

Daxtra Parser について

<概要編>

Daxtra Technologies Japan 合同会社



1. Daxtra Parser（通称 CVX）とは

「世界でただ一つの日本語対応の履歴書自動解析 Solution」

Daxtra Parser は、世界で唯一日本語対応ができている履歴書の自動解析サービスで、かつ世界中のどの履歴書解析ソフトウェアや履歴書パーザーより多くの言語で、より正確に豊富な情報を抽出することができます。また独自の AI エンジンと学習機能を搭載しており、3 か月に一回、より解析精度を向上させたエンジンを提供させていただきます。この強力なツールを使用して、履歴書情報の登録効率、求職者検索精度、マッチング、分析、レポート作成が格段に向上します。



抽出される情報は以下の項目となります。（項目としては約 150 項目を抽出します）

- (1) 個人情報：氏名、性別、国籍、住所、連絡先など
- (2) 学歴情報：学校名、入学・卒業年月、学部など
- (3) 職歴情報：会社名、入社・退職年月、部門、役職、業務内容など
- (4) スキル情報：業界スキル、IT スキル、言語スキル、就労場所など
- (5) その他：現職および希望給与、就労可能時期、退職理由や応募理由、人柄など

一回の処理は平均 7 秒となります。 また日本国内でのご利用の場合は日本の個人情報保護法に準拠し、国内の信頼できるデータセンターを利用して処理をさせていただきます。

弊社は ISO9001 および ISO27001 を取得しており、また GDPR にも準拠しておりますので、安心してご利用いただけます。

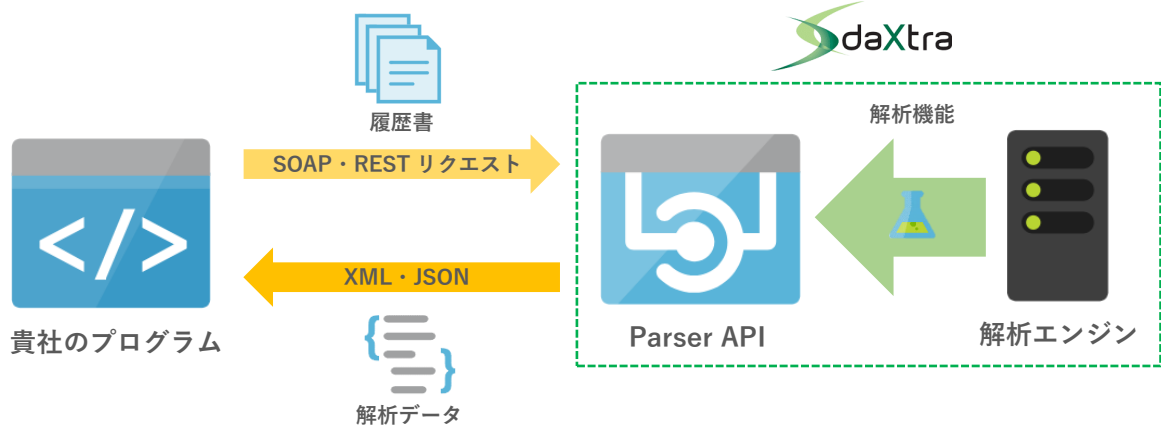
もちろん日本語履歴書以外の履歴書（英語や中国語など）も対応しておりますので、今後増えていく外国人採用などのニーズにもすぐに対応可能です。

