

ARAYA BILINGUALE TERMINOLOGIE EXTRAKTION



Terminologieextraktion als Dienstleistung

In Übersetzungsprojekten ist es sehr wichtig eine konsistente Terminologie zu verwenden. Diese Terminologie muss gewartet und gepflegt werden, insbesondere mit neuer Terminologie angereichert und verglichen werden.

Im Rahmen unserer Terminologie-Extraktionsdienstleistung bieten wir Ihnen die automatische, auf statistischen Verfahren beruhende Extraktion von zweisprachigen Begriffen (Termen) aus TMX Dateien an.

Die Qualität der gefundenen Übersetzungen hängt natürlich von der Anzahl der Einträge in Ihrer TMX Datei ab, je mehr Einträge enthalten, desto mehr und desto bessere Ergebnisse werden erzielt.

Sie erhalten von uns als Ergebnis eine TBX- oder csv Datei, die die extrahierten zweisprachigen Terme Ihrer TMX Datei enthält. Wir können Ihnen aber auch gerne andere Formate anbieten.

Die Terminologie-Extraktion arbeitet sehr rasch und wir können Ihnen in den meisten Fällen die ermittelten Begriffe innerhalb eines Tages zusenden. Bei Bedarf bereinigen wir die Listen von Termen, die Sie schon in Ihrem Terminologiesystem gespeichert haben.

Sie erhalten damit eine rasche und einfache Methode Ihre Terminologie aus Ihren Übersetzungen zu gewinnen. Sie erleichtern und beschleunigen Ihre Terminologearbeit, befreien sich von Routinearbeiten und dem mühsamen manuellen Durchsuchen Ihrer Übersetzungen.



Die Bilinguale Terminologieextraktion

Der bilinguale Terminologie-Extraktor ist ein einfach zu bedienendes Werkzeug, mit dem Sie aus Ihren TMX Dateien schnell und bequem relevante zweisprachige Termpaare erzeugen können. Die automatische Terminologieextraktion erzeugt dabei Vorschläge, die mit einer Güteschätzung bewertet werden. 1,0 bedeutet dabei eine sehr hohe Wahrscheinlichkeit, dass das Termpaar einer Übersetzung entspricht. Es werden nicht nur Einwortbegriffe extrahiert, sondern auch Mehrwortbegriffe.

Das Extraktionswerkzeug ist auch die ideale Ergänzung zu den Araya Übersetzungswerkzeugen. Es unterstützt die mit den vom Araya XLIFF oder TMX Editor erzeugten TMX Dateien optimal, aber auch die TMX Varianten und Formate andere Hersteller.

Höhepunkte

- Einfach zu bedienende Oberfläche.
- Jedes vorgeschlagene Termpaar wird mit einem Qualitätsmaß (Gütekriterium) bewertet.
- Farbliche Markierung der einzelnen Termpaare je nach Güte.
- Export in verschiedene Formate möglich, z.B. TBX, csv.
- Termpaare können als valide markiert werden und optional nur die validierten Einträge exportiert werden.
- Bekannte Terminologie kann von der Extraktion ausgeschlossen werden.
- TMX Unterstützung
- Verschiedene Parameter um die Extraktion zu steuern wie Häufigkeit der Begriffe, Anzahl der zu extrahierenden Übersetzungen, Wortanzahl für einen Begriff, Groß/Kleinschreibung.

KONTAKT:

Heartsome Europe GmbH
Friedrichstr. 17
D-90574 Roßtal
T: +49 (0) 9127 579001
F: +49 (0) 9127 951178
info@heartsome.de
www.heartsome.de

FN 9098 Amtsgericht Fürth
UID: DE225881142

Geschäftsführung



Dr. Klemens Waldhör
Geschäftsführer
klemens.waldhoer@heartsome.de

Über Heartsome Europe

Die Heartsome Europe GmbH wurde 2002 gegründet. Der Gründer und Geschäftsführer, Dr. Klemens Waldhör, ist schon lange mit der Entwicklung von Übersetzungswerkzeugen vertraut. Diese Erfahrungen stammen aus seiner Zeit in der Forschungsabteilung bei TA Triumph Adler und als Gesellschafter und Geschäftsführer der EP Electronic Publishing Partners GmbH. Unter seiner Leitung wurde das Übersetzungsunterstützungssystem EURAMIS für den Sprachendienst der Europäischen Union in Luxemburg entwickelt. Basierend auf diesen Erfahrung entwickelte er die Araya Übersetzungswerkzeuge.

Die Kernkompetenz von Heartsome ist die kundenspezifische Anpassung der Araya Werkzeuge an die Kundenbedürfnisse. In einer intensiven Beratungsphase werden die Kundenanforderungen ermittelt, optimiert, Araya konfiguriert und anschließend in die Prozesse des Kunden integriert.

Neben den Übersetzungswerkzeugen bietet Heartsome auch Beratung im Tourismus / eTourismus an.

