

Octoparse

オクトパス

コード不要で、ウェブデータの収集とプロセス
自動化を簡単に実現する



目次

- 1 Octoaparseとは
- 2 活用事例
- 3 機能・料金

01

Octoparse とは

会社概要

Octoparseは、世界各国の企業や団体に対して、データ収集技術と業界別カスタマイズソリューションを提供し、データサービスをサポートしています。

ビッグデータ・AI技術開発に特化

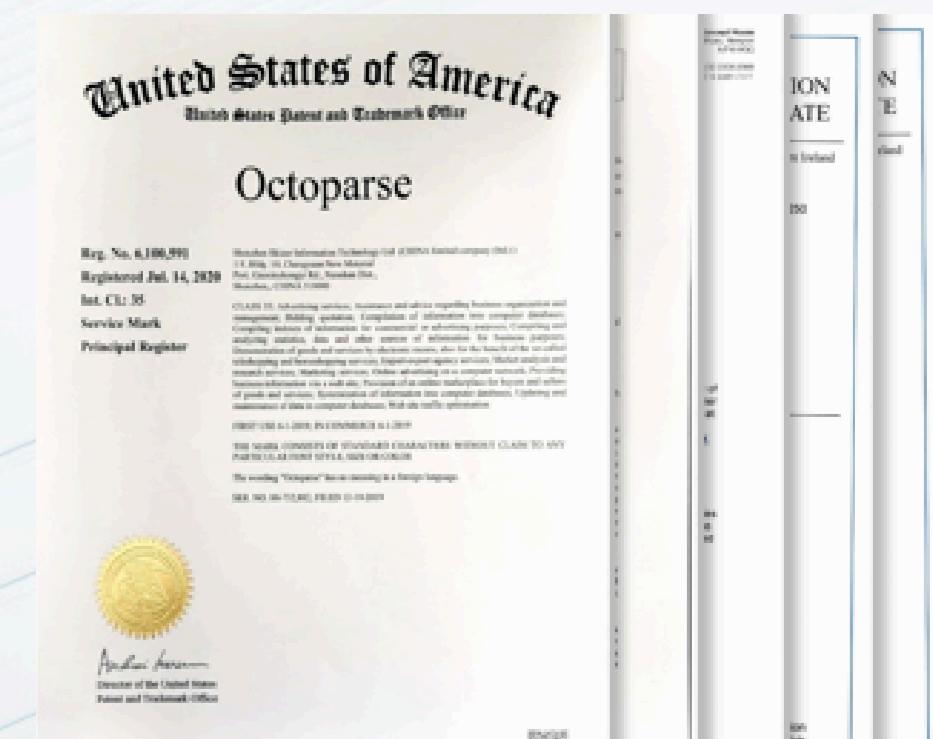
- Octoparse (スクレイピングツール)
 - データ収集代行
 - Octoparse AI (RPAツール)

データ分野における約10年の経験

- 全世界約450万ユーザー、1万社以上の企業・組織
 - 本社：アメリカ、深セン（上海、パリなどに支社）
 - 英語、日本語、スペイン語、フランス語、ドイツ語など10以上の言語をカバー

組織体制

- 創業者: Keven Liu (ビッグデータ10年経験)
 - RedpointVenturesなどから20億円超の投資獲得
 - 社員300名超、研究開発70%以上



