メディア情報リテラシーと「自信過剰の罠」

― 生成 AI時代の誤情報対策設計 ―

山口 真一｜国際大学グローバル・コミュニケーション・センター 教授

要約

誤情報が選挙・市場・公衆衛生に及ぼす影響が拡大するなか、メディア情報リテラシー教育は誤情報対策の重要な柱の 1 つとして位置づけられている。本稿は、筆者が重ねてきた一連の調査研究を起点に、Yamaguchi et al.（2025）を主要な手がかりとして、リテラシーと誤情報行動との関係を整理する。研究蓄積を総合的にみると、客観的に測定したリテラシーは誤情報の「拡散」を明確に抑制し、「真偽判断」にも一定の役割を果たす。他方で、主観的自信の過剰は真偽判断の失敗と拡散を促す逆説的構造がある。これらの知見はリテラシー教育の意義を裏付けるとともに、効果を最大化する設計指針を与える。本稿では、こうした知見を金融市場・政策・企業活動に接続し、生成 AI 時代のリテラシー教育と補完策を論じる。

山口 真一

博士（経済学）。専門は計量経済学、社会情報学。フェイク情報や生成 AI など、情報社会の課題を研究。電気通信普及財団賞など受賞多数、メディアにも多数出演・掲載。著書に『嘘で満ちていく社会』『炎上で世論はつくられる』など。早稲田大学ビジネススクール兼任講師、東京都デジタル広報フェロー、シエンプレ株式会社顧問のほか、内閣府「AI 戦略専門調査会」をはじめ、総務省、厚生労働省、文部科学省などの政府有識者会議委員や座長を務める。

1．はじめに―なぜ誤情報リテラシーをいま問い直すのか―

近年、誤情報が社会に与える影響は無視できない水準に達している。2016 年の米国大統領選挙では、トランプ氏に有利な誤情報が約 3,000 万回、クリントン氏に有利な誤情報が約 760 万回 Facebook 上でシェアされたと推計され（Allcott & Gentzkow, 2017）、選挙で使われた政治コンテンツ拡散用のボット（bot）はその後 2017 年のフランス大統領選挙でも観察された（Lazer et al., 2018）。同年のドイツ連邦議会選挙では、Twitter 上で共有された政治関連リンクにおいて専門ニュース 4 本に対しジャンクニュース 1 本程度が共有されたと指摘される（Neudert et al., 2017）。

経済面の波及も小さくない。2008 年 9 月に、ユナイテッド航空親会社の 2002 年の破産記事が、新たな破産を伝える誤情報として再浮上すると、株価は一時 76% 下落した（Carvalho et al., 2011）。逆に好意的な誤情報の対象となった企業が、異常に高いリターンと出来高を経験したケースも報告されている（Ullah et al., 2014）。新型コロナウイルスをめぐる誤情報は、適切な治療を妨げただけでなく、ウイルス自体の拡散を後押しした可能性も指摘されている（Pennycook et al., 2020）。

MIT の研究チームによる Twitter 上の 12 万件以上のニュース拡散の分析では、虚偽の情報は事実に比べて約 6 倍のスピードで拡散していたと報告されている（Vosoughi et al., 2018）。「嘘は事実よりも速く走る」という構造そのものが、誤情報問題の手強さを物語っている。

日本も例外ではない。筆者らが Google Japan と継続的に実施している Innovation Nippon の調査（山口他 , 2023, 2024）からは、誤情報への接触率の高さと、誤りと気付かないまま受け止めている人の多さが繰り返し示されてきた。実際に広く拡散された誤情報に関する真偽判断を尋ねた調査では、誤りと正しく判断していた人は接触者全体のごく一部にとどまり、半数を超える人々が「正しい」と信じていたという結果も得られている。これらの傾向は年齢や属性を問わず観察されており、誤情報をめぐる課題が特定の層に限られた話ではないことを示している。

そしていま、生成 AI の普及がこの問題に新たな次元を加えつつある。文章、画像、音声、動画のいずれも、ごくわずかな労力で一見もっともらしいコンテンツを大量に生成できるようになり、本物と偽物を見分ける手掛かりが急速に失われている。筆者はこうした状況を「with フェイク 2.0 時代」と呼んでいる。フェイク情報が耳目を集めた 2016 年の米大統領選以降、私たちは誤情報と共に生きる「with フェイク時代」を歩んできたが、生成 AI の登場によって、その規模もスピードも質的に変化した。これまでは誤情報の制作に一定のコストと時間がかかっていたが、いまや誰でも、専門家風の語り口を備えた動画コンテンツをすぐに作り出せる。情報空間の風景は、根本から変わりつつある。

このような状況において、私たちは何を手がかりに情報空間と向き合えばよいのか。日本政府も主たる対抗策の一つとして位置づけているのが、メディア情報リテラシーを軸とする教育的アプローチである。ただし、その効果を最大限に引き出すためには、どのような場面でどのリテラシーがどのように効くのかをエビデンスベースで見定めていく必要がある。

誤情報を「見抜く」局面と「広めない」局面で、リテラシーの効き方は同じなのか。教育を設計するうえで踏まえるべき構造は何か。本稿は、筆者がこれまで重ねてきた一連の調査研究、とくに山口（2022）や Yamaguchi et al.（2025）などの知見を踏まえて、こうした問いに向き合うものである。

