

⑩ Pig Data

業界別
WEBスクレイピング事例
100 選

2022-08版



こんな悩みならPigDataにお任せ

Web上のデータから
自社に適した
営業リストを作成したい

P.08へ

EC
プラットフォームから
競合・自社製品の
情報を自動収集したい

P.11へ

競合サイトの
リリースや
コンテンツ更新時に
情報を自動的収集したい

P.07へ

不動産
ポータルサイトから
データを自動収集したい

P.31へ



まずは目次から自分の業界を探す

目次へ



目次

部署別（業界問わず）

P.05 DX推進・IT関連

P.05 マーケティング

P.07 営業

P.07 経営企画

P.08 事業開発

業界別

P.10 EC関連業界

P.12 Webサービス関連業界

P.14 Webマーケティング関連業界

P.18 メーカー関連業界

P.22 金融関連業界

目次

業界別

- P.24 芸能/メディア関連業界 P.31 不動産・建設関連業界
- P.26 交通/旅行関連業界 P.33 学術・研究機関
- P.28 人材関連業界
- P.29 政府・官公庁
- P.30 店舗保有企業



DX推進・IT関連

“スクレイピング人材内製化に向けた
コード提供・マニュアル作成・育成支援”

- ・企業戦略上Webデータ活用が重要になったため、スクレイピング可能な人材を育成
- ・サンプルコードとコードの説明マニュアルを作成し企業側へトレーニングを実施

DX推進・IT関連

“システム連携のため、
自社システムをスクレイピング”

- ・自社のWebシステム間のデータ連携がうまく行えていない
- ・APIもなく情報シスに要望をだしても実装まで時間がかかる
- ・システムからデータをスクレイピングし整形、分析も行った上で別のシステムへ連携するRPAを作成
- ・今ある環境を乱さずに実装できる点がポイント

マーケティング

“レシピサイト構築に向けたYoutubeからの
レシピ動画メタ情報収集”

- ・レシピサイト運営事業に向けたYoutubeデータの収集
- ・概要欄に記載されている手順や材料を自動収集
- ・動画認識を行い、作業工程情報をラベル化することも行った

マーケティング

“競合企業の関連情報をあらゆるサイトを
監視することで常に市場の驚異を社内共有”

- ・競合企業の公式・ブランドサイト、SNS、ポータル、EC、レビューサイトなど競合に関する様々なデータを収集
- ・サイトの更新情報、商品価格の推移、口コミのポジネガなど競合に関するあらゆるデータを収集することで、企業動向を多角的に把握可能に



マーケティング

“グループ企業・協会のニュース情報を自動収集、かつHPへ自動更新”

- ある協会は、全国にある会員のサイトからニュースリリースを手作業でまとめている
- 各サイトの開発会社とのやり取りを介さずに協会サイトへ新着情報を自動でまとめるスクリーピングシステムを開発

マーケティング

“競合のマーケティング費用を計算、効率的な営業活動に利用”

- 広告掲載可能な検索プラットフォームにおいて広告のランク、広告費を計算
- 推定される広告費をもとに営業活動を効率化

マーケティング

“収集データをWordpressで構築されたHPに自動反映”

- ユーザーへ情報提供を行うコンテンツマーケティングを行っている
- 市場の情報を自動収集して自動でワードプレスサイトに掲載可能なアドオンの作成

マーケティング

“近隣店舗のデータを収集してダイナミックプライシングを実現”

- ホテルや、ネット通販等で価格が表示されている商品についてダイナミックに価格を変更したい
- 自社製品との紐づけを行い、自動で価格比較を行う
- 無駄な値下げをする必要がなくなり、利益率向上



営業

“保有テキストデータからメールアドレス・電話番号・ページURLを抜き出す”

- ・文章形式で営業リストとしては活用できないデータをもっている
- ・人物、メール、電話番号を判定してリストを作成

営業

“BASEなどの特定ドメインで作られているサイトのリストを作成”

- ・特定のサービスを利用している事業者向けに営業を行っている
- ・ドメインにターゲットサービスの文字列が含まれるもの収集

営業

“PRTIMESを監視することで企業の新たな動きを察知し効果的な営業活動へ利用”

