情報は、事実確認がないまま人々の不安をあおる形で広がっていく。

さらに、救助要請を装ったデマもある。前述したように、「車に閉じ込められた」「友人が家から出られない」といった投稿が、#SOS や #拡散希望 のハッシュタグを伴って拡散される。真偽が不明なまま広がるため、実際の救助活動に支障をきたす危険がある。能登半島地震の際は非常に多かったケースだ。人々の善意を逆手に取る寄付詐欺も定番だ。電子マネーや暗号資産での募金を求めるアカウントが災害直後に現れ、多くの人を惑わせる。

そして極めつけは陰謀論である。能登半島地震の際には「人工地震だ」とする主張が一日で 25 万件以上投稿され、中には 850 万回も閲覧されたものがあった。根拠がないにもかかわらず、「本当かもしれない」と感じさせる言葉は驚くほどの速さで広がっていく。

これらはすべて、人々の強い感情に訴える。恐怖、不安、怒り、あるいは「誰かを助けたい」という思いが拡散を後押しする。

### 3.3 生成 AI による with フェイク 2.0 時代の到来

そして近年、生成 AI の登場がこの問題をさらに複雑にしている。数年前までは専門的な知識がなければ作れなかった偽画像や偽動画が、今では一般の人でも簡単に作れるようになった。2022 年の静岡県の水害では、SNS 上で拡散された「ドローンで撮影された水害の様子」が、実は AI によって作られた架空の映像だったことが明らかになった。投稿した本人は、特別な技術者ではなく、家でスマートフォンを操作していた一般の利用者だった。「少し話題になれば」と軽い気持ちで作った画像が、現場の混乱を加速させた。いまや、誰もが“ディープフェイクの制作者”になりうる時代である。

いまや、ディープフェイクは特別なものではない。無料のツールを使えば、数分でもっともらしい被害映像を作ることができる、「ディープフェイクの大衆化が起こった」。こうして作られた偽画像や偽動画が SNS に投稿されると、それを見た人が「本物だ」と信じ、さらに拡散してしまう。問題は、こうしたフェイクが単なる誤報にとどまらず、社会的混乱を引き起こす点にある。

今後このようなフェイクコンテンツはネット上に増えていくだろう。私はこの時代を with

フェイク 2.0 時代と呼んでいる。2016 年の米国大統領選挙以降、私たちはフェイクと共に生きてきた。それが 2.0——次なるステージに行くということだ。

### 3.4 災害時のフェイク情報をもたらすもの

災害時のデマは、実際の被害よりも早く人々の行動に影響を与えることがある。避難経路が誤って伝わり、危険地域に人が集まる。物資不足の噂が広まり、必要以上の買い占めが起こる。特定の地域や集団に責任を押しつける投稿が拡散され、被災者同士の対立を生む。あるいは、「救援物資が届いている」といった誤情報が流れ、支援の遅れを招くこともある。善意で拡散されたデマが、結果として被災地の作業を妨げる。

社会的な分断も深刻だ。誤情報を信じる人と信じない人の間に溝ができ、互いに非難し合う構図が生まれる。特定の人々や団体を“加害者”と断定するような投稿が広まると、偏見や差別が助長される。フェイク情報は、人と人との信頼を削り取っていく。科学的根拠の乏しい「自己防衛策」が拡散されれば、人々をむしろ危険にさらすことにもつながる。災害時のフェイク情報とは、単なる誤解ではなく、社会の絆を揺るがす“二次災害”なのだ。

## 4. データが示すフェイク情報の実態：「誰が」「なぜ」騙されるのか

### 4.1 「多くの人々が騙される」という不都合な現実

フェイク情報というと、「一部の人が騙されるものだ」と考える人は少なくない。だが、それは現実とは程遠い。私が Google Japan の支援を受けて 2019 年から継続的に行ってきた研究プロジェクトでは、日本におけるフェイク情報の受容実態を、数年間にわたって丹念に追ってきた。

2024 年に発表した最新の調査では、2022 年から 2023 年にかけて実際に国内で広く拡散した 15 件のフェイク情報を使用し、人々がそれらをどう受け止め、どのように判断し、さらには広めているのかを精密に分析した。そこで明らかになったのは、「フェイク情報は誰でも簡単に信じてしまう」という不都合な現実である。