以下、第 2 節で先行研究と本稿が参照する分析の枠組みを概観し、第 3 節で真偽判断と拡散をめぐる三つの構造を提示する。第 4 節では金融市場・政策・企業活動への含意を、第 5 節では生成 AI 時代に向けたリテラシー教育と補完的施策の設計を論じ、第 6 節で全体を結ぶ。

2. 先行研究と本研究の位置づけ

2.1 リテラシーと誤情報をめぐる研究の蓄積

リテラシーと誤情報をめぐる先行研究は、近年急速に蓄積されてきた。情報リテラシーが高い人ほど誤情報を適切に見分けられることが報告され（Jones-

Jang et al., 2021)、分析的思考のスコアが高い人ほど誤情報を見抜きやすいこと（Pennycook & Rand, 2019; Pennycook & Rand, 2020)、ニュースリテラシーを重視する人ほどソーシャルメディアの情報品質に懐疑的になりやすいこと（Vraga & Tully, 2021）も示されてきた。誤情報問題の深刻さが増すなかで、リテラシー教育が政策の中心に据えられてきた背景には、こうした研究知見の蓄積がある。

一方で、リテラシーと誤情報の関係は、種類や局面によって効き方が異なることも分かってきた。Jones-Jang et al.（2021）は、米国データを用いた分析で、情報リテラシーが誤情報の真偽判断に効きやすい一方、メディアリテラシーは必ずしもそれに直結しないことを示した。リテラシーをひとくくりに語るのではなく、内容を分けて捉える必要性が浮かび上がっている。さらに、誤情報研究の文脈では、Vosoughi et al.（2018）が示した「嘘は事実よりも約6倍速く拡散する」という構造的問題や、Allcott and Gentzkow（2017）が描いた選挙期間中の誤情報の影響の大きさが、リテラシー対策の重要性と限界を同時に浮き彫りにしてきた。

ただし、これら先行研究の多くは、特定のリテラシー指標に焦点を絞るか各種リテラシーをまとめて扱うかにとどまり、真偽判断と拡散という二つの局面を切り分けて包括的に分析した例は限られていた。自己申告型のリテラシー指標と客観テスト型のリテラシー指標を併用し、両者の効果を対比した研究もそれほど多くはない。リテラシー教育の効果を最大化し、適切なリテラシー教育を検討するためには、どの場面でどのリテラシーが効くのかを、精緻に把握する必要がある。本稿が主に参照する Yamaguchi et al.（2025）は、こうした研究上の空白に応えるものとして位置づけられる。

２.２ これまでの研究の流れと本稿が参照する分析

筆者は、日本における誤情報行動を理解するため、複数の調査を積み重ねてきた。本稿の議論はその系譜の上に位置づけられる。最初のまとまった分析は、山口（2022）『ソーシャルメディア解体全書』に収められている。ここでは Jones-Jang et al.（2021）に倣い、メディアリテラシー、ニュースリテラシー、デジタルリテラシー、情報リテラシーという四つの指標を用い、日本で広く拡散されていた８分野にわたる誤情報９件への行動を回帰分析で検証した。その結果、ニュースリテラシー・デジタルリテラシー・情報リテラシーの三つは、いずれも低い人ほど誤情報を誤りと判断できない傾向にあったが、メディアリテラシーについては真偽判断段階では効果が見られなかった。また、拡散段階ではニュースリテラシーと情報リテラシーが低い人、自己評価が高い人、政治的に極端な人ほど誤情報を気付かずに拡散しやすいことが示された。

続く Yamaguchi and Tanihara（2025）では、新型コロナウイルスワクチンと政治分野の誤情報 12 件を対象に、誤情報の拡散行動について、二つの仮説に整理して検証している。第一の仮説は、誤情報を真と信じてしまった人は、信じない人よりもそれを拡散しやすいというものである。第二の仮説は、メディア情報リテラシーが低い人ほど、誤情報を拡散しやすいというものである。分析の結果、二つの仮説はいずれも支持され、しかも拡散手段がソーシャ

ルメディアであっても、家族・友人との対面の会話であっても、同様の傾向が頑健に確認された。なお、同調査では、誤情報の最も多い拡散経路は「家族、友人、知人との直接の会話」であったことも報告されており、ネット上の拡散だけに目を向けていては問題の全体像をつかめないことを指摘している。

さらに、ファクトチェックに関する研究（Tanihara & Yamaguchi, 2023）からは、リテラシー、とりわけ情報リテラシーが低い受け手ほど、ファクトチェック結果（訂正情報）を正しく受け止めにくいことが示されている。具体的には、誤情報を否定するファクトチェック結果を提示しても、なお元の誤情報を真実だと判断し続けるだけでなく、当初は誤情報だと判断していた、あるいは判断を保留していた人が、提示後にかえってそれを真実だと判断するというパラドキシカルな現象も確認されている。これら一連の研究は、誤情報問題が「真偽の見極め」だけで完結しない多面性を持つことを、繰り返し示してきたといえる。

