



FOLT: Vorstellung, Ziele und Projekte



Mannheim, 19. April 2008, Michael Schneider

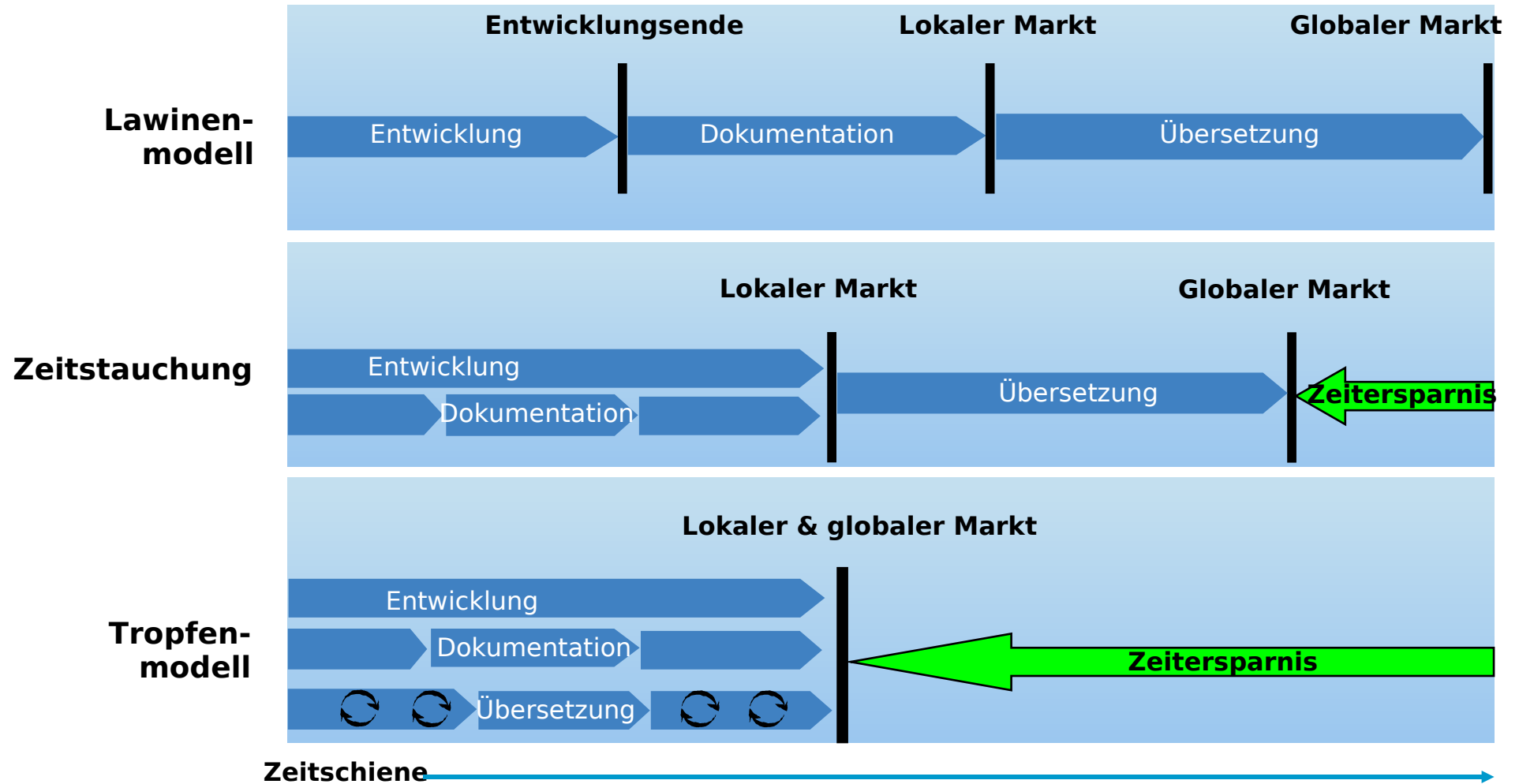
1. Zum Übersetzungsprozess
2. Position und Motivation
3. Forum Open Language Tools
4. Translation Memory Open Source System (TMOSS)