調査では、該当する 15 件のフェイク情報を一度でも見聞きした 3,700 人を対象に、真偽判断を尋ねた。その結果、「これは誤りだ」と正しく判断できていた人はわずか 15%程度にとどまった。逆に、「正しいと思っていた」と回答した人は半数を超え、残る人々は「どちらとも言えない」と判断を保留していた。つまり、フェイクに触れた多くの人々が、その情報を「誤りだ」とは気づかないまま受け止めていたことが分かったのである(山口ほか, 2024)。

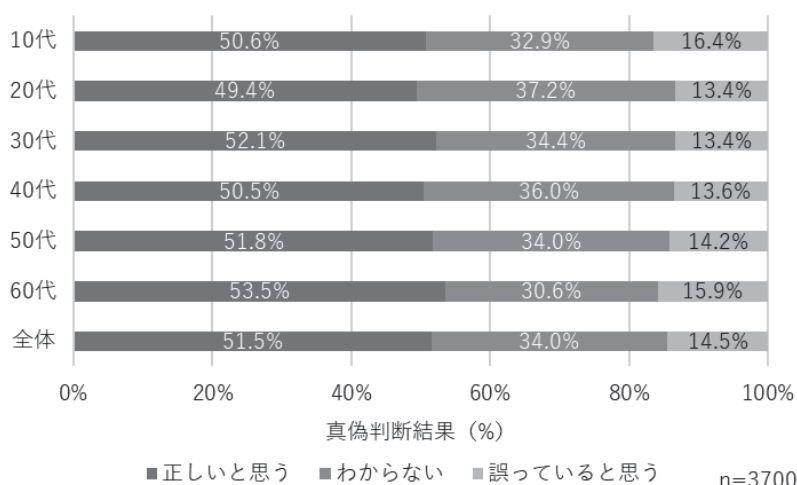


図1 フェイク情報の真偽判断結果（年代別）

ここで見逃してはならないのは、この傾向が特定の層に集中していたわけではなかった点だ。若者だけでも、高齢者だけでもない。中年層も含め、どの年代でもほぼ同じ割合で誤情報を信じていた。ニュースを日頃から欠かさず見ている層も、「自分は大人だから大丈夫」と考える層も例外ではない。つまり、フェイク情報に惑わされるリスクは、老若男女を問わず平均的に存在している。

#### 4.2 フェイク情報に騙される背景

では、なぜ私たちはこれほど容易く誤情報を誤情報と見抜けないのか。その理由は複雑だが、核心には三つの力学がある。

第一に、ソーシャルプルーフ（社会的証明）である。SNS上で何千もの「いいね」やシェアがついた投稿を見ると、私たちは中身を吟味する前に「これだけの人が支持しているなら事実に違いない」と錯覚しやすい。数字の大きさそのものが信頼の裏付けに見えてしまい、虚偽が真実のように見える。同調圧力が誤情報に“真実らしさ”を与えてしまうのだ。

第二に、訂正情報の遅れがある。フェイクが拡散してからファクトチェックが公表されるまでには、どうしても時間差が生じる。その間に最初の誤情報が強く記憶に残り、「継続影響効果」と呼ばれる心理作用によって、後から訂正を見ても誤解が消えにくくなる。訂正は常に後追いであり、どうしてもフェイクの後塵を拝してしまう構造的な弱さを抱えている。

第三に、「らしさ」の力がある。ニュースサイト風のデザイン、専門家らしい肩書き、統計グラフのような図表。体裁が整っているだけで、私たちは内容の真偽を疑わなくなる。まして写真や動画となれば、その“証拠らしさ”は圧倒的である。実際、過去の災害映像が「最新の被害」と偽って投稿された際、多くの人を受け入れてしまった。映像は論理ではなく直

感に訴えかけ、瞬時に信じ込ませてしまう力を持っている。

こうした力学が重なり合うことで、私たちは知らず知らずのうちに真偽の境界を見失っていく。SNS を眺めているだけで、微かな誤解が積もり、ある時ふと、その誤解を前提にものを語ってしまうことすらある。

フェイク情報に騙されるのは一部の人ではなく、「多くの人」であり、そして「誰にでも起こりうる」——この点を出発点に据えなければ、フェイク情報対策はいつまでたっても絵空事に終わる。これは、社会全体が向き合うべき現実の問題なのである。

#### 4.3 フェイク情報は情報空間全体の問題

フェイク情報の拡散といえば、多くの方は SNS のシェアやリポスト機能を思い浮かべるだろう。しかし調査を重ねてみると、拡散の経路はオンラインに限られていなかった。