- ・派遣や転職営業・ソフトウェアの営業において、ターゲット企業の資金調達情報は重要な情報である
- ・PRTIMES等を監視しておくことでいち早く営業活動に移ることが出来るよう情報収集を自動化

経営企画

“地方公共団体から補助金情報を収集”

- ・活用可能な補助金をもれなく情報収集しておきたい
- ・地方の補助金データはまとまっていることが多いため対象サイトを監視

業界問わず

[目次へ戻る](#)

経営企画

“機能過多で高額なSNS分析ツールからシンプルなデータ収集サービスへリプレイス”

- ・海外の汎用的なSNS分析ツールを使っている企業
- ・実際に利用している機能が限定的で、PigDataの機能で代替可能であることが判明
- ・必要機能に限定することで、コストカットに寄与

経営企画

“業界ニュースを速報化して社内へ自動共有することで全体の知見を高める”

- ・世界の業界情報をいち早くキャッチし、社内へ通達する業務をおこなっている
- ・深夜に収集し、出勤時に共有されているような仕組みを構築
- ・各領域の新規事業に関連する情報も合わせて収集することで、事業開発にも役立てる事が可能

経営企画

“自社サイトのクローリングをすることで、個人情報流出をチェック”

- ・自社HPのクローリングを行いそれぞれの問題を解決
- ・何らかの原因で電話番号や個人名に当たるものが掲載されていないかどうかを毎日自動チェック
- ・個人情報など、本来掲載してはいけないデータが出ていないかをチェック

事業開発

“ふるさと納税まとめサイト構築に向けたポータルサイトからのデータ収集”

- ・ポータルサイトから野菜や果物の返礼品情報を集め、同列商品を比較できるようなまとめサイトを構築したい
- ・まとめサイトのバックエンド機能として、複数の大手ふるさと納税ポータルサイトから返礼品情報を収集する継続データ収集システムをご提案



事業開発

“ニュースまとめサイトの構築”

- ・ ニュース記事をスクレイピングしサイト上にまとめることで、ニュースサイトを構築
- ・ 特定用語に合致したニュースのみを収集することで、サイトのテーマに沿ったコンテンツを用意することが可能



DX推進・IT関連

“

楽天・Amazon・Yahooなどの複数モールデータを統合したDB作成

”

- ・モニタリング対象の商品カテゴリーを複数のモールサイトからデータ収集
- ・各モールの価格帯や出品数などを分析することで、最適な出品戦略を構築

DX推進・IT関連

“

ECサイトの価格調整ダッシュボード開発

”

- ・クライアントはDX推進を行う部署
- ・同一商品なのに競合ECサイトの方が安く売られている商品を検知したい
- ・10社の競合ECサイトを24時間365日監視し、ダッシュボード上で価格調整ができるシステムを開発

マーケティング

“

Amazon内のSEO分析

”

- ・Amazon内のSEOランキングを取得し、競合の分析を行った
- ・指定ASINについてキーワード群でAmazon上の検索順位が何位か分かるシステムを構築

マーケティング

“

Amazonにおける特定商品の在庫数監視

”

- ・特定商品の仕入れタイミングを数値的に決定したい
- ・在庫情報を監視し、需給バランスが崩れたときにアラートを上げるシステムの開発



マーケティング

“Amazonの販売者データを収集して
競合分析

- ・自社製品と似た商品を販売している業者の情報を収集
- ・さらに日々監視することで競合の動向を分析



DX推進・IT関連

“介護施設紹介用のWebページ構築に向けた施設情報の収集”

- ・介護事業所紹介事業を行う中で、自社Webサイトに仲介対象の介護事業所情報を掲載したいが、情報収集に手間がかかる
- ・介護施設ポータルサイト（みんなの介護）から、介護事業所データを取得し、自社Webサイトに必要情報を反映

マーケティング

“ゴルフ場予約サイトから予約残枠数を収集し、最適な市場価格を算出”

- ・ゴルフ場の残枠数と価格を毎日収集
- ・収集した時系列価格データから時期別の最適な値付けを設定

事業部

“既存の飲食向けSaaSサービスにWebデータを活用して上位プランを作成”

