

# 生成 AI の学術、ニュース、法律分野におけるコンテンツ検索と引用能力の差異

Gemini Deep Research

## はじめに：コンテンツ検索における生成 AI の台頭と引用正確性の重要な必要性

近年、生成 AI は、学術、ニュース、法律といった多様な分野において、情報検索のあり方を大きく変貌させています。これらの AI モデルは、膨大なデータセットに基づいてテキストを生成する能力を持ち、ユーザーは自然言語で質問することで、関連性の高い情報や洞察を得ることが可能です。しかし、これらの分野においては、情報の信頼性と正確性が極めて重要であり、生成 AI が提供するコンテンツの引用の正確性は、その有用性を評価する上で不可欠な要素となります。不正確な引用は、学術研究においては信頼性の失墜、ニュース報道においては誤情報の拡散、法律分野においては法的な誤判断につながる可能性があります。

本報告書では、生成 AI が学術、ニュース、法律の各分野においてコンテンツを検索し、引用する能力にどのような違いがあるのかを詳細に調査します。各分野における引用の正確性や信頼性、一般的に使用される情報源、そして主要な生成 AI モデル（ChatGPT、Claude、Gemini など）の性能比較を行います。さらに、生成 AI がコンテンツを引用する際の正確性と信頼性に影響を与える要因、そして各分野で正確な引用を行う上での課題と、その解決に向けた取り組みについても考察します。本報告書は、この急速に進化する技術の現状を理解し、各分野における生成 AI の適切な利用を検討する上で、専門家や研究者にとって有益な情報を提供することを目的としています。

## 学術分野における生成 AI のコンテンツ検索と引用

### 学術研究における引用の正確性と信頼性

学術研究においては、引用は単に情報源を示すだけでなく、研究の透明性、信頼性、そして既存の研究との関連性を示すための基盤となります。厳格な引用スタイル（APA、MLA、シカゴスタイルなど）の遵守、参考文献の正確性、そして誤情報の検出能力は、学術論文の質を保証する上で不可欠です<sup>1</sup>。しかし、生成 AI が学術的なコンテンツを検索し引用する能力には、いくつかの重要な課題が存在することが研究によって示されています。

複数の研究が、生成 AI が学術的な文脈において、驚くほど高い割合で架空または不正確な参考文献を生成する傾向があることを示唆しています<sup>2</sup>。ある研究では、生成 AI

によって生成された参考文献のわずか7%が、真正かつ正確であったと報告されています<sup>2</sup>。別の研究では、AIが提案した参考文献の存在率はわずか14%であったとされています<sup>2</sup>。さらに別の研究では、59件の参考文献のうち69%が捏造されたものであったという結果も報告されています<sup>2</sup>。これらの調査結果は、生成AIが学術コンテンツを引用する際の信頼性に深刻な懸念を投げかけています。研究間で捏造された参考文献の割合にばらつきがあることは、使用されたプロンプト、データセット、AIモデルのバージョンが異なる可能性を示唆しており、問題の大きさが一定ではないものの、一貫して重大であることを示しています。

特に注目すべき事例として、ChatGPTが部分的に正確でありながら最終的には捏造された引用を生成した例が挙げられます<sup>3</sup>。この例では、実在する教授と学術雑誌の名前が使用されているにもかかわらず、論文のタイトルとページ番号は存在しないものでした。このような巧妙な誤情報は、その分野に詳しくないユーザーにとっては検出が難しく、厳密な検証の必要性を強調しています。AIは、著者や雑誌に関するメタデータにアクセスできるものの、それらを特定の記事に正確に結びつけることに失敗している可能性があり、真の学術文献の理解というよりも、パターンマッチングに基づいたアプローチを採用していることが示唆されます。

生成AIにおける「ハルシネーション」と呼ばれる現象も、学術的な引用の正確性を損なう大きな要因です<sup>2</sup>。ハルシネーションとは、AIがもつともらしいものの、実際には不正確、捏造、または無意味な情報を生成することを指します。これには、存在しない事実の捏造や、架空の参考文献の作成が含まれます<sup>4</sup>。法律分野におけるハルシネーションの例として、AIツールが実在しない判例を引用するケースが報告されています<sup>4</sup>。学術分野においても同様に、AIが架空の著者、論文タイトル、または出版情報を生成する可能性があります。AIは意図的に嘘をついているわけではなく、単にパターンに基づいてテキストを予測していますが、生成された情報の正確性や真実性を理解しているわけではありません<sup>4</sup>。

このような状況を踏まえ、学術的な目的で生成AIを利用する際には、ユーザーが批判的かつ調査的なアプローチを採用することが不可欠です。AIによって生成された情報は、信頼できる情報源（大学図書館のカタログや検索エンジンなど）を用いて独立して検証する必要があります<sup>2</sup>。AIはあくまで出発点として捉え、最終的な情報の信頼性は人間の専門家によって評価されるべきです。

### 学術分野で一般的に使用される情報源の種類

学術研究は、多様な情報源に基づいて行われ、それぞれの情報源は、その性質、信頼性、および情報の深さにおいて独自の特徴を持っています<sup>6</sup>。学術分野で一般的に使用