現在約 40 言語の履歴書を認識、解析いたします。



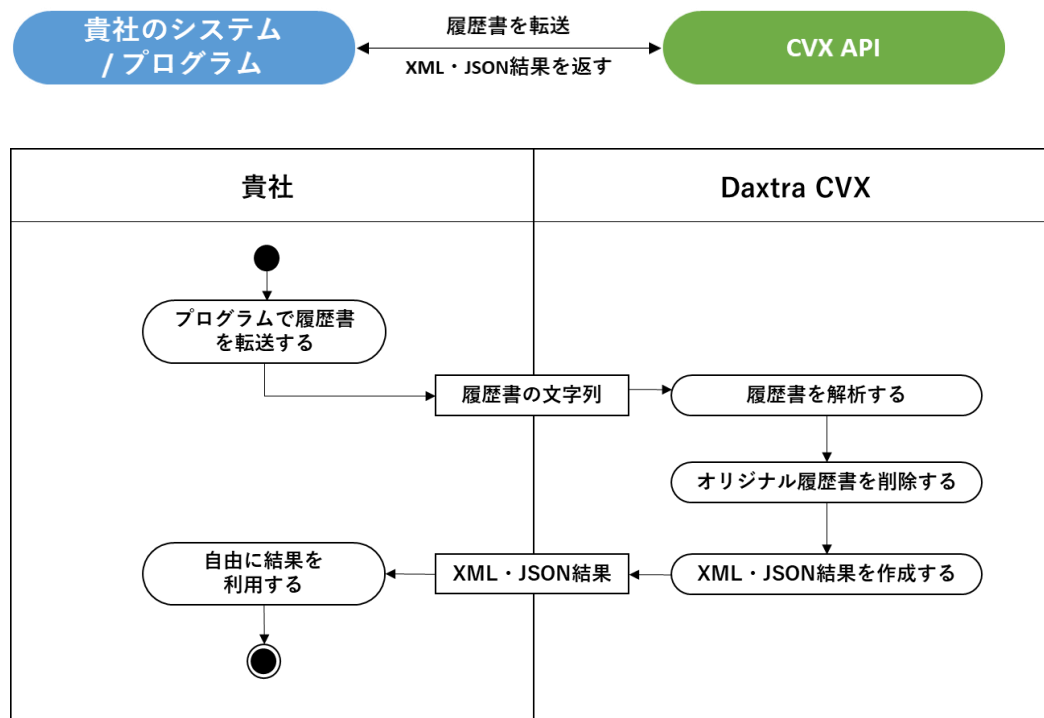
2. CVX との接続形態

「一つのリクエストしか要らない便利性」



- * 対応するファイル形式：DOC、DOCX、RTF、PDF、HTML、XLSX などテキストファイル形式
またイメージ PDF などテキストに変換し解析いたします。
- * その他のファイルに関しては事前に検証を行わせていただきます。
- * CVX は基本的にクラウドサービスで提供されます。世界共通で AWS のサービスを利用しており、セキュアな状態で運用されています。また日本のお客様は日本の AWS にてサービス提供されますので、貴重な情報が処理のために海外に抜けることもございません。この AWS を利用する料金やメンテナンス料金はサービスの利用料金に含まれております。
- * 貴社から解析の依頼や結果を受け取るための API は各種言語で用意しておりますので、それらを利用して簡単にサービスを利用することが可能です。API の詳細やサンプルスクリプトなどは以下 URL に格納されています。
<http://cvxdemo.daxtra.com/cvx/#home>
- * CVX はファイルの解析を行い、その結果を戻すのみで、Daxtra 側には一切の個人情報などは残りません。Daxtra は ISO9001 および 27001 を取得しており、また最も個人情報の取扱いに厳格な GDPR にも準拠しておりますので、安心してご利用いただけます。
- * CVX クラウドサービスはご契約いただく各社様毎に専用の領域を設け、個別に管理運用されますが、解析エンジン自体は共有となっております。どうしても自社専用の解析エンジンをご希望される場合には貴社専用サーバの提供も可能です。その際には別途実費が発生いたします。またオンプレ型のサービス提供にも対応しております。ただしその場合にはメンテナンス用に Daxtra 側がアクセスできる口をご用意いただく必要がございます。
- * サービス利用に関してはすべて HTTPS 経由で行いますので、セキュアな環境で情報漏洩の心配はございません。
- * 解析結果は XML もしくは JSON で受け取ります。CSV 形式での受け取りを希望する場合には有償オプションにて弊社で解析結果を CSV に変換してお返しします。

3. CVX の処理フロー



* 解析結果は有償オプションにて CSV 変換もいたします。

4. CVX の抽出項目について

「解析項目 100 個以上」

解析項目は「個人情報」「職歴」「学歴」「プロジェクト歴」「資格・スキル」に大別され、計 100 以上の項目を含んでいます。詳しくは「XML/JSON 解析表」をご参照ください。

「貴社独自の資格やスキルの登録も可能」

特定の資格、またスキルリストをお持ちの場合は、抽出対象として個別の登録も可能です。これによってより競争力のある解析結果を得ることができます。貴社独自のスキル登録に関しては数が多い場合には別途有償となりますので、事前にご相談ください。(目安：1 スキル登録 1000 円)