こんなお悩みございませんか？



取得したいサイト数が百以上あるけれど、コーディングなしで1つのプラットフォーム上で設定できる方法を探している…



百万、千万規模のデータをスケジュール設定して、短時間で取得したいのですが…



サイトの構造や内容が頻繁に変わるので、自動で最新データを取得できるツールはないか…



取得したデータをリアルタイムで社内システム上で確認したいが、その機能は実現可能か…



ツールがうまく使えず、突然データが取得できなくなった場合、すぐに対応してもらえるかが心配…



Octoparseの強み



通用性

百万以上のサイトからクローリング
実績



可用性

AI自動化と充実テンプレートで
簡単操作



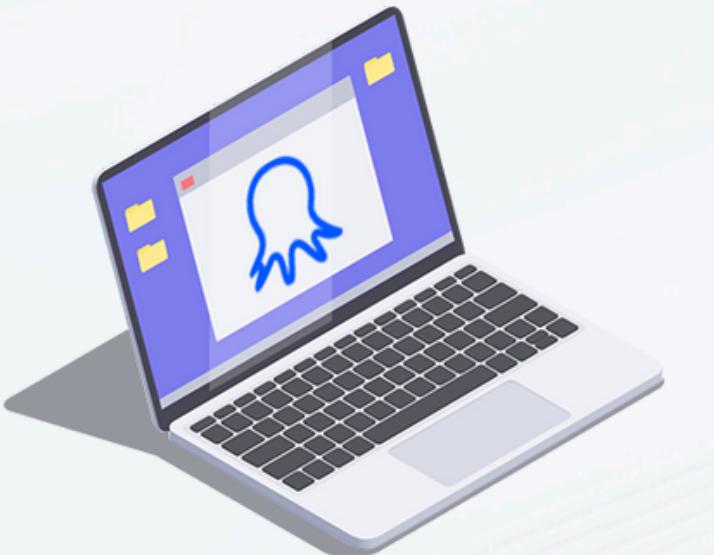
チームワーク

チーム連携と権限管理の強化
サポート



高速化

5000台以上のサーバーで
24時間並列処理



信頼性

高安定性プロトコルとデータ
品質保証システム



柔軟性

SaaS、プライベート、
ハイブリッド環境対応



インテグレーション

API・DB・Zapierなど幅広い
連携に対応

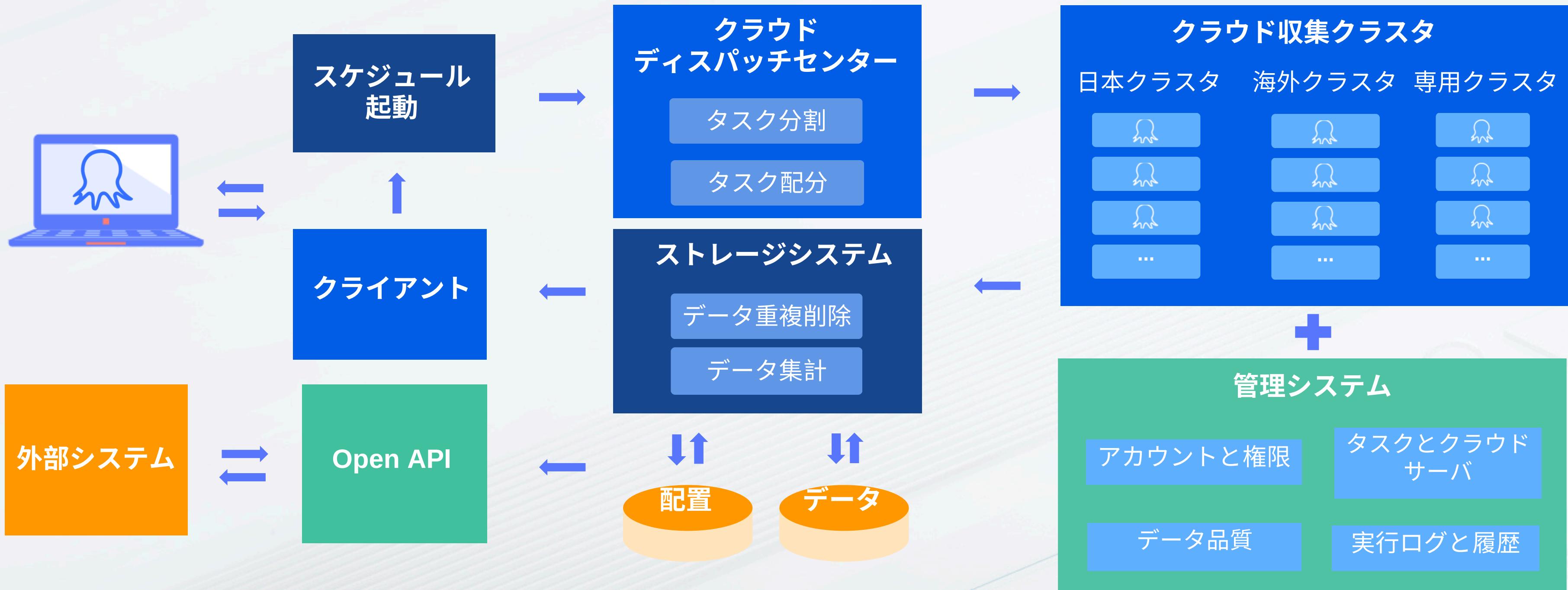


安全性

暗号化通信と多層的なセキュリティ

データ収集メカニズム

本プラットフォームはC/Sアーキテクチャを採用し、**AI認識アルゴリズム・視覚化設定・テンプレート機能**で操作性を強化。クラウド収集・高匿名プロキシIP・自動キャプチャ対策により高精度データを取得し、Google Sheets・Excel・HTML・CSV・JSON・DB出力と**API自動連携**を実現します。



