



DaXtra Parser (CVX) Technical Specifications

Deployment

Scalable cloud rings. On-site deployment is also available.

Integration

RESTful and SOAP APIs. Packaged Java and .Net integrations.

Security

SSL, token-based authentication and stateless processing (does not store your data).

Scalability

Uses cloud elastic rings and dynamic load balancer. Current throughput on DaXtractor ring is more than 1 million documents a day.

Configurability

The client is provided with tools to configure and extend the extraction including client-specific skill taxonomies.

Input Formats

Accepts input documents in every commercial resume/CV document format including:

| | |
|---|---|
| Microsoft Word (<i>doc, docx</i>) | Microsoft Excel (<i>xls,xlsx</i>) |
| Microsoft PowerPoint (<i>ppt</i>) | Text file (<i>txt</i>) |
| Mac Pages (<i>pages</i>) | PDF including graphical document (<i>pdf</i>) |
| HTML (<i>htm,html</i>) | Rich Text (<i>rtf</i>) |
| OpenOffice 2.+ (<i>odt, ods</i>) | Microsoft Office HTML (<i>htm,html</i>) |
| HTML Archives (<i>mht, mhtml</i>) | Email (<i>eml, msg, tnef</i>) |
| XML Paper Specification (<i>xps</i>) | Outlook vcard (<i>vcf</i>) |
| Image files* (<i>jpg, jpeg, png</i>) | Word XML Markup Language (<i>wml</i>) |
| Compressed files (<i>zip, rar, gzip, tar</i>) | Microsoft Works (<i>wps</i>) |
| Word Perfect (<i>wpd</i>) | |

*Parser supports image file processing with OCR (Optical Character Recognition) function.

Throughput

Average parsing times are about 2 seconds for a 2-page Word document or equivalent file on 2 phases integration and you can run transactions simultaneously for even greater throughput.

Output Formats

Output available as JSON or XML. For XML, the Parser outputs in either DaXML (DaXtra's custom XML schema) or HR-XML (HR-XML Consortium's schema with DaXtra extensions). Both XML and JSON schema versions are backwards-compatible.

Multi-Language Support

Performs automatic language and locale detection; this is also manually configurable per transaction. We support parsing in the following languages:

| | | | | |
|----------------------|-----------|-----------|-------------|------------|
| Albanian* | Croatian* | German | Lithuanian* | Serbian |
| Bahasa Indonesian | Czech | Greek* | Malay | Slovak |
| Belarusian* | Danish | Hebrew | Moldovan* | Slovenian* |
| Brazilian Portuguese | Dutch | Hungarian | Norwegian | Spanish |
| Bulgarian | English | Italian | Polish | Swedish |
| Catalan | Estonian* | Japanese | Portuguese | Thai* |
| Chinese Simplified | Finnish* | Korean | Romanian | Turkish |
| Chinese Traditional | French | Latvian* | Russian | Ukrainian |
| | | | | Vietnamese |

*These languages are currently available with limited support.

We are continually working on expanding this supported and developing language list.

Multi-Locale Support

Performs automatic locale detection and every locale can be configured per transaction.

Locale will be used for parsing address, telephone, salary currency, etc.

Explicit support for the following locales and most major cities of the world:

| | | | | |
|---------------------------|---------------------|------------------|-------------------|-------------------------------|
| AE - United Arab Emirates | CZ - Czech Republic | JO - Jordan | PE - Peru | SY - Syria |
| AR - Argentina | DK - Denmark | JP - Japan | PH - Philippines | TH - Thailand |
| AT - Austria | DZ - Algeria | KR - South Korea | PK - Pakistan | TN - Tunisia |
| AU - Australia | ES - Spain | MA - Morocco | PL - Poland | TR - Turkey |
| BE - Belgium | FR - France | MD - Moldova | PT - Portugal | TW - Taiwan |
| BG - Bulgaria | HK - Hong Kong | MM - Myanmar | RO - Romania | UA - Ukraine |
| BR - Brazil | HU - Hungary | MX - Mexico | RS - Serbia | UK - United Kingdom |
| CA - Canada | ID - Indonesia | MY - Malaysia | RU - Russia | US - United States of America |
| CH - Switzerland | IE - Ireland | NG - Nigeria | SA - Saudi Arabia | VE - Venezuela |
| CL - Chile | IL - Israel | NL - Netherlands | SE - Sweden | VN - Vietnam |
| CN - China | IN - India | NO - Norway | SG - Singapore | YE - Yemen |
| CO - Colombia | IT - Italia | NZ - New Zealand | SK - Slovakia | ZA - South Africa |
| | | | | ZW - Zimbabwe |

Supported Resume/CV Sections

| | | | | |
|---|--------------------|---------------------------|---------------------|---------------|
| Achievements | Employment History | Legal Document | Objective | References |
| Available | Executive Brief | Licenses & Certifications | Patent History | Revision Date |
| Competency | Executive Summary | Management Record | Position Type | Salary |
| Contact Info | Experience Summary | Marital Status | Preferred Location | Type of Hours |
| Date | Languages | Military History | Project History | User Fields |
| Education History (includes Training History) | Security Clearance | Nationality | Publication History | Will Relocate |

Supported Job Order Data

| | | |
|-----------------------------|----------------------|----------------------|
| Company Description | Job Description | Position Status |
| Compensation | Min Experience | Position Title |
| Competency | Offer Description | Position Type |
| Contact Description | Position Category | Posting Org Industry |
| Contact Information | Position Description | Profile Description |
| Degree | Position ID | Type Of Hours |
| Equal Opportunity Statement | Position Location | User Field |

Resume/CV Data Normalized

| | | |
|-------------------------|--------------|--------------------|
| Aircraft Make and Model | Degree Type | Qualification Name |
| Certification Name | GPA | Security Clearance |
| County Code* | Language | Skill |
| Degree Name | Phone Number | |

Other fields could also be normalised with a skilling project.

*US, Canadian and Australian state; Irish county; Indian state and union territory.

Skills Taxonomy

We have an extensive inbuilt skills taxonomy covering 57 industries. This taxonomy currently includes over 16,000 skills, which expand to a total of nearly 100,000 aliases (synonyms). Some skills have aliases in multiple languages, including French, German and Japanese. These inbuilt skills are parsed by default when a resume/CV or vacancy is processed. Additionally, clients can have their own skills list applied, either by requesting a skilling project from DaXtra or by creating and/or editing a skills file themselves with the aid of our SkillEditor tool.

Geocoder

Geocoding is available upon request. We can supply your own preferred geocoder such as Bing or Google.