

Bernburg
Dessau
Köthen



Hochschule Anhalt
Anhalt University of Applied Sciences

Bachelorarbeit
zur Erlangung des akademischen Grades
Bachelor of Science (B.Sc.)

Nutzung online verfügbarer Ressourcen für die Darstellung von
Aufbau, Funktionsweise und Einsatzszenarien
maschineller Übersetzung

eingereicht an der

Hochschule Anhalt
Hochschule für angewandte Wissenschaften
Fachbereich Informatik und Sprachen

eingereicht von

Xuanying Li

xuanying.li@student.inf.hs-anhalt.de

Matrikel-Nr. 4055041

1. Prüfer/in:
2. Prüfer/in:

Prof. Dr. Uta Seewald-Heeg
Prof. Dr. Axel Schneider

Ausgabetermin:
Abgabetermin:

16.10.2014
28.11.2014

Affidavit – Eidesstattliche Erklärung

Ich versichere, dass ich die Arbeit ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Quellen angefertigt habe.

Die Arbeit in gleicher oder ähnlicher Form wurde noch keiner anderen Prüfungsbehörde vorgelegt und von dieser als Teil einer Prüfungsleistung angenommen.

Alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, sind als solche gekennzeichnet.

Xuanying Li

Köthen, den 28. November 2014

Danksagung

Es ist ein lobenswerter Brauch:
Wer was Gutes bekommt,
der bedankt sich auch.
Wilhelm Busch

An dieser Seite möchte ich mich ganz herzlich bei meinen beiden Betreuern,
Frau Prof. Dr. Uta Seewald-Heeg und Herr Prof. Dr. Axel Schneider,
die mir während der Fertigstellung dieser Arbeit zur Seite standen, bedanken.
Ohne die fachliche Betreuung und Geduld wäre diese Arbeit nicht möglich gewesen.

Weiterer Dank gilt allen, die mich während meines Studiums in Freundschaft begleitet
haben.

Nicht zuletzt gilt das größte ‚Dankeschön‘ meinen Eltern sowie meinen Freunden. Die
Unterstützung, die ich von Euch in jeglicher Hinsicht erfahren durfte, hat mein Studium
möglich gemacht. Herzlichen Dank!

Xuanying Li, November 2014

Inhaltsverzeichnis

Abstracts	1
1. Einleitung	2
1.1 Motivation	2
1.2 Zielgruppen	4
1.3 Aufbau der Arbeit.....	5
2. Grundlagen	6
2.1 Begriffserklärung und Entwicklungsgeschichte der maschinellen Übersetzung	6
2.2 Ansätze der maschinellen Übersetzung.....	10
2.2.1 Regelbasierte MÜ	11
2.2.2 Statistikbasierte MÜ.....	13
3. Nutzung online verfügbarer Ressourcen für die maschinelle Übersetzung.....	15
3.1 Suchverlauf und Suchkriterien online verfügbarer Ressourcen	15
3.2 Voraussetzungen für die Nutzung online verfügbarer Ressourcen	16
3.3 Klassifizierung online verfügbarer Ressourcen nach Inhaltskategorien	17
3.2.1 Informationsquellen	17
3.2.2 MÜ-Systeme	26
3.4 Vor- und Nachteile bei der Nutzung online verfügbarer Ressourcen	33
4. Fazit	35
Quellenverzeichnis	39
Abbildungsverzeichnis	44
Glossar	45
Anhang	47

Abstracts

Machine translation has been developed after the invention of computers combined with linguistics. It is a sub-field of artificial intelligence and usually used in the localization process to reduce translation and localization costs. It is clear that machine translation has nowadays a strong and significant market presence. Therefore, more and more people are interested in structure, function, scenarios of machine translation, and online available resources for machine translation. This work offers an overview of utilization scenarios and methods of machine translation and possibilities of using online available resources for machine translation. So that translators can become familiar with this technology and the possibilities and limits of it. All the information sources and resources presented in this work can also be used for distance learning or e-learning to learn basic knowledge of machine translation. The possible task field and the potential combination of integrating new technologies in the near future will also be projected in this work.

Nach der Erfindung der Computer in Kombination mit der Sprachwissenschaft entstand die maschinelle Übersetzung als Teilbereich der künstlichen Intelligenz. Maschinelle Übersetzungssysteme werden heutzutage oft im Lokalisierungsprozess eingesetzt, um die Übersetzungs- und Lokalisierungsaufwände zeitlich und kostenmäßig zu reduzieren. Es ist zu sehen, dass heutzutage maschinelle Übersetzung starke und eindeutige Marktpräsenz besitzt. Daher gibt es immer mehr Menschen, die sich für die Darstellung von Aufbau, Funktionsweise und Einsatzszenarien maschineller Übersetzung und für die Nutzung online verfügbarer Ressourcen für maschinelle Übersetzung interessieren. Die vorliegende Arbeit gibt einen Überblick über die Anwendungsbereiche, Anwendungsverfahren und die Möglichkeiten der Nutzung online verfügbarer Ressourcen der maschinellen Übersetzung, so dass sich im Beruf stehende Übersetzer mit dieser Technologie und ihren Möglichkeiten und Grenzen vertraut machen können. Auch im Rahmen eines Fernstudiums können die in der Arbeit vorgestellten Informationsquellen und Ressourcen eingesetzt werden, um die Grundlagen der MÜ zu erlernen. Zu diesem Zweck werden die weiteren Aufgabenfelder und Kombinationsmöglichkeiten in naher Zukunft prognostiziert.

1. Einleitung

1.1 Motivation

Internationalisierung und Lokalisierung zählen heute zu den Schlüsselbegriffen einer globalen Marktwirtschaft.¹ Will sich beispielsweise Unternehmen erfolgreich gleichzeitig auf unterschiedlichen Märkten behaupten, müssen Produkte und Dienstleistungen so entwickelt werden, dass sie schnell und einfach an die Anforderungen des Zielmarktes sowohl sprachlich als auch kulturell angepasst werden können. Mit Hilfe der computer-gestützten bzw. maschinellen Übersetzungstechnologie werden die Übersetzungs- und Lokalisierungsaufwände im Lokalisierungsprozess zeitlich und kostenmäßig erleichtert und reduziert. In zahlreichen Unternehmen wird maschinelle Übersetzung (MÜ) bereits im Lokalisierungsprozess eingesetzt. Dafür ist die Webseite von *Microsoft Translator*² ein typisches Beispiel. Alle Daten auf der Webseite wurden nur in englischer Sprache verfasst, können aber in anderen unterstützten Sprachen abgerufen werden, da sie maschinell in der Microsoft Datenbank übersetzt werden. Durch Klick auf die Menüoption „Diese Seite übersetzen“ (rot markiert in Abb. 1) oben rechts auf der Webseite wird diese Seite in einer der unterstützten Sprachen maschinell und automatisch übersetzt.



Abb. 1: Ins Deutsche maschinell übersetzte Homepage von Microsoft Translator³

¹ [Schmitz 2005] K. Schmitz; A. Ottmann; E. Schildhauer. Internationalisierung und Lokalisierung von Software, Lokalisierung von Softwareoberfläche, Lokalisierung der Dokumentation und der Online-Hilfe. Einführung in die Softwarelokalisierung, Tübingen, 2005:1

² Microsoft Translator, <http://www.microsoft.com/translator/default.aspx> [28.11.2014]

³ Microsoft Translator, <http://www.microsoft.com/translator/default.aspx> [28.11.2014]

Der Einsatz der MÜ ist aus wirtschaftlicher Sicht von Bedeutung: schnelle Kostenreduzierung und Verfügbarkeit einerseits, deutliche Leistungsverbesserung andererseits. Im Vergleich zur begrenzten Produktivität eines menschlichen Übersetzers weist die maschinelle Übersetzung eine deutlich höhere Leistung auf. Mit dieser Technologie ist es möglich, die zu übersetzenden Texte schnell zur Verfügung zu stellen. Und wenn man MÜ mit anderen Verarbeitungsmethoden wie Pre- und Post-Edition oder Kontrollierte Sprache zusammensetzt, kann eine hohe Übersetzungsqualität erreicht werden. Außerdem ist die Anzahl der menschlichen Übersetzer in der Gegenwart zu klein, um den wachsenden Bedarf an Übersetzungen zu decken. Da man viele Jahre braucht, eine neue Sprache zu beherrschen. Aus solchen Gründen sind die Aufwände für eine rein menschliche Übersetzung höher als für die maschinelle Übersetzung.

Aber was ist eigentlich maschinelle Übersetzung (MÜ)? Welche Online-Ressourcen gibt es heutzutage? Wo und wie kann man Materialien für MÜ im Rahmen eines Fernstudiums finden, wenn man die Grundkenntnisse über diese Technologie erlernen möchte? Und für welche Zielgruppen sind die gefundenen online verfügbaren Ressourcen geeignet? Dazu betrachtet die vorliegende Bachelorarbeit die Grundlagen der maschinellen Übersetzung und die Nutzung online verfügbarer Ressourcen für die maschinelle Übersetzung. Sie untersucht auch die Anwendungsgebiete der maschinellen Übersetzung, um die praktische Verwendung der maschinellen Übersetzung deutlich aufzuzeigen und die online verfügbaren Ressourcen für MÜ zu klassifizieren.

1.2 Zielgruppen

Diejenigen, die sich für online verfügbare Ressourcen zur MÜ interessieren und diese möglicherweise im Rahmen eines Fernstudiums nutzen, lassen sich verschiedene Gruppen zuordnen. Im folgenden Teil werden die in der vorliegenden Arbeit betrachteten Zielgruppen genau definiert. Im Hinblick auf die Nutzer online verfügbarer Ressourcen der MÜ lassen sich 3 Gruppen unterscheiden:

Anfänger oder interessierte Laien: Menschen, die auf dem Fachgebiet der MÜ oder in mit der MÜ verbundenen Fachgebieten keine oder ungenügende Kenntnisse haben und diese erlernen und üben möchten.

Studenten in Fächern, in denen Übersetzen eine Rolle spielt: Menschen, die auf Gebieten wie Übersetzen, Softwarelokalisierung, Internationale Kommunikation, usw. die entsprechenden Fachkenntnisse haben und sich in diesem Fachgebiet genau auskennen und sich die Grundkompetenz der MÜ aneignen möchten und ihre Kenntnisse über die MÜ-Technologie zu vertiefen suchen.

Lokalisierungsdienstleister oder freiberuflicher Übersetzer: Lokalisierungsdienstleister sind diejenigen, die eine Dienstleistung im Geschäftsfeld der Lokalisierung von Softwarekomponenten bieten; freiberufliche Übersetzer sind diejenigen, die einen Text von einer Sprache A in eine Sprache B berufsmäßig übertragen und diese Tätigkeit selbständig ausüben.

1.3 Aufbau der Arbeit

Die vorliegende Bachelorarbeit entstand im Rahmen des Studiengangs Softwarelokalisierung an der Hochschule Anhalt.

Nach dieser Einleitung werden die Grundlagen der Arbeit aufgezeigt. Zunächst werden die Kernbegriffe der MÜ erklärt und ihre Entwicklungsgeschichte (2.1) wird vom Anfang bis heute allgemein und kurz erläutert. Weiterhin wird ein kurzer Überblick über die Ansätze der maschinellen Übersetzung (2.2) gegeben.

Kapitel 3 beschreibt den wichtigsten Teil der gesamten Arbeit – die Nutzung unterschiedlicher online verfügbarer Ressourcen für MÜ im Rahmen eines Fernstudiums. Der Suchverlauf und die Kriterien für die Suche nach online verfügbaren Ressourcen werden im ersten Teil dieses Kapitels beschrieben (3.1). Daraufhin folgen die Voraussetzungen (3.2) für die Nutzung online verfügbarer Ressourcen. Dazu gehört das gewisse Vorwissen, über das man verfügen muss, um sich die Ressourcen der MÜ aneignen zu können. Im dritten Teil wird die Ressourcen-Sammlung⁴ für MÜ nach Inhaltskategorien (3.3), indem man geeignete Materialien speziell für maschinelle Übersetzung im Internet finden kann, dokumentiert. Dazu werden die Vorgehensweise, wie man unterschiedliche Ressourcen recherchiert und diese verwendet, dargestellt. Ressourcen werden nach zwei Hauptkategorien (Informationsquellen und MÜ-Systeme) gesammelt und klassifiziert. Darauf werden die Vor- und Nachteile bei der Nutzung online verfügbarer Ressourcen (3.4) im Bereich der MÜ zusammengefasst.

Zuletzt wird ein Fazit sowie ein Ausblick auf weitere Aufgabenfelder und Kombinationsmöglichkeiten der MÜ mit anderen Technologien gegeben. Dazu werden die Ergebnisse der Arbeit über online verfügbare Ressourcen, d.h., die Bewertung der Ressourcen, gezeigt und erläutert.

⁴ Alle gefundenen online verfügbaren Ressourcen wurden im Anhang aufgelistet und auf einer separaten CD-ROM archiviert. Die CD-ROM wurde der Arbeit beigelegt.