される主要な情報源には、以下のようなものがあります。

- **学術雑誌 (Scholarly Journals)** : 特定の学術分野における専門家によって書かれた論文が定期的に掲載される出版物です<sup>6</sup>。これらの論文は通常、査読 (Peer Review) と呼ばれる厳格な審査プロセスを経ており、その内容の信頼性が保証されています<sup>7</sup>。学術雑誌は、最新の研究成果や深い分析を提供する上で重要な情報源です<sup>6</sup>。
- **査読付き論文 (Peer-Reviewed Articles)** : 学術雑誌に掲載される論文の多くは査読を受けています。査読とは、同じ分野の他の専門家が論文の内容を評価し、その質、方法論、および結論の妥当性を検証するプロセスです<sup>7</sup>。査読を経た論文は、学術的な信頼性が高いとされています。
- **学術書籍 (Academic Books)** : 特定のトピックに関する包括的かつ詳細な情報を提供する書籍です<sup>6</sup>。学術書籍は、人文科学分野の研究において特に重要であり、社会科学や自然科学分野においても深い背景情報を提供します<sup>6</sup>。
- **会議録 (Conference Proceedings)** : 学術会議やシンポジウムで発表された研究論文をまとめたものです<sup>7</sup>。会議録は、最新の研究動向や未発表の研究成果を知る上で貴重な情報源となります。特に、科学技術分野においては、研究が最初に会議で発表されることが多いため、その重要性は高いと言えます<sup>7</sup>。
- **博士論文・修士論文 (Theses and Dissertations)** : 大学院の学位取得のために行われた研究成果をまとめたものです<sup>6</sup>。これらの論文は、特定のトピックに関する詳細な調査と分析を含んでおり、新しい学術的貢献を示すものとして評価されます<sup>6</sup>。
- **研究データベース (Research Databases)** : 学術雑誌の記事、書籍、会議録、論文などの情報を検索可能な形式で集めたものです<sup>6</sup>。JSTOR、Google Scholar、PubMed Central などは、学術研究に不可欠なツールであり、特定の研究テーマに関する文献を効率的に見つけることができます<sup>9</sup>。
- **参考図書 (Reference Works)** : 辞書、事典、ハンドブックなど、特定の情報や背景知識を迅速に得るための資料です<sup>6</sup>。これらは、研究の初期段階でトピックの概要を把握したり、専門用語の意味を確認したりする際に役立ちます<sup>6</sup>。

これらの情報源は、それぞれ異なる目的と特徴を持っており、研究者は自身の研究テーマや目的に合わせて適切な情報源を選択する必要があります。生成 AI がこれらの多様な情報源を理解し、正確に引用するためには、高度な自然言語処理能力と、各情報源の特性に関する深い知識が求められます。

## ChatGPT、Claude、Gemini の学術引用における比較分析

主要な生成 AI モデルである ChatGPT、Claude、Gemini は、言語処理能力と知識ベー

スにおいてそれぞれ異なる特徴を持っています<sup>12</sup>。これらのモデルが学術的な文脈でどのようにコンテンツを検索し、引用するのかを比較分析することは、その能力と限界を理解する上で重要です。

ChatGPT は、会話タスクや創造的な文章作成に優れており、広範な知識を持つことで知られています<sup>12</sup>。Gemini は、マルチモーダルなインタラクションとリアルタイムのデータ分析に強みがあり、Google の豊富なデータエコシステムとの統合が特徴です<sup>12</sup>。Claude は、高度な推論と詳細な分析作業に焦点を当てており、複雑な問題解決能力が高いとされています<sup>12</sup>。Gemini は、Google の膨大なデータへのアクセスにより、事実に基づいた回答を提供する能力が高いとされています<sup>12</sup>。

学術分野におけるこれらのモデルの性能を直接比較した研究は限られていますが、関連する研究からいくつかの示唆が得られます。例えば、心血管薬理学に関する質問応答の精度を評価した研究では、ChatGPT-4 が多肢選択式と短答式の両方の質問において最も高い精度を示し、Gemini は複雑な多肢選択式問題と短答式問題において限界が見られました<sup>15</sup>。この結果は、ChatGPT が特定の学術分野において、質問応答の精度が高い可能性があることを示唆していますが、これは引用の正確性を直接示すものではありません。

別の研究では、地理空間タスクにおける ChatGPT-4、Gemini、Claude-3、Copilot の性能を比較し、モデル間で正答率に大きな差があることがわかりました<sup>17</sup>。この研究は、AI モデルの性能がタスクの種類によって大きく異なる可能性を示唆しており、学術的な引用という特定のタスクにおいても、モデルごとに得意不得意が存在する可能性があります。

ChatGPT に統合された SearchGPT は、リアルタイムの検索機能と明確な引用を提供することを謳っています<sup>18</sup>。このような機能は、学術的な引用の精度向上に貢献する可能性があります。既存の研究で示されているような高い割合の捏造された参考文献の問題をどの程度解決できるかは、さらなる検証が必要です。

これらの比較分析から、現時点では、特定の学術分野における質問応答の精度において ChatGPT が比較的高い性能を示す可能性があるものの、学術的な引用の正確性に関しては、どのモデルも課題を抱えていることが示唆されます。特に、既存の研究で指摘されているようなハルシネーションの問題は、依然として大きな懸念事項です。

### 生成 AI による正確な学術引用における課題と取り組み

生成 AI が学術的な文脈で正確な引用を行う上での最大の課題の一つは、ハルシネーション、つまり捏造された情報や参考文献の生成です<sup>19</sup>。現在の生成 AI モデルは、トレ

ーニングデータに含まれるパターンに基づいてテキストを予測するよう設計されており、必ずしも事実の正確性を保証するものではありません<sup>19</sup>。ChatGPT のようなモデルは、明示的かつ検証可能なデータ表現を持たず、学習したパラメータの組み合わせを通じて情報にアクセスするため、真の「知識」を持っているわけではありません<sup>19</sup>。この根本的な違いが、不正確な引用や架空の参考文献の生成につながる可能性があります。

また、多くの生成 AI モデルは、インターネットに接続されていないか、最新の情報でトレーニングされていないため、最新の研究成果や情報にアクセスできないという制約もあります<sup>21</sup>。これにより、古い情報や不完全な情報に基づいて回答したり、存在しないはずの最新の参考文献を生成したりする可能性があります。

さらに、生成 AI モデルは、トレーニングデータの偏りを反映する可能性があり、特定のトピックや情報源に偏った引用を行うことがあります<sup>19</sup>。また、引用スタイルに関する深い理解や、文脈に応じた適切な引用形式の選択も、現在の AI モデルにとっては難しい課題です。

これらの課題に対処するため、いくつかの取り組みが行われています。一つは、より大規模で、より厳選されたデータセットで AI モデルをトレーニングすることです。これにより、モデルがより正確な情報を学習し、ハルシネーションの発生を抑制することが期待されます。また、学術研究に特化した AI ツールの開発も進められています<sup>5</sup>。これらのツールは、信頼性の高い学術情報源に焦点を当てるように設計されており、より正確な引用を支援する可能性があります。例えば、Consensus や Elicit といったツールは、研究に基づいた情報源のみに基づいて回答を提供することを目指しています<sup>5</sup>。

しかし、現時点では、生成 AI が生成する学術的な引用を完全に信頼することはできません。そのため、ユーザーは常に批判的な目で AI の出力を評価し、図書館のカタログや学術データベースなどの信頼できる情報源を用いて、引用の正確性を検証する必要があります<sup>5</sup>。学術的な誠実さを保つためには、AI を研究プロセスの一部として活用する際に、その利用方法を透明にすることが重要です。