Definition

- Ein Übersetzungsspeicher (auch Übersetzungsarchiv; engl. translation memory, abgekürzt TM) ist eine (in der Regel die Haupt-) Komponente von Anwendungen zur rechnergestützten Übersetzung (CAT = computer aided translation).
- Der Übersetzungsspeicher speichert die Übersetzungen – in der Regel menschlicher – Übersetzer und bietet dem Übersetzer diese bereits vorhandenen Übersetzungen später zur Wiederverwendung an. Hierbei können nicht nur identische, sondern auch ähnliche Ausgangstextsegmente erkannt werden. TM-Systeme dienen vor allem zur Sicherung der terminolo-gischen und stilistischen Konsistenz (Einheitlichkeit) von Übersetzungen. Ein weiterer Aspekt ist die Arbeitserleichterung und -beschleunigung und die damit verbundene Kostenreduktion.
- Besonders sinnvoll ist der Einsatz von TM-Systemen zur Übersetzung von Gebrauchstexten, in denen immer wiederkehrende identische oder ähnliche Formulierungen vorkommen (etwa Bedienungsanleitungen für Serien ähnlicher Produkte).
- Die einzelnen Einheiten der Datenbank werden Segmente genannt. Sie umfassen in der Regel je einen Satz oder Absatz. Der Zugriff auf und die Arbeit mit Übersetzungsspeichern erfolgt über Translation Memory-Systeme.



Prozessanforderung



Herstellerzentrik

- alle Prozesse
- alle Werkzeuge
- alle Aufgaben
- alle Rollen
- → kein Austausch
- → Funktionsreduktion
- → Bewertung Kriterien für Übersetzerwahl

Werkzeuge

- viele
- viele Stärken
- viele Schwächen
- viele Formate
- keine Standards

Herausforderung

Lawinen- vs. Tropfenmodell

- Redaktion : Übersetzung
- Redak : Über : tion : setzung
- Red : Über : ak : set : ti : zung : on

Grundhaltungen

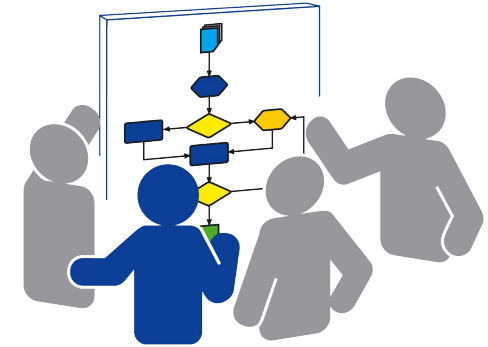
- Methodik
- Textart, -stil
- Termin
- Qualitative Erwartung



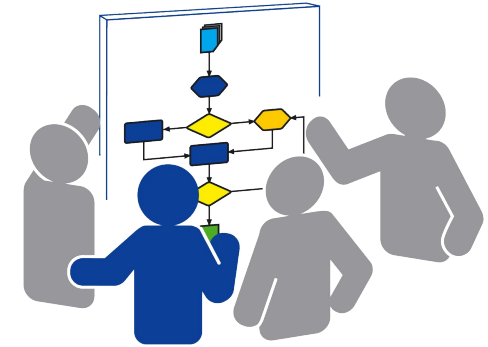
- Unabhängiger Expertenverbund seit 2005
 - Sprachdienstleister
 - Unternehmen
 - Hochschulen
 - Behörden

- Ziele
 - Erfahrungsaustausch: typische Produktionsprobleme
 - Standardisierung von Austauschformaten
 - Erprobung neuer Technologien und Methoden

- www.folt.de

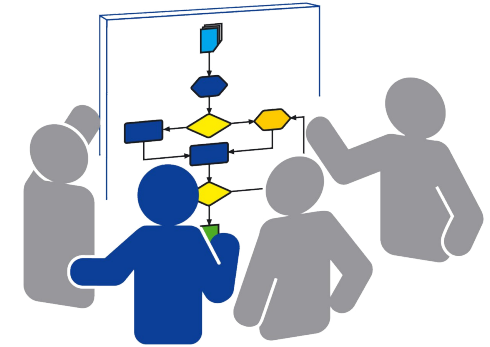


- TMOSS – ein Open Source Translation Memory System
- Förderung von XLIFF
- Regelmäßiger Erfahrungsaustausch
- Geplante Themen:
 - Wortzählung
 - Vereinheitlichte Abläufe in der Terminologearbeit
- Vorträge, Kongresse, Messeauftritte...