2. Grundlagen

Heute wird die maschinelle Übersetzung (MÜ) in vielen Bereichen eingesetzt. Sie hat sich im Laufe der Zeit entwickelt. Das folgende Kapitel gibt dazu einen Überblick über die Grundlagen dieser Arbeit. Neben den Begriffserklärungen wird auch die Entwicklungsgeschichte der maschinellen Übersetzung in den 50er Jahren bis heute kurz erläutert. Außerdem wird im zweiten Teil dieses Kapitels der Aufbau und die Funktionsweise von zwei wichtigen Ansätzen der MÜ beschrieben: regelbasierte MÜ (RBMT⁵) und statistikbasierte MÜ (SMT⁶). Andere Ansätze zur MÜ wie beispielbasierte MÜ (EBMT⁷) werden nicht in dieser Arbeit dokumentiert.

2.1 Begriffserklärung und Entwicklungsgeschichte der maschinellen Übersetzung

Begriffserklärung

Der Begriff „*Maschinelle Übersetzung*“⁸ bezeichnet im engeren Sinne den Prozess oder Vorgang, Wörter, Sätze oder Texte aus einer natürlichen Sprache (Ausgangssprache) in eine andere Sprache (Zielsprache) mit Hilfe von Computerprogrammen vollautomatisch oder automatisch zu übersetzen. Unter „*Vollautomatische Maschinelle Übersetzung*“⁹ (FAMT¹⁰) versteht man eine Übersetzung, die ohne Eingriff des Menschen von einem Übersetzungswerkzeug erzeugt wird. „*Maschinelle Übersetzung mit menschlicher Hilfe*“ (HAMT¹¹) spielt auch eine wichtige Rolle bei den Formen der maschinellen Übersetzung. Dabei werden Wörter, Sätze oder ganze Texte eben von der Übersetzungssoftware automatisch übersetzt, allerdings hilft der Mensch dabei durch Vor- und Nachbearbeitung (Pre- bzw. Post-Edition) oder Verwendung von Kontrollierter Sprache (Controlled Language) des zu übersetzenden Textes. Daher spricht man momentan bei MÜ im engeren Sinne nur von der Verwendung von FAMT und HAMT.¹²

⁵ Die Abkürzung RBMT steht für „Rule-Based Machine Translation“

⁶ Die Abkürzung SMT steht für „Statistical Machine Translation“

⁷ Die Abkürzung EBMT steht für „Example Based Machine Translation“

⁸ Vgl. [Luckhardt 2010]

⁹ Vgl. [Luckhardt 2010], [Reinke 2005:2]

¹⁰ Die Abkürzung FAMT steht für „Fully Automatic Machine Translation“

¹¹ Die Abkürzung HAMT steht für „Human-Aided Machine Translation“

¹² Vgl. [Ramlow 2009:19-20], [Luckhardt 2010]

Entwicklungsgeschichte der Maschinellen Übersetzung

Um die Nutzung online verfügbarer Ressourcen für MÜ angemessen beurteilen zu können, ist es notwendig, über Hintergrundwissen zu deren Entwicklungsgeschichte zu verfügen. Daher wird im Folgenden ein historischer Überblick über die Entwicklung der maschinellen Übersetzung gegeben.

Die ersten Versuche, den Übersetzungsprozess von einer Sprache in die andere auf einer „elektromechanischen“¹³ Maschine automatisch durchzuführen, wurden im Jahre 1933 unternommen. Der Franzose Georges Artsrouni erhielt 1933 für seine elektromechanische Maschine ein Patent. Und gleichzeitig meldete im selben Jahr der Sowjetrusse Petr P. Smirnov-Troyanskij ein Patent für seine Übersetzungsmaschine an. Beide „maschinengestützten“¹⁴ Systeme basierten auf dem maschinellen Vergleich von Einträgen in Grundformen.

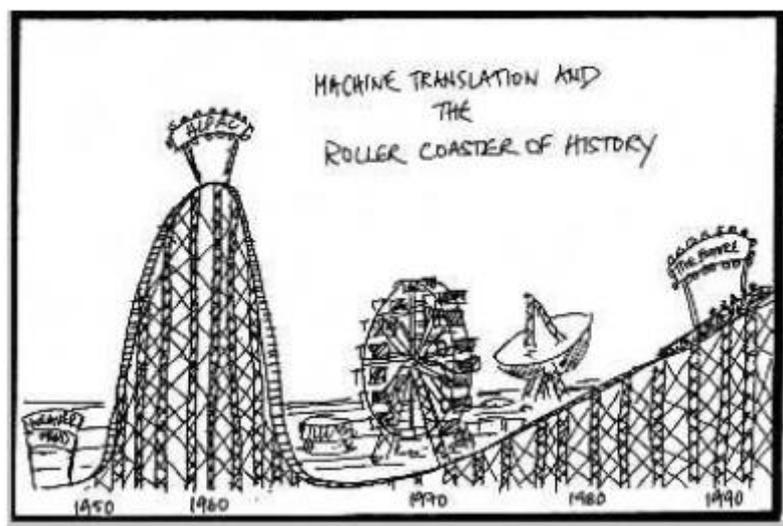


Abb. 2: Entwicklungsgeschichte der MÜ von 1950 bis 1990¹⁵

Die Geschichte der MÜ fing quasi kurz nach dem Zweiten Weltkrieg an. Einige Jahre später nach der Erfindung der Computer gaben im Jahre 1946 Warren Weaver und Andrew Donald Booth den Ansporn, Computer zur Entwicklung der maschinellen Übersetzungssysteme einzusetzen. Daher entwickelte 1948 Booth in Zusammenarbeit

¹³ Vgl. [D.O.G. 2006]

¹⁴ Vgl. [Ramlow 2008:28]

¹⁵ [Arnold 1994] D.J. Arnold; L. Balkan; S. Meijer; R. Lee. Humphreys; L. Sadler. Machine Translation: an Introductory Guide, ISBN: 1855542-17x. Hg. NCC Blackwell Ltd, London, 1994:15

mit Richard H. Richens ein Übersetzungssystem, das auf einem Wörterbuchvergleich in zwei Sprachen beruhte. Dieses System konnte außerdem „Wortstämme“ und „Flexionsendungen“ in beschränktem Umfang vergleichen und zwar eine Analyse auf Satzebene durchführen.¹⁶ Mit Hilfe dieses Übersetzungssystems ermöglicht es, grobe „Wort-für-Wort-Übersetzungen“¹⁷ für wissenschaftliche Texte zu verwenden.

Das Memorandum von Warren Weaver wurde im Jahre 1949 herausgegeben. Daher wurden die Entwicklungen der MÜ außerhalb des Fachgebiets bekannt gemacht. Nach Ramlow¹⁸ war Weavers Auffassung zu „optimistisch“, da Weaver der Meinung war, die vorliegenden Probleme lösen zu können, vor allem die Sprachmehrdeutigkeit zu verhindern, sei möglich. Allerdings war sein Bericht auch „zukunftsweisend“. Darin schlug er verschiedene Methoden zur MÜ vor, z.B. „informatische und linguistische Methoden bei der MÜ-Forschung zu verbinden“ (d.h., man setzt computertechnische und sprachliche Verfahren bei der Forschungsarbeit ein).

Yehoshua Bar-Hillel war der erste Wissenschaftler im Fachgebiet der MÜ und organisierte 1952 am MIT¹⁹ die erste Konferenz über MÜ. Auf der Konferenz wurde die positive MÜ-Einschätzung bestätigt und es wurden Ziele für die zukünftige MÜ-Forschung erklärt und gesetzt. Die Konferenz führte zu einem schnellen Aufstieg der MÜ. 20 Millionen Dollar wurden in den folgenden zehn Jahren zur MÜ-Forschung beigetragen. Im Jahre 1959 drückte Bar-Hillel in seinen *«Report on the state of machine translation in the United States and Great Britain»* seine Zweifel an der Möglichkeit der vollautomatischen Übersetzung aus. Er glaubte, dass die hochqualitative vollautomatische Übersetzung ohne menschlichen Eingriff auf Dauer unmöglich sein würde. Linguistische Probleme im maschinellen Übersetzungsprozess wie Mehrdeutigkeit wurden dann deutlich.²⁰

Die MÜ-Forschung für das System METAL mit Sprachpaaren Deutsch-Englisch ging ursprünglich auf die Gründung des Sprachforschungszentrums (LRC²¹) an der Univer-

¹⁶ Vgl. [Ramlow 2008:29]

¹⁷ Vgl. [Ramlow 2008:29]

¹⁸ Vgl. [Ramlow 2008:30]

¹⁹ Die Abkürzung MIT steht für „Massachusetts Institute of Technology“. Das ist eine Technische Hochschule und Universität in Cambridge, Massachusetts in den USA.

²⁰ Vgl. [Ramlow 2008:31]

²¹ Die Abkürzung LRC steht für „Linguistics Research Center“

sität in Texas in Austin im Jahr 1961 zurück.²² Die Komplexität linguistischer Probleme im maschinellen Übersetzungsprozess wurde zunehmend erkannt.²³

1954	Wissenschaftler der Georgetown University und der Firma IBM stellen das erste maschinelle Übersetzungssystem (Russisch – Englisch) vor.
60er Jahre	Beginn der Entwicklung von METAL für die deutsch-englische Übersetzung an der University of Texas. Die Komplexität natürlicher Sprache und linguistischer Probleme wird zunehmend deutlich, die Begeisterung für maschinelle Übersetzung (MÜ) sinkt.
1966	ALPAC-Report: Negative Bewertung der Möglichkeiten maschineller Übersetzung.
1977	Fertigstellung des kanadischen MÜ-Systems Taum-METEO. Erneut Hinwendung zur MÜ.
heute	MÜ-Werkzeuge im professionellen Einsatz. Regelbasierte und statistische Systeme.

Abb. 3: Historischer Überblick über die Entwicklung der MÜ²⁴

Im Jahre 1966 wurden die Möglichkeiten maschineller Übersetzung im ALPAC-Bericht in Bereichen „Einsatznotwendigkeit“, „schnelle Verfügbarkeit“, „schnelle Kostenreduzierung“ und „deutliche Leistungsverbesserung“ negativ beurteilt. Da der Bedarf am Beispiel der Übersetzung im Sprachpaar Chinesisch-Englisch für die Einsatznotwendigkeit sehr niedrig war, wurde die Einsatznotwendigkeit der MÜ mit „unnötig“ eingeschätzt. Allerdings wurde der ALPAC-Bericht einerseits als „engstirnig“, „voreingenommen“ und „kurzsichtig“ kritisiert, andererseits wurde Geld für die MÜ-Forschung sichtbar gekürzt.²⁵

In den 70er Jahren wurde die MÜ wegen einer großen Nachfrage nach MÜ-Systemen weltweit wiederbelebt.²⁶ Die Forschungsgruppe TAUM²⁷ hat im Jahre 1977 die erste Version des kanadischen MÜ-Systems METEO mit dem Sprachpaar Eng-

²² Vgl. [Hutchins 1992A], [Seewald-Heeg 2011:71]

²³ Vgl. [Seewald-Heeg 2011:71]

²⁴ Vgl. [Seewald-Heeg 2011:71]

²⁵ Vgl. [Ramlow 2008:32], [Seewald-Heeg 2011:71]

²⁶ Vgl. [Ramlow 2008:33]

²⁷ Die Abkürzung TAUM steht für „Traduction Automatique à l'Université de Montréal“

lisch-Französisch in Betrieb genommen.²⁸ Das METEO-System hat zur maschinellen Übersetzung der Wettervorhersage vom Englischen ins Französische für das nationale Umweltministerium von Kanada gedient.²⁹

MÜ-Systeme werden heute im professionellen und beruflichen Umfeld eingesetzt, um die Wirtschaftlichkeit im Übersetzungsprozess sicherzustellen. Die wichtigsten Vertreter der MÜ-Ansätze sind die regelbasierte MÜ und die statistische MÜ. Die beiden MÜ-Ansätze werden im folgenden Abschnitt aufgezeigt.

2.2 Ansätze der maschinellen Übersetzung

Um MÜ-Systeme effizient nutzen zu können, ist es notwendig, sich über den Aufbau und die Funktionsweise der MÜ-Ansätze zu informieren. Im Laufe der Zeit haben sich viele verschiedene Ansätze zur MÜ entwickelt. Die maschinelle Übersetzung ist für uns ein Werkzeug, das Übersetzungsproduktivität beschleunigt. Bei einer Typologie der MÜ lassen sich heute die 2 wichtigsten Ansätze unterscheiden: regelbasierte MÜ (RBMÜ) und statistikbasierte MÜ (SBMÜ). Alternative Ansätze wie beispielbasierte MÜ (EBMT³⁰), kontextbasierte MÜ (CBMT³¹), wissensbasierte MÜ (KBMT³²) und hybride MÜ werden in der Arbeit nicht erwähnt.

