

# Was passiert eigentlich mit meinem zu übersetzenden XLIFF-Dokument aus Oracle Application Express?

Daniel Brockmann, SDL Language Technologies

Damit die Zusammenarbeit zwischen allen Beteiligten am oftmals fragmentierten Übersetzungs- und Lokalisierungsprozess vereinfacht wird, hat die Organization for the Advancement of Structured Information Standards (OASIS, [www.oasis-open.org](http://www.oasis-open.org)) vor einigen Jahren einen Standard zum Austausch von Dokumentendaten verabschiedet, den Technologieunternehmen wie IBM, Novell, Oracle und Sun Microsystems entwickelt haben. Dieser Standard heißt XML Localization Interchange File Format (XLIFF).

Im Oracle-Umfeld erzeugen beispielsweise Application Express (Apex) oder auch der Oracle Business Intelligence Publisher XLIFF-Dateien. Nachdem der Standard nun schon einige Zeit verfügbar ist, bieten immer mehr Hersteller von Übersetzungsmanagement-Systemen eine mehr oder weniger umfangreiche Unterstützung für XLIFF an. Was aber genau geschieht mit XLIFF-Dateien aus dem BI Publisher beziehungsweise aus Apex? Wie können die Prozessbeteiligten mit entsprechender Technologieunterstützung mit XLIFF umgehen? Der Artikel beantwortet diese Fragen und zeigt zudem auf, welches Optimierungspotenzial im Übersetzungsprozess stecken kann. Doch zunächst eine Begriffsklärung:

Ein Translation Memory (TM) ist eine Datenbank, die für das Speichern von Übersetzungen optimiert ist und daher oft auch Übersetzungsspeicher heißt. Hier wird der Ausgangstext zusammen mit dem Zieltext in Form von Segmenten – meist einzelnen Sätzen – während des Übersetzens gespeichert. Mithilfe einer entsprechenden Software zur computergestützten Übersetzung – oft auch einfach als „Übersetzungsumgebung“ bezeichnet – wird ein TM befüllt, woraus sich zwei entscheidende Vorteile für mit einer TM-Software erstellte Übersetzungen ergeben – verbesserte inhaltliche und formale Qualität sowie reduzierte Übersetzungskosten. Wird die zu übersetzende XLIFF-Datei nun in der Übersetzungsumgebung

geöffnet, sind vormals getätigte identische Übersetzungen zu sehen. Eine Mehrfachübersetzung entfällt somit, genauso wie die Gefahr, dass identische Inhalte in unterschiedlichen Varianten entstehen. Allerdings sollte an dieser Stelle explizit darauf hingewiesen werden, dass ein Translation Memory gefüllt werden muss, bevor es vollständig genutzt werden kann, wozu jede in einer TM-Software erstellte Übersetzung beiträgt. Für die initiale Befüllung können aber auch Altdaten, die in verschiedenen Sprachversionen vorliegen, automatisiert übernommen werden.

## Die Qualität auf Knopfdruck verbessern

Im Folgenden wird auf einzelne Funktionen einer Translation-Memory-Software eingegangen, um zu verdeutlichen, wie der Einsatz eines solchen Werkzeugs für Übersetzungen im XLIFF-Format oder aber auch in beliebigen anderen Formaten wie Microsoft Office, Adobe InDesign und FrameMaker, XML, HTML etc. die inhaltliche und formale Qualität verbessert.