## ニュース分野における生成 AI のコンテンツ検索と引用

### ニュース報道における引用の正確性と信頼性

ニュース報道において、引用は事実の裏付けを示すとともに、情報源の透明性を確保し、読者の信頼を得るための重要な要素です<sup>24</sup>。正確な情報源の特定、適切な文脈での引用、そして偏りのない報道の実現は、ジャーナリズムの倫理と質を維持する上で不可欠です。しかし、生成 AI がニュースコンテンツを検索し引用する能力には、学術分野

と同様に、いくつかの課題が存在します。

コロンビア・ジャーナリズム・レビューの Tow Center for Digital Journalism が行った最近の研究では、ニュース検索に使用される 8 つの AI 駆動型検索ツールをテストした結果、60%以上のクエリでニュースソースに関する不正確な情報が生成されたことが明らかになりました<sup>24</sup>。この研究では、AI モデルがしばしばリンクを捏造したり、配信された記事やコピーされたバージョンの記事を引用したり、さらには出版社が Web クローラのアクセスを拒否する robots.txt プロトコルを無視する傾向があることが指摘されています<sup>24</sup>。

特に問題視されているのは、AI モデルが不確かな情報を持っている場合に、回答を拒否するのではなく、もっともらしいものの不正確または推測的な回答を提供する傾向です<sup>24</sup>。テストされたプラットフォーム間でエラー率にはばらつきがあり、Perplexity が 37%のクエリで不正確な情報を提供したのに対し、Grok 3 は 94%という非常に高いエラー率を示しました<sup>24</sup>。ChatGPT Search も 67%のクエリで誤った記事を特定しました<sup>26</sup>。

さらに、AI チャットボットの応答はしばしば自信に満ちており、ユーザーが正確な情報と不正確な情報を区別することを困難にする可能性があります<sup>24</sup>。このような根拠のない自信は、ユーザーに信頼性と正確性の錯覚を与える危険性があります<sup>24</sup>。

AI が信頼できるブランド（BBC など）を情報源として引用する場合、たとえその情報が不正確であっても、読者はその回答を信頼しやすいという問題も指摘されています<sup>24</sup>。これは、不正確な AI の引用が、信頼できるニュース機関に対する信頼を損なう可能性があることを示唆しています。

また、AI 検索ツールは URL を捏造する傾向があり、ユーザーが情報源を検証することを困難にする可能性があります。Grok 2 は、特定の記事ではなく、出版社のホームページにリンクすることが多く、Gemini と Grok 3 の応答の半数以上が、エラーページにつながる捏造または破損した URL を引用していました<sup>24</sup>。

### ニュース分野で一般的に使用される情報源の種類

ニュース分野では、速報性、信頼性、そして幅広い読者へのリーチが重要となるため、特定の種類の情報源が一般的に使用されます<sup>28</sup>。

- **新聞 (Newspapers)** : 地域、国内、および国際的なニュースや出来事を報道する主要な情報源です<sup>28</sup>。ニューヨーク・タイムズ、ワシントン・ポスト、ウォール・ストリート・ジャーナルなどの全国紙や、各地域の地方紙が含まれます<sup>28</sup>。新聞

は、最新の出来事に関する詳細な情報、分析、および意見を提供します<sup>6</sup>。

- **ニュース雑誌 (News Magazines)** : タイム、ニューズウィーク、エコノミストなどのニュース雑誌は、週刊または月刊で発行され、特定の出来事やトピックに関するより深い分析や背景情報を提供します<sup>11</sup>。
- **通信社 (News Agencies)** : AP 通信、ロイター、共同通信などの通信社は、世界中のニュースを収集し、それを新聞社、放送局、オンラインニュースサイトなどのメディア機関に配信します<sup>30</sup>。通信社は、速報性のあるニュースを迅速かつ広範囲に提供する役割を担っています。
- **オンラインニュースウェブサイト (Online News Websites)** : CNN.com、BBC.com、TheGuardian.com などの主要なニュース機関のウェブサイトは、最新のニュース、ビデオ、オーディオコンテンツを提供し、世界中の読者にアクセス可能です<sup>28</sup>。デジタル専用のニュース出版物も増加しています<sup>31</sup>。
- **放送局 (Broadcasting Stations)** : テレビやラジオのニュース放送は、視覚的および聴覚的な形式でニュースを提供し、広範な視聴者や聴取者にリーチします<sup>11</sup>。
- **ニュース監視サービス (News Monitoring Services)** : これらのサービスは、さまざまなニュースソースからの情報を集約し、特定のトピックやキーワードに関する最新の情報を追跡するために使用されます<sup>28</sup>。
- **政府ウェブサイトとデータソース (Government Websites and Data Sources)** : 政府機関のウェブサイトは、政策、統計、およびその他の公式情報を提供し、データジャーナリズムにおいて重要な情報源となります<sup>32</sup>。Data.gov や Google Public Data Explorer などが例として挙げられます<sup>32</sup>。

ジャーナリストは、これらの多様な情報源を利用してニュース記事を作成し、事実確認を行い、読者に正確で信頼できる情報を提供することを目指しています<sup>30</sup>。生成 AI がニュース分野で効果的に機能するためには、これらの情報源の信頼性を評価し、適切な文脈で引用する能力が不可欠です。

### ChatGPT、Claude、Gemini のニュース引用とファクトチェックにおける比較分析

ニュース分野における ChatGPT、Claude、Gemini の性能を比較した研究では、これらのモデルがニュース記事の検索と引用において一様に課題を抱えていることが示されています<sup>24</sup>。Tow Center for Digital Journalism の研究では、テストされたすべてのチャットボットがニュースを正確に引用することに苦労しており、高いエラー率と、捏造された URL のような問題が共通して見られました<sup>24</sup>。

ただし、各モデルにはいくつかの違いも見られます。Gemini は、Google 検索を通じてリアルタイムの情報にアクセスし、検索結果と一貫性のある回答を提供できるとされています<sup>14</sup>。この機能は、理論的には最新のニュースを検索し引用する上で有利に働く

