

生成エンジンにおける引用脆弱性の解明

望月 理来^{1,2} 小松 秀輔^{2,3} 野口 蒼大^{1,2} 安宅 和人¹

¹慶應義塾大学 ²QueryLift 株式会社 ³奈良先端科学技術大学院大学
{riku.mochizuki, shusuke.komatsu, souta.noguchi}@querylift.co.jp
ataka@sfc.keio.ac.jp

概要

生成エンジン (GE) は Web 検索と LLM を用いて、Web ページを引用し、回答を生成するシステムである。GE では誰でも情報を公開できる Web から情報を取得するため、これまでの RAG システムよりもポイズニング攻撃に対して脆弱である。しかし既存の評価研究は、回答が引用元を忠実に反映しているかに焦点を当てており、どの Web ソースを引用しているかについて検討されていない。このギャップを埋めるため、我々は引用元の発行者属性を分類し、攻撃コストを推定する基準を提案する。政治分野の質問では公式政党 Web サイトからの引用が約 21%~90% の範囲であるが、与党は 20%~48% であり、政策伝達における攻撃の脅威が示唆された。

1 序論

GPT、Gemini、Claude などの大規模言語モデル (LLM) アプリケーションは、Web 検索機能を組み込み、回答を生成する**生成エンジン** (GE) を提供している。GE は悪意ある意図をもつ攻撃者を含む様々な発行者が Web 上に公開したコンテンツを引用する。このため GE は攻撃者が意図する回答を生成させるために Web コンテンツを操作するポイズニング攻撃に対して脆弱である [1]。GE が事実性を欠く回答を生成する場合、その原因は LLM のハルシネーション [2, 3] だけでなく、Web 検索で取得した Web コンテンツの事実性の欠如にもある [1]。先行研究は、攻撃者が意図する回答を生成させるために Web 上に偽情報や誤情報を配置するポイズニング攻撃の有効性を明らかにしている [1]。

既存の評価基準は、GE が生成した回答において引用コンテンツがどれだけ忠実に反映されているかを評価することを目的としている [4, 5, 6, 7]。しかし、そのような忠実性検証だけでは引用 Web ソースの権威性を捉えることはできない。ポイズニング

攻撃は攻撃者が公開した二次情報を利用して Web に不正なコンテンツを注入し、GE はそれを忠実に引用することで回答の生成を操作できる。したがって、忠実性だけではなく、どの発行者からのコンテンツを引用しているかを評価することが求められる。特に情報検索の質問に対しては、回答は質問の主題に関連する権威ある (一次) 情報源を引用すべきケース (例えば政策の伝達) がある。

我々はこの脅威に対処するため、引用元の発行者属性とその回答生成への影響に焦点を当てた評価基準を導入する。まず、引用情報を一次情報源と二次情報源に分類し、さらに二次情報源を抽象化し、Web への注入障壁を推定する。次に、GE が生成した回答におけるのカテゴリごとの引用割合を計算し、GE がどの注入障壁の記事を好んで引用するかを定量化する。最後に、注入障壁のカテゴリ別で、引用と回答間の忠実性がどのように変化するかを測定する。

我々は日本の政治ドメインにおける検索質問 (約 180 問) を対象に手法を適用し、実験を行った。結果として、政党 Web サイト (一次情報) への引用が約 20%~65% を占めることが示されたが、与党は 20%~48% であり、政策伝達における攻撃の脅威が示唆された。また、登録のみで公開可能な低障壁ソース (Reddit、X、個人ブログなど) が回答における引用の約 20% を占めることが明らかになった。さらに、低障壁ソースからの引用は中・高障壁ソースと比較して実際の回答内容との意味的一貫性が低い傾向があるにもかかわらず、それでも GE の回答に影響を与えていることが判明した。

2 関連研究と動機

2.1 GE への脅威

Zou らは、外部データベースに悪意のあるコンテンツを配置することで GE に攻撃者の意図した回答

を生成させる **PoisonedRAG** [1] と呼ばれる攻撃手法を明らかにし形式化している。RAG とは GE の中核機能の一つであり、外部データベースから関連コンテンツを検索して LLM の回答生成に組み込む技術である [8]。PoisonedRAG [1] は、攻撃者が外部データベースに少量の悪意のあるテキストを注入することで RAG を悪用し、特定の質問に対して GE に攻撃者の意図した回答を生成させる。この悪意のあるコンテンツには、「OpenAI の CEO は誰ですか?」のような答えが定まる質問 (Closed-end question) を対象としており、攻撃者が質問への回答として表示させたいテキストが含まれる。