Die Grundfunktion einer TM-Software besteht darin, vorhandene Übersetzungen mit neu zu erstellenden abzugleichen. Besonders interessant ist dieses Vorgehen für den Übersetzungsprozess von technischen Dokumenten mit hohen Wiederholungsraten, aber auch im Marketing und im Vertrieb können hier sehr gute Ergebnisse erzielt werden, da man davon ausgehen kann, dass Textmaterial wie Bau- oder Betriebsanleitungen für neuere Versionen eines Produkts zu 70 bis 90 Prozent

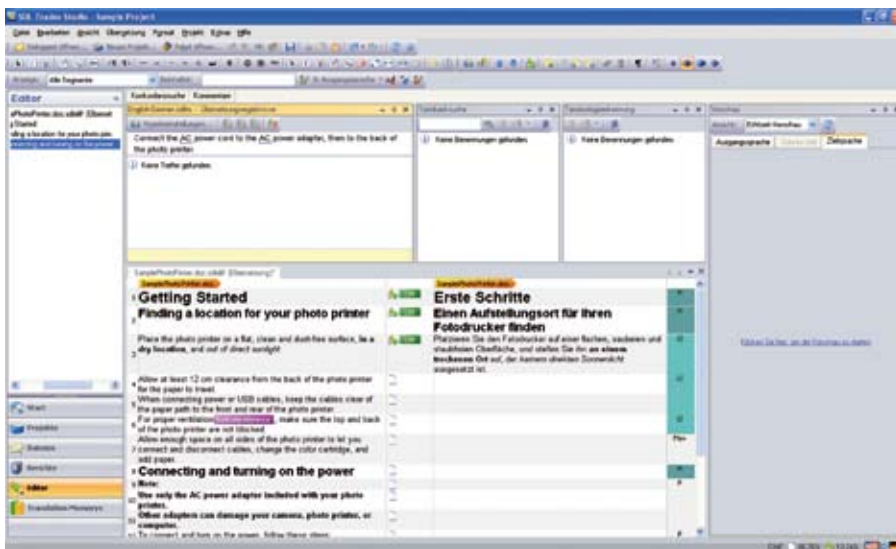


Abbildung 1: Ansicht einer Translation-Memory-Software-Oberfläche mit Kontext-Matches (als CM markiert)

aus Wiederholungen (im Software-Jargon „100%“- oder „Exact“-Matches genannt) bestehen. Jedoch können in diesem Zusammenhang auch produktübergreifende Übersetzungen betrachtet werden: So kann man sich beispielsweise leicht vorstellen, dass die gängige Formulierung „Anweisungen zu Rückgabe und Recycling“ nicht nur bei verschiedenen Produktreihen eines Herstellers zu finden ist, sondern so oder so ähnlich auch herstellerübergreifend eingesetzt wird. Eine weitere, noch genauere Match-Variante ist der Kontext-Match. Hier werden nicht nur einzelne Segmente inklusive Formatierung verglichen, sondern das vorangehende Segment muss ebenfalls mit einer bereits erstellten Übersetzung im TM übereinstimmen, um einen Kontext-Match-Treffer zu erhalten.

Eine Übersetzungsumgebung zeigt aber nicht nur diese „100%“- oder „Exact“-Matches an, sondern bietet auch Vorschläge zu ähnlichen Übersetzungen, sogenannte „Fuzzy-Matches“ an, damit der Übersetzer nur Teile des Segments übersetzen muss. Eine weitere „Vorstufe“ vor der eigentlichen Übertragung können automatisiert erstellte („maschinelle“) Übersetzungen sein. Der Zugriff auf Systeme darauf kann inzwischen direkt aus der Übersetzungsumgebung heraus erfolgen, sodass eine Rohübersetzung ohne das Zutun des Übersetzers erstellt wird. Diese kann dann unterstützend zur menschlichen Übersetzung eingesetzt oder dazu verwendet werden zu entscheiden, dass solch eine nicht vonnöten ist, weil beispielsweise der Text zwar nicht fehlerfrei, aber verständlich genug für seine Nutzung ist.

Während des Übersetzens stehen dann in einer Übersetzungsumgebung weitere Funktionen zur Verfügung, die die Erstellung beschleunigen können. So ist es beispielsweise möglich, dass einzelne Wörter oder Satzfragmente (im Jargon der Branche mitunter auch „Subsegmente“ genannt) nach Eingabe eines ersten Buchstabens, basierend auf dem verwendeten Translation Memory und dem zu übersetzenden Ausgangstext, vorgeschlagen und per Mausklick beziehungsweise Tastendruck eingefügt werden können.