本稿が主要な手がかりとする Yamaguchi et al.（2025）では、リテラシーの捉え方を一段精緻化している。具体的には、批判的思考について、テスト形式で測る「客観スコア」と、自己申告で測る「主観的態度」を切り分け、それぞれが誤情報行動に対してどう作用するかを検証した。先行研究では両者がしばしば混在して扱われてきたが、能力と意識は別の概念であり、効果も異なりうる。後述するように、本稿の中心的な発見はこの分離によって初めて見えてきたものである。

2.3 調査の枠組み

Yamaguchi et al.（2025）では、日本国内で実際に広く拡散された誤情報 15 件を分析対象としている。具体的な手順としては、まず 2023 年に日本で広く流布した誤情報のうち、ファクトチェック機関による検証記事が存在するものを網羅的に収集した。そのうえで、先行研究を参照しつつ、これらの誤情報を、(1) 国内政治（保守寄り）、(2) 国内政治（リベラル寄り）、(3) 医療・健康、(4) 戦争・紛争、(5) 多様性（外国人や性的マイノリティ等）の五分野に分類し、各分野からとくに広く拡散した 3 件ずつを選んだ。筆者が関与した先行研究では 9 件、12 件と対象を変えながら知見の頑健性を確認してきたが、今回の 15 件はその系譜に位置づけられ、分野・拡散規模ともにばらつきを持たせている。特定ジャンルに偏ると、そのジャンル固有のリテラシーが結果を左右しかねないため、分野横断の設計が要点となる。

回答者には、リテラシーに関わる四つの指標を測定する設問に答えてもらった。第一に、メディアリテラシー。小寺（2017）と坂本（2022）を参照し、メディアメッセージの構築性、メディアによる「社会的現実」の構築、メディアの商業性、イデオロギーや価値の伝達、メディアのスタイルと言語、受け手による解釈の非一様性という、メディアの六つの性質に関する設問を「全くそう思わない」から「強くそう思う」の四段階で問い、その平均値をスコアとした（1 ～ 4 点）。第二に、情報リテラシー。Jones-Jang et al.（2021）の枠組みに基づき、日本在住者の情報を読み解く力に焦点を当てるよう修正した 5 問のテスト形式で、正答数をスコアとした（0 ～ 5 点）。第三に、批判的思

考スキル。論理的推論を問う問題への正答率を測る客観テストで、Watson-Glaser の批判的思考評価を日本語化した久原他（1983）の 20 問のテストに基づき、文章から導かれる確からしさを評価させるかたちで測定した（0 ～ 20 点）。第四に、批判的思考態度。中西他（2006）の社会的批判的思考態度尺度に基づき、批判的思考に対する傾向、経験、自己認識の 3 観点をカバーする 27 問について、「全くできない」から「非常にできる」の七段階で自己評価してもらった平均値をスコアとした（1 ～ 7 点）。

このうち批判的思考態度のみが回答者の自己評価を直接尋ねる主観的な指標であり、他の三つは設問への回答内容から能力を測定する客観的な指標と整理できる。本稿では便宜的に、批判的思考態度を「主観的自信」、他の三つを「客観リテラシー」と呼ぶ。

分析では、回答者がそれぞれの誤情報について、「正しいと信じる」「わからない」「誤っていると思う」の三択で真偽を判断してもらい、あわせて、ソーシャルメディア、メッセージアプリ、対面の会話、メール等の手段別に拡散したかどうかを尋ねた。

そのうえで、二つのロジットモデルで推定した。一つめは、誤情報を「誤りと適切に判断できた」かどうかを被説明変数とし、四つのリテラシー指標、メディア利用時間、個人属性などを説明変数とするモデルで、真偽判断行動を分析した。二つ目は、誤情報を「拡散したかどうか」を被説明変数とし、リテラシー指標等に加えて真偽判断の結果（正しい・誤り）も説明変数に含めるモデルで、拡散行動を分析した。すなわち、真偽判断結果ごとに拡散傾向がどう異なるかを統制したうえで、リテラシーの効果を取り出す設計となっている。

調査は 2024 年 2 月に、性年代別の人口比に割り付けられたオンラインモニタを対象に実施し、事前調査で約 2 万件の有効回答を得たうえで、誤情報に接触した経験のある回答者を抽出するスクリーニングを経て、本調査では分析に用いるサンプル 5,000 件を確保した。分析手法の詳細やロジットモデルの推計結果は原論文（Yamaguchi et al., 2025）に譲り、本稿では結果が示す構造を中心に論じる。

3．分析結果－真偽判断と拡散を分ける三つの構造－

3.1 真偽判断段階－研究蓄積が示す客観リテラシーの役割

まず真偽判断段階の結果を、研究蓄積全体のなかで位置づけながら見ていこう。

研究を総合的にみれば、客観的なリテラシーは誤情報の真偽判断に対して、概ね一定の役割を果たしていることがうかがわれる。山口（2022）では、ニュースリテラシー、デジタルリテラシー、情報リテラシーが低い人ほど誤情報を誤りと判断できない傾向が確認された。海外でも、情報リテラシーが誤情報の真偽判断に効くことや、分析的思考スキルが高い人ほど誤情報を見抜きやすいことが報告されている（Jones-Jang et al., 2021; Pennycook & Rand, 2019）。リテラシーが高い人ほど誤情報を冷静に評価できるという基本的な構