- ・クライアントは食べログ、ぐるなびなどに出店している飲食店向けのデータ統合サービスを提供している
- ・そちらに他社の予約情報や、売上をKPIとした広告費の最適化機能などを追加実装
- ・顧客のアップセルに成功し売り上げ拡大に貢献

事業部

“複数の学習サービスのスコア情報を収集してワンストップで管理”

- ・複数の学習サービスや読書・トレーニング管理サービスをまとめて成長を記録したい
- ・APIなどが公開されていないため各サービスへ自動でログインし必要な情報を収集



事業部

“**地図サービスの情報更新のための
データ収集自動化**”

- ・ 介護事業所紹介事業を行う中で、自社Webサイトに仲介対象の介護事業所情報を掲載したいが、情報収集に手間がかかる
- ・ 介護施設ポータルサイト（みんなの介護）から、介護事業所データを取得し、自社Webサイトに必要情報を反映



マーケティング

“Instagramにおける注目ユーザーの最新投稿を収集・監視”

- 注目アカウントの投稿状況を監視することで
 - 信頼性をチェック
 - インフルエンサーとしての活動状況を把握

マーケティング

“リストティング広告における特定キーワードの出稿会社監視”

- 自社のターゲットキーワードに対する出稿会社の抽出
- 各キーワードにおける出稿会社数を算出することで、競合性の高いキーワードを抽出

マーケティング

“D2Cメーカー直販商品の転売防止のためのAmazon出品監視”

- 自社商品を直販しているD2Cメーカーが、Amazon内における自社商品の転売を防ぐためのECサイト監視システム
- 自社以外でその商品を出品した業者がいたら監視システムが検知し担当者に通知が行く
- 各商品ページを30分に1回巡回し、出品の有無を検知

マーケティング

“Amazon内におけるコンサルティング企業支援ツール開発”

- AmazonにおけるEC運用の成果を振り返るため広告、レビュー、競合の観点で分析を行うツール開発をサポート
- 自社広告効果の閲覧と同時に他社の広告やレビューなども閲覧可能なダッシュボードへ
- 顧客毎のレポート作成もできコンサルティング業務を効率化



マーケティング

“B2Bコンテンツマーケティングにおける競合分析”

- ・ 競合サイトから掲載ホワイトペーパー、ブログ、事例などのデータを抽出
- ・ 更新をリアルタイムで監視することでタイムリーに把握
- ・ 自社と比較することで、弊社が追加で作成するべきコンテンツを検討できる

マーケティング

“競合のキャンペーンLPを監視することで、競合の動きに対してもいち早く対策を講じる”

- ・ 競合他社のキャンペーンLPデータを毎日監視する
- ・ 変更があった際に通知担当にされ、自社商品のディスカウント率や訴求条件を競合に対して調整可能

マーケティング

“競合の経営やマーケティングの動きを早期発見”

- ・ 競合のブログデータを収集
- ・ 取得したデータを分析し、頻出ワードや減少ワードを発見し、さらにその動きを分析することで競合の狙いを予測

マーケティング

“ターゲット企業の広告費を計算して、有利な営業活動をサポート”

- ・ クライアントはプラットフォーム型のサイトを運営
- ・ 競合のプラットフォームで広告配信をしている企業とそのランクを判定
- ・ かけている広告費を計算し、狙った企業にコストメリットがあるような営業活動をかける



マーケティング

“EC商品カテゴリー別ランキングの異常値を抽出し、出品戦略に反映”

- ・商品カテゴリー別のランキングデータを収集
- ・ランクインする商品＝売れ筋商品のトレンドを把握し、出品戦略に反映
- ・さらに、予兆なくランクインorランクダウンした商品を把握することで市場の乱れにいち早く気づくことが可能に

営業

“Instagramからハッシュタグ別の投稿データを収集することでトレンドを把握”

- ・指定のハッシュタグにおける投稿ボリュームを定期的に検索し時系列データとして収集
- ・各データを用いてソーシャルキャンペーン成果の振り返りや次の企画立案に活用

マーケティング

“TikTokデータを活用したバズ把握”