ユーザー向け機能

AI認識アルゴリズム搭載

動的ウェブコンテンツとデータストレージを自動識別・可視化し、ページダウンロード機能を標準実装

アンチブロック対策

高可用性プロキシIPプールとブラウザプロファイル自動切替により、持続的なデータ収集を保証

非構造データの統合処理

正規表現を用いたテキスト・メディアデータの構造化変換と統一フォーマット出力を実現

スケジュール型自動収集

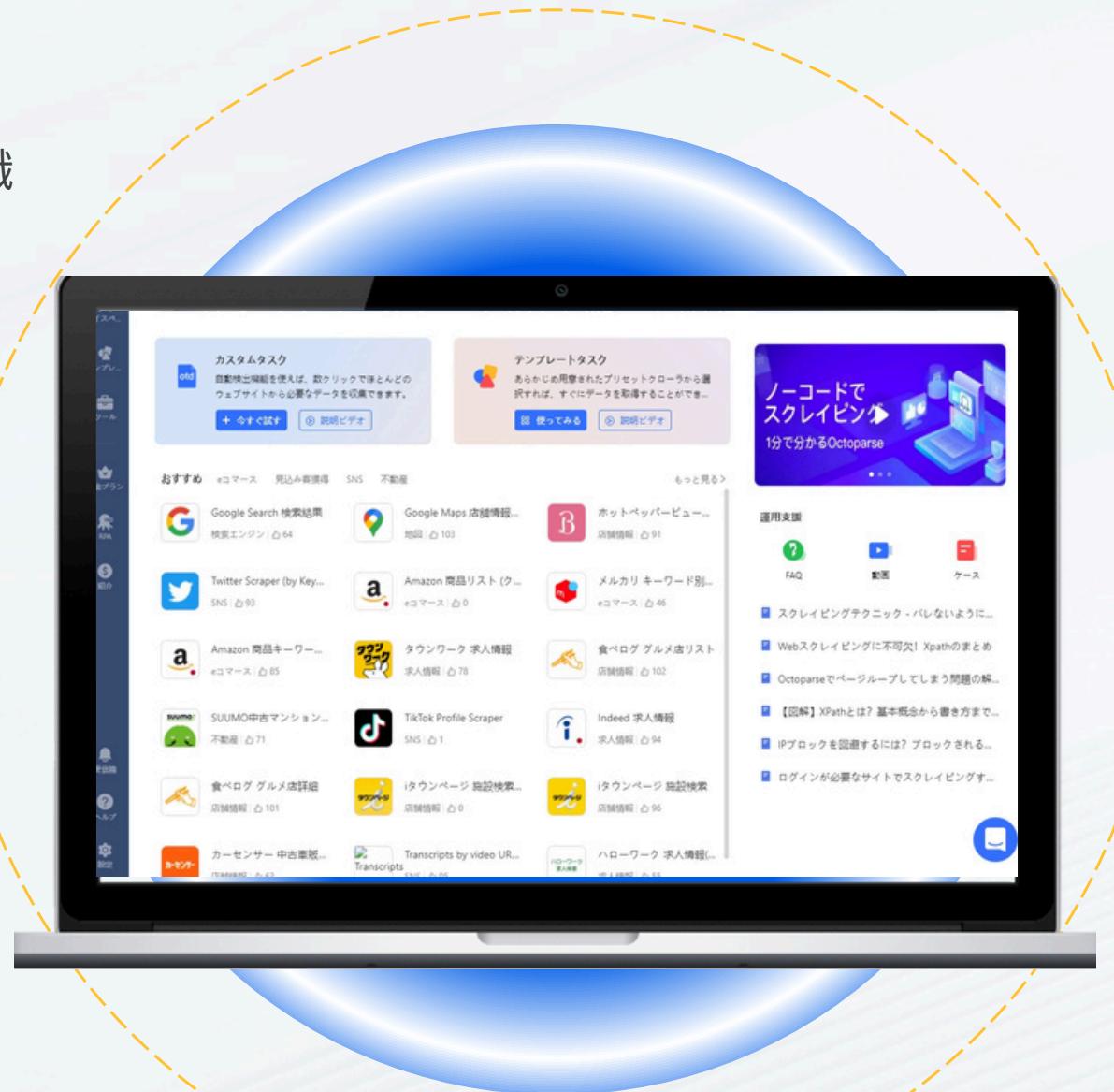
分/時/日単位の柔軟な収集間隔設定とクラウド自律実行で24時間無人運用を実装

グローバル分散処理基盤

グローバル7000+サーバーによる高並列処理で、100万件/分のスループット性能を確保

マルチフォーマット添付対応

テキスト/画像/音声/ドキュメントを指定ディレクトリへ一括ダウンロード・自動整理保存



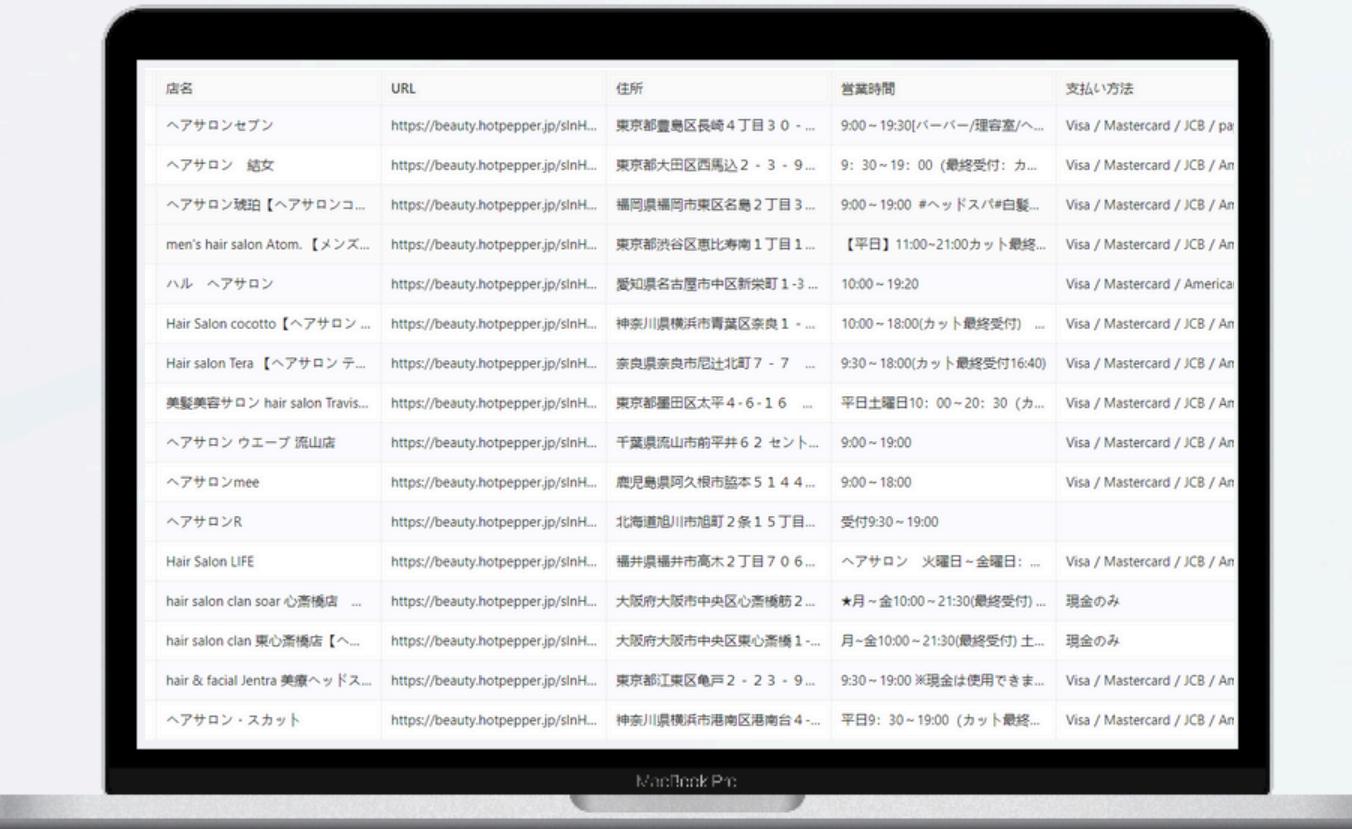
グローバルインフラ基盤

7,000+ノードのマルチリージョン展開により日次10億件の処理能力を実現。完全匿名プロキシ統合型インフラを米欧・東南アジアに戦略配置し、安定性と高速性を両立したグローバルサービスを構築しています。



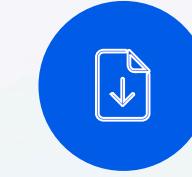
バックエンド管理プラットフォーム