可能性がありますが、実際のテスト結果では、Gemini も他のモデルと同様に引用の正確性に課題を抱えていることが示されています<sup>24</sup>。

ファクトチェックと誤情報識別の能力に関しては、AI は有望なツールとなり得る可能性が示唆されています<sup>33</sup>。AI は、信頼できる情報源と情報を迅速に照合することで、事実確認を支援することができます<sup>33</sup>。政治的な誤情報の検出に関する研究では、ファインチューニングなしでも GenAI モデルが中程度の性能を発揮することが示されています<sup>34</sup>。しかし、別の研究では、LLM が誤った見出しを正確に識別できたとしても、必ずしもユーザーの見出しの正確性を識別する能力が向上するわけではなく、場合によっては悪影響を及ぼす可能性も指摘されています<sup>35</sup>。

Gemini は、最新の情報へのアクセスと優れた引用能力を持つと一部で報告されています<sup>36</sup>。しかし、これは Tow Center の研究結果とは矛盾しており、AI モデルの能力に関する評価にはばらつきがあることを示しています。

全体として、現時点では、ChatGPT、Claude、Gemini を含む主要な生成 AI モデルは、ニュース分野におけるコンテンツの検索と引用において、依然として大きな課題を抱えていると言えます。特に、引用の捏造、不正確な情報源の特定、そして自信に満ちた誤った回答の提供は、ニュースの信頼性を損なう深刻な問題です。

### 生成 AI による正確なニュース引用における課題と解決策

生成 AI がニュース分野で正確な引用を行う上での主な課題は、学術分野と同様に、ハルシネーションや不正確な情報の自信に満ちた提供です<sup>23</sup>。AI モデルは、誤った情報であっても、あたかも事実であるかのように提示する傾向があり、ユーザーがその誤りを認識することを困難にする可能性があります<sup>23</sup>。

また、AI モデルが発行者の robots.txt の指示を無視し、アクセスを許可されていないコンテンツから情報を取得したり、元の情報源ではなく、配信されたコンテンツを引用したりする問題も指摘されています<sup>24</sup>。これは、著作権や情報源の適切な帰属に関する倫理的な懸念を引き起こします。

さらに、AI は人間のようなメディアリテラシーを持たないため、ニュースソースの信頼性や偏りを評価することが難しいという課題もあります<sup>26</sup>。人間のジャーナリストは、情報源の評判、過去の正確性、潜在的な偏りなどを考慮して情報を評価しますが、現在の AI モデルは、このような複雑な判断を行うことができません。

これらの課題に対する解決策としては、ジャーナリズムの用途に特化した AI モデルの開発、引用の正確性に焦点を当てたトレーニングデータの改善、そして人間のファクト



チェッカーを AI のワークフローに統合することが考えられます<sup>33</sup>。AI は、人間のジャーナリストの仕事を完全に置き換えるのではなく、その活動を支援するツールとして活用されるべきです<sup>33</sup>。

最終的には、ニュースの消費者自身が、AI によって生成された情報に対して懐疑的な姿勢を持ち、複数の信頼できる情報源から情報を検証することが重要です<sup>23</sup>。AI はニュースの検索や要約に役立つ可能性がありますが、その出力は常に批判的に評価されるべきです。

## 法律分野における生成 AI のコンテンツ検索と引用

### 法律研究と実務における引用の正確性と信頼性

法律分野においては、正確かつ精密な法律の権威（判例、法令、規則など）の引用は、極めて重要です。誤った引用は、法的な議論の根拠を揺るがし、訴訟の結果に重大な影響を与える可能性があります<sup>4</sup>。生成 AI が法律コンテンツを検索し引用する能力は、その潜在的な有用性ととも、重大な責任と倫理的考慮事項を伴います。

法律分野における生成 AI の利用において最も懸念される問題の一つは、ハルシネーション、特に架空の判例の引用です<sup>39</sup>。2023 年の *Mata v. Avianca* 事件では、弁護士が AI に盲目的に依存した結果、完全に捏造された判例を含む訴訟文書を提出し、裁判所から前例のない虚偽の法的権威の提出として厳しく非難されました<sup>41</sup>。また、スタンフォード大学の教授が、AI を使用して作成した専門家の意見書に架空の引用を含めていた事例も報告されています<sup>40</sup>。これらの事例は、法律分野における AI の不正確な引用が、専門家の信頼性を損ない、法的手続きに深刻な悪影響を与える可能性を示しています。

法律分野におけるハルシネーションには、「グラウンデッドネス (groundedness)」の問題も含まれます<sup>4</sup>。これは、AI が事実としては正しい情報を提供し、関連性の高い判例を引用したとしても、その判例の内容を誤って解釈したり、現在の法的問題に不適切に適用したりするケースを指します。法律の引用においては、単に正しい情報源を見つけるだけでなく、その内容を正確に理解し、適切に適用することが不可欠です。

弁護士は、AI ツールを使用する際に、能力 (competence)、注意義務 (diligence)、正確性 (accuracy)、裁判所に対する誠実さ (candor to the tribunal)、および顧客情報の守秘義務 (confidentiality) を含む多くの倫理的義務を負っています<sup>4</sup>。AI によって生成された情報を検証せずに裁判所に提出したり、誤った法的助言を提供したりすることは、これらの義務に違反する可能性があります<sup>39</sup>。特に、AI が生成する判例や法令の引用は、必ず人間が確認し、その正確性を保証する必要が

あります<sup>39</sup>。

生成 AI は、契約書や訴訟文書の作成、法律調査の要約、契約レビューなど、さまざまな法律業務を支援する可能性があります<sup>4</sup>。しかし、現在の AI は、高度なテキスト予測に基づいており、法律概念の真の理解や論理的推論の能力には限界があります<sup>42</sup>。そのため、AI の出力を盲信することなく、人間の専門家による監督と検証が不可欠です。