Translation Memory (TM): Eine Übersetzungstechnologie, die es ermöglicht, Segmente (Sätze, Absätze oder Phrasen) von Dokumenten basierend auf der Suche nach ähnlichen Segmenten in Datenbanken als mögliche Übersetzungen zu verwenden.

XLIFF (XML Localisation Interchange File Format) ist ein offener XML basierter Standard, der zum Austausch von Lokalisierungsinformationen, insbesondere Dokumentenformaten unterschiedlicher Hersteller, entwickelt wurde und auf XML basiert.

TMX steht für Translation Memory eXchange und ist ein anbieterneutraler, offener Standard zum Speichern und Austauschen von Translation Memories in XML, die mit CAT (Computer Aided Translation)- und Lokalisierungsprogrammen erstellt werden. TMX erleichtert den Austausch von Translation Memory-Daten zwischen Programmen und/oder Übersetzern ohne bei diesem Vorgang Daten zu verlieren. TMX wurde auf Initiative des OSCAR (Open Standards for Container/Content Allowing Re-use) Ausschusses - einer Interessensgruppe in LISA (Localisation Industry Standards Association) - erarbeitet.

Araya Bilingual Term Extractor 1.0 C:\araya\test\demo.tmx.csv

Nr	Wert	SL	LS	Freq 1	Freq 2	Ausgangsbegriffe	Zielbegriffe	Validi...
6	1,00	5	1/1	9	6	Server	Server	<input type="checkbox"/>
8	1,00	3	1/1	3	3	Zusätzlich	addition	<input type="checkbox"/>
9	1,00	3	1/1	4	3	integriert	integrated	<input type="checkbox"/>
3	0,97	15	1/1	21	20	Araya	Araya	<input type="checkbox"/>
13	0,88	3	1/2	3	4	Übersetzungsstatusanalyse	status analysis	<input type="checkbox"/>
0	0,80	3	1/1	3	5	Anfrage	support	<input type="checkbox"/>
12	0,80	3	1/2	5	3	Übersetzungeditor	translation editor	<input type="checkbox"/>
10	0,69	4	1/1	8	10			<input type="checkbox"/>
11	0,69	4	1/1	8	9			<input type="checkbox"/>
7	0,68	3	1/1	9	8			<input type="checkbox"/>
4	0,67	2	1/1	4	3			<input type="checkbox"/>
5	0,58	2	1/1	5	4			<input type="checkbox"/>

Einfache tabellenorientierte Benutzeroberfläche mit Farbmarkierungen für die unterschiedlichen Qualität der extrahierten Termpaare, in der letzten Spalte wird angezeigt, ob der Begriff bereits validiert wurde.

Extrahiere Terme aus Datei

Extraktionsdatei

Quelldatei: D:\araya\test\zielh-abbegg\Araya3.tmx

Extraktionsdatei: D:\araya\test\zielh-abbegg\Araya3.tmx.csv

Extraktionsdateieigenschaften

Ausgangssprache: de German (Standard) Zielsprache: en English

Minimale Wortanzahl: 1 Maximale Wortanzahl: 3

Minimale Häufigkeit: 3 Maximale Häufigkeit: 100

Maximale Übersetzungen: 3 Ausgangsterme in Kleinschreibung Zielterme in Kleinschreibung

Öffne Extraktionsdatei nach Extraktion

Validierte Terminologie zum Ignorieren

Terminologiedatei: C:\araya\test\demo-valid.tmx.csv

C:\araya\test\demo.tmx.csv
 0/0/0.8/Anfrage/support/unapproved
 3/0/0.96875/Araya/Araya/unapproved
 4/2/0.666667/Datenbank/addition/unapproved

Mit wenigen Eingaben werden aus einer TMX Datei die relevanten Terme und Übersetzungen erzeugt.

```
nr;score;status;term1.LangCode;term1.wordGroup;term1.wordGroup
Len;term1.wFreq;term2.LangCode;term2.wordGroup;term2.wordGroup
Len;term2.wFreq;sentLinked
```

Beispielzeile

Startzeile = Markierung der Extraktionsdatei

```
0;1.0;approved;de;DIN;1;15;en;DIN;1;15;11
```

Systemanforderungen

- auf Java™ basierende Anwendung.
- Software-Anforderungen: Java >= 1.5.
- Betriebssysteme: Windows™ | Linux | Solaris™ | Mac™.

Preise und Lizenzmodelle

- Einzelbenutzerlizenz: € 800,-; zzgl Mwst.
- Mehrbenutzerlizenz: Auf Anfrage.
- Terminologieextraktionsdienstleistung: Auf Anfrage, mengenabhängig

Bestellformular

Ich bestelle __ Araya Bilinguales Extraktionswerkzeug zum Preis von € 800 + Mwst. pro Stück/Einzelbenutzerlizenz.

Firma:
Name:
Strasse:
Ort:
E-mail:

Unterschrift:

Bitte schicken Sie die Bestellung an:

- Fax: +49 9127 95 11 78
- oder
- Heartsome Europe GmbH
- Z. Hdn. Hr. Dr. Klemens Waldhör
- Friedrichstr. 17
- D-90574 Roßtal