造は、複数の研究で繰り返し確認されてきたといえる。

一方、本稿が主要な手がかりとする Yamaguchi et al.（2025）では、客観リテラシー（メディアリテラシー、情報リテラシー、批判的思考スキル）について、真偽判断段階で統計的に明確な効果は確認されなかった。後述するように、同研究では主観的自信（批判的思考態度）が真偽判断を妨げる方向に効果を示しており、こちらが結果の前面に表れている。

ここで注意しておきたいのは、統計的に有意でないことは、効果がまったく存在しないことを意味するわけではない、という統計的解釈の基本である。研究はあくまで一つのサンプル、一定の設計のもとで得られた推定結果であり、効果がゼロであることを「証明」するものではない。研究ごとに、対象とする誤情報の分野や件数、扱うリテラシー指標の組み合わせ、調査時期や対象者の属性構成などが異なれば、特定のリテラシーが見せる効果の現れ方も変動しうる。たとえば Jones-Jang et al.（2021）では、メディアリテラシーが真偽判断に効かなかったが情報リテラシーは効いていた。

研究ごとの違いはあるものの、これらをあわせて読めば、客観リテラシーが真偽判断に対して概ね肯定的な役割を果たしているという全体像は、研究蓄積を通じて変わっていないと整理できる。同時に、研究によって効果の表れ方に違いが出るという事実は、真偽判断という行動が、リテラシーだけで決まる単純なものではないことも示している。

真偽判断が難しい背景には、誤情報そのものの形態の多様化もある。健康・美容分野でしばしばみられる、専門家風の語り口や研究論文を装った数字の引用を伴う誤情報は、その典型例といえる。「ある成分ががん予防に効く」「この食品は免疫力を劇的に高める」といった主張は、形式的に「らしさ」を備えるほど受け手の警戒を解いてしまう。リテラシー教育を中核に据えつつ、ファクトチェックや情報設計、技術的補助といった補完的施策を組み合わせていくことが、対策設計の鍵となる。

3.2 拡散段階ーリテラシー教育の防波堤効果ー

拡散行動について、研究蓄積はより明瞭で頑健な姿をみせる。Yamaguchi et al.（2025）では、メディアリテラシー、情報リテラシー、批判的思考スキルといった客観リテラシーが高い人ほど、誤情報を拡散しにくい傾向が明確に確認された。これは、Yamaguchi and Tanihara（2025）や山口（2022）など、筆者の過去の研究で繰り返し示されてきた傾向と整合的である。客観的なリテラシーを高めることが、誤情報の社会的拡散を抑える防波堤として機能することは、複数の研究を通じて確認されており、頑健な発見といえる。

真偽判断段階よりも拡散段階のほうが客観リテラシーの効果が前面に出やすいという点については、一つの見方が成り立つ。真偽判断は、特定の知識やドメイン理解、誤情報の形態に対する感度など、文脈依存性が高い。これに対し、拡散はより一般的な「立ち止まる習慣」や「自分の判断を他者に押しつけることへの慎重さ」によっても抑制されうる。リテラシー教育を通じて育まれる思考の習慣が、その情報を他者へと運ぶ手前で人を踏みとどまらせる、社会的なブレーキの役割を果たしているといえよう。

媒体別の分析からも興味深い知見が得られている。Yamaguchi et al.（2025）は、SNS、メッセージアプリ、対面の会話のそれぞれについて、真偽判断結果が拡散行動とどう結びつくかを分析している。その結果、ソーシャルメディアでは「正しい」と判断された情報のほうが有意に拡散される傾向がみられた一方、メッセージアプリや対面の会話では「誤り」と判断された情報のほうが有意に拡散される傾向が確認された。

後者は「これは誤りだから気を付けて」という注意喚起としての拡散と解釈でき、一見、望ましい行動のようにも思える。しかし、受け手のリテラシーが低い場合、「誤りだ」というメッセージそのものが理解されず、誤情報のほうが記憶に残ってしまうリスクも指摘されている。これは、ファクトチェックの効果に関する先行研究（Tanihara & Yamaguchi, 2023）とも整合する。否定的に提示された情報であっても、繰り返し触れることで「事実」として定着してしまう、いわゆる継続影響効果のリスクがあるためである。

加えて、Yamaguchi et al.（2025）では動画共有サービスの利用時間が長い人ほど誤情報を適切に誤りと判断しにくい傾向も報告されている。動画は文字情報よりも直感的に受け手の心に届くため、検証行動を経ずに受け入れられやすい性質を持つといわれる。短尺動画やショート動画の利用が拡大している現在、長時間の動画視聴が習慣化している層に対しては、別途の働きかけが必要になると考えられる。媒体特性ごとに、効くアプローチと効きにくいアプローチが異なるという認識は、対策設計にとって重要な出発点である。

3.3 主観的自信がもたらす逆説的リスク

もっとも興味深いのが、主観的自信に関する発見である。Yamaguchi et al.（2025）では、批判的思考態度、すなわち自分は批判的に物事を考えられているという自己認識が高い人ほど、誤情報を誤りと判断できない傾向が観察され、しかも誤情報を拡散しやすいという結果が示された。客観的なリテラシーと主観的な自信が、逆方向の効果を持つというある種の逆説である。「能力」を磨くことに加えて、「自信」のあり方そのものに目を向けることの重要性が、ここからみえてくる。