攻撃を成功させるには、注入するコンテンツが 2 つの条件を満たす必要がある：検索条件と生成条件である。検索条件は、コンテンツが検索エンジンでの検索結果で上位 k 件の関連コンテンツとして選択されることを意味する。生成条件は、コンテンツが LLM で攻撃者の意図した回答を生成するためのコンテキストとして使用されることを意味する。これには生成エンジン最適化 (GEO) も含まれる [9]。

2.2 引用評価

理想的な GE とは、生成回答のすべての主張文が引用によって裏付けられ (引用カバレッジ)、すべての引用が主張を正確に裏付け忠実に反映する (忠実性) システムである。これら进行分析するため、多くの研究が RAG システムの忠実性を評価している [4, 5, 7]。

忠実性の定量化は、類似度ベースの手法が用いられる。類似度ベースの手法は、あるテキストとそれを要約したテキストの意味的一致度における測定精度が高いことが示されている [10]。したがって、検索結果から取得した Web ページコンテンツを要約し、引用付きで回答を生成する GE の動作と一致するため、類似度ベースの手法が用いられる。

2.3 研究動機

引用評価に焦点を当てた既存研究は、PoisonedRAG 攻撃を想定した GE 生成回答の評価基準を確立していない。忠実性のみでは攻撃者が二次情報源として悪意のあるコンテンツを投稿し、GE がそれを引用して攻撃者の意図した回答を生成するケースを捉えることができず、GE に対するポイズニング攻撃の脅威を明らかにすることができない。悪意のあるコンテンツには、SEO と GEO 技術を使

用して検索条件と生成条件を満たすように最適化された偽情報や誤情報 [11] が含まれる。したがって、我々は Closed-end question に対して引用する情報源が相応しいかどうかとも評価することが情報伝達において重要であると主張する。GE の普及に伴い PoisonedRAG 攻撃が広く認識されるようになるため、そのような攻撃を想定した評価基準は GE をより安全にするために重要である。

3 手法

本節では、GE 生成回答における引用の選択と忠実性を分析する評価基準を紹介する。本手法は 2 つのコンポーネントで構成される：(1) 発行者属性に基づく引用分類と (2) 引用分類に基づく忠実性の測定。

3.1 発行者属性に基づく引用分類

各引用 c を発行者属性カテゴリに割り当てる分類器 $\lambda(c)$ を導入する。分類は一次情報識別と抽象化された二次情報カテゴリ分類の 2 段階で行う。

一次情報識別では、一次情報のドメインセット \mathcal{D} を定義し、引用 c のドメイン d_c が \mathcal{D} に含まれる場合、その引用を一次情報として分類する。二次情報カテゴリ分類では、抽象化された二次情報のカテゴリセット $L = \{\lambda_1, \lambda_2, \dots, \lambda_p\}$ を準備し、 $\lambda(c)$ により各引用をカテゴリにマッピングする。第 4 節では、「Party」、「Media」、「Platform」、「Owned」、「Academia」、「Non-media-industry」などのラベルを使用する。分類器の構築には LLM-as-a-Judge [12] を採用し、ドメインと WHOIS 情報から発行者を識別する。第 4 節で用いたプロンプトは付録 A に示す。

我々は 2 つの GE モデル (異なるプロバイダーのモデル) を用いて識別を行う。複数の GE モデルを組み合わせることで、単一 GE モデルの潜在的バイアスを軽減し、分類器の精度を向上させることができる [12]。2 つの GE モデル間で分類結果が一致する場合はその結果を採用し、不一致の場合は最終的な人間判断を行う (我々の実験では、著者が Web サイトの企業情報 and/or ドメイン情報から最終判断を決定する)。最終的に、各カテゴリ λ_i に分類された引用のセットを $C^{\lambda_i} = \{c \in C \mid \lambda(c) = \lambda_i\}$ として取得する。