- ・TikTokのハッシュタグ別の動画本数やライク数を抽出することでいち早くトレンドを把握
- ・データの伸び率からトレンドの勢いを判定することで、効率良くTikTokの動画企画を立案

営業

“競合のWebサービスを使っている企業をコードを利用してリスト化”

- ・競合である、Q&Aサービスを簡単に作成出来るサービスを利用している企業向けの営業をおこなう企業
- ・対象となりそうなサイトをGoogleなどでリスト化
- ・それらのサイトをクローリングし、競合サービスのソースコードがあるかを確認することでリスト作成



営業

“TVの出演効果をTwitterの投稿データから
広告費を試算”

- ・ 製品をTVで取り上げた場合の効果が視聴率ベースで計算されていた
- ・ 近年SNSでの波及効果も考慮したいという要望があった
- ・ 番組に取り上げられる前後のTwitterでのツイート数やリツイート数を用いて広告効果を数値で計算

Twitterデータを商用利用する際には、NTTデータのサービス（※）を利用する必要があります。
※NTTデータのサービス：「[Twitterデータ提供サービス](#)（インテグレーション）」



DX推進・IT関連

“自社のDXを推進するため、
DX事例をPR TIMESからデータ収集

- DXを導入したいと思っているが何から手をつければ良いか分からず、まずは事例を収集するためにPR TIMESより企業が行っているDXに関するデータを全て収集
- 各DX事例データに対して業界、業種などで分類して活用できる事例をピックアップ

マーケティング

“自社製品がECモール上で値崩れをしていないか常時監視

- ブランドを守るために、不当な値下げが行われていないか監視したい
- amazon、楽天、ビックカメラ等の主要なECサイトで販売されている自社商品のプライスをモニタリング

マーケティング

“電気自動車の充電スポットマーケットの
調査データを収集

- DXの導入事例を収集、リスト化
- リストの傾向分析を行い、DX導入事例のトレンド可視化

マーケティング

“基盤製品の見積ロジックを分析するために
見積もり結果の収集

- 工業製品の自動見積もりサイトにおけるロジックを分析したい
- 様々な条件を変えてデータを収集することで分析用教師データとして利用



マーケティング

“大手ECサイトから自社製品の販売価格をモニタリング”

- ・1日1回価格を各プラットフォームから自動収集
- ・各サイトのポイント還元を差し引いた金額も自動計算することで、より高精度な価格を抽出
- ・一定金額以下になった際に自動的にアラートを送信

マーケティング

“コールセンターの記録にWebレビューを加えた商品改善データの収集”

- ・コールセンターでの記録やWeb上の商品レビューデータを収集し統合
- ・収集したデータを商品別に振り分け社内の商品改善用データとして蓄積
- ・社内で展開し、お客様の声を全社共有できるようインフラ化

経営企画

“法改正情報を自動収集することで、情報を早期に把握し、無駄な生産コストを削減”

- ・インターネット官報より日々の法令改正などの情報を取得し、特定のキーワードを含む場合、アラートメールを配信するシステムを開発

経営企画

“ブランドを守るために、横流し商品の把握”

- ・美容院限定商品がメルカリなどに横流しされている
- ・あるブランドラインの商品をメルカリなどに横流しをしている店舗をあぶり出す必要がある
- ・事前にホットペッパーの店舗情報を収集し、事業規模を計算し危険な店舗の予測を行う



経営企画

“海外展開に向けた渡航規制情報の自動収集”

- ・新型コロナウイルスによって入出国が各国によって大幅に情報が変化しているため、海外出張などの行動計画に影響を与えるが手動での情報収集は大変
- ・スクレイピングと機械翻訳を組み合わせて情報を収集し翻訳を実施

経営企画

“経営が安定している企業を判定し、営業リスト作成”

- ・保育士の求人サイトから施設のデータを抽出(スクレイピング)し、リスト化
- ・リフォームを検討している施設への営業リスト
- ・給与や従業員数の増減で、リフォームを検討しているかどうかを測れないかの検証可能

研究

“自社製品の故障原因の分析に活用する気象データの収集”