## 法律分野で一般的に使用される情報源の種類

法律分野における情報源は、主に一次資料と二次資料に分類されます<sup>45</sup>。

- **一次資料 (Primary Sources)** : 法そのものを記したものであり、以下のものが含まれます<sup>45</sup>。
  - **判例 (Case Law)** : 裁判所の判決や意見。連邦裁判所および州裁判所の判例が含まれます<sup>45</sup>。
  - **法令 (Statutes)** : 議会や州議会によって制定された法律<sup>45</sup>。連邦法典や各州の法律が含まれます。
  - **規則・規制 (Regulations)** : 行政機関によって制定された規則や規制<sup>45</sup>。
  - **憲法 (Constitutions)** : 連邦憲法および州憲法。
  - **条例 (Ordinances)** : 地方自治体によって制定された条例<sup>45</sup>。
- **二次資料 (Secondary Sources)** : 法律を解説、分析、批評するものであり、以下のものが含まれます<sup>45</sup>。
  - **法律百科事典 (Legal Encyclopedias)** : 法的トピックに関する包括的な概要を提供します。Corpus Juris Secundum (C.J.S.) や American Jurisprudence 2d (Am. Jur.2d) などがあります<sup>46</sup>。
  - **法律論文 (Law Reviews)** : 法律の専門家や学者が書いた、特定の法律トピックに関する詳細な研究論文が掲載されます<sup>45</sup>。
  - **法律解説書 (Legal Treatises)** : 特定の法律分野に関する深く掘り下げた解説を提供します<sup>45</sup>。
  - **法律辞典 (Legal Dictionaries)** : 法律用語の意味や用法を解説します。Black's Law Dictionary などがあります<sup>46</sup>。
  - **法律の再記述 (Restatements of the Law)** : 特定の法分野における一般的な法原則をまとめたものです<sup>45</sup>。
  - **ナッツシェルとホーンブック (Nutshells and Hornbooks)** : 特定の法律トピックに関する簡潔な概要や入門書です<sup>45</sup>。

法律研究者は、これらの情報源を利用して法的問題を分析し、法的議論を構築します。

生成 AI が法律分野で効果的に機能するためには、これらの情報源の種類と階層構造を理解し、それぞれの信頼性と重要性を評価する能力が求められます。

## ChatGPT、Claude、Gemini の法律引用と法的推論における比較分析

ChatGPT、Claude、Gemini は、法律分野においてもさまざまな応用が試みられています<sup>4</sup>。これらのモデルは、契約書の作成支援、法律調査の要約、契約レビュー、法律翻訳などのタスクを支援する可能性があります<sup>4</sup>。ChatGPT に統合された SearchGPT は、リアルタイムの検索機能を提供し、明確な引用を伴う回答を生成することを謳っています<sup>18</sup>。Lexis+ AI は Claude 2 を使用し、Westlaw Precision は GPT-4 を使用するなど、法律情報プラットフォームはこれらのモデルを活用しています<sup>39</sup>。

しかし、法律分野におけるこれらのモデルの引用精度と法的推論能力には、いくつかの懸念があります<sup>52</sup>。一般的な AI モデルは、法律固有のワークフローや、大量の PDF ドキュメントの取り込み、複雑な法的分析の構造化された作成には最適化されていません<sup>52</sup>。また、最新の法律データベースへのアクセスがない場合や、情報の正確性を検証するツールがない場合、ハルシネーションのリスクが高まります<sup>52</sup>。

ある意見では、一般的な AI モデルは、法律文書の作成において、正確な参照や特定の法律形式への準拠が難しく、存在しない判例や法令を捏造するリスクがあると指摘されています<sup>52</sup>。また、管轄区域固有の専門知識も欠けている可能性があります<sup>52</sup>。

法律調査ツールとして開発された Lexis AI や Westlaw CoCounsel などの専門的な AI ツールは、より信頼性の高い引用機能を提供すると考えられます<sup>39</sup>。これらのツールは、広範な法律データベースと統合されており、法律専門家の独自の要件を理解しているため、不正確な応答の頻度を減らすことができるとされています<sup>39</sup>。

法的推論に関しては、現在の生成 AI モデルは、事実の重要性を文脈に応じて判断したり、仮説的な状況を分析したりする能力に限界があります<sup>53</sup>。法律的な議論の文脈で結論を導き出す演繹的推論能力もまだ発展途上であると言えます<sup>42</sup>。

## 生成 AI による正確な法律引用における課題と倫理的考慮事項

生成 AI が法律分野で正確な引用を行う上での最大の課題は、ハルシネーションによる架空または不正確な法律引用の生成であり、これは深刻な法的影響を及ぼす可能性があります<sup>41</sup>。法律分野は、判例という拘束力のある先例に大きく依存しているため、捏造された判例の引用は、法的議論全体の信頼性を損なうこととなります<sup>41</sup>。

弁護士は、AI を使用する際に、能力、守秘義務、裁判所に対する誠実さなどの倫理的義務を遵守する必要があります<sup>41</sup>。AI によって生成されたすべての法律コンテンツ

は、裁判所に提出する前に、弁護士自身が綿密に検証し、正確であることを保証しなければなりません<sup>41</sup>。これには、AI が引用した判例や法令の存在と内容を確認することが含まれます<sup>41</sup>。

AI ツールに入力されたクライアントの機密情報が、AI プロバイダーによってトレーニングデータとして使用されるリスクも、重要な倫理的考慮事項です<sup>41</sup>。弁護士は、クライアントデータを匿名化したり、法律専門家向けに設計された、データセキュリティを優先するツールを選択したりするなど、適切な対策を講じる必要があります<sup>41</sup>。

法律分野における AI の利用は、法的および規制上の監視の対象となる可能性が高く、裁判所や弁護士会は、AI の責任ある利用を確保するためのガイドラインや規則を策定しています<sup>40</sup>。弁護士は、これらの進化する規制を常に把握し、遵守することが求められます。

最終的に、法律分野における生成 AI の利用は、人間の専門家の知識と判断を補完するものであり、代替するものではありません<sup>44</sup>。AI は、法律調査や文書作成の効率を高める可能性があります。その出力は常に人間の弁護士によって批判的に評価され、検証される必要があります。

## 分野横断的比較：学術、ニュース、法律分野における生成 AI の引用能力の比較分析

学術、ニュース、法律の各分野における生成 AI の引用能力を比較すると、いくつかの重要な違いと共通点が見られます。

領域	一般的な正確性/信頼性 (研究に基づく)	主な課題	利用に関する示唆
学術	研究によると、捏造された参考文献の割合が高い <sup>2</sup> 。ハルシネーションが一般的 <sup>2</sup> 。	特定の引用スタイルの遵守、架空の情報源の捏造、誤情報の検出 <sup>1</sup> 。	厳格な検証が必要。信頼できる情報源との照合が不可欠 <sup>2</sup> 。AI は出発点として捉えるべき <sup>5</sup> 。
ニュース	調査によると、60%以上のクエリで不正	自信に満ちた誤った回答、発行者の	情報源の多角的な検証が重要 <sup>23</sup> 。AI の出