Befinden sich in einem Ausgangsdokument verschiedene Formatierungen wie Kursivschreibung und unterschiedliche Schriftarten oder -größen, wird auch für diese Fälle eine qualitätssichernde Hilfestellung von Softwareseite angeboten. Dies funktioniert sogar für das eher Layout-neutrale XLIFF-Format, da auch hier gewisse einfache Formatierungen wie „kursiv“ oder „fett“ hinterlegt sein können. Das schnelle Übernehmen von Formatierungen erfolgt, wenn der betreffende Teil eines übersetzten Segments im Zieltext markiert wird. Mittels Tastenkombination wird die im Ausgangstext verwendete Formatierung angezeigt und kann mit einem Klick eingefügt werden, sodass mühsames Nachformatieren der Vergangenheit angehört. Funktionen zur Qualitätskontrolle (etwa die automatische Prüfung der Übersetzung auf überflüssige Leerzeichen oder falsche Zeichensetzung) und Dokumentenvorschau,

um zu überblicken, ob der übersetzte Text in sein künftiges Layout passt, sind selbstverständlich ebenso Standard bei einer guten Übersetzungsumgebung wie eine Rechtschreibprüfung.

Ist die Übersetzung schließlich fertiggestellt, durchläuft sie oft noch weitere Stufen wie das Lektorat oder eine endgültige Freigabe. Auch hier bieten gängige Übersetzungsumgebungen umfangreiche Unterstützung an, um den Status der Übersetzung bis zur endgültigen Fertigstellung nachzuvollziehen. In unserem Beispiel kann die XLIFF-Datei aus Oracle Application Express dann als sogenannte „Zieldatei“ fertig übersetzt, korrekturgelesen und gespeichert werden. Danach wird sie in Oracle Application Express zurück eingelesen und dort als zielsprachliche Variante entsprechend weiterverwendet.

Die konsistente Verwendung fach- oder unternehmensspezifischer Begriffe spielt eine entscheidende Rol-