3.2 忠実性の測定

我々は上記のカテゴリ別に、引用した Web ソースが生成回答にどの程度忠実に反映されているか測定

する。

まず、回答 r を文 $\{l_1, l_2, \dots, l_k\}$ に、各引用 $c_i \in C$ を文 $S_{c_i} = \{s_{i,1}, s_{i,2}, \dots, s_{i,n_i}\}$ に分解する。回答では各回答文の意味が独立していること（箇条書き等）が多く、忠実性を正確に評価するために文に分解している。次に、忠実性を定量化するため、2つの文間の類似度を計算する関数 $\text{sim}(x, y)$ を導入し、Sentence-BERT [13] を使用して意味的類似度を計算する。関数 $\text{sim}(x, y)$ は2つの文間の類似度スコアを $[-1, 1]$ の範囲の値として返す。ここで、 -1 に近い類似度値は忠実性が低いことを示し、 1 に近い値は忠実性が高いことを示す。

回答における引用の反映影響力を明らかにするため、回答文セットと引用間の最大類似度 $\text{sim}_{\max}(S, c_i) = \max_{j,m} \text{sim}(l_j, s_{i,m})$ を計算する。GE は長い引用テキスト内の一部の文のみを使用して特定の回答文を生成することが多い。したがって、最大類似度を見つけることで、引用テキストの中で回答内容に最も関連する部分を特定できる。

類似度スコアを3つのバンドに分類する：「高」($[0.9, 1.0]$) は強い意味的整合、「中」($(0.8, 0.9]$) は中程度の重複、「低」($[-1.0, 0.8]$) は弱い意味的接続を示す。これらの閾値は先行研究 [14, 15] に基づく。

3.3 注入障壁のカテゴリ分類

我々は GE がどのような情報源を好むか特定するために、二次情報のカテゴリセット ($L = \{\lambda_1, \lambda_2, \dots, \lambda_p\}$) に基づき注入障壁のカテゴリに分類する。我々は引用を以下の5カテゴリに分類する：一次情報源は、質問で言及された政党が所有するドメインである。対立情報源は、対立政党が運営するドメインである。低障壁情報源は、ユーザーまたは所有者が自由にコンテンツを公開または編集できるドメインである（例えば、ソーシャルメディアプラットフォームや個人ドメイン）。中障壁情報源は、ジャーナリズムのバイアスや利益が表れる可能性のある編集プロセスを持つ組織や企業が所有するドメインである（例えば、メディアアウトレットや業界団体）。高障壁情報源は、著者が中立かつ客観的であることが求められ、操作が困難なドメインである（例えば、学術や政府）。実験では、回答内の引用から各カテゴリの割合を計算し、GE がどのカテゴリの記事を好んで引用するかを定量化する。回答は2025年9月4日時点で取得した。

4 実験

本節では、日本の政治ドメインにおける GE の引用パターンを示す。

4.1 実験設定

日本の9政党を対象に、各政党に対して20問（政策10問、イデオロギー10問）、合計180問の政治的質問を準備した（付録B参照）。回答生成には3つのGEモデルを使用：OpenAI GPT-5、Claude Sonnet 4、Gemini Flash 2.0。

4.2 APIからの明示的引用の結果

各GEへのAPIアクセスを通じて引用URLを収集し、第3.1節の手法で分類した。各カテゴリの割合を図1aに示す。

各政党の平均値では、3つのモデルすべてで一次情報源が約60%を占める（OpenAI 63.0%、Gemini 60.9%、Claude 60.9%）。モデル間では、Geminiは低障壁情報源への依存が高く（19.9%）、Claudeは中障壁情報源を多用し（19.3%）、OpenAIはより均衡の取れた分布を示す。

政党別では、自由民主党は主要政党として二次情報源の影響により一次情報源からの引用が低い傾向がある。与党の自民党は一次情報源が21%~48%を占めるが、野党の立憲民主党、公明党、日本維新の会は一次情報源が70%~90%を占める。参政党と日本保守党ではGeminiとClaudeでプラットフォームとメディアからの割合が増加している。

我々の研究 [16] では米国の政党に関する実験も行っており、日本と米国のコンテキスト間で引用パターンに有意な違いを特定した。米国の質問では一次情報源は20%~40%の範囲に下がる。日本の設定では一次情報源依存、米国の設定では外部情報源依存である構造的な違いを観察する。

