

# ChatGPTからの送客および送客後遷移の判定について

株式会社リタ / WebPita Labo

## 1. ChatGPT経由アクセスのWeb訪問ログについて

WebPitaでは、ChatGPT → Webサイト、および ChatGPT → Webサイト → Webサイト のアクセス遷移を接続ログから観測しています。

従来のアクセス解析では、ChatGPTからの単発アクセス判定に留まるケースが多く、送客後のサイト内遷移を把握することは困難でした。WebPitaでは、ChatGPTからWebサイトへ訪問したアクセスと、その後のサイト内回遊をあわせて記録することで、ChatGPT経由ユーザーの関心導線を可視化・把握しています。

### ※トラフィック傾向の注意点

ChatGPT経由アクセスは、通常の検索エンジン型の流入と比べて発生件数が少ない傾向があります。そのためWebPitaでは、単純な流入数のみを計上するのではなく、ChatGPTからWebサイトへ直接遷移したファーストアクセスと、その後に継続して閲覧された内部アクセスを厳密に切り分けて記録しています。

## 2. ChatGPT → Webサイトの判定方法

ChatGPTからWebサイトへ直接遷移（初回流入）したアクセスは、以下の情報をもとに多角的に判定しています。

- アクセス先URLパラメーターに `utm_source=chatgpt.com` が含まれる
- リファラー（HTTP参照元）に `chatgpt.com` が含まれる
- リファラー（HTTP参照元）に `openai.com` が含まれる

これらの条件のいずれかに該当するアクセスを、ChatGPTまたはOpenAI関連ページからの「AI送客」として扱います。

## 3. ChatGPT → Webサイト → Webサイトの判定方法

ChatGPT経由でサイトへ訪問した「その後」のサイト内遷移についても追跡記録を行っています。例えば、ChatGPT内のリンクをクリックして「ページA」にランディングし、そこからさらに「ページB」へと遷移するような一連の回遊行動（ChatGPT → ページA → ページB）です。

この送客後遷移の判定には、以下の複合条件を利用しています。

- リファラーに自サイト（WebPita等）のURLを含む
- かつ、リファラーに `utm_source=chatgpt.com` のパラメーターを引き継いでいる

これにより、AIを起点として流入したユーザーがどの程度サイトに関心を持ち、どのような導線で回遊したかを正確に把握できます。

## 4. ChatGPT経由アクセスの詳細分類

WebPitaでは、検知したChatGPT関連アクセスをその挙動・性質に応じて以下のセグメントに分類しています。

分類コード	定義・内容
Visit_us01	ChatGPT検索経由でWebサイトへ訪問（パラメーター・リファラーベース）
Visit_us02	ChatGPT経由で訪問した後のサイト内遷移（セッション回遊）
Visit_us04	<code>chatgpt.com</code> の直接的なドメインリンク経由での訪問
Visit_us05	<code>openai.com</code> の関連ドメインリンク経由での訪問
Visit_bot01~03	人間ではなく、自動巡回等のbotらしいアクセスとして判定されるもの

## 5. ChatGPTリンク表示および挙動の実証検証

WebPitaでは、ChatGPT画面上でWebPitaのURLを通常のWebリンクとして表示させ、実地検証を行いました。検証時には、ChatGPT画面上にWebPita公式サイトへの外部リンクが表示されていることを確認し、そのリンクをクリックした際の接続ログを実証ベースで取得・解析しました。

以下は、実際にChatGPT画面上でWebPitaのURLが通常の外部リンクとして正常に表示・出力された際の検証キャプチャ（エビデンス画像）です。



図1：ChatGPTがWebPita URLを通常の外部リンクとして生成・表示した検証画面

## 【検証データ：実際の接続ログ】

### 1回目：ChatGPT → Webサイトのログ

```
DATE : 2026-05-20 13:48:39
REQUEST_URI : /?utm_source=chatgpt.com
QUERY_STRING : utm_source=chatgpt.com
HTTP_REFERER : https://chatgpt.com/
HTTP_SEC_FETCH_SITE : cross-site
HTTP_SEC_FETCH_DEST : document
```

### 2回目：ChatGPT → Webサイト → Webサイト（サイト内遷移）のログ

```
DATE : 2026-05-20 13:48:49
REQUEST_URI : /labo/webPitaLabo_list.html
HTTP_REFERER : https://www.webpita.com/?utm_source=chatgpt.com
HTTP_SEC_FETCH_SITE : same-origin
HTTP_SEC_FETCH_DEST : document
```

## ログの分析：

1回目のアクセスでは、`https://chatgpt.com/` をリファラーに持ち、`HTTP_SEC_FETCH_SITE : cross-site`（クロスサイト遷移）が記録されていることから、ChatGPTから外部サイトであるWebPitaへの明確な送客が発生したことが確認できます。

2回目のアクセスでは、1回目に付与された `utm_source=chatgpt.com` を保持したリファラーから、`same-origin`（同一サイト内）への遷移が行われており、AI送客から離脱せず、サイト内回遊に至っていることが実ログから証明されています。

## 6. 観測対象ログの選別と精査

AIやAI経由のユーザーが「実際にコンテンツを消費したか」を正しく評価するため、WebPitaでは解析対象となるログを以下のように絞り込んでいます。

### ■ 判定対象とするアクセス

- HTTPステータスが `200` `206` `304`（正常応答、またはキャッシュ利用）のアクセス
- `html` `php` `pdf` 等のドキュメント・コンテンツファイルへのアクセス
- 各種ディレクトリ・フォルダURLへの直接アクセス

### ■ 判定対象から除外するアクセス

- 画像ファイル（jpg, png, svg等）、CSS、JavaScriptなどのリソース読み込み
- リダイレクトアクセス（301, 302などによる内部処理）、404エラー（存在しないページ）

## 7. 偽装アクセスの判定と排除

ChatGPT関連アクセスを模したUser-Agentの偽装や、悪質なbotトラフィックが混在するケースがあります。WebPitaでは以下の除外・分類処理を多層的に行うことで、純粋な「AI送客・閲覧」の分析精度を担保しています。

- **robots.txt / sitemap.xml アクセス判定：**

サイトのクローラー向け設定ファイルである `robots.txt` や `sitemap.xml` にアクセスしたIPについては、その後のページ巡回も含め、一律「自動クローラー（非人間）」として処理し、一般のAI送客ログから除外します。

- **User-Agent (UA) によるフィルタリング：**

`bot`、`crawler`、`spider`、`fetcher`、`headless`、`python`、`curl`、`wget` などの文字列を含むUAは、人間によるブラウジングではないと判定し、別分類として管理します。

- **ISP・ネットワーク判定：**

接続元IPアドレスを解析し、主要なレンタルサーバーやクラウド事業者、データセンターからの接続である場合はBot判定とします。海外IPについては、AbuseIPDB等の情報を組み合わせて精査します。

## 8. まとめと今後の対応

User-Agentやリファラー、IP帯域などの情報は、OpenAI側の仕様変更やサードパーティによる偽装のリスクに常に晒されています。特に生成AI分野の進化は早く、アクセス方式やパラメーターの付与形式は継続的に変更されることが予想されます。

WebPitaでは、日々蓄積される実際の観測ログと最新の公開仕様を照らし合わせ、判定ロジックのアップデートを継続的に実施してまいります。