私の研究では、フェイク情報を見た後に「誰かに伝えた」と答えた人は全体の 6 人に 1 人ほどいたが、その中で最も多かった手段は意外にも日常の会話だった（山口ほか, 2024）。拡散者の約半数は、家族や友人、知人とのやり取りの中でフェイク情報を共有していたのである。食卓で「今日こんなニュースを見たよ」と誰かが話し、それを聞いた家族が職場や学校で別の人に伝える。そこで広がった噂が再び SNS に戻り、別の文脈で再生産される。この往復運動を通じて、フェイク情報はオンラインとオフラインを行き来しながら急速に広まっていく。

ここで重要なのは、人々が「誰の言葉を最も信じるか」という点である。同じ調査では、家族や友人、知人といった身近な人からの情報は、新聞やテレビと同程度に信頼されていた。専門家でも記者でもない、ただ隣に座っているだけの相手の言葉が、ときに既存メディアと同じ重みを持つ。

この構造は「悪意なき拡散」を生む。本人は良かれと思って伝えているだけだが、それが誤っていれば結果的にフェイクの増幅に加担することになる。親しい間柄ゆえに受け手は疑いを挟まず、内容をそのまま取り込んでしまう。

現代は、オンラインだけでなくオフラインの場を介してフェイク情報が増幅される時代である。ネットと現実が往復することで噂は強固になり、訂正はますます追いつかない。フェイクの拡散源は「無数のシェアボタン」だけではなく、「私たち自身の口」でもあるのだ。

#### 4.4 自信がある人ほど騙される

フェイク情報に騙されやすいのはどのような人かと問われれば、多くの方は「情報リテラシーが低い人」や「普段ニュースをきちんと見ない人」と答えるだろう。しかし、調査を重ねると答えはそれほど単純ではなかった。

私はこの疑問を明らかにするため、四つの「力」に着目した。第一はメディアリテラシー、第二は情報リテラシー、第三はテストによって測定した批判的思考スコア、そして第四が

「自分は批判的に物事を考えている」とどの程度感じているかを尋ねる自己評価型の批判的思考尺度である。前の三つが客観的な能力を測る数値だとすれば、最後の一つは自己認識に近い。

分析の結果、メディアリテラシー、情報リテラシー、批判的思考スコアが低い人ほどフェイク情報を拡散してしまう傾向が見られた。逆に、これらの力が高い人は怪しい情報に出会ったとき「本当にそうだろうか」と立ち止まる余裕を持つといえる。この点は直観的にも理解しやすい。

しかし興味深いのは、自己申告による批判的思考尺度だけがまったく異なるパターンを示したことである。「自分は批判的に考えている」と強く答える人ほどフェイク情報を見抜けず、かえって拡散に加担していたのである (Yamaguchi et al., 2025)。

これは「自分は騙されない」という過信の危うさを示している。冷静に判断しているつもりでも、実際には思い込みに引きずられ、情報検証を怠り、都合のよい情報をそのまま受け入れてしまう。いわば「自分は賢い」という錯覚が、逆説的にフェイク情報への脆弱性を高めていた。似た傾向は海外でも確認されており、アメリカの研究では「ニュースの真偽を見抜く力に自信がある人ほど、実際の判定テストでは成績が低い」という結果が報告されている (Lyons et al., 2021)。これは「知識の錯覚」と呼ばれる現象に近い。知識や判断力があると信じることで検証を怠り、虚偽に足をすくわれるのである。

この知見は、フェイク情報への対策に重要な示唆を与える。メディア情報リテラシー教育や批判的思考のトレーニングは確かに重要だが、それ以上に欠かせないのは「自分も騙されるかもしれない」という謙虚さである。フェイクの拡散に加担するのは一部の無知な人ではなく、「自分は間違えない」と信じて疑わない人こそが危うい参加者になり得る。

結局のところ、フェイク情報に対する最大の防御は「自分も例外ではない」と受け止める姿勢である。どれほど経験を積んでもフェイク情報から完全に自由にはならない。その事実を認め、常に一步引いて情報を見る姿勢こそが、嘘に満ちた社会を生き抜くために必要なのだ。