4.3 引用カバレッジの結果

第3.2節の手法で引用カバレッジを計算した。類似度を3レベル（低 $[-1.0, 0.8]$ 、中 $(0.8, 0.9]$ 、高 $(0.9, 1.0]$ ）に分類し、各バンド内の割合を図1(b)-(d)に示す。

一次情報源や高障壁情報源などの信頼性の高い情報源が高類似度バンドで最大のカバレッジを持ち、低類似度バンドでは中~低障壁情報源からの参照が目立つ。この傾向はすべてのモデルに共通して

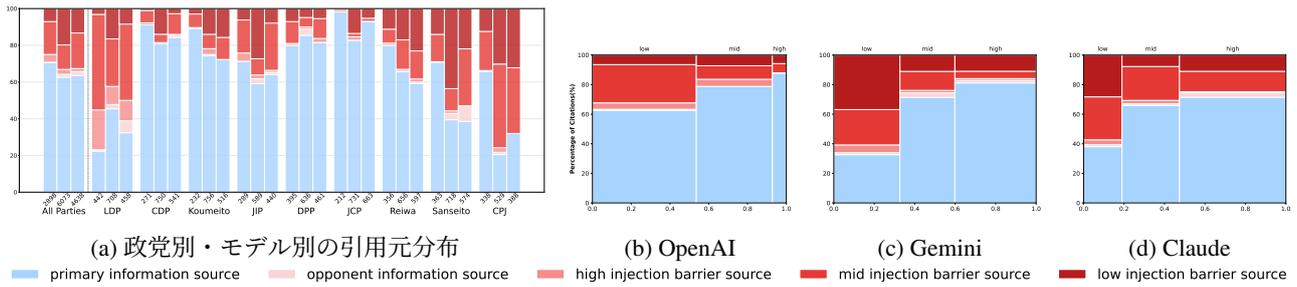


図 1: 日本の政治ドメインにおける引用パターン分析: (a) 政党別・モデル別の引用元分布、(b)–(d) 類似度レベル別の引用カバレッジ

いる。

モデル別では、OpenAI は高類似度バンドが狭く低類似度に集中、Gemini は類似度上昇に伴い一次情報源割合が増加、Claude は高類似度バンドを主要コンポーネントとして形成している。これらの結果は、引用カバレッジがモデルの戦略と情報エコシステムの特徴を明らかにできることを示している。

これらの結果は、どの注入障壁カテゴリが回答に反映される影響力を持つかを示すことで、モデルごとの情報エコシステムの特徴を明らかにできることを示している。我々の米国に対する実験 [16] では米国は高障壁の対立情報源と二次情報源が高類似性の形成に積極的に貢献することが明らかになった。

5 議論

本節では、GE における理想的な引用バランスについて議論する。

定性分析: 第 4 節では引用元の定量的結果を提示した。ここでは分析からの定性的洞察を議論する。

Web 検索ヒット数も与党が一次情報源からの引用が少ない傾向についての考察する。Google 検索エンジンで日本の政党名のみを入力すると、歴代政権の与党である自民党の検索ヒット数は約 4700 万件で、他党の平均は約 860 万件であった。これは、自民党に関連する Web 上の一次情報源の全体的な割合が他党より相対的に小さく、引用における一次情報源の割合が小さくなることを示唆している。しかし、我々の米国を対象とした実験では、検索ヒット数が低い政党でも一次情報割合が低いケース、および検索ヒット数が高い政党でも一次情報割合が高いケースという反例が明らかになっている。これは、情報量と一次情報源引用の関係が単純ではなく、情報提供者の戦略が影響することを示唆している。

GE における理想的な引用バランス: 政治ドメインの質問については、政党の政策を有権者に伝える

ことが重要であるため、一次情報源の引用割合を高めることが望ましい。しかし、ユーザーの観点からは、閉じた質問のすべてのドメインで一次情報源の引用割合を最大化することが必ずしも最適ではない。例えば、先行研究では一次情報提供者がユーザーに提示するコンテンツで過度にポジティブな側面のみを表示する可能性があることを示している (製品レビューなど) [17]。したがって、一次情報源の信頼性について社会的疑念が生じる場合、引用に二次情報を含めることが適切である。