「日本語を含む多言語対応が可能」

日本語の履歴書や職務経歴書を Parsing できるのは現在（2020 年 10 月時点）では Daxtra の CVX のみです。また並列で多言語対応しておりますので、日本語履歴書、英文履歴書、中国語履歴書など無作為に送ってもすべて同じ品質で結果を戻しますので、今後の外国人採用などに際してもすぐにご利用いただけます。



5. CVX の設定に関して

以下の項目は設定が可能な項目の一覧となります。設定は導入後でも変更できますが、初期設定時に行うのが最も効率が良いです。

No.	項目	備考
1.	複数ファイルの解析	* ZIP、或いはメールにある複数のファイルを同時に解析する機能です。 * デフォルトは「ON」です。
2.	OCR	* 画像の履歴書を文字として読み込む機能です。 * 解析スピードが下がる可能性があります。 * デフォルトは「ON」です。
3.	優先の解析言語	* デフォルトは「ON」で、優先言語は「日本語」となります。
4.	言語での分割	* 候補者情報は一つのファイルに明確に複数の言語で書かれた場合には、エンジンが自動的にいくつかのファイルに分けて解析する機能です。 * 一般的に A3 と併用します。デフォルトは「ON」です。
5.	Span	* 解析結果に文字をグループ化することができるタグ Span を加えます。 * Styler をご利用いただく場合には、変更できず「ON」となります。 * デフォルトは「ON」です。
6.	電話番号を国際フォーマットに変換	* 電話番号の国を認識して、国際フォーマットに変換します。 * デフォルトは「OFF」です。
7.	解析文字の制限	* 解析の範囲を制限することで、解析スピードをアップさせます。 * デフォルトは「OFF」です。
8.	個人情報のみの抽出	* デフォルトは「OFF」です。
9.	メール本文の除外	* 履歴書が添付されたメールの本文を除いて、解析します。 * デフォルトは「ON」です。
10.	解析したくない言語	* 選択した言語は解析しません。 * デフォルトは「OFF」です。
11.	ターボ	* 解析スピードはアップしますが、画像 PDF の解析は不可となります。 * デフォルトは「OFF」です。
12.	写真のファイル名	* 履歴書に写真があれば、写真のファイル名は結果に入れます。 * デフォルトは「ON」です。
13.	写真のエンコーディング	* 履歴書に写真があれば、写真を Base64 に変換して結果に入れます。 * デフォルトは「ON」です。
14.	解析エンジンのバージョン	* 解析エンジンのバージョンを指定する機能です。 * デフォルトは「OFF」です。
15.	XML アウトプット	* HRXML、或いは DaxML フォーマットで XML 結果を作成します。 * デフォルトは「ON」で、フォーマットは「DaxML」となります。
16.	プロジェクト歴	* プロジェクト歴を職歴として抽出する機能です。 * デフォルトは「ON」です。



また 2020 年 10 月には最新機能として以下の項目は設定が可能な項目の一覧となります。設定は導入後でも変更できますが、初期設定時に行うのが最も効率が良いです。

6. 記載のない項目の判断機能について

履歴書に記載がない事項についても、一部、他の情報から推測し、自動で判断する機能がございます。

No.	項目	備考
1.	国コード	＊都道府県名、州名、地域名、都市名から判断します。（各国、各言語対応） ＊同一都市名がある都市が国名なしで書かれていた場合、電話番号やそのほかの文脈から判断します。（各国、各言語対応） ＊身分証明書情報などから判断します。（各国、各言語対応；日本語未対応）
2.	都市名	＊地区名などから判断します。（一部、各国、各言語対応）
3.	州コード	＊都市名から判断します。（アメリカ、カナダなどに対応）
4.	携帯電話、 固定電話の分類	＊電話番号のパターンから判断します。（各国対応）
5.	スキルの業界分類	＊スキル名から予測します。（各言語対応） 例：平面研削 → engineering-manufacturing（エンジニアリング-製造）、 例：レジオネラ分析 → health-safety（安全衛生）
6.	性別	＊候補者氏名（名前の終わり方など）から予測します。（各言語対応）
7.	取得できる学位	＊学校名から予測します。（各国、各言語対応）