バックエンド管理プラットフォームでは、タスク・リソース・アカウント・APIを一元管理できます。収集ノードのリソース配分は手動・自動で調整でき、並列タスク数も最適化されます。管理者はインターフェース上でタスクを制御し、収集数や進捗、リソース使用状況をリアルタイムで把握できます。



API連携

収集データはExcel、CSV、HTML、JSONなどの形式でエクスポートできるほか、APIを通じてGoogle Sheets、SQL Server、MySQL、Oracle、DM8などのデータベースへ直接連携できます。



リソース管理

収集リソースの統計・調整が可能です。システムは収集負荷に応じて動的にスケール調整を行い、管理者は高品質IPや認証コードリソースを手動で割り当て、収集時のリソース消費状況を把握できます。



タスク管理

個別の収集ノードを起動、再起動、停止でき、実行中の収集タスクには影響を与えません。また、各タスクの収集件数や進捗状況をリアルタイムで確認可能です。



アカウント・権限管理

組織内の役割に応じてアカウントを発行し、適切なタスク設定や管理権限を付与できます。これにより、データ収集のセキュリティを強化できます。

02

活用事例

活用場面

活用場面

価格情報

- 価格監視
- 価格調査
- 競合他社の調査
- 製品トレンドの調査
- 市場調査

競合分析

- オンライン世論調査
- 競合他社の調査
- センチメント分析（感情分析）
- ブランドモニタリング
- リードジェネレーション

マーケティング