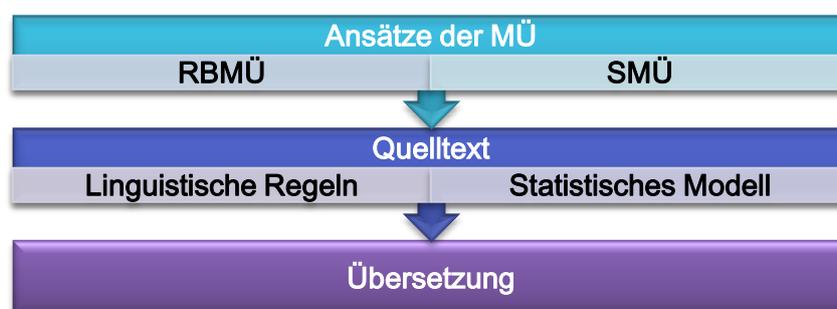


Abb. 4: Regelbasierte MÜ vs. statistikbasierte MÜ im Überblick

²⁸ Vgl. [Seewald-Heeg 2011:71]

²⁹ Vgl. [Hutchins 1992B], [Reinke 2005:3]

³⁰ Die Abkürzung EBMT steht für „Example Based Machine Translation“

³¹ Die Abkürzung CBMT steht für „Context Based Machine Translation“

³² Die Abkürzung KBMT steht für „Knowledge Based Machine Translation“

formen) analysiert, um die Transformations-Regeln anzuwenden und deren Sprachstruktur wurde direkt in die Struktur des Zieltexts überführt.³⁶ Dazu wird ein bilinguales Wörterbuch verwendet. Regelbasierte direkte MÜ-Systeme beschränken sich auf begrenzte Anwendungsbereiche. Es liegt daran, dass es für eine große Gruppe von Wörtern, die mehr als eine Bedeutung haben, mehrere mögliche Übersetzungen gibt. Auch die Übersetzung von Mehrwortausdrücken, wie z.B. ‚ins Gras beißen‘ oder ‚überall seinen Senf dazu geben müssen‘, können Probleme verursachen.³⁷

Transfer-System

Unter „Transfer“ versteht man eine Übertragung, die zwischen den Interface-Strukturen der Ziel- und Quellsprachen eingesetzt wurde:³⁸ die Analyse des Ausgangstexts produziert eine abstrakte Repräsentation des Ausgangstexts (ausgangssprachliche Interface-Struktur); die abstrakte Repräsentation des Zieltexts (zielsprachliche Interface-Struktur) ist für die Synthese bzw. Generierung des Zieltexts erforderlich.³⁹ Diese drei Schritte wurden bei den Transfer-Ansätzen durchgeführt, z.B. ein deutscher Quelltext wird zuerst syntaktisch geparkt; die Analyse führt zu einer abstrakten Repräsentation des deutschen Textes; für die Übertragung werden bilinguale Regeln benutzt, um die Strukturen der Quellsprache mit Strukturen der Zielsprache zueinander in Bezug zu setzen; schließlich wird der Zieltext aus der ausgangssprachlichen Interface-Struktur syntaktisch generiert. Beim Transfer-System können neben syntaktischen Regeln morphologische und semantische Informationen berücksichtigt werden.⁴⁰ Es scheint in der Praxis, dass es interne Konflikte und widersprechende Regeln aus allen drei Quellen gibt. Neue Fehler werden produziert.⁴¹ Allerdings ist der Aufwand dafür so groß, dass man für n Sprachen $n * (n-1)$ Transfer-Module benötigt, d.h., für 1 zusätzliches Sprachpaar sind 4 weitere Transfer-Module erforderlich.⁴² Und für die EU sind alle Berichte, Protokolle, Erlasse in 552 Sprachrichtungen (Stand 2014: 24 offizielle EU-Sprachen) zu übersetzen.

³⁶ Vgl. [Reinke 2005:7], [Seewald-Heeg 2011:76]

³⁷ Vgl. [Seewald-Heeg 2009:8], [Seewald-Heeg 2011:93-99]

³⁸ Vgl. [Reinke 2005:9]

³⁹ Vgl. [Reinke 2005:9], [Seewald-Heeg 2011:78]

⁴⁰ Vgl. [Seewald-Heeg 2009:8]

⁴¹ Vgl. [Seewald-Heeg 2009:8]

⁴² Vgl. [Seewald-Heeg 2011:79], [Reinke 2005:9-10]

Interlingua-System

Unter „Interlingua“ versteht man eine Repräsentation, die mehreren Sprachen gemeinsam ist.⁴³ Diese abstrakte universelle und sprachunabhängige Repräsentation ist die Grundlage eines Interlingua-Übersetzungssystems. So würde sie funktionieren: ein Quelltext wird vollständig analysiert und ein neuer Zieltext, der vom Ausgangstext völlig unabhängig aber gleichwertig ist, wird generiert.⁴⁴ Das heißt, dass man mit den Interlingua-Ansätzen eine universalsprachliche Repräsentation mit beliebigen Sprachen verknüpfen kann. Diese ideale universalsprachliche Repräsentation ist vor allem bis heute noch utopisch. Allerdings lassen sich nicht alle Nuancen der Sprache damit abbilden: z.B. ‚Man bringt einen neuen Artikel auf den Markt‘, ‚Auf den Markt bringt man einen neuen Artikel‘ und ‚Einen neuen Artikel bringt man auf den Markt‘. Der Satzteil wird in diesen drei Sätzen unterschiedlich betont. Außerdem ist die Gleichwertigkeit⁴⁵ der Sprachpaare problematisch: z.B. heißt der englische Satz ‚I went to the cinema.‘ auf Deutsch ‚Ich ging ins Kino.‘ oder ‚Ins Kino ging ich.‘ Diese beiden deutschen Sätze unterscheiden sich minimal.

2.2.2 Statistikbasierte MÜ

Die statistische MÜ ist ein auf Wahrscheinlichkeit basierender MÜ-Ansatz.⁴⁶ Beim statistischen MÜ-Ansatz wird ein großes zweisprachiges Textkorpus verwendet, um einen Text zu übersetzen. Als Datenbasis nutzt man die bereits bestehenden Übersetzungsbeispiele, die als Korpora bezeichnet werden. Daher gehört „statistische maschinelle Übersetzung“ bzw. „statistikbasierte maschinelle Übersetzung“ zu den korpusbasierten⁴⁷ MÜ-Ansätzen. Wörter und grammatische Formen werden aufgrund der Häufigkeit einander zugeordnet: auf diesem Weg werden Wörterbuch und Grammatikübertragungsregeln extrahiert und dadurch wird der Ausgangstext mittels eines statistischen Analysemodells in umfangreichen Textsammlungen automatisch übersetzt.⁴⁸ Die Typen der SMÜ lassen sich in zwei Analyseebenen unterteilen: die Ebene von Wörtern und die

⁴³ Vgl. [Reinke 2005:8]

⁴⁴ Vgl. [Seewald-Heeg 2009:8]

⁴⁵ Vgl. [Reinke 2005:9]

⁴⁶ Vgl. [SDL 2013]

⁴⁷ Vgl. [Ramlow 2008:42]

⁴⁸ Vgl. [Seewald-Heeg 2011:74], siehe auch „Maschinelle Übersetzung“ in Wikipedia, http://de.wikipedia.org/wiki/Maschinelle_%C3%9Cbersetzung [28.11.2014]

Ebene von Phrasen.⁴⁹ Aufgrund des wachsenden Bedarfs an Übersetzungen in und aus verschiedenen Sprachen werden die Informationsmengen immer größer. Im Hinblick auf die großen Datenmengen (Übersetzungsbeispiele) haben SMÜ-Systeme gegenüber RBMÜ-Systemen Vorteile: neue Sprachpaare schnell zu erstellen, Sprachpaare einfach für bestimmte Fachgebiete zu trainieren und Übersetzungsqualität in viel kürzerer Zeit und kontinuierlicher zu verbessern.⁵⁰

Wortbasierte SMÜ⁵¹

Der ursprüngliche Ansatz der SMÜ analysiert die Trainings- und Testmaterialien auf der Ebene von Wörtern. Das heißt, ein Wort im Ausgangstext muss mindestens einem Wort im Zieltext entsprechen. Es wäre mit der wortbasierten SMÜ möglich, dass ein Wort im Ausgangstext mehreren Wörtern im Zieltext entspricht, z.B. Deutsch ‚Schnapsidee‘ => Englisch ‚pipe dream‘. Die Umkehrrichtung ist jedoch mit der wortbasierten SMÜ unmöglich, da jedem Wort in der Quellsprache mindestens ein Wort in der Zielsprache entsprechen muss. Ein weiteres Problem bei der wortbasierten SMÜ besteht darin, dass zusammengesetzte Wörter, vor allem bei Verben mit abtrennbaren Präfixen in der deutschen Sprache, nicht zusammen übersetzt werden können.

Phrasenbasierte SMÜ⁵²

Um eine Lösungsmöglichkeit für die Probleme der wortbasierten SMÜ zu finden, wurde nach neuen Ansätzen zur SMÜ auf der Ebene von Phrasen geforscht. Diese phrasenbasierten SMÜ-Systeme sind jedoch nicht „linguistisch motiviert“. Im Gegensatz zu den wortbasierten SMÜ-Systemen werden die Trainings- und Textmaterialien in den phrasenbasierten SMÜ-Systemen maschinell in Gruppen eingeteilt und es müsste daher die Reihenfolgen der Wörter gekennzeichnet werden. Auf diese Art und Weise ist es möglich, mehrere Wörter mit einem Wort zu übersetzen und umgekehrt. Ein weiterer Vorteil der phrasenbasierten SMÜ legt nahe, dass mehrere Wörter mit erweitertem Kontext übersetzt werden könnten, um die Ambiguität von Wörtern zu beseitigen. Beispielsweise ist die phrasenbasierte SMÜ in der Lage, eine Entscheidung für die Übersetzung des englischen Worts ‚pretty‘ in den Kontexten ‚pretty sure‘ und ‚pretty flower‘ zu treffen.

⁴⁹ Vgl. [Seewald-Heeg 2009:11]

⁵⁰ Vgl. [SDL 2013]

⁵¹ Vgl. [Seewald-Heeg 2009:11-12]

⁵² Vgl. [Seewald-Heeg 2009:12]

3. Nutzung online verfügbarer Ressourcen für die maschinelle Übersetzung

Um die online verfügbaren Ressourcen für die maschinelle Übersetzung im Rahmen eines deutschen Fernstudiums effizient nutzen zu können, werden die Informationsquellen und Ressourcen nach Inhaltskategorien klassifiziert und als *Hyperlink* im folgenden Kapitel gesammelt. Zu diesem Zweck werden der Suchverlauf und die Kriterien für die Suche nach online verfügbaren Ressourcen im ersten Teil dieses Kapitels beschrieben. Dann folgen die weiteren technischen oder sprachlichen o.a. Voraussetzungen zur Nutzung dieser Technologie. Der dritte Teil dieses Kapitels ist die Ressourcen-Sammlung. Die Vor- und Nachteile bei der Nutzung online verfügbarer Ressourcen werden zuletzt aufgezeigt.

3.1 Suchverlauf und Suchkriterien online verfügbarer Ressourcen

Zahlreiche Ressourcen für die MÜ wurden bereits im Internet zur Verfügung gestellt und der Zugriff auf diese ist mit Hilfe von Suchmaschinen meistens möglich. Alle in der Arbeit gesammelten online verfügbaren Ressourcen werden durch die *Google-Suche*⁵³ mit Suchwörtern in deutscher Sprache „Maschinelle Übersetzung“ bzw. in englischer Sprache „machine translation“ oder mit Schlüsselwörtern wie „Maschinelle Übersetzung Ressourcen“ o. ä., die mit dem Kernbegriff „Maschinelle Übersetzung“ verbunden sind, gefunden. Die Ressourcen sind meistens deutsch oder englisch verfasst. Als Materialien eines Fernstudiums gibt es auch Ressourcen in chinesischer Sprache mit englischen Untertiteln zur MÜ, da es viele chinesische Studierenden an deutschen Hochschulen oder Universitäten gibt. Alle Ressourcen konnten in unterschiedlichen Formen, wie z.B. Folien, E-Books, Filme, Software, usw. archiviert werden. Deshalb werden die ausgewählten inhaltsähnlichen Ressourcen in zwei Hauptkategorien „Informationsquellen“ und „MÜ-Systeme“ untergeteilt.

Vor allem müssen alle ausgewählten Ressourcen im Rahmen eines Fernstudiums hilfreich, nützlich und frei zugänglich für die zulässige Verwendung der MÜ sein. Ressourcen mit beschränkter Bedienbarkeit oder mit eingeschränkten Nutzungsbedingungen sollten genau dokumentiert werden. Um die gefundenen Ressourcen langfristig zu behalten, werden sie auf einer separaten CD-ROM archiviert. Die CD-ROM wird der Arbeit beigelegt. Die archivierten Ressourcen werden als freiverfügbare Materialien für

⁵³ Google Suche, <https://www.google.de/> [28.11.2014]

ein Fernstudium nach „Informationsquellen“ und „MÜ-Systemen“ eingeordnet, so dass man die Materialien nach Inhalt schnellstmöglich finden kann. Nicht alle der Ressourcen, die das gleiche Thema oder einen ähnlichen Inhalt haben, werden in der Arbeit dokumentiert, werden aber trotzdem zum Nachschlagen im Anhang und auf der CD-ROM archiviert.