	確な情報源が示される <sup>24</sup> 。URLの捏造や配信記事の引用も問題 <sup>24</sup> 。	robots.txtの無視、メデアリテラシーの欠如 <sup>24</sup> 。	力を批判的に評価する必要 <sup>23</sup> 。人間のジャーナリストの役割が不可欠 <sup>26</sup> 。
法律	架空の判例引用による深刻な事例が存在 <sup>39</sup> 。判例の誤解釈も問題 <sup>4</sup> 。	法的権威の正確かつ精密な引用、ハルシネーションによる捏造、倫理的義務（能力、守秘義務、誠実さ）の遵守 <sup>4</sup> 。	すべてのAI生成コンテンツに対する人間の監督と検証が不可欠 <sup>39</sup> 。法律専門家向けの専門ツールの利用が推奨される <sup>39</sup> 。進化する規制への対応が必要 <sup>40</sup> 。

学術分野では、厳格な引用スタイルと参考文献の正確性が求められるのに対し、生成AIは架空の参考文献を高い割合で生成する傾向があります。ニュース分野では、速報性が重視される一方で、AIは情報源を誤って引用したり、URLを捏造したりする問題が顕著です。法律分野では、引用の誤りが法的判断に直接影響を与えるため、AIによるハルシネーションは特に深刻な問題となります。各分野において、AIが正確な引用を行う上での難しさは、それぞれの分野の情報源の特性、引用の目的、そして求められる信頼性のレベルによって異なります。

## 生成AIの引用の正確性と信頼性に影響を与える要因

生成AIの引用の正確性と信頼性は、いくつかの要因によって影響を受けます。

- **トレーニングデータの質と性質:** AIモデルが学習するデータの質と量が、その出力の正確性に直接影響します<sup>54</sup>。偏ったデータや不完全なデータでトレーニングされたAIは、不正確な情報や偏った引用を生成する可能性があります<sup>54</sup>。ChatGPTのようなモデルは、膨大な量のテキストデータでトレーニングされていますが、そのデータには誤った情報や古い情報が含まれている可能性もあります<sup>54</sup>。GoogleのGeminiは、「Infiniset」と呼ばれる大規模なデータセットを使用していますが、その詳細についてはあまり公開されていません<sup>54</sup>。
- **基盤となるアルゴリズムとモデルアーキテクチャ:** 生成AIのテキスト予測は、統計的な重み付けに基づいており、必ずしも論理的推論や概念の理解を伴うものではありません<sup>4</sup>。この根本的な仕組みが、AIが真の理解なしにパターンに基づいてテキストを生成し、結果として不正確な引用につながる可能性があります<sup>42</sup>。
- **ユーザーのプロンプトと指示:** ユーザーがAIに与えるプロンプトの明確さや具体

性も、引用の正確性に影響を与える可能性があります<sup>4</sup>。明確な指示を与えることで、AI はより関連性の高い情報を検索し、適切な形式で引用する可能性が高まります。しかし、プロンプトだけで AI の根本的な限界を克服することはできません。

- **文脈理解、ニュアンス、および引用スタイル固有の要件の理解における現在の AI の限界:** AI は、人間が自然に行うような文脈の理解や、微妙なニュアンスの解釈に苦勞することがあります<sup>4</sup>。引用スタイルには、情報源の種類や引用の位置によって複雑なルールが存在しますが、現在の AI モデルは、これらのルールを完全に理解し、正確に適用することが難しい場合があります<sup>55</sup>。

## 結論と今後の方向性：より正確で信頼性の高い AI 駆動型コンテンツ検索と引用に向けて

本報告書では、生成 AI が学術、ニュース、法律の各分野においてコンテンツを検索し引用する能力の違いを調査しました。その結果、各分野において引用の正確性と信頼性には大きな課題が存在し、その原因は AI モデルの根本的な限界、トレーニングデータの質、そして各分野固有の引用要件の複雑さに起因することが明らかになりました。

学術分野では、生成 AI は高い割合で架空の参考文献を生成する傾向があり、厳格な引用スタイルの遵守も課題です。ニュース分野では、AI は情報源を誤って引用したり、URL を捏造したりする問題が顕著であり、信頼できる情報源に対する信頼を損なう可能性があります。法律分野では、引用の誤りは法的判断に直接影響を与えるため、AI によるハルシネーションは特に深刻な問題となります。

これらの課題に対処するためには、AI モデルの改善、専門分野に特化したツールの開発、そして AI の利用に関する明確なガイドラインとベストプラクティスの確立が必要です。また、生成 AI の出力を常に批判的に評価し、人間の専門家による検証を行うことが不可欠です。

将来的には、より高度な AI モデル、より質の高いトレーニングデータ、そして分野固有の知識と引用ルールを組み込んだ専門的な AI ツールの開発により、生成 AI のコンテンツ検索と引用能力は向上する可能性があります。しかし、現時点では、生成 AI はあくまで人間の知識と判断を補完するツールとして捉え、その出力を盲信することなく、常に批判的な視点を持つことが重要です。特に、学術、ニュース、法律といった信頼性が重視される分野においては、AI の利用は慎重に進められるべきであり、人間の専門家による厳格な監督と検証が不可欠です。