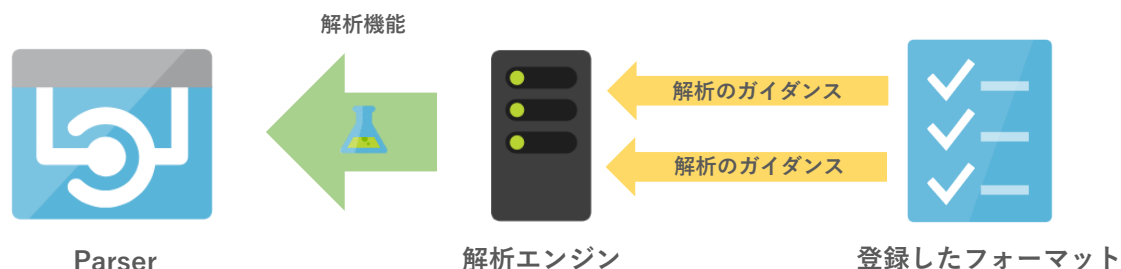
＊履歴書に記載がある場合には、記載事項が優先されます。



7. その他のオプション製品

Formalism：「特定の履歴書フォーマットを登録し、解析精度アップ！」

Daxtra Formalism (有償 Option) は、特定の履歴書フォーマットをあらかじめ登録することで、特定書式の履歴書に対する解析精度を高めるツールです。この機能を使用した解析結果の精度はほぼ 100% 正確になります。



(使用例 1)

エージェント推薦状解析：紹介会社からの専用推薦状を確実に解析します。

Formalism を利用して人材紹介会社からの複雑なキャリアシート書式や、推薦状内の貴重な情報の取得が可能なオプションです。現在は以下の書式、内容に対応しております。

人材紹介会社からの推薦が多い事業会社には必須のオプションとなります。

User Field		En-Japan	MyNavi	Persol	Recruit	WorkPort	一般書式
Career Sheet	キャリアシート			0	0	0	
推薦状抽出	Career_Summary	職歴要約	0	0	0	0	
	Changing_Job_Reason	転職理由	0	0	0	0	
	Interview_Impression	面接での印象	0		0		
	Personality	性格	0	0	0	0	
	Recommending_Reason	推薦理由	0	0	0	0	
	Other_Information	その他の情報				0	
	Recommended_Company	推薦先企業	0		0		
	Recommending_Agent	推薦エージェント	0	0			
	Recommended_Occupation	推薦先役職	0				
	Work_Status	就業状況	0				
	Closest_Train_Station	最寄り駅	△	△	△	△	0
	Commute_Time	通勤時間	△	△	△	△	0
	Closest_Train_Startion	最寄り駅沿線	△	△	△	△	0
		0	対応済				
		△	要望あれば対応可能				
		無印	該当項目なし				

(使用例 2)

転職サイトの登録者のプロフィール解析：転職サイトの登録者情報をすぐにデータ化可能

転職サイトや自社の Web サイトからの応募者もしくはスカウト候補者の情報を PDF もしくは HTML で直接 CVX に投げることで内容を解析します。

現在、主要な転職サイトはすでにいくつか登録されていますので、詳細は弊社にお問い合わせください。 貴社の募集サイトに関しても Formalism 登録（有償作業になります）をさせていただければ、そのまま解析が可能となります。転職サイトを利用されている紹介会社や、今後ダイレクトソーシングを強化されるご予定の事業会社人事には必須の機能になると思います。

(注意点)

Formalism は書式を事前に登録して解析するため、イメージデータ（画像データ）は解析できません。 PDF ファイルでも、Word や Excel をそのまま PDF 保存したものは問題ありませんが、いったん複合機などでスキャンした後のファイルの場合は Formalism は適用外となりますので、複合機でスキャン保存などをされる前の状態で解析をされるようお願いします。

Widget：「スマホ環境からも簡単に Parser を利用できます！」

Daxtra Widget（有償 Option）を Web サイトに埋め込むことで、登録者などのエンドユーザーにも簡単に Daxtra Parser のご利用を可能にするツールです。とくにユーザーがスマホ環境で貴社求人などを閲覧している場合に、大きな効果が期待できます。 Daxtra Widget を利用して履歴データを自動的に取り込むことで、スマホから個人情報の入力に費やす時間や工数を削減が可能になります。アップロード方法はデバイスに保存されているファイルばかりか、SNS の個人情報からクラウドにあるファイルまでアップロードできます。 もちろんスマホ以外の PC 環境でもご利用いただけます。転職サイト、Web からの登録が多い人材紹介会社や事業会社向けのオプションとなります。