■ Konsortiale Entwicklung

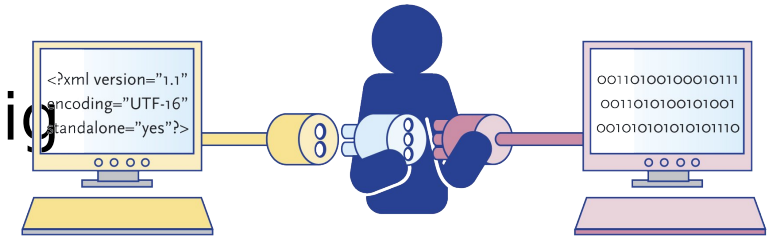
- Wettbewerber unterschiedlicher Größen
- Steuerungskreis
- Hochschulen
- Behörden



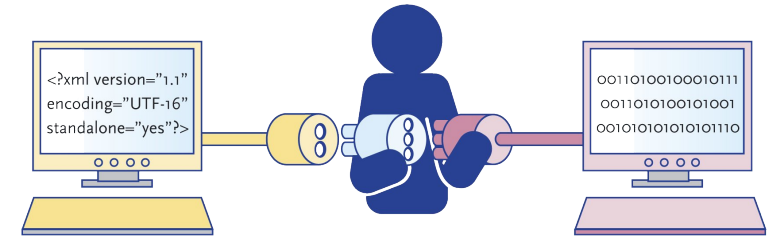
■ Aktivitäten

- Workshops, Spezifikation www.tmooss.org
- Mitgliedschaft bei LiSoG e.V.
- Networking (Partner, Hersteller, Vorträge, Konferenzen)
- Community Building (Mailing-Listen und Newsletter für Interessierte und aktive Unterstützer)
- TMOSS-, Entwicklerkongress in 2008 (Vorveranstaltung)
- Umfrage (wissenschaftlich begleitet), Systemvergleich

- Modularer Aufbau
 - Web-basiert, datenbankunabhängig
 - Terminologieunterstützung
 - Schnittstellen, Standards, Austauschformate
 - Hoher Sicherheitsstandard
-
- CPL/EPL (Common Public License / Eclipse Public License)
 - Service Level Agreement
-
- Vorteile aus Sicht des Sprachdienstleisters
 - Einfachere Einbindung in Produktionsabläufe (extern/intern)
 - Reduzierung von Pflegeaufwand
 - Reduzierung von allg. IT-Kosten



- 100% FOSS
- XLIFF als internes und externes Datenformat
- Integration von Terminologieunterstützung („kleine Lösung“)
- Anbindung an bestehende Terminologielösungen (z.B. von Acolada: „große Lösung“)
- Unterstützung des ganzen Übersetzungs-Workflows
- Berücksichtigung des ganzen Dokumentations-Umfelds (Redaktionssysteme, CMS, Authoring Memory)



- **X**ML **L**ocalization **I**nterchange **F**ile **F**ormat
- OASIS-Standard seit 2002
(Organization for the Advancement of Structured Information Standards)
- XML-basiertes Datenformat für den Austausch von Dokumentendaten im Redaktions- und Übersetzungsablauf

- Übersetzer-Unterstützung
- Import und Export von Terminologiedaten (zunächst zweisprachig) in Standardformaten
 - TBX (ISO DIS 30042)
 - MARTIF (ISO 12200)
 - Text-Dateien (CSV)
 - keine proprietären Formate

- Pflege
 - Definitionen, weitere Informationen/Metadaten
 - Negativlisten, Abkürzungen, Akronyme
 - später: Bilder/Graphiken
 - Freigabe-Workflows Zugriff auf bestehende elektronische Wörterbücher
- Automatische Extraktion
- Anbindung an Standardwerkzeuge für Rechtschreibkontrolle

- Automatisierte Terminologiekontrolle
- Schnittstellen zu „QA-Systemen“
(CLAT, Acrocheck.....)
- Anbindung an Workflow- bzw. Projektmanagement-Systemen
- Anbindung an Redaktionssysteme, CMS, Web-CMS
- Integration in OpenOffice
- Dokumentenfilter, Dokumentenfilter, Dokumentenfilter

- FOLT: www.folt.org
- TMOSS: www.tmoss.org
- OASIS: www.oasis-open.org

Vielen Dank für Ihre
Aufmerksamkeit!



michael.schneider@beo-doc.de