### 4.5 フェイク情報は事実の6倍のスピードで広まる

フェイク情報の拡散力は、私たちの直感を大きく裏切る。2018年、米マサチューセッツ工科大学 (MIT) の研究チームが発表した論文はその衝撃を数量的に示したものだ (Vosoughi et al., 2018)。彼らは Twitter (現 X) 上で 2006 年から 2017 年までに流れた 12 万件以上のニュースを分析し、真実と虚偽がどのように広がっていくかを詳細に比較した。その結果、虚偽のニュースは事実のニュースよりも約 6 倍のスピードで広まり、到達範囲でも大きく上回っていた。特に政治分野でこの差は顕著だった。

なぜ嘘はこれほど速く、そして遠くまで届くのか。研究チームの分析では、虚偽のニュースは「新規性」が高く、「驚き」「嫌悪」「怒り」といった強い感情を喚起しやすいことが指

摘されている。たとえば「ある有名人が秘密の不正をしていた」といった情報は、人々の“まさか”という感情を刺激し、さらに怒りや不安と結びつけば、「誰かに知らせたい」という共有欲求を誘発する。対して事実は、多くの場合、地味で平凡だ。「感染者数が微増した」「政策が粛々と進んでいる」といった報道は、生活にとって重要であっても拡散の衝動は弱い。人間の心理は、退屈な正確さよりも、刺激的な“嘘の物語”に強く動かされるのである。

さらに注目すべきは、この拡散の主体が「ボット」ではなく私たち人間自身だった点である。自動プログラムによる機械的な拡散が疑われていたが、分析の結果、フェイク情報を広めていたのは主に一般の利用者であり、その多くは悪意ではなく善意だった。「誰かの役に立つかもしれない」「これは知らせておいた方がいい」——そんな気持ちが嘘を真実より速く走らせていたのである。

これは私自身の調査とも符合する。前節で示したように、フェイク情報を見た人のうち6人に1人が「誰かに伝えた」と答えていたが、その多くはSNSだけでなく、家族や友人との会話の中で広めていた。また、不安や怒りといった感情や、社会や他の人のためになると思ったという動機が多かった。悪意ではなく悪気のない共有が、結果として虚偽の増幅に加担していたのである。

こうした研究成果を踏まえると、嘘が事実を凌駕するスピードで広がるのは偶然ではなく、人間の心理そのものに根ざしていることが分かる。しかしここで自然に浮かぶ疑問がある。「ならばフェイク情報が広がったとき、すぐ訂正すればいいのではないか」。だが、これが簡単ではない現実もまた、研究結果が示している。

第一に、訂正には時間がかかる。虚偽の情報は根拠も裏付けも不要なため思いついた瞬間に拡散できるが、訂正は調査や検証を経て事実を確認しなければならない。災害の直後に「死者が何百人にのぼるらしい」といった根拠のない数字が一気に広がる一方、正確な被害状況を把握して公表するには数時間から数日を要する。この時間差こそがフェイク情報の優位をつくり出す。

第二に、人の心理の問題がある。人は最初に触れた情報を強く記憶し、その後に訂正を見ても印象が完全には上書きされない。「継続影響効果」と呼ばれるこの現象は多数の研究で確認されており、一度誤情報を信じ込んだ人は訂正を読んだ後でさえ、その誤情報を前提に判断や推論をしてしまう傾向がある。つまり訂正が届いた時点で、すでに手遅れの部分が生じているのである。

第三に、訂正はフェイクほど心を揺さぶらない。虚偽のニュースは驚きや怒りを喚起し、共有欲求を直接刺激するが、訂正記事は淡々とした説明が中心で広がりにくい。SNSでもフェイク情報が何万回と拡散される一方で、訂正はほとんど目に触れず、「フェイク情報だけが広く知られ、訂正は一部にしか届かない」という非対称が生じる。

こうして見ていくと、フェイク情報に対する訂正は、速度・心理・態度の三つの面で圧倒的に不利な条件からスタートしていることが分かる。嘘は瞬時に広がり、訂正は後から地道

に追いかける。しかしその努力は多くの場合、完全な修正に至らない。フェイク情報を減らすには、単に「誤りを正す」だけでは足りない。なぜ訂正が届かないのか、なぜ人は嘘を拡散してしまうのか——その構造を理解することこそが、私たちが最初に向き合うべき課題なのである。

### 5. 社会的対処：信頼の担保された社会のために

ここまで見てきたように、災害とテクノロジーが結びついた現在の情報環境では、フェイク情報の問題は単なる「ネット上のトラブル」では済まない。河川の氾濫、道路の寸断、避難所の混雑状況といった命に直結する情報が、真実と虚偽の区別もつかないまま SNS 上に洪水のように押し寄せてくる。