- ・全国に空調製品を提供している企業様
- ・故障の原因を分析するためにセンサーデータと合わせて、気象データも必要としていた
- ・気象データの収集に加え、すぐに分析で使えるようデータの保存環境の構築も行った

研究

“商品の乗り心地改善に向けたレビューデータの収集”

- ・生産しているバイク製品の改善分析のため、バイクの「乗り心地」に関する情報を集めたい
- ・複数候補のある中、海外のバイクレビューサイトから「乗り心地」に関するレビューデータを収集



研究

“自社開発スマホアプリの改善分析のための
テキストデータ収集”

- ・ 電化製品管理アプリのレビューデータを収集し、自社アプリと競合アプリのレビューを比較分析
- ・ AppstoreとGooglePlayからレビューデータを収集し、テキストマイニング可能なデータセットにして納品
- ・ アプリのUIUX改善に向けた報告書の根拠材料として活用

研究

“ECサイトのレビューデータを活用した競合商品と自社製品の独自性分析”

- ・ クライアントはメーカーの分析部門
- ・ 自社、他社製品のそれぞれのレビューを収集し、共通しているキーワードや1社のみが使っているキーワードをあぶり出すことで、ユーザーの商品に対する認識を分析

研究

“自社製品販売チャネル拡大に向けた飲食店、
宿泊施設のコロナ対策情報収集”

- ・ ホテルレビューサイト内にある空調に関するレビューや、飲食店のコロナ対策情報を収集し、自社空調製品のターゲット層を抽出したい
- ・ Hotpepperとじゃらんからレビューデータを収集し、分析可能なデータセットにして納品



営業

“PayPay・楽天Pay等の電子決済利用可能店舗の収集”

- ・電子決済可能店舗に向けに営業活動を行っている
- ・地域ごとに導入店舗、導入時期を含めて収集しリスト化

営業

“中小企業の社長一覧をWebから収集し営業リスト活用”

- ・ターゲットとしている規模の企業から社長一覧を収集
- ・CRMにターゲットとして組み込むことで営業リスト作成を効率化

経営企画

“Twitter、5ちゃんねるを監視し自社の情報漏洩、炎上案件を検知”

- ・アルバイトの情報流出、取り扱い商品のSNS上での炎上をいち早く気付きたい
- ・アカウント監視、キーワード検索を日々行い、ポジネガエンジンを用いて危険な投稿は優先的に報告するシステムを構築

研究

“OpenWorkから上場企業のレビューデータを収集し、投資評価”

- ・クライアントは投資コンサルティングを行っている企業
- ・上場企業の従業員のコメントなどの企業データを用いて分析を実施
- ・内情を知っている社員の評価を加味することでより精度の高い投資判断を実施



研究

“有価証券報告書データの自動収集および自動整形”

- ・ 有価証券報告書が掲載されているコーポレートやポータルサイトからPDFを探しダウンロード
- ・ PDF収集後の手作業での修正作業も自動化

事業部

“最新情報獲得のため
学術カンファレンス情報を収集”

- ・ 投資情報を提供するサービスを運営している
- ・ 学術カンファレンスにおいて新たに発表された情報が投資行動に影響するため、いち早く情報を収集したい
- ・ 自動で収集し、サービスに組み込むことに成功



マーケティング

“Web版タレント名鑑にSNSデータを連携することで、サービスの質を向上”

- SNSデータをブッキング時の判断材料としている
- Twitterから関連ツイートを収集、以下の分析APIを提供
 - ポジネガ判定（ポジティブツイートの割合を数値化）
 - 10ワード機能（タレント名と一緒に使われる単語を10個ピックアップ）

Twitterデータを商用利用する際には、NTTデータのサービス（※）を利用する必要があります。
※NTTデータのサービス：「[Twitterデータ提供サービス](#)（インテグレーション）」

営業

“Youtubeや楽曲投稿サイトを監視して、伸びしろがあるクリエイターをいち早く発見”

- キーワードやトレンドワードを月次で更新
- チャンネル登録者数に対して再生回数や日時の再生数の伸びを分析
- 早い段階でスカウティング活動が出来るようサポート