3.2 Voraussetzungen für die Nutzung online verfügbarer Ressourcen

Zur Anwendung der MÜ-Systeme gibt es bestimmte Voraussetzungen, die erfüllt werden müssen. Die Voraussetzungen für die Nutzung online verfügbarer Ressourcen im Rahmen eines Fernstudiums zur maschinellen Übersetzung lassen sich nach folgenden Aspekten kategorisieren:

- ◆ Sprachliche Grundkompetenz: Vor allem muss die sprachliche Kenntnisse der deutschen und englischen Sprache als notwendige Bedingungen für die Nutzung der gesammelten online verfügbaren Ressourcen vorausgesetzt werden. Wenn man etwas nicht verstehen kann, ist es unmöglich, etwas zu lernen und zu üben.
- ◆ Technische Anforderungen: Um auf die Online-Ressourcen zugreifen zu können, muss man einen Internetzugang besitzen. Die Nutzung eines MÜ-Systems ist mittels Computer oder mobiler Geräte möglich. Um das MÜ-System effizient nutzen zu können, muss die Leistungsfähigkeit der PCs den Anforderungen des MÜ-Systems entsprechen. Um online verfügbare Ressourcen in einem bestimmten Format (wie z.B. Folien im PDF-Format) nutzen zu können, muss man entsprechende Softwareanwendung auf seinem PC installieren.
- ◆ Kostenmäßige Beschränkung: Diese Voraussetzung bezieht sich auf die Kosten für den Erhalt einiger online verfügbarer Ressourcen. Die meisten im Rahmen der Arbeit zusammengestellten online verfügbaren Ressourcen sind frei zugänglich, um diese im Studium ohne finanzielle Aufwände nutzen zu können. Es gibt jedoch Ressourcen, die kostenpflichtig sind oder deren Nutzung an bestimmte Nutzungsbedingungen geknüpft ist.

- ◆ Fachliche Vorkenntnisse: Grundkenntnisse auf den Fachgebieten Sprachwissenschaft, Computerlinguistik und Informatik sind für einige gesammelten Ressourcen erforderlich.
- ◆ Regionale Zugänglichkeit: Der Zugriff auf die gesammelten Ressourcen im Internet kann in anderen Staaten oder Regionen möglicherweise eingeschränkt sein. Der Zugriff auf die Ressourcen wurde nur in Deutschland geprüft. In China besitzt man beispielsweise aus politischen oder gesellschaftlichen Gründen kein Recht, auf *YouTube*⁵⁴ zuzugreifen.

3.3 Klassifizierung online verfügbarer Ressourcen nach Inhaltskategorien

Die herausgefundenen Ressourcen⁵⁵ werden in zwei Hauptkategorien klassifiziert

In der Kategorie „*Informationsquellen*“ kann man allgemeine Informationen über die Geschichte, verschiedene Ansätze und die Einsatzszenarien der MÜ finden und genau betrachten. Hier beziehen sich die Einsatzszenarien der MÜ auf die Post-Edition und kontrollierte Sprache. Der Einsatz unterschiedlicher MÜ-Systeme wurde in der folgenden Kategorie „MÜ-Systeme“ erklärt. Daher kann man sich mit den gegenwärtigen und vielfältigen Anwendungsbereichen der MÜ vertraut machen.

In der Kategorie „*MÜ-Systeme*“ lassen sich die MÜ-Systeme in „online“ und „lokal“ unterscheiden. Auf dem globalen Markt werden die allgemeinen online und lokalen MÜ-Systeme weithin erkannt. Online verfügbare MÜ-Systeme werden darauf erwähnt und erklärt. Lokale MÜ-Systeme beziehen sich auf diejenigen, die aus dem Internet heruntergeladen und auf dem eigenen PC nachträglich installierbar sind.

3.2.1 Informationsquellen

Quellen mit Informationen über die MÜ werden in drei Inhaltsaspekten eingeteilt: Geschichte, Ansätze der MÜ-Systeme und Einsatzszenarien.

⁵⁴ YouTube, <http://www.youtube.com> [28.11.2014]

⁵⁵ Die Bewertung der einzelnen ausgewählten Ressourcen befindet sich im Fazit.

Geschichte

- Nach Eingabe der Suchwörter „machine translation introduction“ bei *Google* findet man auf der Seite von W. John Hutchins⁵⁶ u.a. das Buch «[An introduction to machine translation](#)»⁵⁷ aus dem Jahre 1992 im PDF-Format finden. Das Buch wurde in englischer Sprache verfasst. Jedes Kapitel kann in einer PDF-Datei aufgerufen werden. Einen historischen Überblick kann man direkt im Kapitel 1.3 „[Brief history of MT](#)“⁵⁸ finden. Dies ist eine umfangreiche Einführungsliteratur zur MÜ, da man sich sowohl Hintergrundwissen über die MÜ (Einführung, historischer und linguistischer Hintergrund, technische Aspekte, grundlegende Strategien, Problematik und Analyse und Generation inklusiv) als auch die Evaluierung für die oft verwendeten MÜ-Systeme aneignen kann.
- Nach Eingabe der Suchwörter „machine translation an introductory guide“ bei *Google* kann man das Buch «[MACHINE TRANSLATION: An Introductory Guide](#)»⁵⁹ von D. Arnold, L. Balkan, S. Meijer, R. Lee Humphreys und L. Sadler im PDF-Format finden. Das Buch wurde in englischer Sprache verfasst. Einen historischen Überblick kann man direkt im Kapitel 1.4 „A Bit of History“ finden. Dies ist auch ein umfangreiches Literaturwerk zur MÜ, da man systematische Informationen über MÜ (Einführung und Überblick, MÜ in der Praxis, linguistisches Wissen und dessen Architektur, Terminologie, Übersetzungsprobleme und Evaluierung der MÜ-Systeme inklusiv) finden kann. Im Vergleich zum Buch «*An introduction to machine translation*» ist dieses Buch ausführlicher und systematischer in einigen Aspekten. Beide Werke sind keine rein historischen Bücher und enthalten ausführliche Informationen zur MÜ.
- Ein rein historischer Artikel «[Die Übersetzungstechnologien und ihre Pioniere](#)»⁶⁰ wird nach Eingabe der Suchwörter „Maschinelle Übersetzung Pioniere“ bei *Google* gefunden. Der Artikel wurde in deutscher Sprache verfasst und auf der Unternehmenswebseite von der D.O.G. GmbH veröffentlicht. Ein kurzer historischer Überblick über die Entwicklungsgeschichte und die Pioniere der MÜ wird in diesem Ar-

⁵⁶ Vgl. [Hutchins 1992C]

⁵⁷ Inhaltsübersicht unter: <http://www.hutchinsweb.me.uk/IntroMT-TOC.htm> [28.11.2014]

⁵⁸ Vgl. [Hutchins 1992D]

⁵⁹ Vgl. [Arnold 1996]

⁶⁰ Vgl. [D.O.G. 2006]

tikel gegeben. Diese Ressource hat im Gegensatz zu den oben erwähnten Büchern und Folien den Vorteil, die Entwicklungsgeschichte der MÜ in einen kurzen Überblick darzustellen.

- Nach Eingabe der Suchwörter „machine translation“ bei *Google* kann man zwar Materialien eines Fernstudiums zum Thema „[machine translation](#)“⁶¹ finden. Alle Materialien zu diesem Fernstudium liegen in englischer Sprache vor. Informationen über die Einführung und Entwicklung der MÜ kann man durch Klicken auf das erste Thema „Introduction“ in der Registerkarte „Syllabus“ erhalten. Nach dem Klicken wurde die Webseite direkt zur Seite „[Introduction](#)“⁶² geführt. Hierzu wurden Folien im PPT-Format gesammelt. Der Unterricht wurde anhand vieler Beispiele mit dem Sprachpaar Englisch-Chinesisch erteilt. Auf der Webseite gibt es sehr viele praktische Aspekte der MÜ, die anhand vieler Beispiele mit verschiedenen Sprachpaaren in unterschiedlichen Formaten gesammelt wurden. Für diejenige, die sich für die Vertiefung unterschiedlicher MÜ-Modelle (vor allem „Translation Models“) in verschiedenen Sprachpaaren interessieren, ist diese Ressource geeignet. Daraufhin stehen auch viele Übungen für eingeschriebene Studenten zur Verfügung. Die Einschreibung erfolgt kostenfrei per E-Mail-Bestätigung. Allerdings kann man denselben Vortrag zum Thema „[Machine Translation: Models, Search, and Evaluation](#)“⁶³ auch auf der Website [Videlectures.net](#) finden, wo er ebenfalls frei zugänglich ist.

Ansätze der MÜ-Systeme

- Nach Eingabe der Suchwörter „Maschinelle Übersetzung“ bei *Google* kann man den Beitrag «[Maschinelle Übersetzung – ein Überblick](#)» aus dem Jahre 2009 von Daniel Stein⁶⁴ im PDF-Format finden. Der Beitrag wurde in deutscher Sprache verfasst und in der Fachzeitschrift JLCL⁶⁵ zum Thema «[Maschinelle Übersetzung – von Theorie zur Anwendung](#)“⁶⁶ veröffentlicht. Auch die anderen Beiträge der Zeitschrift sind für

⁶¹ Fernstudium zum Thema „machine translation“ unter: <http://mt-class.org/jhu/index.html> [28.11.2014]

⁶² Vgl. [Lopez 2014A]

⁶³ Vgl. [Lopez 2010]

⁶⁴ Vgl. [Stein 2009], [Seewald-Heeg 2009:5-18]

⁶⁵ Die Abkürzung JLCL steht für „Journal for Language Technology and Computational Linguistics“. JLCL-Portal, <http://www.jlcl.org/> [28.11.2014]

⁶⁶ Vgl. [Seewald-Heeg 2009]

Interessierte empfehlenswert. Im Kapitel 2 „Typologie“ der Arbeit Steins werden verschiedene Ansätze zur MÜ dargestellt. Daraufhin werden sowohl der Aufbau und die Funktionsweise der zwei wichtigsten Vertreter – regelbasierte MÜ und statistikbasierte MÜ als auch der Aufbau der anderen Ansätze zur MÜ wie beispielbasierte MÜ, kontextbasierte MÜ, wissensbasierte MÜ und hybride MÜ resümiert. Dazu bietet der Beitrag die Möglichkeit, verschiedene Ansätze zur MÜ auf einen Blick zu betrachten.

- Nach Eingabe der Suchwörter „machine translation mooc“ bei *Google* kann man das MOOC⁶⁷ zum Thema „[Principles and Practice of Computer-Aided Translation](#)“⁶⁸ auf der Webseite von „[MOOC 学院](#)“⁶⁹ finden. Im Feld oben rechts auf der Website wird ein kurzer Lehrfilm zum Abstract angezeigt (siehe Abb. 6). „MOOC 学院“ ist das größte chinesische online Studienforum. „学院“ bedeutet im Chinesischen „Hochschule“ oder „Universität“. Die Filme wurden von einem Professor an der Peking Universität erstellt und in chinesischer Sprache mit englischen Untertiteln verfasst. Die Inhalte der Onlinekurse sind das Grundkonzept computerunterstützter Übersetzungstechnologie und die Verwendung computerunterstützter Übersetzungswerkzeuge. Diese Ressource ist für eingeschriebene Studenten frei zugänglich. Die Einschreibung erfolgt kostenfrei durch Eingeben von Namen und E-Mail-Adressen. Während des MOOCs können die Studenten mit dem Lehrer kommunizieren, um Fragen zu lösen.



Abb. 6: Lehrfilm zum Abstract auf der Website von „MOOC 学院“⁷⁰

⁶⁷ MOOC ist die Abkürzung von „Massive Open Online Course“ und bezeichnet Onlinekurse, die für alle eingeschriebenen Studenten frei zugänglich sind.

⁶⁸ Vgl. [Yu 2014]

⁶⁹ MOOC 学院, <http://mooc.guokr.com/> [28.11.2014]

⁷⁰ Vgl. [Yu 2014]

Wenn man diesen Onlinekurs abgelegt hat, kann man ein Zertifikat (siehe Abb. 7) bekommen. Im Gegensatz zu dem Beitrag von Daniel Stein sind diese Onlinekurse praktischer, da man während des Kommunizierens mit dem Lehrer die Probleme mit der praktischen Verwendung der CAT-Tools lösen kann. Zu den Unterschieden und Zusammenhängen zwischen MÜ und CAT⁷¹ ist der Beitrag [«CAT und MÜ – Getrennte Welten?»](#)⁷² empfehlenswert. Andererseits ist dieser Onlinekurs so zeitlich und thematisch beschränkt, dass man dem Stundenplan immer folgen muss und sich Wissen nur im CAT-Umfeld aneignen kann. Allerdings eignet sich dieser Onlinekurs für Doktoranden und Studenten im Fachbereich Übersetzen und Dolmetschen.

