

## c-rex.net DTS – Konvertierung von Datenformat A nach B, automatische Metadatenextraktion und mehr

---

c-rex.net **Data Transformation Service** ist ein Baustein der c-rex.net-Datendienste. Hier geht es um die Aufbereitung von Inhalten, die in bestimmten Dateiformaten vorliegen.

### Informationsextraktion

c-rex.net DTS extrahiert automatisch Inhalte und Metadaten aus unstrukturierten und alten Dokumenten und hilft Ihnen, einen Schritt in Richtung einer Welt der Struktur zu machen.

Konvertieren und migrieren Sie Ihre Dokumente in DITA XML, DocBook oder jede benutzerdefinierte XML-Variante bis hin in Ihr CCMS oder in automatisierte Content Delivery Workflows.

Oder konvertieren Sie einfach Dokumente von einem Format in ein anderes.

### Übersetzungsmanagement

c-rex.net DTS hat eine Translation Engine, mit der komplette Übersetzungsprozesse automatisiert werden.

Die Engine bereitet Inhalte auf und generiert standardisierte XLIFF-Dokumente, die in jedem Translation-Memory-System genutzt werden können, ohne dass eine TMS-Konfiguration erforderlich ist.

## c-rex.net DTS – Automatisierung durch regelbasierte Content-Strukturierung

---



## Wie funktioniert c-rex.net DTS?

c-rex.net DTS (Data Transformation Service) ist eine regelbasierte Engine, um aus unstrukturierten Daten strukturierte Inhalte zu gewinnen. Wir sprechen hier von Informationsextraktion.










Es findet eine Datenprozessierung statt, die Inhalte aus einem Format in ein anderes Format überführt und die Inhalte dabei, wenn nötig, mit zusätzlichen Informationen anreichert.

## Welche Formate werden unterstützt?

### Eingangsformate, Quelldatenformate

c-rex.net unterstützt mehr als 20 Datenformate, darunter gängige Formate wie **Microsoft Word**, **Adobe FrameMaker**, **Adobe InDesign** aber auch exotische Formate wie **Quicksilver/Interleaf** sowie Formate aus verschiedenen Content-Management- und Redaktionssystemen.

All diese Formate und auf Wunsch auch neue Formate können als Eingangsformate verarbeitet und umgewandelt werden.













<p><small>doc, docx, wml</small></p> <p><b>Microsoft Word Documents</b></p>  <p>Word Binary Documents Word XML Documents (2003) Word Documents from version 2007</p>	<p><small>mif</small></p> <p><b>Adobe FrameMaker Documents</b></p>  <p>MIF Format from version 8 (Unicode Support)</p> <p>Get in touch for support of earlier versions and binary documents!</p>	<p><small>md</small></p> <p><b>Markdown</b></p>  <p>Markdown documents md</p>
<p><small>xls, xlsx</small></p> <p><b>Microsoft Excel Documents</b></p>  <p>Excel Binary Documents Excel XML Documents Excel Documents from version 2007</p>	<p><small>odf</small></p> <p><b>Open Document Format</b></p>  <p>OASIS Open Document Format for Office Applications ODF</p>	<p><small>chm</small></p> <p><b>Compiled Help</b></p>  <p>Compiled HTML Help chm</p>
<p><small>ildoc</small></p> <p><b>QuickSilver ASCII Documents</b></p>  <p>ASCII Format from version 8.6 (Unicode-Support)</p>	<p><small>idml</small></p> <p><b>InDesign Documents</b></p>  <p>Adobe InDesign Interchange Dokument (IDML packages)</p>	<p><small>pdf</small></p> <p><b>Portable Document Format</b></p>  <p>Only in case you lost your source documents. Quality strongly depends on quality of the PDF. (No scanned PDFs)</p>

*Eingangsformate von c-rex.net DTS*

## Ausgangsformate, Zielformate

Im Ausgang, also als Zielformate, unterstützen wir verschiedene semantische XML-Strukturen wie **DITA**, **DocBook** und verschiedene proprietäre XML-Formate internationaler Hersteller von **Redaktionssystemen** (CCMS).