- 市場の価格設定
- 市場トレンド分析
- 製品開発
- 製品リサーチ
- ニュース&コンテンツ監視

処理

- ラベリング
- データクリーニング
- データフロー
- データ監視

収集

- Octoparse
- Python
- サードパーティ製品

SNS

Twitter、Youtube
やTikTokなど

ニュース&掲示板

Yahooニュース、Google
やPRTimesなど

リード獲得

タウンページ、食べログ、
Google Mapsなど

Eコマース&小売

Amazon、eBay、楽天や
メルカリなど

不動産

SUUMO、ホームズ、
いい部屋ネットなど

求人

Indeed、doda、
マイナビ転職など

活用事例①

Eコマース業界 - Y社

1 課題

毎月変化する家電製品の価格トレンドを追跡する際、各オンラインショップを手作業でチェックすることに限界を感じていた。特に、夏セール、年末年始セールでは、競合他社の価格変動を即座に把握することが難しく、マーケティング戦略の立案に支障をきたしていた。

2 対応方法

- ・スタンダードプランを導入
- ・[クラウド採集機能](#)を活用し、Amazon、Yahoo!ショッピングなどから、商品画像、商品リンク、価格情報を自動的に収集
- ・[定期自動採集設定](#)により、24時間継続的なデータ更新を実現
- ・[IPローテーション機能](#)で検出回避し、サイトからの遮断を防止。