マーケティング

“国内外のマンガデータを収集することで、翻訳化がどの程度進んでいるか市場分析”

- 出版業者が新たなマンガ制作の戦略を立てている
- 日本、韓国の2つのマーケットで翻訳が進んでいる作品とそうでない作品の違いを分析

事業部

“鉄道や航空機の運行状況をリアルタイムでTV放送に連携”

- 緊急速報は目視で確認し、入力を行っている
- 主要な交通機関のサイトを監視することで、番組中に自動で発信できるように設定



事業部

“芸人のYoutubeデータをTVキャストアサインの評価軸として利用”

- Youtubeの登録者数の伸び、実数、コメント数をもとに本当に人気がある芸人を正しく把握
- 感覚的に行っていたキャストのアサインを定量的に行うこと で世間の評判にあった番組作りが可能に



マーケティング

“観光戦略構築に向けたTrip Advisorからの言語別口コミデータ収集および分析”

- ・観光地として力を入れていくべきアクティビティやコンテンツの戦略を立てたい
- ・中国語、韓国語、英語、ベトナム語などの口コミと評価をそれぞれ収集し言語ごとの傾向を分析

マーケティング

“楽天トラベルやVELTRAから競合予約情報を収集”

- ・クライアントはホテルを経営しており、楽天トラベルやVELTRA上での戦略を立案したい
- ・競合ホテルのプラン情報や予約状況を収集競合の特徴や、予約状況から売り上げの柱をデータから読み取ることに成功

マーケティング

“法規制に関する各国政府の発信情報を定期的に収集”

- ・法規制に関するデータを収集
- ・取得したデータを分析し、「民泊」、「環境」、「DX」などの分野に各国は力をいれているか、もしくはいれるつもりかを把握

マーケティング

“競合ホテルの価格変動ロジックを分析”

- ・ダイナミックプライシングの値段変動を分析したい
- ・ホテル検索サイト「じゃらん」のプラン情報を複数日で取得
- ・複数のデータポイントから価格変更のロジックを仮説検証



マーケティング

“人身事故発生時におけるSNSのトピック分析”

- SNS上の投稿データを自動取得し自然言語処理を用いて分析することで、人身事故発生時のトピックや態度変容を算出
- 分析データは今後のマーケティング活動や人身事故発生時のメッセージングに活用

マーケティング

“人気旅行先における競合の企画型旅行商品データ収集”

- 人気旅行先における募集型旅行企画の情報を収集
- 費用、選択可能な宿泊先、そして旅行行程を抽出することで、自社商品の価格や内容の競合優位性を確認

経営企画

“特定地域の旅行情報を収集してWebページにコンテンツとして追加”

- 特定地域におけるSNS投稿や旅行関連ブログのデータを収集してポータルサイトを作成
- 自治体が作成している特定観光地のWebサイトページにコンテンツとして追加することでUGCとして活用

経営企画

“ホテル予約サイトデータ収集による価格変動モニタリング”

- ターゲットとなるホテルチェーンの予約サイトから毎日自動で価格やプランのデータを収集
- 時系列で価格を算出することで、シーズンごとの価格設定などを把握



DX推進・IT関連

“市場の需給バランスを分析するため、大手求人検索サイトから募集情報を収集”

- ・大手求人サイトから定期的に人材募集データを抽出した
- ・テキストデータを募集企業、募集日、募集ポジションなど使いやすい形でデータを整形市場においてどの業種人材が枯渇しているか、などをリアルタイムで把握

DX推進・IT関連

“薬学生向け就活サイト構築に向けた募集掲載企業のホットな営業リスト作成”

- ・薬学生向けの求人サイト立ち上げを検討掲載企業募集に向けて新卒薬剤師求人を出している企業
- ・情報収集したいIndeedからリアルタイムで収集し、ホットだと思われる企業へ優先的に営業

DX推進・IT関連

“重要顧客の自社求人サイトへの掲載漏れを防ぐデータ収集”

- ・自社の求人情報を掲載している企業様がクライアント
- ・重要顧客の他社採用関連サイトへの求人掲載情報を自動収集
- ・他社に掲載されているが、自社にされていない求人を早期発見し営業の効率化が可能になった