この知見は筆者の過去の研究とも符合する。山口（2022）では、一般的な自己評価が高い人ほど誤情報を気付かずに拡散しやすいことが報告されていた。Yamaguchi et al.（2025）で測定された批判的思考態度は、自分の思考力に対する自己評価という、より直接的な指標である。海外の研究でも、ニュースを見抜く力に自信がある人ほど実際の真偽判定テストの成績が低いという「過信」研究の知見がある（Lyons et al., 2021）。これらをあわせると、ひとつのパターンが浮かび上がる。自分は誤情報に騙されないと信じる人こそが、むしろ騙されやすく、しかも他者にそれを広げやすい、という構図である。

この構図は、具体事例に当てはめても理解しやすい。たとえば、政治家の発言を一部だけ切り取った短い動画が、政治的信条に強くコミットする層の間で拡散される。発言全体の文脈が省かれ、ときには真逆の意味に編集された動画が、断定的な見出しとともに広まる。山口（2022）の分析でも、政治的に

極端な人ほど誤情報を気付かずに拡散しやすいことが示されていた。「自分は政治に詳しい」「この候補者の本質を見抜いている」という確信が、検証をスキップさせる方向に働く可能性がある。これらの事例に共通するのは、拡散者の多くに悪意がない点である。むしろ「家族や知人に役立つかもしれない」「この事実を知らせなければ」という、ある種の善意や正義感が動機となっている。

前述したように、筆者の調査では、誤情報の拡散経路として、「家族・友人・知人との直接の会話」が最も多いことが分かっている。食卓やオフィスの雑談で交わされる一言が、噂を強固な「共有された事実」へと変えていく。SNS 上で広まった誤情報が、家庭や職場の会話のなかで「みんなが知っている話」へと変質し、そこからまた SNS に戻って再生産される。誤情報の流通は、ネット空間に閉じた現象ではなく、現実の生活と往復しながら進行する社会的プロセスなのである。この往復が「自信過剰の罠」と結びつくことで、誤情報は単なるネット上の話題を超えて、社会の認識そのものに影響を与える力を持つようになる。

以上を整理したのが図表 1 である。①客観的に測定される三指標が拡散段階で抑制的に働く一方、②真偽判断段階で明確な効果は確認されず、③自己申告で測定される主観的自信（批判的思考態度）が真偽判断を妨げ拡散を促す方向に作用するという、本稿の中核的な構造を示している。なお、真偽判断段階での客観リテラシーの効果については、本節冒頭で論じたとおり、研究蓄積全体としては一定の役割が示されている点を踏まえる必要がある。

図表1　四つのリテラシー指標が誤情報行動に及ぼす効果の整理

リテラシー指標	測定方法	真偽判断段階 （誤情報を見抜く）	拡散段階 （誤情報を広めない）
メディアリテラシー	客観的指標 （4 段階尺度）	△ 効果は明確には現れず	○ 抑制効果あり
情報リテラシー	客観的指標 （5 問テスト）	△ 効果は明確には現れず	○ 抑制効果あり
批判的思考スキル	客観的指標 （20 問テスト）	△ 効果は明確には現れず	○ 抑制効果あり
批判的思考態度 （主観的自信）	主観的指標 （自己評価）	× 真偽判断を妨げる	× 拡散を促す

出所）Yamaguchi et al.（2025）の知見を踏まえ筆者作成。真偽判断段階の客観リテラシーについては、過去研究では一定の効果が示されている点に留意。

4．金融市場・企業活動への含意

4.1 投資家行動と「自信過剰の罠」

金融市場の文脈では、自信過剰と誤情報拡散の結びつきは、行動ファイナンスの長年の関心事と接続する。投資判断における自信過剰は古典的なテーマで

あり、自信のある投資家ほど取引頻度が高く、結果としてパフォーマンスが劣るという指摘は繰り返し行われてきた。本稿の知見は、この古典的な論点に、誤情報拡散という新しい補助線を引くものである。

ミーム株現象を分析した Matsumoto et al.（2025）は、ソーシャルメディア情報に依存する個人投資家のなかで、利益を得るのは中心的な発信者に限られ、多くの投資家はむしろ損失を被りやすいことを示した。情報源からの距離が遠い投資家ほど、不利な立場に置かれる構造である。ここに自信過剰のリスクが重なる。「自分は SNS で投資の本質をつかんでいる」と感じる個人投資家は、検証を経ない情報をもとに行動しやすい。市場の急変時に誤情報がより速く広く伝播するなかで、こうした投資家ほど不利な意思決定を重ねやすい可能性がある。

経済・金融分野の誤情報は、しばしば数字やグラフを伴うことで形式的な「らしさ」を帯びる。「ある銀行が経営難に陥っている」「特定の通貨が近く暴落する」「ある上場企業の経営陣に不正がある」といった情報は、SNS 上で繰り返し拡散されてきた。生成 AI の普及によって、こうした「らしさ」を備えた誤情報の作成コストはさらに下がりつつある。とりわけ問題なのは、専門知識のある人ほど自分の判断に自信を持ちやすい点である。「自分は経済を理解している」「数字を読める」と思っている人ほど、最初の情報を疑わずに行動してしまう余地がある。