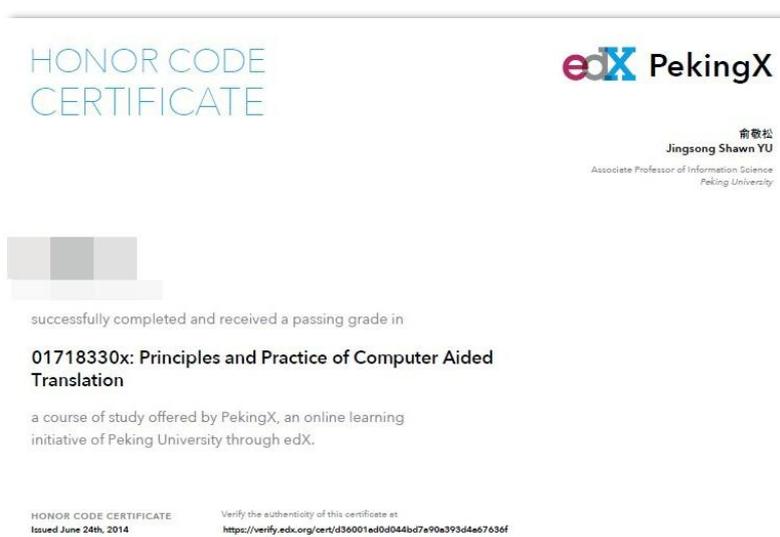


Abb. 7: Zertifikat bei „MOOC 学院“⁷³

- Das *SDL Download-Center*⁷⁴ wurde in den Suchergebnissen von *Google* mit dem Schlagwort „Maschinelle Übersetzung“ angezeigt. Das ist eine Online-Archivierung von SDL-Anwendungen und Datenbanken. Informationen können mittels 3 Filter und den zusätzlichen Filteroptionen (rot markiert in Abb. 8 auf der nächsten Seite) in verschiedenen Dokumentationsarten⁷⁵, Themen⁷⁶ und in Bereichen des Produkts⁷⁷

⁷¹ Die Zusammenhänge und Unterschiede zwischen MÜ- und CAT-Systemen werden nicht in dieser Arbeit dokumentarisch dargestellt.

⁷² Vgl. [Seewald-Heeg 2009:19-36]

⁷³ Vgl. [Yu 2014]

⁷⁴ SDL Download-Center, <http://www.sdl.com/de/resources/> [28.11.2014]

⁷⁵ Herunterladbare Dokumentationsarten: Aufgezeichnetes Webinar, Branchenforschung, Dienstleistungsbeschreibung, E-Books, Fallstudie, Infographik, Lösungsübersicht, Podcast, Präsentation, Produktbeschreibung, Schulungsinformationen, Technische Dokumente, Video und Whitepaper

und Services⁷⁸ gesucht werden. Die Suchergebnisse erscheinen in Sekundenschnelle auf dem Bildschirm.

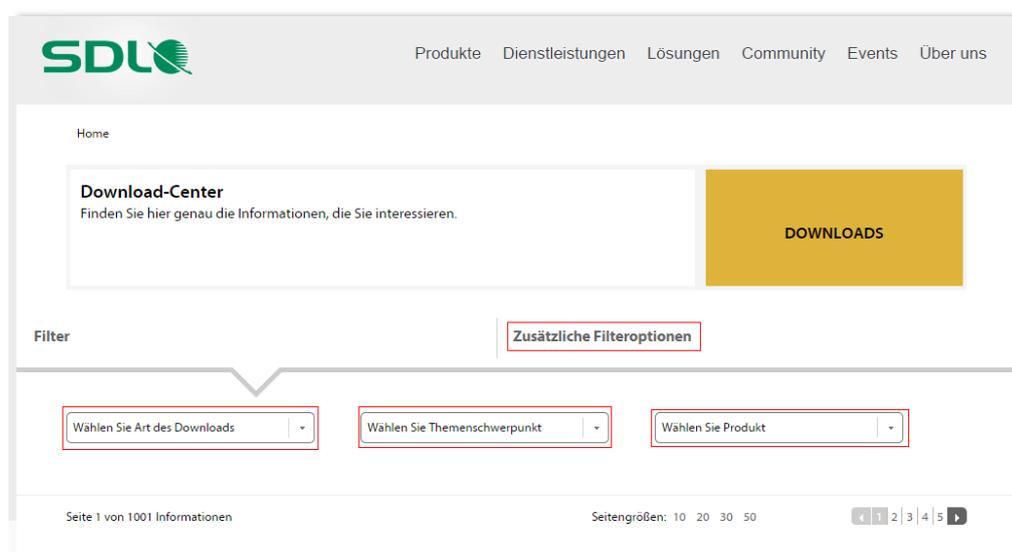


Abb. 8: Filteroptionen des SDL Download-Centers⁷⁹

Die gefilterten Ressourcen⁸⁰, die nach dem Themenfeld „Maschinelle Übersetzung“ herausgefunden werden, eignen sich für diese Arbeit und zum Selbststudium der MÜ. Es werden insgesamt 15 Seiten Dokumentationen zur MÜ in den Suchergebnissen gefunden. Alle vorhandenen Dokumentationen sind frei herunterladbar. Zum Herunterladen der gewünschten Ressourcen (siehe Abb. 9 auf der nächsten Seite) ist das Ausfüllen bestimmter personenbezogener Daten⁸¹ (Vorname, Nachname, Unternehmen, Land und E-Mail) erforderlich. Wie zum Beispiel, das erste Dokument *«Regelbasierte und statistische Ansätze»*⁸² in den Suchergebnissen bei SDL Download-Center, das aus dem Jahre 2013 in deutscher Sprache und im

⁷⁶ Themenfelder: Analysen und Optimierung, DITA, Echtzeit-Targeting, Globalisierungsberatung, Integriertes Global Information Management, Kampagnen- und E-Mail-Management, Language Services, Maschinelle Übersetzung, Medienverwaltung, Multi-Channel-Bereitstellung, S1000D, Schreiben für globale Märkte, Social Intelligence und Consulting, Structured Content – DITA und S1000D, Terminology management, Übersetzungsproduktivität, Übersetzungsmanagement, Web Content Management und XML-Publishing

⁷⁷ Produktbereiche: 290 Alle SDL-Produkte, Analytics, Campaigns, Documentation, eCommerce, Language, Social und Web. Außer dem Bereich „290 Alle SDL-Produkte“ könnte man weitere Optionen aussuchen und eingrenzen.

⁷⁸ Service zum Auswählen: 010 Beratung, 020 Sprachdienstleistungen, 030 Implementierungsservices und 040 Ausbildung und Zertifizierung. Alle Services sind mit Unteroptionen gestaltet.

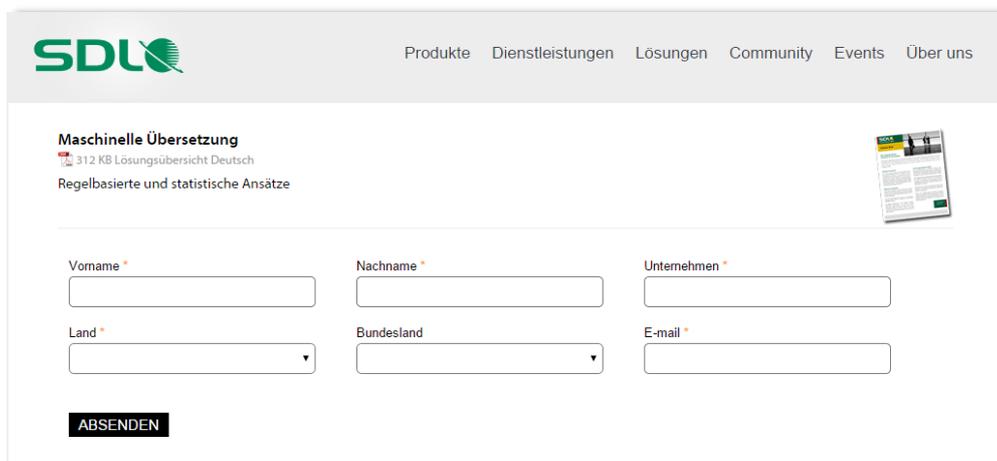
⁷⁹ Vgl. [SDL 2014A]

⁸⁰ Vgl. [SDL 2014A]

⁸¹ Pflichtangaben sind mit „*“ gekennzeichnet

⁸² Vgl. [SDL 2013]

PDF-Format archiviert wurde, ist erst nach der Eingabe personenbezogener Daten kostenlos erhältlich. Weitere nützliche Dokumentationen zur MÜ kann man selbst im SDL Download-Center finden.



The screenshot shows the SDL website's registration form for machine translation resources. The header includes the SDL logo and navigation links: Produkte, Dienstleistungen, Lösungen, Community, Events, Über uns. The main content area is titled 'Maschinelle Übersetzung' and includes a sub-header '312 KB Lösungsübersicht Deutsch' and the text 'Regelbasierte und statistische Ansätze'. The form consists of several input fields: 'Vorname *', 'Nachname *', 'Unternehmen *', 'Land *', 'Bundesland', and 'E-mail *'. A black 'ABSENDEN' button is located at the bottom left of the form area.

Abb. 9: Pflichtfelder zum Erhalten von SDL-Ressourcen⁸³

Einsatzszenarien

- Nach Eingabe der Suchwörter „Post-Editing machine translation“ wurden die Richtlinien «[Machine Translation Post-Editing Guidelines](#)» von TAUS⁸⁴ in den Suchergebnissen bei *Google* angezeigt. Diese Ressource wurde im Jahre 2010 in englischer Sprache verfasst und man kann sie im PDF-Format⁸⁵ kostenlos herunterladen. Sie umfassen Posteditionsverfahren bei der MÜ (*Recommendations, Guidelines for Post-Editing, Guidelines for achieving a “sufficient” quality, Guidelines for achieving comparable with human quality translations*)⁸⁶. Auch weitere freizugängliche Berichte kann man mittels der Navigationskategorie⁸⁷ (siehe Abb. 10 auf der nächsten Seite) „Resources“ => „Publications“ => „Reports“ => „Free Reports“ erhalten, wie z.B. der Bericht «[Putting Machine Translation to Work](#)»⁸⁸ aus dem Jahre 2006. Das Résumé „Best practices in Post-Editing MT“ zum Thema

⁸³ Vgl. [SDL 2013]

⁸⁴ Vgl. [TAUS 2014A]
Die Abkürzung TAUS steht für „Translation Automaton User Society“

⁸⁵ TAUS-Richtlinien zur Post-Editition [TAUS 2014A] im PDF-Format:
<https://evaluation.taus.net/images/stories/guidelines/taus-cn-gl-machine-translation-postediting-guidelines.pdf>,
[28.11.2014]

⁸⁶ Vgl. [TAUS 2014A]

⁸⁷ Vgl. [TAUS 2014B]

⁸⁸ Vgl. [TAUS 2006]

Post-Edition ist für Post-Editoren und für Menschen, die im Geschäftsfeld der MÜ tätig sind, interessant. Freie Ressourcen wie Glossare, Artikel, Multimedia, Videos, usw. sind auf der Website von TAUS verfügbar.



Abb. 10: Ressourcen auf dem TAUS-Portal⁸⁹

- Mit dem englischen Suchbegriff „controlled language MT“ kann man die Arbeit *«[Postediting Experiments with MT for a Controlled Language](#)»*⁹⁰ zum Thema „Kontrollierte Sprache“ in den Suchergebnissen von *Google* finden. Die Arbeit wurde in englischer Sprache verfasst und man kann sie im PDF-Format herunterladen. Ziel dieser Arbeit ist es, zu überprüfen, ob der Einsatz neuer kontrollierter Sprache sowohl die menschliche als auch maschinelle Übersetzung für notfallbezogene Texte erleichtern kann. Eine Evaluierung für den unterschiedlichen Zeitverbrauch der menschlichen Übersetzung und Post-Edition wurde gegeben, um abzuschätzen, ob die kontrollierte Sprache Auswirkungen auf die menschliche und maschinelle Übersetzung hat. Die zeitlichen Unterschiede zwischen den Texten mit und

⁸⁹ Vgl. [TAUS 2014B]

⁹⁰ Vgl. [Temnikova 2009]. Diese Arbeit wurde im Rahmen des MESSAGE-Projekts (unter dem vollständigen Titel: Alert Messages and Protocols), das von der EU (JLS/2007/CIPS/022) finanziert worden ist, produziert. Weitere Informationen über das MESSAGE-Projekt unter: <http://clg.wlv.ac.uk/projects/Message/>, [28.11.2014]

ohne Post-Edition wurden auch gemessen.⁹¹ Dadurch kann man sich viele Kenntnisse über die Auswirkungen von Post-Edition aneignen.