その中で「誰か一人が悪い」「どこか一社が責任を取れば解決する」といった単純な構図はすでに崩れている。国や企業、自治体、メディア、そして市民一人ひとりが、何を守り、どのように振る舞うのか。その判断の積み重ねが、災害時の情報空間の質を左右しているのである。法、プラットフォームと技術、教育、この三つの軸がそれぞれの役割を果たし、その上に日常の実務と市民の判断が重なってはじめて、「情報に強い社会」はかたちを持つ。

#### 5.1 制度的対応の功罪

まず法制度から見ておきたい。2022 年のプロバイダ責任制限法改正により、被害者が発信者情報の開示を求めやすくなり、侮辱罪の厳罰化によって悪質な誹謗中傷に一定の歯止めがかかるようになった。2024 年には「情報流通プラットフォーム対処法」が成立し、SNS などの事業者に対して違法情報への迅速な対応や透明性の確保が義務づけられた。2025 年 4 月から施行されている。削除基準の明示、対応状況の公表、発信者への通知といった要素が制度の中に組み込まれ、従来は各企業の裁量に委ねられていた部分に一定の共通ルールができたと言える。災害時に人命を危険にさらす明白なデマや、被災者や自治体職員への悪質な攻撃が放置されにくくなったという点で、こうした法改正には一定の意義がある。

とはいえ、法規制は常に諸刃の剣である。誹謗中傷やフェイク情報を抑制する目的で導入された規制が、時間の経過とともに対象を広げ、政権批判や調査報道など、本来守られるべき言論にまで影を落とす危険性は現実には存在する。実際、他国では「フェイクニュース対策」の名のもとに、政府への批判的言説が処罰の対象となっている例がいくつもある（山口、2022）。フェイク情報という概念は本質的にあいまいであり、どこまでを許容し、どこからを規制するかで判断は大きく分かれる。

重要なのは、「フェイクだから消せ」という感情に流されるのではなく、あくまで限定的かつ抑制的に法を運用し、その妥当性を社会が継続的に監視できる仕組みを併せて整えることである。信頼を守るはずの制度が、いつのまにか信頼を損なう装置になってしまうことだけは避けなければならない。

その中で私が注目しているのが、「マネタイズ」に着目したアプローチである。災害のような注目を集める事件が発生すると、アテンション・エコノミーの中で、お金儲けのために偽の救助要請や偽の被害動画などが投稿される可能性がある。実際、能登半島地震では大きな問題となった。

そこで、大きな災害が発生した際に、収益化を一時的に止めるような施策が考えられる。実際、私が委員として参加している総務省の「デジタル空間における情報流通の課題に関する検討会」は、2025年10月に中間まとめ案を発表している。その中で、災害時に虚偽情報が拡散されて広告収益が発生する構造を断つため、災害期間中は広告配信を停止する「マネタイズ停止」の提案が検討対象となっている。言葉を変えれば、フェイクを“儲ける機会”ごとに取り除く制度設計だ。

この良い点は、表現の自由と両立することである。災害に関連して自由に情報発信はできる。しかし、それによってお金を得ることはできない。このような取り組みをプラットフォーム事業者に要請していくことが考えられ、場合によっては制度化もありうる。

## 5.2 プラットフォーム事業者求められる責任と対応

次に、プラットフォーム事業者の責任である。Facebook や Instagram は外部のファクトチェック機関と提携し、虚偽の可能性が高い投稿に注意喚起のラベルを付ける取り組みを続けている。TikTok や YouTube などの動画共有サービスは AI 生成コンテンツに自動的に表示されるラベルの導入をしており、X (旧 Twitter) の「コミュニティノート」は、ユーザー自身が誤解を生みやすい投稿に補足情報を書き込む仕組みとして注目されてきた。災害時にも、こうした仕組みを通じて、明らかに根拠のない「人工地震説」や、存在しない被害情報に注意が促されれば、被害の拡大を一定程度は抑えられるかもしれない。

しかし忘れてはならないのは、これらのサービスが、私たちにとっては半ば社会インフラのように感じられる一方で、現実には一民間企業の事業に過ぎないという事実である。経営者が変われば方針も変わる。政権が変われば、プラットフォームが重視する価値のバランスも揺らぐ。X が買収後にコンテンツ管理を大幅に緩め、トランプ氏が 2024 年の米国大統領選挙で勝利した後に Meta が第三者ファクトチェックの取り組みを米国で停止したように、サービス提供者側の事情によって、私たちが前提としてきた安全策がある日突然消えてしまうこともありうる。災害時の極度に不安定な状況のなかで、地域住民の命綱となる情報流通を、このような揺れ動く土台にだけ頼るのは危うい。