市場全体のリスクという観点からは、誤情報のスピードと規模が、流動性リスクや信用リスクの新たな源泉となりうる点に注意が必要である。前述したユナイテッド航空の事例（Carvalho et al., 2011）のように、誤った経営破綻情報が株価を急落させることがある。生成 AI 時代には、本物と区別のつかない経営者の偽動画や、もっともらしい財務データの捏造が、瞬時に市場を動かすリスクも視野に入る。市場の頑健性を支えるためには、こうした誤情報の影響を吸収できる情報インフラと、投資家の自己過信を抑える制度設計の双方が求められる。

投資家保護の観点からは、リテラシー教育と並行して、自己過信を抑える設計、たとえば過去の投資判断の正答率を可視化する仕組みや、情報源の偏りを表示する UI といったナッジ的アプローチも検討に値する。金融機関の顧客対応や投資家教育においても、「自分は分かっている」と感じる顧客ほど誤った情報をもとに行動しやすいことを踏まえ、リスク説明の伝え方やフィデューシャリー・デューティーの実践に自己過信への配慮を組み込むことが、今後より重要性を増すと考えられる。

4.2 企業・プラットフォーム事業者に期待される役割

企業、特にプラットフォーム事業者にも、できることは多い。具体的には、広告枠などを用いた啓発活動、インフルエンサーとの連携によるリテラシー教育コンテンツの普及、SNS 上で情報の出所や信頼性を可視化する UI の工夫などが挙げられる。Yamaguchi et al.（2025）は、動画共有サービスの利用時間が長い人ほど誤情報を真偽判断しにくい傾向があることを指摘しており、こうした媒体特性を踏まえた啓発設計は、特に効果が期待できる。

また、誤情報の経済的動機（広告収益）を断つ仕組みも引き続き重要である。広告主や広告ネットワークが誤情報配信メディアへの広告出稿を抑制する取り組みは、世界的にも進展してきた。プラットフォーム事業者が業界横断で連携することで、こうした取り組みの効果は一段と高まると考えられる。コミュニティガイドラインの透明化、第三者ファクトチェック機関との連携、生成 AI で作成されたコンテンツへのラベル付け、選挙や災害時の関連コンテンツの収益化停止、拡散前にリンク先を読んだかどうかの確認、信頼性スコアの開示など、講じうる施策は多岐にわたる。

プラットフォーム事業者の責任のあり方をめぐる議論は、海外では既に制度化が進んでいる。欧州連合のデジタルサービス法（DSA）は、大規模プラットフォームに対し、システミックリスクへの対応や透明性報告を義務づけている。日本でも、情報流通プラットフォーム対処法が 2025 年 4 月 1 日に施行され、制度整備が進んだ。表現の自由と安全のバランスに目を向けつつ、事業者の自主的取り組みと、適切な公的枠組みの双方が、誤情報対策の基盤をかたちづくっていくことになる。

5．生成AI時代に向けた対策の再設計

5.1「らしさ」を量産する時代と求められる構え

ここまで論じてきた誤情報の構造は、生成 AI の登場によって質的に変化しつつある。生成 AI が普及したことで、誤情報の制作コストは劇的に下がった。専門家風の語り口、もっともらしい数値、論文調の構成を備えたコンテンツが、わずか数秒で大量に生成できる。本物と区別がつかない画像や動画が、選挙や市場混乱の場面で繰り返し問題化している。特に懸念されるのが、有名人や経営者の偽動画を用いた詐欺投資の広告、政治家の発言を改変した偽音声、大量生成された「水増しレビュー」や「らしさ」のある偽ニュースサイトなどである。これらは個別の事案として目立つだけでなく、情報空間全体への信頼を継続的に損なっていく性質を持つ。

この変化は、本稿で論じてきた「自信過剰の罠」をさらに深刻化させる方向に働きうる。生成 AI が作り出す「らしさ」のあるコンテンツは、専門家風の体裁を備えるほど、自信のある受け手の検証行動を解除してしまうからである。これまでも誤情報を見抜くのは難しかったが、生成 AI 時代にはその難度がさらに上がっていく。対策の設計は、この前提を受け入れたうえで組み立てていく必要がある。

5.2 リテラシー教育の進化ー謙虚さを組み込んだ設計へー

生成 AI 時代の到来によって、リテラシー教育の重要性はこれまで以上に高まっている。同時に、本稿でみてきた知見は、教育の効果を最大化するために、設計の段階で押さえるべきポイントがあることを示している。

第一に、知識の伝授だけでなく、「自分も誤情報に騙されうる」という認識、すなわち「謙虚さ」を意図的に組み込むことが重要である。具体的には、自分

の真偽判定の正答率をフィードバックする教材、自信と実際の能力のギャップを体感させるワークショップ、ファクトチェックの限界を含めて教える啓発などが考えられる。「正解を覚える」型の教育と並行して、「自分の判断を疑う視点」を育てる教育を組み合わせていくことで、リテラシー教育はより深い効果を発揮するようになる。

第二に、リテラシー教育を進めるなかで、能力の伸びに見合わない自信だけが先行してしまうことのないよう、設計面での配慮も求められる。Yamaguchi et al.（2025）の知見は、能力を伴わない自信の上昇が拡散リスクを高める可能性を示唆している。教育の効果をより確実なものにするために、自己評価とフィードバックの仕組みを丁寧に組み込んでいくことが望まれる。これは教育を否定するものではなく、むしろ教育の質を高めるための論点である。