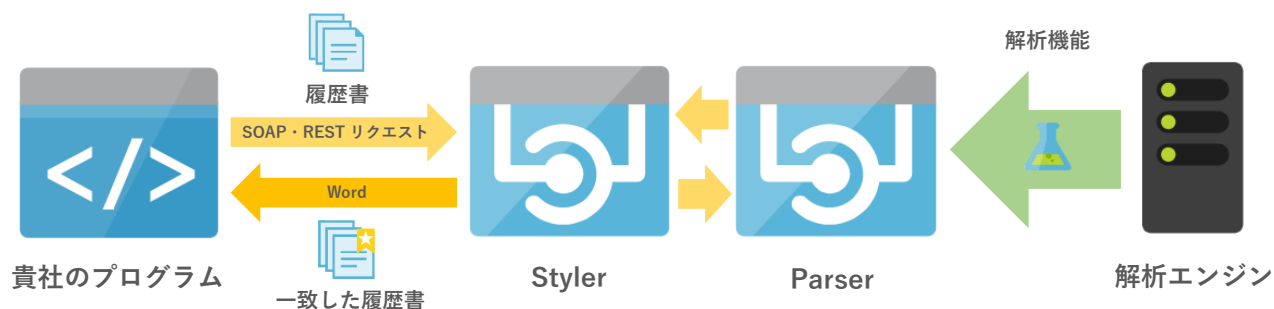


* Widget から返すデータは XML と同じです。

* 埋め込みの API 情報やデモは <https://applywith-tok.daxtra.io/doc> にてご確認ください。

Styler: 「履歴書のフォーマットを統一して、よりプロフェッショナルになる！」

Daxtra Styler（有償 Option）は、様々な履歴書を登録されたフォーマットに変換するツールです。ご要望によって、履歴書に書かれている候補者の個人情報を非表示化したり、自社のロゴを入れたり、フォーマットの変更などが可能となります。企業へ候補者を紹介する際に統一した形で推薦することで、より貴社のイメージ向上にもつながります。人材紹介会社ではかなり利用価値の高いオプションとなります。