営業

“企業サイト採用ページのリアルタイム監視による求人ニーズ把握”

- ・企業サイトの採用ページを常時監視求人が掲載された際にアラートを送信掲載日時を元にタイムリーな営業リストを作成



DX推進・IT関連

“中古車情報カーセンサーのデータ収集し自動車の適正価格を逸脱していないか監視”

- ・消費者に信頼される車の販売を推進している官公庁がお客様
- ・消費者が車を選ぶ際に必要な価格や品質等の適正な情報の提供を行うために、国と連携し不当な表示等を禁止している
- ・そのためにカーセンサーから自動車の価格、品質データを収集し、異常値を検知した場合に通知を受け取る仕組みを確立

マーケティング

“スマホアプリ市場の動向調査のための、スマホアプリ開発会社リストの作成”

- ・市場取引調査のため、スマホアプリ開発会社にアンケート調査を行いたいが、業界をまたいで様々な企業がアプリ開発を行っているため、通常のリスト作成業者では対応不可
- ・アプリレビューサイトとアプリのマーケットプレイスからアプリ開発会社情報を収集



DX推進・IT関連

“買取価格値付けのための中古ホビー商品価格データ検索システムの構築”

- 中古ホビーの買取価格調査のため、大手中古ホビー販売サイトの特定商品の価格を定期的に取得し、自社内に中古ホビー価格検索システムを構築したい
- 複数の販売サイトから、特定商品を収集しクライアントのDBに格納、検索可能な継続データ収集システムをご提供

マーケティング

“UberEatsのデータを活用して中食業界のマーケット調査”

- 新規事業としてUberEatsの出店を行いたい
- 口コミ数を活用してマーケットサイズを計算

営業

“営業アタックリストとしての新規出店店舗の自動抽出”

- 飲食・美容業界向けにサービスを提供している企業の営業用アタックリストを作成
- 店舗が掲載されているプラットフォームから毎週新規店舗だけを抽出しリスト化

営業

“新規出店・閉店店舗の最新情報を反映させた営業リストの作成”

- 食べログやぐるなびから最新の飲食店の新規出店・閉店情報を取得
- 営業リストを常に最新の出店情報に更新することで、閉店店舗への営業の無駄足を防ぎ、さらに新規出店店舗へのアプローチ機会を逃さない



DX推進・IT関連

“不動産物件情報の投資判定を自動付与することで、投資判断を効率化”

- ・売買不動産のデータを日々収集し、独自の判定基準で購入を検討している
- ・データの収集～購入判定までを自動的に行い、業務の効率化を行った

DX推進・IT関連

“営業成約率を飛躍的に上昇させるための、補助金データ自動収集”

- ・顧客が申請できそうな補助金を探し、自社サービスと併せて提案したい
- ・複数サイトから補助金情報を手動収集しているが、より多くの情報を収集しつつ手作業の手間を減らしたい
- ・補助金データを複数サイトから自動収集・整形しExcelで提供

マーケティング

“自社が扱っていない物件を自動でHPに更新”

- ・賃貸物件は日々更新していくので更新が大変
- ・不動産ポータルサイトからデータを収集しHPに自動で掲載
- ・自社のみで保有している情報も付加することが可能

営業

“不動産の管理や修繕費用データを活用した営業リストの自動作成”

- ・不動産ポータルから管理費や修繕費を収集している物件リストを抽出
- ・近隣地区の物件よりも相対的に高い費用を取っている物件を判定し、営業の優先ターゲットに設定



営業

“工事進捗を把握するために、SNSデータを活用”

- ・工事進捗に応じてサービスの営業をかけている
- ・SNSやWebから工事現場に関する画像を集めて進捗の把握を行った

営業

“地図会社のサービスから駐車場データを取得し営業リストに活用”

- ・駐車場データを収集することで、駐車場が足りていない地域を可視化
- ・今後の不動産営業における重点エリアの算出に活かす

経営企画

“周辺情報と物件価格を組み合わせた出店戦略の構築”

- ・不動産情報サイトから空き物件の情報を取得
- ・さらに周辺にある物件情報を組み合わせることで、コンビニや塾などの店舗展開をしたい際に収益性が担保されるか判断材料に活用