Selbstverständlich verarbeiten wir auch andere Formate, je nach Use Case **iiRDS**, aber auch den Weg wieder zurück nach Word oder in eine kontextsensitive **HTML5-Hilfe**.

<p>xml</p> <p><b>Custom Flavor XML</b></p>  <p>Your custom specific XML files</p> <p>Get in touch for support your XML standard</p>	<p>xml</p> <p><b>DITA XML</b></p>  <p>DITA Topic, Concepts, Tasks, References, Maps and specialized topic types</p>	<p>doc, docx</p> <p><b>Microsoft Word Documents</b></p>  <p>Word Binary Documents Word Documents from version 2007</p>	<p>html</p> <p><b>Webhelp</b></p>  <p>Responsive HTML5 Webhelp in modern design <a href="#">Take a look at our sample »</a></p>
<p>odf</p> <p><b>Open Document Format</b></p>  <p>OASIS Open Document Format for Office Applications ODF</p>	<p>idml</p> <p><b>InDesign Documents</b></p>  <p>Adobe InDesign Interchange Dokument (IDML packages)</p> <p>Targets text-flow-oriented documents. Depending on requirements post-processing (Scripts) required</p>	<p>iiRDS</p> <p><b>iiRDS package</b></p>  <p>Standard for IoT and Industry 4.0. iiRDS</p>	<p>xls, xlsx</p> <p><b>Microsoft Excel Documents</b></p>  <p>Excel Binary Documents Excel Documents from version 2007</p>
<p>pdf</p> <p><b>Portable Document Format</b></p>  <p>and, of course, standardized or customized PDF documents for print output</p>	<p><b>And more will follow....</b></p>  <p>Let's talk about your requirements</p>	<p>html</p> <p><b>(X)HTML(5)</b></p>  <p>customized HTML(5) outputs</p>	<p>md</p> <p><b>Markdown</b></p>  <p>Markdown documents md</p>

### Ausgangsformate von c-rex.net DTS

## Automatische Extraktion von Metadaten

c-rex.net DTS analysiert die zugehörigen Metadaten und überführt diese automatisiert z.B. in ein iiRDS-Datenpaket.

Nach der automatischen Vergabe der Metadaten werden die Informationen regelbasiert nachbereitet. Bei Bedarf wird eine Plausibilitätsprüfung auf Basis des Gesamtkontexts durchgeführt.

c-rex.net IDS basiert auf RDF und damit auf der **LinkedData-Technologie**. Damit können Informationen beliebig verknüpft werden. Eine hochwertige Suche (Volltext und Facetten) bietet einen Zugang zu weitreichenden Informationen.

Besonders leistungsstark ist das auf dem iiRDS-Standard basierende **Metadatenkonzept**, mit dem in verschiedenen Use-Cases und Zuständen Informationen vorgefiltert werden können.

## Bereitstellung der Informationen in verschiedenen Sprachen

Ein wesentlicher Punkt in Publikationsprozessen ist die Bereitstellung der Informationen in verschiedenen Sprachen. Sprachen, die die Dokumentation von Haus aus unterstützt, aber auch Sprachen, die situationsbedingt kurzfristig benötigt werden.

Dies unterstützen wir durch unsere **Translation Engine**. Dabei integrieren wir Machine-Translation-Systeme und optimieren die Daten inklusive der fachspezifischen **Terminologie**, so dass auch dynamisch benötigte Sprachen in einer möglichst hohen Qualität schnell zur Verfügung gestellt werden.

## Erklärvideo und Blogartikel über c-rex.net DTS

---

In unserem Blogbeitrag [c-rex.net erklärt: Wie Datenkonvertierung leicht gelingt](#) stellen wir unseren Konvertierungsdienst c-rex.net DTS vor.

Schauen Sie mal hinein und sehen Sie DTS im Einsatz.

Erfahren Sie, wie Kurt, Abteilungsleiter Technische Dokumentation, Informationen, die in verschiedenen Datenformaten vorliegen, schnell in ein Kundenportal bekommt.