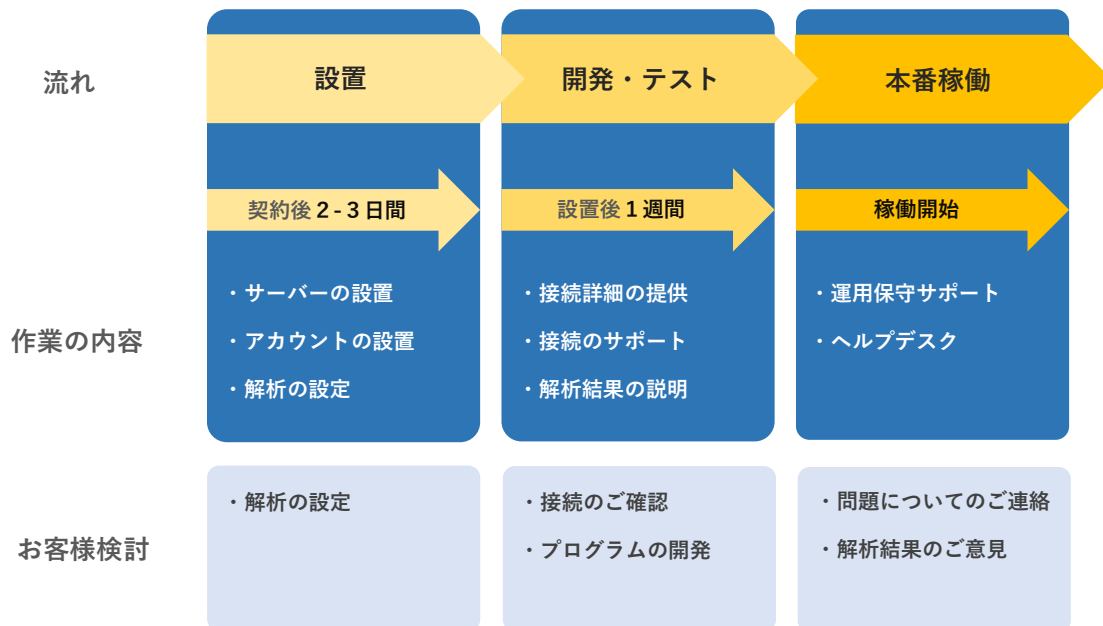


8. ご検討方法

- A. デモアカウントを申請して Daxtra Parser を利用して、解析結果はご評価ください。
* 貴社専用スペースではありませんが、弊社はリクエストを処理するのみで、元データや解析結果は一切保持いたしませんので、外部漏洩の心配はございません。
- B. <https://cvxdemo.daxtra.com/cvx/>を参照して Daxtra Parser との接続方法をご確認ください。
* 複数のプログラミング言語の接続例を掲載しております。
* 何か不明な点等ございましたら、ぜひ Daxtra までご連絡ください。
- C. 上記の接続方法によって、SOAP・REST リクエストができるプログラムを作成してください。
- D. ご契約後に Daxtra から「URL」「アカウント名」を提供いたしますので、Daxtra Parser との実際の接続をご確認ください。
* 貴社専用スペースでお試しいただけます。



9. 導入スケジュール



CVX は解析サービスなので、リクエストを受けて、解析結果を戻す部分のみ対応しております。従いまして上記のスケジュールに加え、貴社側では解析結果を自社のデータベースにロードするためのマッピング作業など別途必要になります。通常はそれらの作業も含め 1 ヶ月程度でご利用可能となります。詳細に関しては弊社にお問い合わせください。

10. よくある質問

Q. 解析データ（XML と JSON）をどう利用したらいいですか。

A. 解析項目は 100 以上ありますが、貴社がそのすべてを必要とはしないはずで。まず「XML 解析表」を参照して、貴社に必要なタグのみを選択し、それらを貴社データベースに流し込むマッピングを行うか、CSV に変換する作業をしていただきます。

Q. XML と JSON の解析結果はどう違うのですか？

A. 基本的な解析結果は同じです。ただし内容的に XML がより詳細に結果を網羅しており、JSON は良く使用するものだけをよりわかりやすいタグ名で表示をしています。弊社からはいずれかでしか解析結果は戻しませんので、貴社の目的に応じてご指定ください。

Q. 解析エンジンはどこにありますか。履歴書の解析は、日本国外で行われることがありますか。

A. 処理は日本の AWS（AmazonWebService）を利用しますので、解析はすべて日本国内で行われます。



Q. 解析の後、Daxtra は履歴書をどう処理しますか。サーバに保存しますか。

A. 履歴書の保存は一切しません。解析結果を返す時点で弊社側からはすべてのデータを消去しております。

Q. サーバの処理能力はどうですか。短時間で数多くの履歴書を解析できますか。

A. 現在ご利用のサーバは 3 つのリクエストを同時に処理できます。さらに、Auto Scaling という技術を使用するので、処理量が最大容量に近づくと、サーバの数を自動で増加します。しかし、増加するには数分間がかかりますので、リクエスト量を少しずつ増やす「待ち行列システム」を推奨しております。

Q. 一人の候補者が複数 File ある場合（履歴書と職務経歴書など）、解析には問題がありませんか。

A. 最大限 3 つの書類をまとめて解析する機能があります。この場合は、まず 3 つの書類を 1 つの ZIP・EML ファイルに圧縮して、その ZIP・EML ファイルをリクエストしてください。

Q. 契約をしてすぐに課金開始ですか。

A. ご契約時に契約書内に両社で合意した課金開始日が記載されます。通常はご契約いただいてから貴社の処理環境を弊社でご用意させていただき、貴社での最終テスト期間なども考慮し、課金開始まで約 1 ヶ月程度の期間を設けております。

課金開始は貴社の実際のサービス開始ではなく、弊社が必要な利用環境を提供できたタイミングとなります。開発のための API はすべて公開しておりますので、契約前でも貴社側の開発を進めることで、効率の良い課金開始が可能となります。