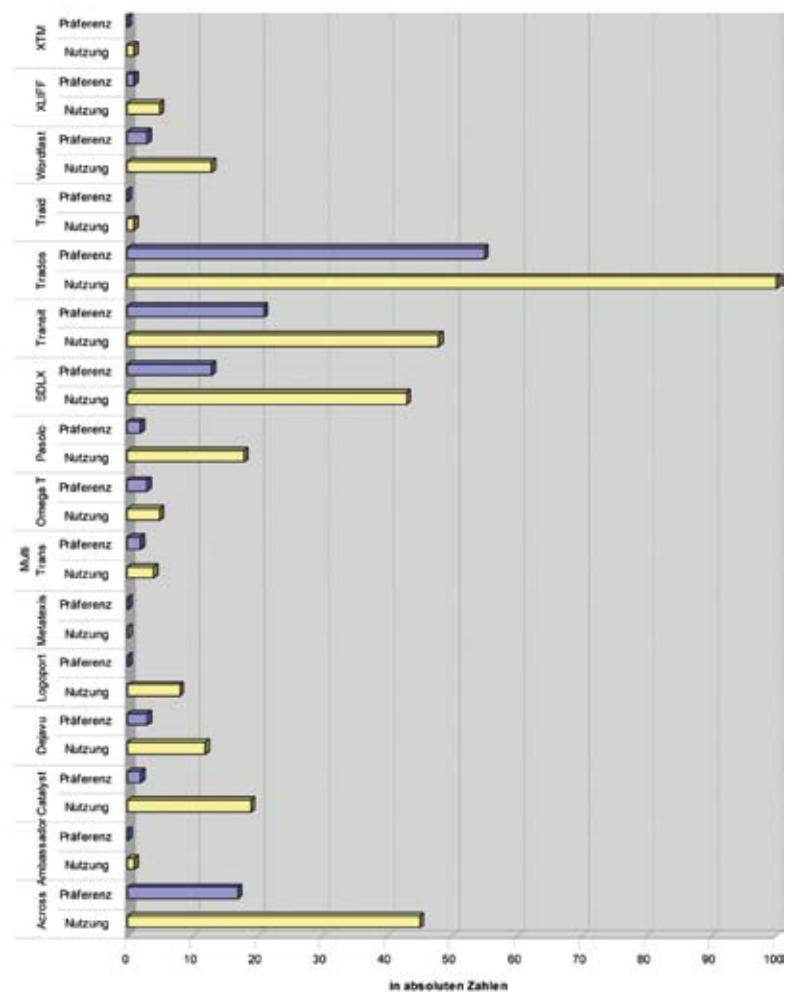


Abbildung 2: Nutzung und Präferenz von TM-Systemen nach einer Studie des Forums Open Language Tools (FOLT)

le in der Übersetzung. Hier kommen zwei Aspekte zum Tragen. Einerseits können aufgrund falscher Benennungen in Betriebsanleitungen mitunter sogar Arbeitsunfälle entstehen, die Schadensersatzklagen zur Folge haben, andererseits können inkonsistent verwendete Begriffe zur Verunsicherung bei Kunden führen und den Markenaufbau bzw. die Markenpflege behindern. Deshalb ist zu empfehlen, bei der Einführung einer Übersetzungsumgebung oder der Beauftragung eines externen Übersetzungsdienstleisters darauf zu achten, dass eine Terminologie-Software zum Übersetzungsprozess hinzugezogen werden kann – oder noch besser bereits fester Bestandteil der Übersetzungsumgebung ist.

Hinter einer Terminologie-Software liegt eine sogenannte „Termbank“, in der fach- oder unternehmensspezifische Benennungen festgelegt, definiert und standardisiert werden. Dies kann dadurch geschehen, dass man eine bestimmte Benennung als Vorzugsbenennung festlegt oder aber Termini bestimmt, die als verboten gelten und folglich von der Nutzung in der Übersetzung ausgeschlossen werden müssen. Während der Übersetzung erhält nun der Übersetzer Vorschläge aus der Termbank, die ihn bei der richtigen Wortwahl unterstützen. Prinzipiell gilt jedoch: Eine Übersetzung kann nur so gut sein wie ihr Ausgangstext, sodass es ratsam ist, auf den korrekten Einsatz von Terminologie bereits bei der Erstellung des Ausgangstextes zu achten.

### **Projektmanagement für Übersetzungen und Softwarelokalisierung**

Neben Funktionen, die die Erstellung einer Übersetzung beschleunigen und ihre Konsistenz verbessern, eignet sich eine Übersetzungsumgebung nicht zuletzt auch für das Management von Übersetzungsprojekten. Ein Projektmanager kann vor der Vergabe einer Übersetzung, intern wie auch extern, mittels des Translation-Memory-Ansatzes beispielsweise prüfen, wie hoch der noch zu übersetzende Anteil des Textes ist und wie viel Text bereits in einem Translation Memory vorliegt, um genaue Angaben über Lieferzeiten

machen zu können. Projektmanagementaufgaben können gestrafft werden, indem Projektvorlagen erstellt werden, um wiederkehrende Aufgaben schneller bearbeiten zu können oder alle relevanten Dokumente, die zur Übersetzung benötigt werden, wie TMs, Termbanken oder Referenzdateien, direkt aus der TM-Software zu versenden, sodass ein Springen zwischen verschiedenen Programmen entfällt. Übersetzungsmanagement funktioniert auf kleiner Flamme bei freiberuflichen Übersetzern und kleinen Übersetzungsbüros in ähnlicher Weise wie bei den Großen – immer geht es darum, wiederkehrende Prozesse zu automatisieren und zu straffen.