DX推進・IT関連

“所属研究者の論文実績をHPに掲載するための論文情報収集自動化”

- 研究所のWebサイトに、スクレイピングを用いて所属する研究員の執筆した論文を掲載していた
- 論文サイトのAPIと研究所が持つ研究者情報DBを紹介し、論文情報を自動で掲載するシステムの開発を実施

研究

“映画の宣伝コピーに使われる単語の心理的効果の分析に利用するデータの収集”

- マーケティングに必要な表現の分析を自然言語解析を行う
- 映画の解説サイトから宣伝文と映画に付与されたタグを収集しお客様へ提供

研究

“クラウドファンディングとSNSの関係性を分析するための情報収集”

- MAKUAKEやCampFireで目標達成（未達成）した案件のデータとTwitter上のデータを収集
- 収集後にそれらのデータを自然言語解析することで必要な表現となる要素を分析
- 加えて目標達成する案件の数値的解析を実行

研究

“言語学研究に用いるための辞書サイトからのデータセットの作成”

- クライアントは言語学の教授
- 研究のため辞書サイト（コトバンク）から特定の辞書の収録語句を収集したい
- コトバンク内の辞書別にデータを収集し、txt形式で納品

Twitterデータを商用利用する際には、NTTデータのサービス（※）を利用する必要があります。
※NTTデータのサービス：「[Twitterデータ提供サービス（インテグレーション）](#)」



研究

“政府のオープンデータから消費に関する統計データを収集・成形”

- ・日本国民の消費動向の研究のため、世帯の消費支出に関する統計情報を収集したい
- ・総務省統計局のポータルサイトから、年次ごとにまとめられている該当のEXCELデータを収集し1つのファイルに整形

研究

“国が行っていない市区町村道の交通量予測分析のためのデータ集計”

- ・小規模な市区町村道の交通量予測アルゴリズムを作成したい
- ・国が集計している市区町村道とその付近の国道の交通量ペアデータを作成し、目的のインプットデータを生成
- ・1TB以上ある交通量データの集計を実施し、インプットデータ作成の工数を大幅削減した

研究

“ネット掲示板文化の変遷を追う社会学研究のため掲示板サイトからのデータの収集”

- ・クライアントはネット掲示板上でのオカルト文化の変遷について研究しており、テキストマイニングツールを使ってネット掲示板に掲載される怪談情報を時系列ごとに分析したい
- ・怪談投稿掲示板サイトをスクリエイピングし、テキストマイニング可能なデータセットにして納品

研究

“過疎地域における商業・公共系施設のアクセス分析のため、施設情報収集”

- ・過疎地域の生活インフラ施設へのアクセス状況の分析のため、全国の商業・公共系の施設の位置データが欲しい
- ・商業・公共系の施設の位置データをNAVITIMEから収集し納品



研究

“新着論文を研究所で相互に共有”

- 香料や着色料など、研究に関連する論文をキーワードベースで収集したい
- 論文の自動収集に加えて、社員（研究員）が情報収集の際に効率的に論文調査ができるようなダッシュボードを作成

Pigdata運営会社のご紹介



インディゴデータ社が運営する
PigDataはデータDXを主軸とした
サービスを展開。
Web上のデータをビジネス目的
に沿った状態に収集し加工する
ことで皆様の事業を支援します。



代表取締役社長 大蔵 陽一

運営会社	株式会社インディゴデータ	代表取締役社長	大蔵 陽一
設立	2019年5月		
本社所在地	東京都中央区入船2丁目7番4号 政光ビル 2F		
備考	NTTデータの孫会社（SMSデータテックの子会社）		
提供サービス		スクレイピング代行 SNS分析 競合監視	
		サイトモニタリング 営業リスト作成 Tableau構築	

お取引先企業



本資料を読まれて少しでもご興味を持っていただけた方
ぜひデータ活用に多数の実績を保有している
インディゴデータまでお問い合わせください

お問い合わせ



<https://pig-data.jp/contact/>

その他の資料



<https://pig-data.jp/whitepaper/>

⑩ PigData