## 引用文献

1. Citation and Attribution - Generative Artificial Intelligence, 3 月 24, 2025 にアクセス、 <https://libguides.brown.edu/c.php?g=1338928&p=9868287>
2. Why does generative AI produce fake references? - FAQs, 3 月 24, 2025 にアクセス、 <https://rasmussen.libanswers.com/faq/403604>
3. AI and Academic Research: Assessing Generative AI Outputs - Research & Course Guides, 3 月 24, 2025 にアクセス、 <https://libguides.stthomas.edu/ai/assessing>
4. Navigating Generative AI in Legal Practice - Oklahoma Bar Association, 3 月 24, 2025 にアクセス、 <https://www.okbar.org/barjournal/december-2024/navigating-generative-ai-in-legal-practice/>
5. Verifying and Citing Generative AI - Generative AI Tools for Students - BU Libraries, 3 月 24, 2025 にアクセス、 <https://library.bu.edu/gen-ai/verifying-citing>
6. Types of Information Resources - Finding Information - Library Guides at Brown University, 3 月 24, 2025 にアクセス、 <https://libguides.brown.edu/findinginformation/types>
7. Different Types of Sources - Understanding & Evaluating Sources - LibGuides, 3 月 24, 2025 にアクセス、 <https://nmsu.libguides.com/sources/types>
8. Primary, Secondary, and Tertiary Sources of Information in the Sciences - LibGuides at University of Wisconsin Stevens Point, 3 月 24, 2025 にアクセス、 <https://libraryguides.uwsp.edu/InformationSourcesInTheSciences>
9. 10 Best Online Academic Research Tools and Resources | Best Colleges, 3 月 24, 2025 にアクセス、 <https://www.bestcolleges.com/blog/best-online-resources-academic-research/>
10. libguides.brown.edu, 3 月 24, 2025 にアクセス、 [https://libguides.brown.edu/findinginformation/types#:~:text=Dictionaries%20and%20glossaries%20\(meanings%20of,atlases%2C%20bibliographies%2C%20and%20more\)](https://libguides.brown.edu/findinginformation/types#:~:text=Dictionaries%20and%20glossaries%20(meanings%20of,atlases%2C%20bibliographies%2C%20and%20more))
11. Types of Sources - Metropolitan Community College - Library, 3 月 24, 2025 にアクセス、 <https://www.mccneb.edu/Community-Business/Library/Research-101/Types-of-Sources>
12. ChatGPT vs Gemini vs Claude: The Best AI Model Compared - Kanerika, 3 月 24, 2025 にアクセス、 <https://kanerika.com/blogs/chatgpt-vs-gemini-vs-claude/>
13. Best AI Chatbots Compared: ChatGPT vs Gemini vs Claude - Ajelix, 3 月 24, 2025 にアクセス、 <https://ajelix.com/ai/best-ai-chatbots-compared-chatgpt-vs-gemini-vs-claude/>
14. Comparison of Generative AI: ChatGPT vs. Claude vs. Gemini - SEO Zoom, 3 月 24, 2025 にアクセス、 <https://www.seozoom.com/best-ai-comparison-chatgpt-gemini-claude/>
15. Artificial intelligence in healthcare education: evaluating the accuracy of ChatGPT, Copilot, and Google Gemini in cardiovascular pharmacology - Frontiers, 3 月 24, 2025 にアクセス、

- <https://www.frontiersin.org/journals/medicine/articles/10.3389/fmed.2025.1495378/full>
16. (PDF) Artificial intelligence in healthcare education: evaluating the accuracy of ChatGPT, Copilot, and Google Gemini in cardiovascular pharmacology - ResearchGate, 3 月 24, 2025 にアクセス、  
[https://www.researchgate.net/publication/389133057\\_Artificial\\_intelligence\\_in\\_healthcare\\_education\\_evaluating\\_the\\_accuracy\\_of\\_ChatGPT\\_Copilot\\_and\\_Google\\_Gemini\\_in\\_cardiovascular\\_pharmacology](https://www.researchgate.net/publication/389133057_Artificial_intelligence_in_healthcare_education_evaluating_the_accuracy_of_ChatGPT_Copilot_and_Google_Gemini_in_cardiovascular_pharmacology)
  17. Correctness Comparison of ChatGPT-4, Gemini, Claude-3, and Copilot for Spatial Tasks Hartwig H. Hochmair\*, Levente Juhász, a - arXiv, 3 月 24, 2025 にアクセス、 <https://arxiv.org/pdf/2401.02404>
  18. Best Legal AI Tools for 2025 - Pocketlaw, 3 月 24, 2025 にアクセス、  
<https://pocketlaw.com/content-hub/legal-ai-tools>
  19. Generative AI in Academic Research: Perspectives and Cultural Norms, 3 月 24, 2025 にアクセス、 <https://research-and-innovation.cornell.edu/generative-ai-in-academic-research/>
  20. Challenges and Possibilities of Generative AI - Generative Artificial Intelligence - LibGuides at University of California San Diego, 3 月 24, 2025 にアクセス、  
<https://ucsd.libguides.com/c.php?g=1322935&p=9734831>
  21. Using Generative AI in Research: Limitations & Warnings, 3 月 24, 2025 にアクセス、 <https://libguides.usc.edu/generative-AI/limitations>
  22. Concerns about Generative Artificial Intelligence | Alliant International University Center for Teaching Excellence, 3 月 24, 2025 にアクセス、  
<https://cte.alliant.edu/concerns-about-generative-artificial-intelligence/>
  23. Fact-Checking Generative AI and Avoiding Plagiarism - Unity Environmental University, 3 月 24, 2025 にアクセス、 <https://unity.edu/commhub/using-generative-ai/fact-checking-generative-ai-and-avoiding-plagiarism/>
  24. AI Search Has A Citation Problem - Columbia Journalism Review, 3 月 24, 2025 にアクセス、 [https://www.cjr.org/tow\\_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php](https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php)
  25. neuron.expert, 3 月 24, 2025 にアクセス、 <https://neuron.expert/news/ai-search-engines-cite-incorrect-sources-at-an-alarming-60-rate-study-says/11700/en/#:~:text=A%20recent%20study%20by%20the,incorrect%20information%20about%20news%20sources.>
  26. AI search engines cite incorrect sources at an alarming 60% rate, study says - Reddit, 3 月 24, 2025 にアクセス、  
[https://www.reddit.com/r/Futurology/comments/ljbvgpb/ai\\_search\\_engines\\_cite\\_incorrect\\_sources\\_at\\_an/](https://www.reddit.com/r/Futurology/comments/ljbvgpb/ai_search_engines_cite_incorrect_sources_at_an/)
  27. Report Finds AI Tools Are Not Good at Citing Accurate Sources | Social Media Today, 3 月 24, 2025 にアクセス、  
<https://www.socialmediatoday.com/news/report-ai-tools-fail-correct-citations-queries/742236/>