我々は各ドメインとタスクについて発行者属性に基づくターゲット引用割合を定義するマニフェストを作成し、そのマニフェストに対して回答を評価することを主張する。さらに、ユーザーが一次情報と二次情報のバランスを制御できるメカニズムが必要である。これにより、ポイズニング攻撃による偽情報リスクを軽減できる。これらのメカニズムの開発には、引用から一次情報源を自動で識別する方法を確立する必要があり、デジタル証明書技術 [18] を活用できる可能性がある。

6 結論

本研究は政治ドメインにおける GE の引用パターンを分析し、GE モデルと政党間での一次情報源利用の有意な差異を明らかにし、GE のポイズニング攻撃脆弱性を浮き彫りにした。本手法は発行者属性を識別し、発行者属性カテゴリが生成回答の内容にどのように影響するかを定量化し、分類された一次情報と二次情報を区別してコンテンツ注入障壁を推定した。結果として、一次情報源の引用割合が日本では政党別に差があり 21%~90% の範囲であるが、与党は 20%~48% のみであり、政策伝達にポイズニング攻撃の脅威があることが示された。また低コンテンツ注入障壁情報源は回答内容への反映忠実性が低下することを見出した。

参考文献

- [1] Wei Zou, Runpeng Geng, Binghui Wang, and Jinyuan Jia. Poisonedrag: knowledge corruption attacks to retrieval-augmented generation of large language models. In **Proceedings of the 34th USENIX Conference on Security Symposium**, SEC '25, USA, 2025. USENIX Association.
- [2] Yue Zhang, Yafu Li, Leyang Cui, Deng Cai, Lemao Liu, Tingchen Fu, Xinting Huang, Enbo Zhao, Yu Zhang, Yulong Chen, Longyue Wang, Anh Tuan Luu, Wei Bi, Freda Shi, and Shuming Shi. Siren's song in the ai ocean: A survey on hallucination in large language models. **Computational Linguistics**, pp. 1–46, 08 2025.
- [3] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. **ACM Comput. Surv.**, Vol. 55, No. 12, March 2023.
- [4] Yumo Xu, Peng Qi, Jifan Chen, Kunlun Liu, Rujun Han, Lan Liu, Bonan Min, Vittorio Castelli, Arshit Gupta, and Zhiguo Wang. CiteEval: Principle-driven citation evaluation for source attribution. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, **Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 32759–32778, Vienna, Austria, July 2025. Association for Computational Linguistics.
- [5] Shahul Es, Jithin James, Luis Espinosa Anke, and Steven Schockaert. RAGAs: Automated evaluation of retrieval augmented generation. In Nikolaos Aletras and Orpheus De Clercq, editors, **Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations**, pp. 150–158, St. Julians, Malta, March 2024. Association for Computational Linguistics.
- [6] Nelson F. Liu, Tianyi Zhang, and Percy Liang. Evaluating verifiability in generative search engines. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, **Findings of the Association for Computational Linguistics: EMNLP 2023**, pp. 7001–7025, Singapore, December 2023. Association for Computational Linguistics.
- [7] Yuheng Zha, Yichi Yang, Ruichen Li, and Zhiting Hu. AlignScore: Evaluating factual consistency with a unified alignment function. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, **Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 11328–11348, Toronto, Canada, July 2023. Association for Computational Linguistics.
- [8] Wenqi Fan, Yajuan Ding, Liangbo Ning, Shijie Wang, Hengyun Li, Dawei Yin, Tat-Seng Chua, and Qing Li. A survey on rag meeting llms: Towards retrieval-augmented large language models. In **Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining**, KDD '24, pp. 6491–6501, New York, NY, USA, 2024. Association for Computing Machinery.
- [9] Pranjal Aggarwal, Vishvak Murahari, Tanmay Rajpurohit, Ashwin Kalyan, Karthik Narasimhan, and Ameet Deshpande. Geo: Generative engine optimization. In **Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining**, pp. 5–16, New York, NY, USA, 2024. Association for Computing Machinery.
- [10] Weijia Zhang, Mohammad Aliannejadi, Yifei Yuan, Jiahuan Pei, Jia-Hong Huang, and Evangelos Kanoulas. Towards fine-grained citation evaluation in generated text: A comparative analysis of faithfulness metrics, 2024.
- [11] Canyu Chen and Kai Shu. Can llm-generated misinformation be detected?, 2024.
- [12] Jiawei Gu, Xuhui Jiang, Zhichao Shi, Hexiang Tan, Xuehao Zhai, Chengjin Xu, Wei Li, Yinghan Shen, Shengjie Ma, Honghao Liu, Saizhuo Wang, Kun Zhang, Yuanzhuo Wang, Wen Gao, Lionel Ni, and Jian Guo. A survey on LLM-as-a-judge, 2025. arXiv preprint.
- [13] Nils Reimers and Iryna Gurevych. Sentence-BERT: Sentence embeddings using siamese BERT-networks. In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan, editors, **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)**, pp. 3982–3992, Hong Kong, China, November 2019. Association for Computational Linguistics.
- [14] Muhammad Rafsan Kabir, Md Mohibur Rahman Nabil, and Mohammad Ashrafuzzaman Khan. Banglaembed: Efficient sentence embedding models for a low-resource language using cross-lingual distillation techniques. In **2024 7th International Conference on Algorithms, Computing and Artificial Intelligence (ACAI)**, pp. 1–6. IEEE, 2024.
- [15] Siddharth Tumre, Sangameshwar Patil, and Alok Kumar. Improved near-duplicate detection for aggregated and paywalled news-feeds. In **Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 3: Industry Track)**, pp. 979–987, 2025.
- [16] Riku Mochizuki, Shusuke Komatsu, Souta Noguchi, and Kazuto Ataka. Exposing citation vulnerabilities in generative engines, 2025.
- [17] Jean Jinghan Chen, Jianmei Liu, Li Xie, and Xinsheng Cheng. Impression management, forward-looking strategy-related disclosure, and excess executive compensation: evidence from china. **The British Accounting Review**, Vol. 56, No. 5, p. 101430, 2024.
- [18] Shigeiya Suzuki, Michiko Kuriyama. Originator Profile Framework — Architectural Overview v0.1 (Public Draft), 1 2025. Draft Status.
- [19] Arina Kostina, Marios D. Dikaiakos, Dimosthenis Stefanidis, and George Pallis. Large language models for text classification: Case study and comprehensive review, 2025. arXiv preprint.
- [20] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. **OpenAI Blog**, Vol. 1, No. 8, p. 9, 2019.