3 效果

- ・日次商品データ収集数17万3千件
- ・データ収集コストを削減
- ・競合分析レポートの自動化を実現

活用事例②

地域情報業界 - X社

1 課題

自社のリード獲得サービスにおいて、収集する企業データの鮮度に大きな課題を抱えていた。特に、小売業の中小企業は、事業統合や業態変更が頻繁で、既存のデータベースでは3~6ヶ月以上前の情報しか得られず、営業効率を大幅に低下させていた。

2 対応方法

- スタンダードプランで、[カスタム爬虫テンプレートを特別開発](#)
- 法人マイナビ、企業情報データベース、地方自治体の公開データから、企業名、連絡先、業種情報を収集
- 検証コード機能で、[CAPTCHAや動的ロード要素を自動突破](#)

3 效果

- 月間営業名簿6万件を収集
- 企業情報の重複率を削減
- リード生成コストを削減
- マーケティングキャンペーンの対象精度を向上

活用事例③

不動産仲介業 - T社

1 課題

関西圏で不動産投資を手がけるにあたり、築30年以上の中古物件の市場価値を正確に評価することに苦戦していた。各不動産ポータルサイトの情報を一つひとつ確認し、エリアごとの価格相場、設備の経年劣化、最寄り駅からの距離などを比較分析するプロセスが、膨大な工数を要していた。

2 対応方法

- ・プロフェッショナルプランを利用
- ・SUUMO、ホームズ、アットホーム、いい部屋ネットから、URL、物件詳細、エリア統計情報を一括ダウンロード
- ・[データベースとのシームレスな連携](#)で、最新情報の一元管理を実現
- ・[カスタム検証モジュール](#)で複雑な物件ページの情報抽出を可能に

3 效果

- ・月間物件情報10万3千件を収集
- ・物件データの空白率を削減
- ・市場分析工数を削減
- ・不動産投資予測精度を向上

活用事例④

デジタルマーケティング - W社

1 課題

化粧品業界のSNSマーケティングにおいて、製品に関する生の消費者の声を大量に収集できず、トレンド分析や製品改良のための質的データが不足していた。特に、若年層ユーザーの生の意見や感情的な反応を効果的に捉えることが困難だった。

2 対応方法

- ・プロフェッショナルプランを利用
- ・Twitter、TikTok、YouTubeから、関連のハッシュタグ、コメント、シェア数、製品レビューを収集
- ・[クラウド採集機能](#)で大量データを並列処理し、[IPローテーション機能](#)で検出回避。
- ・[APIを活用して](#)データをエクスポートし、リアルタイム分析を実現。

3 效果

- ・月間SNS投稿データ32万件を収集
- ・ブランドセンチメント分析の精度を改善
- ・インフルエンサーマーケティングの効率を向上
- ・マーケティング意思決定時間を短縮

活用事例④

教育業界 - A大学教員

1 課題

研究室では、最新の学術論文を追跡し、カリキュラムに反映させることに苦戦していた。研究分野の急速な進化に伴い、各研究データベースから論文を探し出し、手作業で比較分析するプロセスが、膨大な時間と労力を要していた。

2 対応方法

- スタンダードプランを利用
- Google Scholar、CiNii、J-STAGEから、学術論文のメタデータ、引用情報、フルテキストPDFを自動収集
- [クラウド採集機能](#)で大量データを並列処理し、[Excelへ直接出力](#)
- [時間指定自動採集](#)により、深夜帯の効率的なデータ収集を実現