日本語圏に特有の課題もある。大手プラットフォームはグローバル企業であり、世界中の法制度や文化に同時に対応しなければならない。そのなかで、日本の災害とフェイク情報に関する文脈が、彼らの優先順位のトップに来るとは限らない。日本語で流通する災害デマへの対応が後手に回ることも想定しておかなければならない。

だからこそ、自治体、報道機関、ファクトチェック団体が連携し、国内向けの検証ネット

## 特集

ワークと周知ルートを自前で整える必要がある。たとえば、災害時の疑わしい情報を受け付ける窓口を設け、検証結果をまとめて公開するサイトを常設し、平時から住民に周知しておく。プラットフォーム任せではなく、日本社会としての「検証の場」を持つことが、テクノロジー時代の自衛策である。

### 5.3 with フェイク 2.0 時代の災害時の SNS 活用の限界

ただ、ここまで見てきたような法と企業側の取り組みだけでは、災害時の情報環境を安定させるには不十分である。根本的な問題として、「事実と虚偽が入り混じった SNS を前提に、そこから必要な情報を拾い上げる」というやり方そのものに限界があるからだ。アルゴリズムは人々の関心を引くかどうかで投稿を並べ替え、自治体の公式発表も、匿名アカウントの憶測も、同じような見目でタイムラインに流れてくる。AI によって偽の被害状況の画像や動画を容易に投稿できる。住民側に真偽を見抜くリテラシーを持てと要求するだけでは、構造的な不利を覆すことはできない。with フェイク 2.0 時代では、それほどまでに事実と虚偽が入り乱れている。

そこで重要になるのが、SNS 以外のルートで市民が情報提供できる公的なプラットフォームを整備し、運用していくことである。たとえば、自治体が公式に提供する防災アプリやウェブサイト上に、「現場からの報告フォーム」を常設しておく。あるいは、内閣府などが一元化してそういったものを日本国として整備・周知してもよい。道路の冠水や橋の損傷、土砂崩れの兆候、停電や断水の状況などを、位置情報付きの写真や動画とともに送信してもらおう。送られてきた情報は、自治体職員や消防・警察、地元の自主防災組織などが連携しながら確認し、信頼できると判断したものを地図上に可視化する。こうした仕組みがあれば、住民は「とりあえず X に動画を上げる」のではなく、最初から公式の受け皿に情報を届けることができる。

もちろん、そこにも誤情報が紛れ込む可能性は残る。しかし、少なくとも投稿の目的と利用ルールが明確であり、記録が残り、運営主体がはっきりしている場では、検証と説明責任のプロセスを制度として組み込める。SNS のタイムラインのように無数の声が渦巻く空間とは異なり、「誰が、どのような目的で集めている情報なのか」が最初から共有されていること自体が、大きな強みになる。さらに、虚偽の情報であってもそれが一般市民の目に触れることもない。

テクノロジーを災害対応に生かすとは、本来こうした「公式でありながら市民参加型」のインフラを育てることであるはずであり、SNS はその補助的な拡散手段として位置づけ直されるべきである。

### 5.4 メディア情報リテラシー教育や技術による対抗

我々の意識も大切である。フェイク情報を最終的に見抜くのは、法でも AI でもなく、一

人ひとりの判断である。災害時に届く情報を前にして、「本当にそうなのか」「出所はどこなのか」と一度立ち止まること。強い怒りや恐怖をかき立てる投稿ほど、すぐに共有するのではなく、時間を置いてから考えること。こうした小さな行動の積み重ねが、社会全体の「情報の免疫力」を高めていく。

学校教育の中で、体系立ったメディア情報リテラシー教育を行い、情報空間の特性や、発信に伴う責任を教えることも欠かせない。自治体や企業が、講演会やパンフレット、オンライン講座を通じて、地域住民とともに学ぶ場をつくることも重要である。人は知れば変わるし、社会もまた学ぶことによって強くなる。