Ein Thema, das im Zusammenhang mit Übersetzungen immer mehr in den Vordergrund tritt, ist die Lokalisierung von Software- und Web-Oberflächen, weil hier ganz spezifische Anforderungen an das Übersetzungswerkzeug gestellt werden. Auch hier gilt: Eine gute Translation-Memory-Software liefert die Möglichkeit zur Softwarelokalisierung gleich mit. Ein solches Tool kann nicht nur Wörter und Sätze interpretieren, sondern auch den Softwarecode, und zeigt dem Übersetzer in Echtzeit an, wie sich die Übersetzung (zum Beispiel von „edit“ zu „bearbeiten“) auf Dialogfelder, Menüs oder Schaltflächen auswirkt. Gerade im Umfeld der Software-Lokalisierung wird sehr gerne das XLIFF-Format eingesetzt, da es ursprünglich auf diesen Teilbereich eingeschränkt war, mittlerweile aber auch für alle anderen gängigen Formate, so auch Office- und DTP-Formate, eingesetzt werden kann.

### **Der Markt für Sprachtechnologien ist vielfältig**

Mitte der 1980er Jahre war mit Computer-Assisted-Translation (computer-gestützte Übersetzung, CAT) eine neue Technologie auf dem Vormarsch, die der steigenden Nachfrage nach lokalisierten Inhalten Rechnung tragen sollte. Übersetzer konnten mit der Vielzahl von Projekten, die ihre Kunden ihnen zutragen, kaum noch Schritt halten.

Eines der ersten Übersetzungsunternehmen, das sich der Entwicklung

von TM-Systemen annahm, war Trados aus Stuttgart (heute SDL Language Technologies), das im Jahre 1988 mit TED einen sehr frühen, noch recht rudimentären Vorläufer seiner heutigen Übersetzungsumgebung SDL Trados Studio 2009 auf den Markt brachte. Eine ähnliche Entwicklung durchlebte das Schweizer Unternehmen Star, das, ebenfalls 1984 gegründet, ab dem Jahr 1989 seine Übersetzungssoftware Transit vertrieb. Atril gehörte 1993 mit der Software „Déja vu“ ebenfalls zu den frühen Herstellern von TM-Systemen.

Ende der 1990er Jahre und hauptsächlich zu Beginn des 21. Jahrhunderts wurde der Markt vielfältiger mit weiteren Produkten von Herstellern wie Wordfast (USA), Heartsome (Hongkong), Metatexis, Omega T, Across (alle Deutschland) und Kilgray Translation Technologies (Ungarn). Aktuellster Neuzugang war im Sommer 2010 IBM mit einer Wiederauflage seines Klassikers TM/2 als Open-Source-Produkt unter dem Namen „OpenTM2“. Die Stimmen professioneller Übersetzer auf dieses und andere Ereignisse aus der Übersetzungsbranche finden sich in einschlägigen Nutzerforen wie [www.proz.com](http://www.proz.com) oder [www.translatorscafe.com](http://www.translatorscafe.com).

Neben den klassischen Übersetzungsumgebungen für Projektmanager beziehungsweise Übersetzer gibt es mittlerweile auch Großlösungen für das Übersetzungsmanagement in großem Stil im Unternehmen. Die meisten Hersteller bieten hier sogenannte „Enterprise“-Lösungen an, die im Rahmen von Großprojekten im Unternehmen implementiert und nach und nach ausgerollt werden.

Um einen ausführlichen Überblick über die TM-Software-Landschaft zu erhalten, ist die gemeinsame Studie des Forums Open Language Tools (FOLT), des Zentrums für empirische Forschung Steinbeis SMI und der Dualen Hochschule Baden-Württemberg aus dem vergangenen Jahr empfohlen, die unter [www.folt.org](http://www.folt.org) unter dem Stichwort „Umfrage openTMS Ergebnisse, DHBW Juni 2009“ abrufbar ist.

#### **Kontakt:**

Daniel Brockmann  
dbrockmann@sdl.com