- Mit dem englischen Suchbegriff „controlled language MT“ bei *Google* findet man auf der Website von *Microsoft Research*⁹² die Arbeit [*«Impact of Controlled Language on Translation Quality and Post-editing in a Statistical Machine Translation Environment»*](#)⁹³. Diese Arbeit wurde im Englischen verfasst und man kann diese in PDF herunterladen. Sie untersucht die Zusammenhänge zwischen kontrollierten Sprachen und der Qualität der maschinellen Übersetzung und dem Aufwand von Post-Edition (PE) und überprüft auch die Ergebnisse aus früheren Studien. Sie zeigt, die Qualität der MÜ durch die Verwendung von kontrollierter Sprache verbessert wurde. Außerdem wurde das „Dreier-Verhältnis“ zwischen Verwendung von kontrollierter Sprache, Qualitätsverbesserung der MÜ und Reduzierung des PE-Aufwands festgelegt. Anhand dieser Arbeit lässt sich zeigen, dass man Wissen über die Qualitätsverbesserung der MÜ und die Reduzierung des PE-Aufwands mittels der Verwendung kontrollierter Sprache erhalten kann. Weitere Informationen kann man auf der Website der EAMT⁹⁴ erhalten.
- Durch Klicken auf die Schaltfläche „Archives“ in der Navigationsmenüleiste auf der Website der EAMT wird die Seite für die Beschreibung der maschinellen Übersetzungsarchivierungsdatenbank⁹⁵ gezeigt. Ziel der Archivierungsdatenbank⁹⁶ zur MÜ ist es, umfassende englischsprachige Publikationen zu den Themen „MÜ“, „Computer Übersetzungssysteme“ und „Computerbasierte Übersetzungswerkzeuge“ abzudecken. Forschungen und Bücher aus den vergangenen Jahren wurden auch hinzugefügt, um die Entwicklungsgeschichte der MÜ in den 50er Jahren bis 1980 vollständig zu umfassen. Es gab jedoch einige Publikationen auf der Seite der Archivierungsdatenbank, die aus urheberrechtlichen Gründen nicht zur Verfügung gestellt

⁹¹ Vgl. [Temnikova 2009]

⁹² Microsoft Research, <http://research.microsoft.com/en-us/default.aspx>, [28.11.2014]

⁹³ Vgl. [Aikawa 2007]

Abstract dieser Arbeit unter: <http://research.microsoft.com/apps/pubs/default.aspx?id=69483>, [28.11.2014]

⁹⁴ Die Abkürzung EAMT steht für „European Association for Machine Translation“ European Association for Machine Translation, <http://www.eamt.org/>, [28.11.2014]

⁹⁵ Seite für die Beschreibung der maschinellen Übersetzungsarchivierungsdatenbank von EAMT unter: <http://www.eamt.org/archives.php> [28.11.2014]

⁹⁶ Vgl. [EAMT 2014]

werden können. Der Zugriff auf diese [Archivierungsdatenbank](#)⁹⁷ erfolgt über den Link auf der Website der EAMT. Alle archivierten Dokumentationen wurden in Inhaltskategorien⁹⁸ und nach Dokumentationsarten⁹⁹ in englischer Sprache gesammelt und gepflegt.

3.2.2 MÜ-Systeme

Online MÜ-Systeme

Heutzutage gibt es auf dem Verbrauchermarkt schon eine große Menge von online oder lokalen MÜ-Systemen, mit denen Texte in eine Fremdsprache maschinell übersetzt werden können. Die Übersetzung mit dieser Technologie gehört zu einem weiteren Anwendungsbereich der MÜ. Im Folgenden werden die regelbasierten und statistischen MÜ-Systeme anhand beispielhafter Programme oder Anwendungen vorgestellt. Ähnliche Systeme werden auch erwähnt.

- Der regelbasierte Online-Übersetzer [Reverso Online](#)¹⁰⁰ tauchte in den Ergebnissen von *Google* mit Suchwörtern „Kostenlose maschinelle Übersetzung“ auf. *Reverso Online* in deutscher Version bietet die Möglichkeit, deutsche Wörter, Texte, Sätze und Ausdrücke direkt ins Arabische, Chinesische, Englische, Französische, Hebräische, Italienische, Japanische, Niederländische, Portugiesische, Rumänische, Russische, Spanische zu übersetzen. Mit Hilfe von anderen Reverso Softwarelösungen unterstützt dieser Online-Übersetzer (siehe Abb. 11 auf der nächsten Seite) die folgenden Funktionen: Quelltext im Kontext übersetzen¹⁰¹ (*Reverso Context*¹⁰²), nach Wörtern oder Ausdrücken in Online-Wörterbüchern suchen (*Reverso Dictionary*¹⁰³), Konjugationswörterbuch (*Reverso Conjugator*¹⁰⁴) und Grammatik- und Rechtschrei-

⁹⁷ Vgl. [EAMT 2014]

⁹⁸ Inhaltskategorien: Authors, Organizations and affiliations, Methodologies, techniques, applications, uses, Languages and language pairs, Systems and project names, Keyword index of titles

⁹⁹ Dokumentationsarten: Conferences and workshops – contents lists, Journals, Books on MT, Bibliographical resources, Links to other databases and websites

¹⁰⁰ Vgl. [Reverso 2013A]

¹⁰¹ Übersetzen im Kontext erfolgt durch TM-Technologie von Reverso

¹⁰² Reverso Context, <http://context.reverso.net/%C3%BCbersetzung/> [28.11.2014]

¹⁰³ Reverso Dictionary, <http://dictionary.reverso.net/> [28.11.2014]

¹⁰⁴ Reverso Conjugator, <http://conjugator.reverso.net/> [28.11.2014]

bungsprüfung für englische und französische Texte. Das heißt, *Reverso Online* basiert grundsätzlich auf linguistischen Regeln, um Texte zu übersetzen.



Abb. 11: Homepage von Reverso Online in deutscher Version¹⁰⁵

Für den täglichen Bedarf, einfache Texte zu übersetzen, ist *Reverso Online* ein guter freier Online-Übersetzer. Aber nach der Empfehlung von Reservo¹⁰⁶ ist die Übersetzungsqualität eines Übersetzerautomates nicht ausreichend für einen fachlichen Text, wie z.B. im Bereich Medizin. Allerdings beschränkt er sich beim Eingeben des Quelltextes auf 5000 Zeichen. Daher muss man beim Verarbeiten von Texten unbedingt die Texte teilen. Nach der Eingabe des Quelltextes in der linken Zone auf der Website und der Wahl der gewünschten und verfügbaren Übersetzungsrichtung öffnet sich eine neue Seite, auf der die entsprechende Übersetzung von Reverso vorgeschlagen wird. Es ist auch nach der Registrierung möglich, dass man mittels [Reverso Localize](#)¹⁰⁷ Dateien im Word-, PowerPoint-, Excel-, HTML- und Text-Format, Webseite, Anwendungen auf iPhone- oder Android-System, Web-Anwendungen in PHP¹⁰⁸, .NET¹⁰⁹, JAVA¹¹⁰, Flex¹¹¹ und ASP¹¹², Daten in

¹⁰⁵ Vgl. [Reverso 2013A]

¹⁰⁶ Vgl. [Reverso 2013A]

¹⁰⁷ Vgl. [Reverso 2013B]

¹⁰⁸ Die Abkürzung PHP steht für „Personal Home Page Tools“. PHP ist eine Programmiersprache zur Erstellung dynamischer Webseiten oder Webanwendungen. Siehe auch, <http://de.wikipedia.org/wiki/PHP> [28.11.2014]

XML¹¹³-, CSV¹¹⁴-Dateien und Datenbank in 13 verfügbaren Sprachen¹¹⁵ zu lokalisieren. Aber diese Funktion ist kostenpflichtig. Bei der Erstellung eines neuen Kontos hat man erst jedoch das Zugriffsrecht auf die kostenpflichtige Version mit einer Beschränkung auf 8000 Wörtern. Es gab auch viele andere kostenpflichtige Übersetzungssoftwareprodukte¹¹⁶ von Reverso wie *Reverso Prompt* und *Reverso Translator*, die auf lokalen End-Geräten installierbar sind.

Bei Interesse an anderen online verfügbaren regelbasierten MÜ-Systemen kann man selbst weiter recherchieren, wie z.B. bei:

- ♦ [Lucy LT KWIK Translator](#)¹¹⁷: unterstützt die Übersetzung von Kurztexen und Webseiten in 36 Sprachrichtungen mit der Verwendung von allgemeinem Vokabular, allgemeinem sozialem Vokabular und allgemeinem technischem Vokabular (siehe Abb. 12 auf der nächsten Seite).

¹⁰⁹ .NET ist eine von Microsoft entwickelte Software-Plattform zur Entwicklung und Ausführung von Anwendungsprogrammen. Siehe auch, <http://de.wikipedia.org/wiki/.NET> [28.11.2014]

¹¹⁰ JAVA ist eine von Sun Microsystems entwickelte Programmiersprache. Siehe auch, [http://de.wikipedia.org/wiki/Java_\(Programmiersprache\)](http://de.wikipedia.org/wiki/Java_(Programmiersprache)) [28.11.2014]

¹¹¹ Flex ist ein Entwicklungsframework zum Erstellen von Rich Internet Applications (RIAs) auf der technischen Basis der Adobe Flash-Plattform. Siehe auch, http://de.wikipedia.org/wiki/Apache_Flex [28.11.2014]

¹¹² ASP („Active Server Pages“) ist eine von Microsoft entwickelte Programmiersprache und wurde im Jahr 1996 veröffentlicht. Siehe auch, http://de.wikipedia.org/wiki/Active_Server_Pages [28.11.2014]

¹¹³ XML („Extensible Markup Language“) ist eine maschinenlesbare Sprache zur Darstellung hierarchisch strukturierter Daten in Form von Textdateien. Siehe auch, http://de.wikipedia.org/wiki/Extensible_Markup_Language [28.11.2014]

¹¹⁴ CSV („Comma-separated values“) ist ein Dateiformat, das den Aufbau einer Textdatei zur Speicherung oder zum Austausch einfach strukturierter Daten beschreibt. Siehe auch, [http://de.wikipedia.org/wiki/CSV_\(Dateiformat\)](http://de.wikipedia.org/wiki/CSV_(Dateiformat)) [28.11.2014]

¹¹⁵ Hilfe zu Reverso Localize, <http://localize.reverso.net/HowItWorks.aspx?lang=en> [28.11.2014]

¹¹⁶ Reverso by Softissimo, <http://reverso.softissimo.com/en/home> [28.11.2014]

¹¹⁷ Vgl. [Lucy 2014]

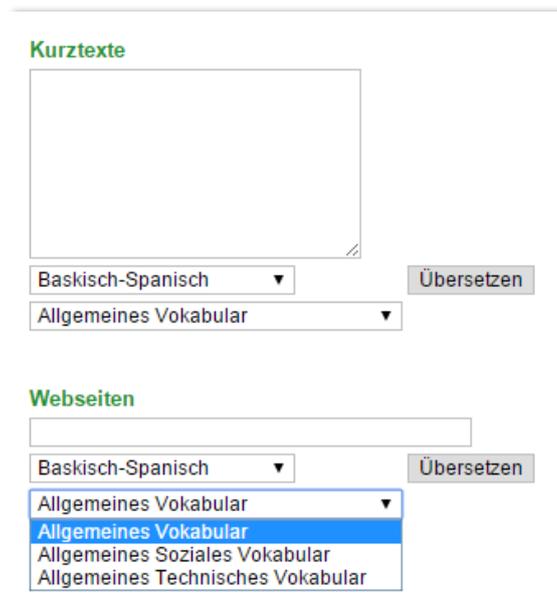


Abb. 12: Lucy LT KWIK Translator¹¹⁸

- Mit den gleichen Suchwörtern zeigt sich das freie statistische MÜ-System [SDL FreeTranslation.com](http://SDLFreeTranslation.com)¹¹⁹ auch in den Suchergebnissen von *Google*. Mit dem statistischen MÜ-System *SDL FreeTranslation.com* kann man Quelltexte in Word, PowerPoint, OpenDocument, PDF und reine Textdateien max. mit einer Dateigröße von 5 MB übersetzen. Nach der Aktivierung der Option „Sofortübersetzung“, die sich unter dem Eingabefeld befindet, werden die Quelltexte automatisch übersetzt und man kann die übersetzten Texte per E-Mail schicken, anhören, drucken und bewerten (rot markiert in Abb. 13).



Abb. 13: GUI von SDL FreeTranslation.com¹²⁰

¹¹⁸ Vgl. [Lucy 2014]

¹¹⁹ Vgl. [SDL 2014B]

¹²⁰ Vgl. [SDL 2014B]

Dieses MÜ-System bietet auch eine kostenpflichtige professionelle Übersetzung (mit Eingriff des Menschen) *SDL Language Cloud Translation Services*¹²¹ für geschäftliche und private Dokumente und eine kostenfreie Übersetzungsanwendung im App Store für Mobilgeräte. Mit *Translation Services* kann man zusätzliche Dateiformate in XLS, ASP, PHP, Flash, InDesign¹²², Quark¹²³ in bestimmten Fachgebieten (Gesetzestexte, technische Handbücher, Marketing-Unterlagen und kontextloser Text) übersetzen. Kontextloser Text ist z.B. eine englische Schaltfläche „Exit“ in einem Softwareprogramm, wenn man das englische Wort „Exit“ ins Deutsche übersetzt, muss man daran denken, welche Bedeutung („beenden“ oder „verlassen“) richtig ist. Außerdem wurde zusätzlich eine Übersetzungs-API¹²⁴ für den Zugriff auf die Übersetzungsplattform *SDL Language Cloud*¹²⁵ zur Verfügung gestellt. Bei *SDL FreeTranslation.com* wurde Englisch als Standard-Ausgangssprache festgelegt, d.h. englische Texte können bei *SDL FreeTranslation.com* in 42 Sprachen¹²⁶ übersetzt werden, aber die Umkehrung gilt nur in den folgenden Sprachpaaren mit eingegrenzten Übersetzungsrichtungen:

- ◆ Arabisch ⇔ Englisch, Französisch, Spanisch
- ◆ Deutsche ⇔ Englisch, Französisch, Spanisch
- ◆ Französisch ⇔ Arabisch, Englisch, Französisch, Spanisch
- ◆ Italienisch ⇔ Englisch, Spanisch
- ◆ Spanisch ⇔ Arabisch, Deutsch, Englisch, Französisch, Italienisch