Q. 解析結果に問題があった場合はどのようにしたらよいですか・

A. 解析に技術的な問題（結果がエラーとなるなど）が生じた場合にはすぐに弊社に連絡をお願いします。弊社で内容を調査し、報告をさせていただきます。また解析結果に関して日本語の抽出等で改善が必要と思われる場合も弊社に連絡をお願いします。

Q. 日本語の精度などには問題はありませんか。

A. 自然言語解析は機械学習をさせることで精度が上がっていきますので、現段階においては 100% の解析結果ではありません。弊社では年間に 3~4 回、日本語の抽出精度を向上させたバージョン更新を行っております。これらのバージョン更新はすべてのご利用企業に提供されます。

Q. 個別のスキル登録の方法を教えてください。

A. 貴社で登録したいスキルと該当する用語の一覧表をいただければ弊社にて貴社専用辞書として登録させていただきます。（別途有償：1 スキル登録 1100 円）ご要望がある場合には別途ご相談ください。



11. CVX の利用費用やご契約に関して

CVX のご利用料金は、1 処理当たりの従量課金となります。課金単価は毎月のご契約使用数によって異なりますが、目安としては、月間 100 枚ご契約であれば 1 処理単価 100 円（月 1 万円）、毎月の処理が 1000 枚だと 1 処理単価 50 円（月 5 万円）になります。月の処理数が増えれば、さらに処理単価は下がりますが、一般的には日本だと平均的に月々の利用料金は 10 万円以内となっています。具体的なお見積りに関しては弊社にお問い合わせください。

ご契約期間は 24 ヶ月契約が標準で、少し単価が上がりますが 12 ヶ月契約も可能です。また契約期間内は処理数の変更は自由に行えます。（例えば当初は月 1000 処理で契約し、処理数の上昇に応じて毎月処理数を上位テーブルに変更するなど）

支払いは向こう 3 ヶ月の利用料金の前払いとなります。その場合は例えば月間 1000 処理でご契約いただいている場合は、3 ヶ月合計で 3000 処理を超えなければ、単月の処理数が 1000 を超えても超過費用は発生しません。3 ヶ月の合計で 3000 処理を超えた場合には、次回の 3 か月分の請求時に超過分を単価割で請求させていただきます。

年間利用料金を一括でお支払いいただいた場合は、10%の割引がございます。また使用数計算が年間枠となりますので、上記の例ですと年間で 12000 処理（1000 処理 x 12 ヶ月）を超えない限りは超過費用は発生しません。

初期セットアップ費用：10 万円（初回のみ）

ご契約をいただくと、AWS 内に貴社専用区画領域を用意させていただきます。そのために初回のみ 10 万円の費用が発生しますが、以降はご契約期間中の AWS に掛かる費用はすべて弊社で負担させていただきます。

Parsing 毎月の使用料計算になります。実際には 2 年契約をしていただき、3 ヶ月ごとの利用料金を前払いしていただきます。

（例：毎月 3000 枚のご契約をいただいた場合には、3 ヶ月の利用料合計が 9000 枚を超えない限りは追加費用は発生いたしません。ただし実数が 9000 枚に満たない場合でも返金はございません）
各テーブル間の移動はご契約期間中いつでも行うことが可能です。



12. 簡易ツール JET のご紹介

毎月の解析数量が少ない場合には、簡易ツール「JET」をご利用いただくことも可能です。

こちらのツールは弊社で用意した Web から解析したいファイルを選択することで、そのまま解析結果のファイル（CXML,JSON、CSV の中から 1 種類）を受け取ることができます。



「JET」は月 100 処理で 1 万円となりますので、毎月の利用数量が少ない場合には非常にお得になります。ただし解析エンジンにリクエストをあげる作業は人間が行う必要がありますので、月間の解析数量が多い場合にはお薦めできません。

13. デモアカウントの詳細

顧客名：

ユーザーインタフェース：<https://cvxdemo.daxtra.com/cvx/#demo>

URL：cvxdemo.daxtra.com

アカウント名：

パスワード：

* このデモアカウントでユーザーインタフェースと API の利用が可能です。

* デモアカウントには解析制限（有効期限最長 30 日、期間中 200 処理上限）がございます。

14. 連絡先

所在地： 150-0002 東京都渋谷区渋谷 1-12-2 CrossOffice 8F 808

電話： 03-6868-8734

URL：<https://www.daxtra.co.jp/>

email：k.yano@daxtra.com