28. WRIT 101 Research Guide: News & Journalism Library Guides at Montana State University, 3 月 24, 2025 にアクセス、  
<https://guides.lib.montana.edu/c.php?g=862440&p=8947108>
29. Find News Sources - Journalism - Research Guides at Temple University, 3 月 24, 2025 にアクセス、  
<https://guides.temple.edu/journalism/news>
30. Chapter 59: Sources of information - The News Manual, 3 月 24, 2025 にアクセス、  
[https://www.thenewsmanual.net/Manuals%20Volume%203/volume3\\_59.htm](https://www.thenewsmanual.net/Manuals%20Volume%203/volume3_59.htm)
31. 6.5 Journalistic Sources – Information Strategies for Communicators - Publishing Services - University of Minnesota, 3 月 24, 2025 にアクセス、  
<https://open.lib.umn.edu/infostrategies/chapter/6-5-journalistic-sources/>
32. Data Sources - Journalism & Media - LibGuides at University of Texas at Austin, 3 月 24, 2025 にアクセス、  
<https://guides.lib.utexas.edu/c.php?g=530875&p=10524341>
33. Artificial Intelligence and Generative AI for Media & Journalism - LibGuides at University of North Carolina at Chapel Hill, 3 月 24, 2025 にアクセス、  
<https://guides.lib.unc.edu/generativeAI/ai-journalism>
34. GenAI vs. Human Fact-Checkers: Accurate Ratings, Flawed Rationales - arXiv, 3 月 24, 2025 にアクセス、  
<https://arxiv.org/html/2502.14943v2>
35. Fact-checking information from large language models can decrease headline discernment | PNAS, 3 月 24, 2025 にアクセス、  
<https://www.pnas.org/doi/10.1073/pnas.2322823121>
36. How Does Claude Compare to ChatGPT and Gemini Advance? : r/ClaudeAI-Reddit, 3 月 24, 2025 にアクセス、  
[https://www.reddit.com/r/ClaudeAI/comments/1fmtr0u/how\\_does\\_claude\\_compare\\_to\\_chatgpt\\_and\\_gemini/](https://www.reddit.com/r/ClaudeAI/comments/1fmtr0u/how_does_claude_compare_to_chatgpt_and_gemini/)
37. Generative AI Reliability and Validity - AI Tools and Resources - LibGuides at University of South Florida Libraries, 3 月 24, 2025 にアクセス、  
<https://guides.lib.usf.edu/c.php?g=1315087&p=9678779>
38. Evaluating Information Sources: Generative AI and ChatGPT - UBC Library Research Guides - The University of British Columbia, 3 月 24, 2025 にアクセス、  
<https://guides.library.ubc.ca/GenAI/Evaluating>
39. Legal research tools in the age of generative AI - ABA Journal, 3 月 24, 2025 にアクセス、  
<https://www.abajournal.com/columns/article/legal-research-tools-in-the-age-of-generative-ai>
40. The Irony – Using Generative AI in a Case About the Dangers of Generative AI, 3 月 24, 2025 にアクセス、  
<https://www.bracewell.com/resources/the-irony-using-generative-ai-in-a-case-about-the-dangers-of-generative-ai/>
41. Generative AI in Legal Practice: A Survey of Professional and Ethical Challenges by Texie Montoya | Idaho State Bar, 3 月 24, 2025 にアクセス、  
<https://isb.idaho.gov/blog/generative-ai-in-legal-practice-a-survey-of-professional-and-ethical-challenges-by-texie-montoya/>

42. Generative AI and the Duty of Competence Conundrum | Insights - Holland & Knight, 3 月 24, 2025 にアクセス、  
<https://www.hklaw.com/en/insights/publications/2025/02/generative-ai-and-the-duty-of-competence-conundrum>
43. The Implications of Generative AI: From the Delivery of Legal Services to the Delivery of Justice | IAALS, 3 月 24, 2025 にアクセス、  
<https://iaals.du.edu/blog/implications-generative-ai-delivery-legal-services-delivery-justice>
44. Generative AI in the legal industry: is accuracy everything? | by Jack Shepherd | Medium, 3 月 24, 2025 にアクセス、  
<https://jackwshepherd.medium.com/generative-ai-in-the-legal-industry-is-accuracy-everything-60c6f3a15947>
45. General Information - Legal Research for Beginners - Guides at Texas State Law Library, 3 月 24, 2025 にアクセス、 <https://guides.sll.texas.gov/legal-research-for-beginners>
46. Law Library Basics: Legal Research - Secondary Sources - LibGuides, 3 月 24, 2025 にアクセス、 <https://nyulaw.libguides.com/c.php?g=773842&p=5551762>
47. Free and Low-Cost Legal Research: U.S. Secondary and Primary Sources, 3 月 24, 2025 にアクセス、 <https://guides.law.stanford.edu/free-lowcost/primary>
48. Free Legal Research Resources - United States: Secondary Sources, 3 月 24, 2025 にアクセス、  
<https://guides.library.harvard.edu/c.php?g=310432&p=2072004>
49. Exploring Generative and Extractive AI in Legal Research: A Comparison of Leading Models, 3 月 24, 2025 にアクセス、  
<https://law.unh.libguides.com/blog/Exploring-Generative-and-Extractive-AI-in-Legal-Research-A-Comparison-of-Leading-Models>
50. Comparing Claude, ChatGPT, and Gemini: Which AI Tool is Best? - Coffee Sprints, 3 月 24, 2025 にアクセス、 <https://coffeesprints.com/blog/claude-chatgpt-gemini/>
51. ChatGPT Search and How it Benefits Lawyers - Clio, 3 月 24, 2025 にアクセス、  
<https://www.clio.com/blog/chat-gpt-search/>
52. Can ChatGPT Give Legal Advice? Why Purpose-Built AI is Reshaping LegalTech, 3 月 24, 2025 にアクセス、 <https://callidusai.com/can-chatgpt-give-legal-advice-why-purpose-built-ai-is-reshaping-legaltech/>
53. Wondering about AI in legal: r/legaltech - Reddit, 3 月 24, 2025 にアクセス、  
[https://www.reddit.com/r/legaltech/comments/li44fw7/wondering\\_about\\_ai\\_in\\_legal/](https://www.reddit.com/r/legaltech/comments/li44fw7/wondering_about_ai_in_legal/)
54. ChatGPT Vs. Gemini Vs. Claude: What Are The Differences? - Search Engine Journal, 3 月 24, 2025 にアクセス、  
<https://www.searchenginejournal.com/chatgpt-vs-gemini-vs-claude/483690/>
55. Working with Artificial Intelligence Series: The Trust Factor in Generative AI (DDN2-A35), 3 月 24, 2025 にアクセス、 <https://www.csps->

[efpc.gc.ca/tools/articles/trust\\_factor-ai-eng.aspx](http://efpc.gc.ca/tools/articles/trust_factor-ai-eng.aspx)