第三に、教育の場面においては、「受信」と「発信」の両面を視野に入れる必要がある。私たちはしばしば、メディア情報リテラシーというと「だまされない力」「見分ける力」だけを思い浮かべがちだが、現代の情報環境では、それだけでは不十分である。なぜなら、私たちは情報の受信者であり、同時に発信者でもあるからだ。SNS やメッセージアプリが当たり前になった社会では、誰もが日常的に情報を送り出している。投稿するつもりがなくても、シェアやリポスト、コメント一つで情報の流れに加担することになる。リテラシー教育は、この両面に対応する形で進化させていく必要がある。

第四に、媒体ごとの特性に応じた設計が重要となる。ソーシャルメディアでは、情報の信頼性スコアや出所情報を UI 上で可視化する施策が効きやすい。X のコミュニティノートはそのような対策の 1 つだろう。「正しいと信じた情報が拡散される」という構造を踏まえれば、信じる前に客観情報を提示することが拡散抑制につながる。また、メッセージアプリや対面の口コミについては、情報の伝え方そのものを工夫する啓発も有効である。Tanihara and Yamaguchi（2023）がファクトチェック提示の逆効果を指摘したように、注意喚起が逆効果になりうる現象も生じうる。こうしたリスクを踏まえ、「事実を肯定形で示す」「誤情報を再生産しない伝え方」を提案するメディア情報リテラシー教育が、家庭や学校の現場でますます重要となる。

5.3 技術と教育の組み合わせ、情報的健康への接続

技術的アプローチとして、コンテンツの来歴情報（Content Provenance）を標準化し、生成 AI によって作られたコンテンツに自動的にラベルを付与する取り組みが、国際的に進展している。C2PA（Coalition for Content Provenance and Authenticity）などの枠組みがその代表例である。ただし、これらは万能ではない。来歴情報は意図的に削除されうるし、検知技術は生成技術との「いたちごっこ」を続けている。さらに、検知結果が「生成 AI の可能性が高い」と表示されただけで、内容の真偽とは無関係に発信者の信頼が損なわれるリスクもある。技術的対策は、真偽を断定する装置としてではなく、立ち止まって考えるための合図として位置づけるのが現実的である。

より広い視座を提供するのが、鳥海・山本（2022）が提唱する「情報的健

康（Informational Health）」という概念である。食品の栄養成分表示のように、情報摂取のバランスを可視化し、ユーザーが自身の情報環境を客観視できる仕組みを社会に組み込むこのアプローチは、本稿のリテラシー観と深く呼応する。誤情報を完璧に見抜くことではなく、情報空間との関係を健全に保つことを目標に据える、新しい枠組みである。情報的健康の発想は、特定のコンテンツの真偽判定に終始してきた従来のファクトチェック中心の議論を、構造的な視点へと押し広げる可能性を持つ。アテンション・エコノミーが感情を刺激する情報を優遇するなかで、利用者が自身の情報摂取の偏りに気付ける仕組みを社会基盤として整えていく。本稿の知見に引き寄せて言えば、これは「自分の判断は偏っているかもしれない」という気付きを支える、社会的なインフラの構築でもある。

政策設計の観点からは、三つの方向が重要となる。第一に、リテラシー教育の積極的な拡充である。日本のメディア情報リテラシー教育は諸外国と比較してまだ十分とは言えず、学校教育・社会人教育の両面で一段の充実が求められる。第二に、プラットフォーム事業者を含むマルチステークホルダー間の連携である。一国・一機関で完結する問題ではないため、国際的な枠組みも視野に入れる必要がある。第三に、技術と教育の組み合わせである。コンテンツの来歴情報や生成 AI 検知技術といった技術的補助と、人間の判断力を育てる教育とを、相互補完的に組み合わせていくことが望ましい。リテラシー教育が中核を担い、それを技術と制度が支える。この三層構造こそが、生成 AI 時代の情報空間を支える基盤になると考えられる。

6．おわりに－謙虚さという社会的資本－

本稿では、筆者がこれまで重ねてきた一連の研究、とりわけ Yamaguchi et al.（2025）の知見を起点に、リテラシーと誤情報の関係を整理してきた。研究蓄積を総合的にみると、メディア情報リテラシー教育が誤情報対策の重要な柱として確かに機能することが浮かび上がる。客観的なリテラシーは、誤情報の拡散を抑える社会的な防波堤として有効に機能し、真偽判断段階でも一定の役割を果たしていることが示唆される。同時に、能力に見合わない自信の高まりは、真偽判断と拡散の双方を悪化させる方向に作用しうるという、逆説的な構造も見えてくる。

これらは、リテラシー教育の意義をあらためて裏付けるものである。そのうえで、教育の効果を一段引き出すための設計指針も与えてくれる。対策設計においては、リテラシー教育を中核に据えつつ、能力の育成と謙虚さの涵養を両立させていく方向が望ましい。「自分は騙されうる」という認識を社会全体が共有する。それは個人にとっても市場にとっても民主主義にとっても、ひとつの社会的資本になりうると、筆者は考えている。