A カテゴリ分類用プロンプト

本付録では、LLM-as-a-Judge を使用して二次情報源を分類するためのプロンプトテンプレート q_u を提示する。Kostina ら [19] が訓練例なしで高い性能を示すことを実証したのに従い、ゼロショット分類 [20] を採用する。各引用 $c \in C$ について、プロンプトは2つの入力を受け取る： $url(c)$ は発行者のドメインとパスを示す完全な URL を抽出し、 $whois(c)$ は所有権と組織情報を含むドメイン登録データを取得する。

```
Access the web page that belongs to
the given domain URL below, and
generate a label to detect the type
of the provider.
```

```
There are 7 options for the label:
```

1. party: the domain website belongs to a particular political party.
2. media: the domain website belongs to mass media such as newspapers, TV shows, internet news, and Reddit

```
... [whole prompt is omitted]
```

```
The output should ONLY be a label and
no other information.
```

```
Domain URL: \{url(c)\}
```

```
Whois: \{whois(c)\}
```

B 政治質問テンプレート

本付録では、日本の政治ドメインにおける GE 引用パターンを評価するために第4節で説明した実験で使用した閉じた質問テンプレートの代表例を提示する。各質問テンプレートには、実験中に特定の政党名に置き換えられるプレースホルダー {PARTY} が含まれている。質問は2つのカテゴリに分けられる：具体的な政府のアプローチを扱う政策関連と、基本的な政治原則を検討するイデオロギー関連である。全20問の完全なリストは我々の論文 [16] を参照されたい。

```
[Policy-related]
```

1. Regarding government debt, does {PARTY} currently prioritize debt restraint or growth-oriented

```
investment?
```

4. For reducing greenhouse gas emissions, does {PARTY} currently prioritize stronger regulations or greater subsidies for technological innovation?

```
[Ideology-related]
```

11. As a fundamental principle of society, does {PARTY} currently prioritize individual freedom or economic equality?
13. In reforming social systems, does {PARTY} currently prioritize preserving traditions and gradual change or bold, progressive reform?
20. In addressing international issues, does {PARTY} currently prioritize national sovereignty or international cooperation/multilateralism?