Daher eignet sich *SDL FreeTranslation.com* besonders für englische Ausgangstexte. Die Gründe dafür liegen daran, dass die Zielgruppen von *SDL FreeTranslation.com* mittelständische und globale Unternehmen sind und Englisch als Geschäftssprache

¹²¹ SDL Language Cloud Translation Services, <https://languagecloud.sdl.com/> [28.11.2014]

¹²² InDesign ist ein von Adobe entwickeltes professionelles Layout- und Satzprogramm für das Desktop Publishing. Siehe auch, http://de.wikipedia.org/wiki/Adobe_InDesign [28.11.2014]

¹²³ Quark ist ein von Quark Inc. entwickeltes Programm für Desktop Publishing. Siehe auch, <http://de.wikipedia.org/wiki/QuarkXPress> [28.11.2014]

¹²⁴ Die Abkürzung API steht für „application programming interface“. API ist ein Programmteil, der von einem Softwaresystem anderen Programmen zur Anbindung an das System zur Verfügung gestellt wird. Siehe auch, <http://de.wikipedia.org/wiki/Programmierschnittstelle> [28.11.2014]

¹²⁵ SDL Language Cloud: die erste vollständige Cloud-Übersetzungsplattform, die sowohl maschinelle als menschliche Übersetzung anbietet. <https://languagecloud.sdl.com/> 28.11.2014

¹²⁶ Für englische Texte verfügbarer Sprachen: Arabisch, Bengalisch, Bulgarisch, Chinesisch (traditionell), Chinesisch (vereinfacht), Dari, Deutsch, Dänisch, Estnisch, Finnisch, Französisch, Griechisch, Hausa, Hebräisch, Hindi, Indonesisch, Italienisch, Japanisch, Koreanisch, Litauisch, Malaiisch, Niederländisch, Norwegisch, Paschtu, Persisch, Polnisch, Portugiesisch, Rumänisch, Russisch, Schwedisch, Serbisch, Slowakisch, Slowenisch, Somali, Spanisch, Thailandisch, Tschechisch, Türkisch, Ukrainisch, Ungarisch, Urdu und Vietnamesisch

weltweit anerkannt ist. Es ist offensichtlich, dass der Bedarf an MÜ für die Lokalisierung, Internationalisierung und Globalisierung von Softwareprodukten und -Dienstleistungen im globalen Geschäft dringend gedeckt werden muss.

Es gibt auch weitere statistische Online-MÜ-Systeme, wie:

- ◆ [Google Translate](#)¹²⁷ unterstützt kostenlose online Übersetzung für Texte, Webseiten und Dateien (durch Hochladen von Dateien) in 80 Sprachen.
- ◆ [Bing Translator](#)¹²⁸ unterstützt kostenlose online Übersetzung für Texte, Webseiten in HTML in 50 Sprachen.
- ◆ [WorldLingo](#)¹²⁹ unterstützt kostenlose online Übersetzung für Texte, Webseiten, E-Mail und Dateien (max. 500 Wörter) in 32 Sprachen mit optionaler Auswahl der Glossare in 20 Fachgebieten. „Professional Translation“ ist kostenpflichtig.
- ◆ [BabelFish](#)¹³⁰ unterstützt kostenlose online Übersetzung für Texte in 15 Sprachen.
- ◆ [Baidu 翻译](#)¹³¹ unterstützt kostenlose online Übersetzung für Texte und Webseiten in 14 Sprachen.

Lokale MÜ-Systeme

Lokale MÜ-Systeme beziehen sich auf die Systeme, die aus dem Internet heruntergeladen und auf einem eigenen PC nachträglich installierbar sind. Die meisten davon sind kostenpflichtig. Im Folgenden wird ein kostenfreies phrasenbasiertes SMÜ-System Moses als Beispiel vorgestellt. Kommerzielle MÜ-Systeme werden auch erwähnt.

- Nach Eingabe der Suchwörter „translation system resources“ wurde das phrasenbasierte SMÜ-System [Moses](#)¹³² in einer vorläufigen Version (August 5, 2013) in den

¹²⁷ Vgl. [Google 2014]

¹²⁸ Vgl. [Bing 2014]

¹²⁹ Vgl. [WorldLingo 2014]

¹³⁰ Vgl. [BabelFish 2014]

¹³¹ Vgl. [Baidu 2014]

¹³² Vgl. [Moses 2013A]

Suchergebnissen von *Google* angezeigt. Wenn man das Suchwort „Moses“ in die Google-Suchmaschine eingegeben hat, wird die Startseite von Moses¹³³ als das erste Ergebnis herausgefunden. Moses ist ein kostenloses statistisches maschinelles Übersetzungssystem, das ermöglicht, Übersetzungsmodelle für jede Sprachrichtung automatisch zu bilden. Alles was man braucht, ist eine Sammlung an übersetzten Texten (Textkorpus). Mit Hilfe des Modells findet das System mit Wortsequenzen schnellstmöglich die Übersetzungen mit der höchsten Wahrscheinlichkeit unter der potentiellen Anzahl von Auswahlmöglichkeiten¹³⁴ – je wahrscheinlicher ein Wort eine Übersetzung eines anderen Wortes ist, desto höher sein Wert. Auf der Webseite von Moses gibt es Schulungsunterlagen, Online-Demos, Übungs- und Referenzmaterialien, mit denen man sich sowohl die Funktionsweise dieses SMÜ-Systems aneignen als auch das Übersetzungsmodell praktisch aufbauen und üben kann.

- [*Personal Translator*](#)¹³⁵ wurde nach den Suchwörtern „maschinelles Übersetzungssystem“ bei *Google* gefunden. Der *Personal Translator* ist ein regelbasiertes MÜ-System und ist eines der führenden Übersetzungsprogramme im Bereich der Übersetzungstechnologie. 7 Sprachpaare¹³⁶ sind im *Personal Translator* verfügbar. Allerdings ist dieses Übersetzungsprogramm zahlungspflichtig.

Es gibt weitere kostenpflichtige MÜ-Systeme, wie:

- ◆ [*Systran 7*](#): Übersetzungssoftware von Systran¹³⁷. Insgesamt sind 4 Systran-Softwareprodukte (*Systran 7 Home Translator*, *Systran 7 Office Translator*, *Systran 7 Business Translator* und *Systran 7 Premium Translator*) vorhanden und orientieren sich an unterschiedliche Zielgruppen. Zum Beispiel, von kleinen und mittelgroßen Unternehmen könnte *Systran 7 Business Translator* begeistert eingesetzt werden.

¹³³ Vgl. [Moses 2013B]
Benutzerhandbuch und Kodierungsrichtlinien für Moses in PDF: <http://www.statmt.org/moses/manual/manual.pdf>
[28.11.2014]

¹³⁴ Vgl. [Moses 2013B]

¹³⁵ Vgl. [Linguec 2014]

¹³⁶ Verfügbare Sprachpaare im Personal Translator: Deutsch ⇔ Englisch, Deutsch ⇔ Französisch, Englisch ⇔ Französisch, Englisch ⇔ Italienisch, Englisch ⇔ Portugiesisch (BR), Englisch ⇔ Spanisch, Englisch ⇔ Chinesisch

¹³⁷ Vgl. [Systran 2014]

- ♦ [translate version 12](#): Übersetzungssysteme von Lingenio¹³⁸. Das Produkt *translate version 12* enthält 3 Versionen (*translatepro*, *translateplus*, *translatequick*). *translatequick* ist beispielsweise für die Übersetzung beliebiger Texte und Webseiten mit garantierter Qualität und Schnelligkeit preiswert.
- ♦ [PROMT](#): Desktop-Übersetzungssoftware von PROMT. PROMT bietet verschiedene Möglichkeiten, Übersetzungssoftware auf unterschiedlichen Betriebssystemen und Mobilgeräten einzusetzen. Testversionen für jedes Produkt von PROMT sind 15 Tage nach erfolgreichem Herunterladen auf dem eigenen PC einsetzbar.¹³⁹

3.4 Vor- und Nachteile bei der Nutzung online verfügbarer Ressourcen

Unabhängig von der MÜ haben die obigen gesammelten Ressourcen für ein Fernstudium gezeigt, dass die Nutzung online verfügbarer Ressourcen eine Vielzahl von Vorteilen aufweist: in kostenmäßigem Aufwand, zeitlicher Einsparung und Wiederverwendbarkeit. Die Nutzung online verfügbarer Ressourcen ist aus wirtschaftlicher Sicht so bedeutend, weil es zahlreiche Ressourcen bereits im Internet gibt und der Zugriff auf diese mit Hilfe von Suchmaschinen meistens kostenlos zugänglich ist. Da man die Möglichkeit hat, wenn man Kenntnisse auf einem Gebiet haben und erlernen möchte, nach freien Ressourcen zu einem beliebigen Thema online zu suchen. Außerdem wird Zeit durch die Nutzung online verfügbarer Ressourcen gespart. Das ist besonders günstig für Menschen, die keine Zeit für ein direktes Studium haben. Alles was man braucht, ist ein Internetzugang, mit dem man verschiedene Online-Ressourcen jederzeit abrufen und diese wiederverwenden kann. Weiterhin haben die Online-Ressourcen keine örtliche Beschränkung, d.h., wenn man die gewünschten Ressourcen auf einem eigenen Datenträger gespeichert hat, kann man mit einem tragbaren Gerät, das das Format der Ressourcen unterstützt, diese an allen Orten öffnen und lesen.

Ein Nachteil der Nutzung online verfügbarer Ressourcen folgt quasi aus ihrem Vorteil, da es schwerwiegende Probleme bei der Stabilität bzw. Verfügbarkeit gibt. Zum Beispiel

¹³⁸ Vgl. [Lingenio 2012]

¹³⁹ Vgl. [PROMT 2014]

waren die Übungsmaterialien von Lopez¹⁴⁰, die zwei Übungen zur SMÜ enthalten, schon am 18.11.2014 nicht mehr verfügbar. Es ist festzustellen, dass die online verfügbaren Ressourcen nicht stabil verfügbar sind. Deshalb muss man die Online-Ressourcen regelmäßig pflegen. Auch wenn man keine Internet-Schnittstelle und kein Mobilgerät besitzt, können die Ressourcen nicht abgerufen werden. Ein weiteres Problem bei der Langzeitarchivierung der Ressourcen ist, dass Datenträger, auf denen die Ressourcen gespeichert wurden, zu einem späteren Zeitpunkt nicht mehr von einem neueren Gerät lesbar sind oder nur fehlerhaft funktionieren, da das Gerät (z.B. Videorekorder) oder der Datenträger (z.B. Diskette) sich nicht mehr auf dem Markt befindet oder defekt ist. In diesem Fall wäre die Nutzung der gespeicherten Online-Ressourcen nicht möglich. Auch bei der Übertragbarkeit der Datei-Formate können Probleme auftreten. Für Anfänger oder interessierte Laien ist die Nachvollziehbarkeit der Ressourcen zu einem bestimmten fachlichen Thema eine knifflige Frage, weil spezifische Fachkenntnisse für manche online gefundenen Ressourcen notwendig sind. Wenn man solche Fachkenntnisse bereits vorher hat, ist es leicht, sich Kenntnisse im Selbststudium anzueignen. Es ist bereits in früherer Studie¹⁴¹ erwähnt worden, dass die Nutzung online verfügbarer Ressourcen Risiken birgt. Die Informationsqualität aus zweifelhaften Internet-Quellen kann nicht garantiert werden, da jeder Internetnutzer seine Meinung frei im Internet eingeben kann. Während der Recherche der Online-Ressourcen ist es auch gefährlich, dass man eine Viren verbreitende Seite aufruft. Die personenbezogenen Daten wie Bankkonto können dadurch verloren und von den anderen illegal verwendet werden. Daher muss jeder darauf achten, dass Ressourcen und Informationsquellen im Internet sorgfältig recherchiert und behutsam ausgesucht werden sollten. Die Autorin hat im Rahmen der vorliegenden Arbeit die Qualität der Quellen und die Vertrauenswürdigkeit der Seiten besonders sorgfältig überprüft.