これからの政策・教育・市場設計に向けては、本稿が示した三つの構造、すなわち真偽判断の難しさ、拡散段階での客観リテラシーの有効性、そして主観的自信のもたらす逆説的リスクを、共通の出発点として共有することが望まし

い。リテラシー教育を一層推進しつつ、それを補完する仕組みを社会のあちこちに組み込んでいく。研究者・行政・企業・教育機関、そして金融機関や市場参加者それぞれが、それぞれの場で「学ぶ仕組み」と「立ち止まる仕掛け」を作っていく。その積み重ねが、情報空間の安全性を支えていく。

生成 AI 時代の情報空間は、これからも複雑さを増していくと予想される。完全な防衛はおそらく不可能であろう。しかし、メディア情報リテラシー教育の地道な蓄積と、それを支える制度・技術・文化の整備によって、社会全体の情報的レジリエンスを高めていくことは、十分に可能なはずである。誤情報問題は、技術や制度だけで解決する課題ではない。一人ひとりの認識のあり方と、それを支える社会の仕組みが問われている。本稿が、そのための小さな手がかりになれば幸いである。

謝辞

本稿は、Yamaguchi et al.（2025）を主たる手がかりとして再構成したものである。同論文の共著者である大島英隆氏、渡辺智暁氏、逢坂裕紀子氏、谷原つかさ氏、井上絵理氏、田邊新之助氏に深く感謝申し上げる。本稿の基礎となった一連の研究は、JSPS 科研費 JP21K12586 の助成、JST ムーンショット型研究開発事業（JPMJMS2215）の支援、JST、RISTEX、JPMJRS23L2 の支援、ならびに Google Japan の研究助成を受けて実施されたものである。また、本稿で参照した Innovation Nippon シリーズの調査では、一般社団法人セーファーインターネット協会のご協力を賜った。最後に、本寄稿の機会を頂いた SBI 金融経済研究所、ならびに本特集をご紹介いただいた関係各位に、心より御礼申し上げる。

参考文献

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives, 31*(2), 211–236.

Carvalho, C., Klagge, N., & Moench, E. (2011). The persistent effects of a false news shock. *Journal of Empirical Finance, 18*(4), 597–615.

Jones-Jang, S. M., Mortensen, T., & Liu, J. (2021). Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American Behavioral Scientist, 65*(2), 371–388.

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science, 359*(6380), 1094–1096.

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), Article e2019527118.

Matsumoto, M., Hashimoto, R., Suzuki, M., Murayama, Y., & Izumi, K. (2025). Impact of

information disparity between individual investors on profits of meme stocks using an artificial market simulation approach. *Journal of Computational Social Science, 8*, Article 25.

Neudert, L.-M., Kollanyi, B., & Howard, P. N. (2017). *Junk news and bots during the German parliamentary election: What are German voters sharing over Twitter?* Project on Computational Propaganda. https://ora.ox.ac.uk/objects/uuid:33f7d85c-a25e-4796-aebf-a0e601fc84c1

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science, 31*(7), 770–780.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39–50.

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality, 88*(2), 185-200.

Tanihara, T., & Yamaguchi, S. (2023). Literacy is necessary to understand fact-checking: An empirical research using survey experiments. *Journal of Socio-Informatics, 16*(1), 33–46.

Ullah, S., Massoud, N., & Scholnick, B. (2014). The impact of fraudulent false information on equity values. *Journal of Business Ethics, 120*(2), 219–235.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151.

Vraga, E. K., & Tully, M. (2021). News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication & Society, 24*(2), 150–166.

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.

Yamaguchi, S., & Tanihara, T. (2025). Relationship between misinformation spreading behaviour and true/false judgments and literacy: An empirical analysis of COVID-19 vaccine and political misinformation in Japan. *Global Knowledge, Memory and Communication, 74*(1/2), 35–53.

久原恵子・井上尚美・波多野誼余夫 (1983)「批判的思考力とその測定」『読書科学』27, 4, pp.131-142.

小寺敦之 (2017)「メディア・リテラシー測定尺度の作成に関する研究」『東洋英和女学院大学人文・社会科学論集』34, pp.89-106.

坂本旬 (2022)『メディアリテラシーを学ぶ：ポスト真実世界のディストピアを超えて』大月書店 .

鳥海不二夫・山本龍彦 (2022)『デジタル空間とどう向き合うか：情報的健康の実現をめざして』日経 BP 日本経済新聞出版 .

中西良文・廣岡秀一・横矢祥代 (2006)「動機づけと社会的クリティカルシンキングとの関連：大学生の「感じる力」と「考える力」」『三重大学教育学部附属教育実践総合センター紀要』26, pp.57-66.

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 .

————・谷原吏・大島英隆 (2023)『Innovation Nippon 2022：偽・誤情報、陰謀論の実態と求められる対策』グーグル合同会社・国際大学グローバル・コミュニケーション・センター , pp.1-191. https://www.glocom.ac.jp/activities/project/8839

————・渡辺智暁・逢坂裕紀子・谷原吏・大島英隆・井上絵理・田邊新之助 (2024)『Innovation

Nippon 2024：偽・誤情報、ファクトチェック、教育啓発に関する調査研究』一般社団法人セーファーインターネット協会・国際大学グローバル・コミュニケーション・センター, pp.1-335. https://www.glocom.ac.jp/activities/project/9439