3 效果

- 年間学術文献250万を収集
- 研究トレンド分析の手動工数を削減
- 新規研究テーマ発見率を向上
- 研究助成金申請の競争力を強化

03

機能・料金

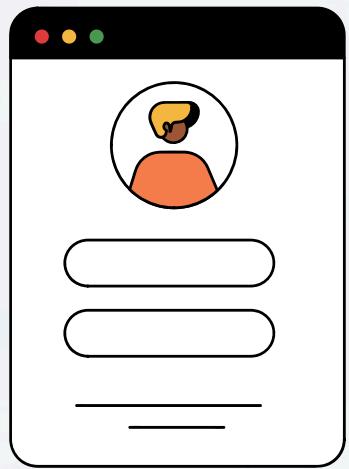
	フリー	プランスタンダード	プロフェッショナル	エンタープライズ
料金 (年間契約)	無料	\$99/月	\$249/月	問い合わせ
料金 (月額プラン)	無料	\$119/月	\$299/月	問い合わせ
タスク	10	100	250	750
エクスポート	1回最大10,000件	無制限	無制限	無制限
デバイス数	1	3	5	5台以上
ローカル収集	2	無制限	無制限	無制限
クラウド抽出	×	最大6件まで	最大20件まで	40件以上
IPローテーション	×	✓	✓	✓
住宅用プロキシ	×	✓	✓	✓
スケジュール実行	×	✓	✓	✓

さらに多くの機能については、公式サイトの[料金プラン](#)をご確認



ご利用の流れ

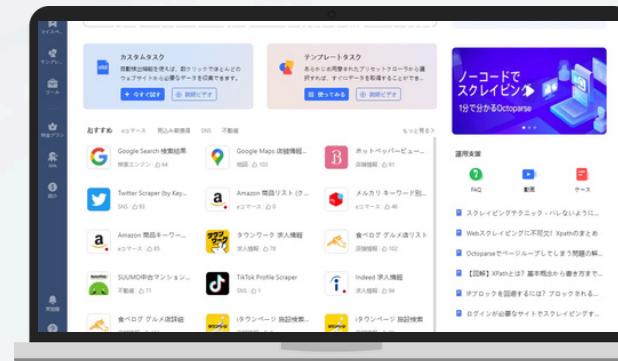
Octoparseをダウンロードし、ログインする



01

STEP.

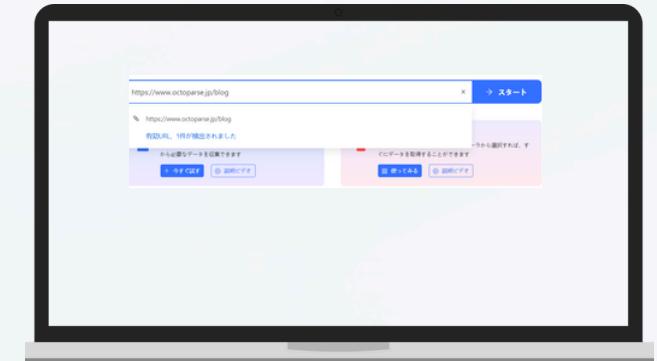
サイトURLまたは関連
テンプレートを選択



02

STEP.

URLを入力し、スタート
をクリック



03

STEP.

サイト情報が自動的に認識され、認識されたフィールドを確認する



04

STEP.

データ収集を実行し、
エクスポートする



05

STEP.

代行サービス

Octoparseでは、ウェブクローラー代行サービスやデータ収集代行サービスをご提供しております。
ビジネスの効率化やデータ活用を支援するため、柔軟なサポートをお約束します。



よくある質問

Q Octoparseの無料試用期間はどのくらいですか？

A. Octoparseには無料プランがあり、基本的な機能を利用できます。また、有料プランの機能を試せる14日間の無料トライアルも提供されています。

Q Octoparseはどのようなデータ形式でエクスポートできますか？

A. Excel、CSV、JSON、HTML形式でのエクスポート、またはデータベース（MySQL、SQL Server）やクラウドプラットフォーム（Googleスプレッドシート、API）への直接同期が可能です。

Q Octoparseのクラウドスクレイピングとは何ですか？

A. クラウドスクレイピングは、Octoparseのサーバー上でデータ収集を行う機能です。これにより、PCを起動せずにスクレイピングを実行し、リソースの消費を抑えることができます。

Q Octoparseの利用にプログラミング知識は必要ですか？

A. 不要です。Octoparseはドラッグ＆ドロップで収集ルールを設定できるノーコードツールです。XPath/CSSセレクターやカスタムスクリプトの活用で、動的ページ処理といった複雑なタスクも効率化可能です。

誰でも簡単に Webスクレイピングを行う

今すぐ利用開始しましょう!

無料トライアル