AI 技術も、正しく位置づければ心強い味方になりうる。人間の目だけではとても追いつかない膨大な投稿をスクリーニングし、危険性の高いものを絞り込む作業は、AI が得意とする領域である。また、AI によって作られた偽画像や偽映像の検証には、AI を使うことが必須だ。ただし、AI も決して万能ではない。虚偽を見逃すこともあれば、正当な批判を誤って弾くこともある。最終的には、AI と人間がそれぞれの強みと限界を理解したうえで役割分担し、相互に補い合う仕組みを整えることが現実的である。

フェイク情報の拡散そのものをゼロにすることは、おそらくできないだろう。しかし、だまされにくく、振り回されにくい社会を目指すことはできる。法が最低限のラインを守り、プラットフォームが一定の責任を果たすと同時に技術で対抗し、教育が人を育てる。その上に、私たち一人ひとりの小さな判断と行動が積み重なる。そのような社会であれば、災害とテクノロジーが絡み合う時代であっても、情報に押し流されずに踏みとどまることができるだろう。

## 6. 自治体が築くべき“情報の耐性”と信頼の発信

フェイク情報が現実の行動を動かす時代に、自治体が直面している課題は、「フェイク情報をどう防ぐか」という一点ではない。より根本的には、「誰の言葉を信頼すべきか」「どこに耳を傾ければよいのか」という、市民の拠りどころの問題である。

災害時、人々が最後に頼りにするのは、自分が暮らす地域の行政である。市役所や町役場、消防、保健所といった組織が、どれだけ落ち着いた声で状況を伝えられるかが、地域全体の不安と行動を左右する。だからこそ自治体には、「情報の耐性」を組織として備えることが求められている。ここでいう耐性とは、デマや批判にさらされても簡単には揺らぐず、噂に振り回されず、事実を発信し続ける力である。

その土台になるのは、最終的には職員一人ひとりの判断である。SNS 上でさまざまな情報が錯綜するなかで、「広く拡散されているから」「市長のアカウントに多数のリプライが来ているから」といった理由だけで動いてしまうと、組織としての軸を失う。一時的な空気に押し流されないためには、「確かな裏付けが取れていない段階では、発信を急がない」という姿勢が必要である。

現場の部署や関係機関に確認し、一次情報を丁寧に集める。そのうえで、住民にとって切実で、かつ確度の高い情報から順に発信していく。災害時には「とにかく早く出せ」という圧力が強まるが、自治体という公共機関には、早さだけでなく正確さを重んじる文化が求められる。

一方で、フェイク情報が広く出回っているにもかかわらず、行政が沈黙を続けることは、別のリスクを生む。「否定しないのは本当だからだ」「何か隠しているに違いない」といった疑念を招き、真実とは異なる物語が独り歩きし始める。SNS が現実社会に大きな影響を及ぼすようになった今日、「無視していればそのうち収まる」という発想は通用しにくい。誤った情報が人々の行動を誤らせる恐れがあると判断したならば、正しい情報が確認できた時点で根拠を添えて説明することが必要である。ここで大切なのは、相手を打ち負かそうとする反論ではなく、「いま分かっていること」「まだ分かっていること」を整理して淡々と共有することである。

こうした発信を行えば、SNS 上のコメント欄に批判や中傷が並ぶこともあるだろう。「炎上したらどうするのか」「職員が傷つくのではないか」という懸念は理解できる。しかし、能動的な発信だけが目に見える SNS の特性上、極端で強い感情を持った人ほど頻繁に投稿し、そのような声ほど目立ちやすい (Yamaguchi, 2023)。事実を読んで「そうだったのか」と静かに納得した人は、わざわざコメントを残さない。画面の上には批判の言葉だけが並ぶが、その背後には圧倒的に多い“黙って見ている人”がいる。このサイレントマジョリティに届く言葉を持つことが重要である。目に入りやすい批判に振り回されるのではなく、長い時間をかけて静かな多数派との信頼関係を育てていく視点が必要だ。

防御だけでは、信頼は育たない。いま自治体に求められているのは、フェイク情報を恐れて口を閉ざすことではなく、誠実さを武器に前へ出る「攻めの広報」である。政策の背景や意図を、専門的な行政用語ではなく、自分たちの言葉で説明する。災害対応の現場では、避難所運営や物資の仕分け、道路の復旧作業など、日々の具体的な仕事が山のようにある。それらのプロセスを写真や短い動画を交えながら伝えることで、「何がどこまで進んでいるのか」「どこがボトルネックになっているのか」を市民と共有できる。行政の仕事はしばしば無機質に見られがちだが、その背後には多くの職員と地域の人々の試行錯誤がある。その姿が見えるだけでも、自治体への見方は大きく変わる。

テクノロジーの使い方も改めて問い直す必要がある。SNS に頼り切るのではなく、先に述べたような「市民参加型の公式プラットフォーム」を中心に据える発想がこれからは求められそうだ。具体的には、自治体が公式に運用する防災アプリや Web ポータルを整備し、平時からイベント情報や防災講座、地域の小さなニュースを通して住民との接点を持っておく。有事には、そのアプリやサイトに「被害状況の報告」「避難所の満空情報」「道路・ライフラインの障害」などの項目を用意し、住民が簡単に投稿できるようにする。投稿された情報を、担当部署が一定のルールに従って確認し、信頼できるものは公式情報として地図上

に表示する。こうした仕組みは、単に情報を集めるだけでなく、「どの経路の情報を優先的に信頼すべきか」を住民に示す役割も果たす。

このような公式プラットフォームが育てば、SNS の役割は少しずつ変わっていくはずである。一次的な情報提供と検証の場合は専用のインフラに集約し、SNS はその内容を周知するための補助的なチャンネルとして用いる。住民が迷ったときに立ち返るべき「情報のホームグラウンド」を自治体が用意すること。それが、災害とテクノロジーが絡み合う時代の自治体広報における、大きな課題である。

平時から、公式アカウントや公式アプリのフォロワー・利用者を増やしておくことも欠かせない。有事になってから急に「こちらを見てください」と呼びかけても、届く範囲には限りがある。日常的に、地域の小さな話題や生活情報、防災訓練の日程、行政サービスの案内などを発信し、「困ったときにはまずここを見る」という習慣をつくっておくことが、結果として災害時の情報伝達力を高める。信頼はキャンペーンで一気に獲得できるものではない。日々の小さな発信、丁寧な問い合わせ対応、間違いが分かったときの素早い訂正。その積み重ねが、自治体と住民のあいだに目に見えない「情報の橋」をかけていく。

「情報の耐性」とは、フェイク情報をはね返す力であると同時に、批判の渦中にあっても説明をやめない強さでもある。心ない言葉に怯えて沈黙するのではなく、事実と向き合い、必要な説明を続けることが、最終的にはもっとも強い防御になる。情報社会における自治体は、単なる行政の窓口ではない。地域の信頼インフラであり、混乱の中で人々の進む方向を示す羅針盤でもある。真偽が曖昧になりやすい時代だからこそ、「ここで発信される情報は信用できる」と感じてもらえるような言葉を持たなければならない。

フェイク情報に満ちた時代に、真実を語り続けることは簡単ではない。誤解されることもあるし、批判を浴びることもある。それでも誰かがそれをやめてしまえば、公共の言葉は静かに痩せ細っていく。信頼とは、華々しい一回の成功ではなく、毎日の仕事と小さな判断の積み重ねから生まれるものである。その重みを最も実感しているのは、おそらく自治体職員自身であろう。テクノロジーに振り回されるのではなく、それを使いこなしながら、誠実な発信を続けることで地域の未来を守る。with フェイク 2.0 時代における自治体の使命は、その一言に尽きるのである。

## 参考文献

- Lyons, B. A., Montgomery, J. M., Guess, A. M., Nyhan, B., & Reifler, J. (2021). Overconfidence in news judgments is associated with false news susceptibility. *Proceedings of the National Academy of Sciences*, 118(23), e2019527118.
- Vosoughi, S., Roy, D.K., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151.
- Yamaguchi, S. (2023). Why are there so many extreme opinions online?: An empirical, comparative analysis of Japan, Korea and the USA. *Online information review*, 47(1), 1-19.

## 特集

Yamaguchi, S., Oshima, H., Watanabe, T., Osaka, Y., Tanihara, T., Inoue, E., & Tanabe, S. (2025). An analysis of literacy differences related to the identification and dissemination of misinformation in Japan. *Global Knowledge, Memory and Communication*, 74(11), 121-139.

山口真一. (2022). ソーシャルメディア解体全書： フェイクニュース・ネット炎上・情報の偏り. 勁草書房.

山口真一, 渡辺智暁, 逢坂裕紀子, 谷原つかさ, 大島英隆, & 田邊新之助. (2024). *Innovation Nippon 2024* 報告書 偽・誤情報、ファクトチェック、教育啓発に関する調査.