¹⁴⁰ [Lopez 2012] A. Lopez. Statistical Machine Translation. Lehrvorträge und Übungsmaterialien zur SMÜ. In: NASSLLI, North American Summer School for Logic, Language and Information, 2012. Online-Quelle: <<http://www.cs.jhu.edu/~alopez/nasslli2012.html>>, 17.11.2014

¹⁴¹ [Näf 2000] M. Näf; P. Streule; W. Hartmann. Risiko Internet? Sicherheitsaspekte bei der Internet-Benutzung. Hg. Orell Füssli Verlag, Zürich, 2000. Online-Quelle: http://www.swisseduc.ch/informatik/internet/internet_sicherheit/docs/sicherheit.pdf [28.11.2014]

4. Fazit

Im letzten Kapitel wird ein Ausblick auf weitere Aufgabenfelder und Kombinationsmöglichkeiten der MÜ mit anderen Technologien dargestellt. Dazu werden die Ergebnisse der Arbeit über online verfügbare Ressourcen, d.h., die Bewertung der Ressourcen, zunächst aufgeführt und erläutert. Die gesammelten Ressourcen werden für einen bestimmten Zweck oder zu einem gewissen Themenbereich der MÜ beurteilt und für einzelne Zielgruppen¹⁴² empfohlen:

Zuerst werden Ressourcen für die Geschichte der MÜ beurteilt. Für Anfänger oder interessierte Laien sind die obengenannten Werke «An introduction to machine translation» und «MACHINE TRANSLATION: An Introductory Guide» zur Einführung in der MÜ empfehlenswert, da beide Bücher ausführliche Informationen zur MÜ umfassen. Außerdem werden keine besonderen Anforderungen an Vorkenntnisse in diesen Werken gestellt. Deshalb sind sie leicht zu nutzen. Der Artikel «Die Übersetzungstechnologien und ihre Pioniere» von der D.O.G. GmbH ist für die Studenten, die bereits Vorkenntnisse haben, sehr nützlich, die Entwicklungsgeschichte der MÜ im Gedächtnis zu behalten. Ein Fernstudium zum Thema „Machine Translation“ hilft sowohl Studenten als auch Lokalisierungsdienstleister oder freiberuflichen Übersetzern dabei, die Modelle, Forschungen und Evaluierung der MÜ auf einen Blick zu betrachten. Studenten können dadurch tiefere Kenntnisse über die MÜ bekommen. Für Lokalisierungsdienstleister oder für einen freiberuflichen Übersetzer sind die Informationen über die MÜ-Modelle und -Evaluierung in zweierlei Hinsicht bedeutend. Zum einen, weil Übersetzer oder Lokalisierungsdienstleister einen Text mit Hilfe eines MÜ-Modells übersetzen können. Der maschinell übersetzte Text kann als eine Rohübersetzung verwendet werden. Dadurch werden die zeitlichen und kostenmäßigen Übersetzungsaufwände reduziert. Wenn Übersetzer Kenntnisse über die MÜ-Modelle besitzen, wird ihnen erleichtert, ein effizientes Modell auszuwählen und aufzubauen. Zum anderen, weil Lokalisierungsdienstleister ein korrektes MÜ-System anhand der Evaluierung im Lokalisierungsprozess festlegen können. Dadurch werden die Kosten für ein Lokalisierungsprojekt reduziert.

Anschließend werden Ressourcen für die Ansätze der MÜ bewertet. Der Beitrag «Maschinelle Übersetzung – ein Überblick» und das MOOC zum Thema „Principles and Practices of Computer-Aided Translation“ werden für Anfänger oder interessierte Laien

¹⁴² Die genaue Definition der Zielgruppen befindet sich im zweiten Teil des 1. Kapitels ([1.2](#)).

empfohlen, die die Ansätze der MÜ-Systeme kennenlernen wollen. Der Grund liegt daran, dass der Beitrag «Maschinelle Übersetzung – ein Überblick» ein leicht zugängliches Werk ist, indem Anfänger oder interessierte Laien sich das grundlegende Wissen über den Aufbau und die Funktionsweise der MÜ-Systeme aneignen können. Mit Hilfe des MOOCs zum Thema „Principles and Practices of Computer-Aided Translation“ ist es möglich, dass Anfänger oder interessierte Laien das Grundkonzept computerunterstützter Übersetzungstechnologie und die Verwendung computerunterstützter Übersetzungswerkzeuge in der Praxis erlernen und üben. Danach werden die Zusammenhänge und Unterschiede zwischen MÜ- und CAT-Systemen auch erkannt. Das *SDL Download-Center* eignet sich für alle drei Zielgruppen. Informationen über MÜ können mittels 3 Filter und den zusätzlichen Filteroptionen zum Selbststudium recherchiert und in Sekundenschnelle gefunden werden. Allerdings sind die Suchergebnisse ohne konkreten Filter so zahlreich, dass man jedes Material in den Suchergebnissen noch genau schauen muss, um die gewünschte Ressource herauszufinden. Dies ist für Anfänger oder interessierte Laien ein bisschen schwer.

Im Hinblick auf die Einsatzszenarien der MÜ werden die ausgewählten Ressourcen im folgenden Abschnitt zusammenfassend eingeschätzt. Freie Ressourcen auf der Website von TAUS zum Thema Post-Editing, die Arbeiten «Postediting Experiments with MT for a Controlled Language» und «Impact of Controlled Language on Translation Quality and Post-editing in a Statistical Machine Translation Environment» leisten einen wertvollen Beitrag zu den praktischen Anwendungsbereichen der MÜ für alle drei Zielgruppen. Man kann sich dadurch viele Kenntnisse über Posteditingsverfahren bei der MÜ, die Auswirkung der Post-Editing und die Zusammenhänge zwischen kontrollierten Sprachen und der Qualität der maschinellen Übersetzung und den Aufwand der Post-Editing aneignen. Die Archivierungsdatenbank der EAMT ist für alle drei Zielgruppen sehr empfehlenswert. Da ausführliche englischsprachige Informationen über „MÜ“, „Computer-Übersetzungssysteme“ und „Computerbasierte Übersetzungswerkzeuge“ in dieser Datenbank gesammelt werden. Und sie werden aktualisiert und gepflegt.

Die Bewertung der obigen gesammelten Ressourcen für die praktische Nutzung der MÜ-Systeme wird im Folgenden gegeben. Der regelbasierte Online-Übersetzer „Reverso Online“ und das statistische MÜ-System „SDL FreeTranslation.com“ sind zwei Vertreter für die praktische Nutzung der MÜ-Systeme im Internet. Dadurch können sich Anfänger oder interessierte Laien mit MÜ-Systemen vertraut machen und zwei unterschiedliche MÜ-Ansätze erleben. Das ausgewählte lokale MÜ-System Moses ist das einzige kos-

tenfreie SMÜ-System in der Gegenwart. Aber zum Selbststudium ist es für alle drei Zielgruppen sehr schwer, vor allem, weil man zunächst ein eigenes MÜ-Modell mit diesem System aufbauen muss. Und es gibt auch viele dringend zu lösende Probleme in diesem MÜ-System. Deshalb wird das lokale SMÜ-System Moses nur für computerlinguistische Lehrveranstaltungen empfohlen.

In der Arbeit werden dann weitere Anwendungsfelder erläutert und die Kombinationsmöglichkeiten in naher Zukunft prognostiziert. Für Maschinelle Übersetzung eröffnen sich zahlreiche Einsatzfelder. Online und lokale MÜ-Systeme werden häufig im Übersetzungs- und im Lokalisierungsprozess eingesetzt, um die Lokalisierungsaufwände zeitlich und kostenmäßig zu reduzieren. Mit dem Einsatz kontrollierter Sprache ist MÜ als Übersetzung von technischer Dokumentation sehr sinnvoll und nützlich. Wenn man die maschinell übersetzten Übersetzungen durch menschliches Editieren (Post-Edition) nachbearbeitet, wird eine hohe Übersetzungsqualität garantiert. Weitere Anwendungsbereiche für MÜ sind auch denkbar, wie z.B. die Verbindung mit Translation Memory (TM), maschinelle Übersetzung von E-Mails oder Online-Chat und cloud-basierte maschinelle und menschliche Übersetzung. Übersetzungsspeicher (TM) sind bereits in vielen Lokalisierungswerkzeugen eingebettet und bieten Nutzern die Möglichkeit, Quelltext nach Konkordanzen oder Fuzzy-Matches zu suchen und als Übersetzung anzubieten. Diese Übersetzung kann vom Nutzer übernommen werden. Falls für den Text noch keine passenden Segmente im TM gefunden werden, wird gegebenenfalls ein Vorschlag, der vom internen oder externen maschinellen Übersetzer geliefert wird, für diesen Text gegeben. E-Mails oder in Echtzeit stattfindende Kommunikation über das Internet ist auch ein weiteres Anwendungsgebiet der MÜ. Zum Beispiel kann es passieren, dass ein Nutzer eine in einer Fremdsprache geschriebene E-Mail erhielt, die er gar nicht lesen und verstehen kann. Mit der MÜ kann er dieses Problem lösen und eine Rohübersetzung dafür bekommen. Umgekehrt ist es auch möglich, eine Antwort zurückzuschicken. Ein hypothetisches Einsatzfeld für MÜ ist die cloud-basierte maschinelle und menschliche Übersetzung. Die cloud-basierte maschinelle und menschliche Übersetzung würde im idealen Fall so funktionieren: jeder Text würde mit einem bestimmten cloud-basierten Online-MÜ-System nach der Auswahl von Fachgebieten und Ausgangs- und Zielsprachen übersetzt und nachfolgend in der Cloud gespeichert. Diese Übersetzungsdaten würden von den Post-Editoren, die Muttersprachler der Zielsprache sind, nachbearbeitet und danach in der Cloud gespeichert. Zwischenzeitlich würde der Text auch Fachgebieten und Sprachpaaren zugeordnet. Jeder registrierte Nutzer besitzt das Zu-

griffsrecht auf das System. Diese Prognose wäre mit der Technologie des Cloud Computing in naher Zukunft möglich.

Zusammenfassend ist zu sagen, dass schnelle Übersetzungsgeschwindigkeit, zeitlich und kostenmäßig reduzierte Übersetzungsaufwände und eine garantierte Übersetzungsqualität durch die Verwendung der MÜ in Zusammenarbeit mit anderen Technologien in- zwischen gewährleistet werden könnten. Wer sich für diese Technologien interessiert, kann anhand der Materialien dieser Arbeit Kenntnisse in diesem Bereich erwerben.

Quellenverzeichnis

[Aikawa 2007] T. Aikawa; L. Schwartz; R. King; M. Corston-Oliver; C. Lozano. Impact of Controlled Language on Translation Quality and Post-editing in a Statistical Machine Translation Environment. Oktober, 2007. Online-Quelle:

<<http://research.microsoft.com/pubs/69483/MTS-2007-Aikawa.pdf>>, 28.11.2014

[Arnold 1996] D.J. Arnold; L. Balkan; S. Meijer; R. Lee. Humphreys; L. Sadler. Machine Translation: an Introductory Guide, ISBN: 1855542-17x. Original Hg. NCC Blackwell Ltd, London, 1994. Kopie im PDF-Format, 1996. Online-Quelle:

<<http://promethee.philo.ulg.ac.be/engdep1/download/bacIII/Arnold%20et%20al%20Machine%20Translation.pdf>>, 28.11.2014

[BabelFish 2014] BabelFish. Free Online Translator. Online-Quelle:

<<http://www.babelfish.com/>>, 28.11.2014

[Baidu 2014] Baidu fanyi. Freie chinesische online Übersetzung. Online-Quelle:

<<http://fanyi.baidu.com/#auto/zh/>>, 28.11.2014

[Bing 2014] Bing Translator. Online-Quelle: <<http://www.bing.com/translator/>>, 28.11.2014

[Duden 2013] Duden Online. Online-Wörterbuch. 2013. Online-Quelle:

<<http://www.duden.de/woerterbuch>>, 28.11.2014

[D.O.G. 2006] Die Übersetzungstechnologien und ihre Pioniere. In: D.O.G. GmbH Dokumentation ohne Grenzen. D.O.G. news, 01/2006. Online-Quelle:

<<http://www.dog-gmbh.de/uebersetzung-lokalisierung/uebersetzung-lokalisierung-fachartikel/die-uebersetzungstechnologien-und-ihre-pioniere.html>>, 28.11.2014

[EAMT 2014] EAMT: European Association for Machine Translation. Machine Translation Archive. Electronic repository and bibliography of articles, books and papers on topics in machine translation, computer translation systems, and computer-based translation tools. Letztes Update: 7. November 2014. Online-Quelle:

<<http://www.mt-archive.info/>>, 28.11.2014

[Google 2014] Google Translate. Online-Quelle: <<https://translate.google.com/>>, 28.11.2014

- [Hutchins 1992A] W. John Hutchins; Harold L. Somers. An introduction to machine translation. METAL. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-15.pdf>>, 28.11.2014
- [Hutchins 1992B] W. John Hutchins; Harold L. Somers. An introduction to machine translation. M é é. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-12.pdf>>, 28.11.2014
- [Hutchins 1992C] W. John Hutchins; Harold L. Somers: An introduction to machine translation. Hg. Academic Press, London, 1992. Inhalts ü bersicht unter: <<http://www.hutchinsweb.me.uk/IntroMT-TOC.htm>>, 28.11.2014
- [Hutchins 1992D] W. John Hutchins; Harold L. Somers: An introduction to machine translation. General introduction and brief history. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-1.pdf>>, 28.11.2014
- [Lingenio 2012] Ü bersetzungssysteme von Lingenio: translate version 12. Online-Quelle: <<http://static.lingenio.de/Deutsch/Produkte/Uebersetzungssysteme.htm>>, 28.11.2014
- [Linguec 2014] Personal Translator. Das preisgekr ö nte Ü bersetzungsprogramm von Linguec. Online-Quelle: <<http://www.linguec.de/products/tr/pt>>, 28.11.2014
- [Lopez 2010] A. Lopez. Machine Translation: Models, Search, and Evaluation. In: videolectures.net. exchange ideas & share knowledge. Online-Quelle: <http://videolectures.net/clsps09_lopez_mt/>, 28.11.2014
- [Lopez 2014A] A. Lopez; M. Post; Y. Cao. Machine Translation. Introduction to Machine Translation. 29.01.2014. Online-Quelle: <<http://www.slideshare.net/alopezfoo/jhu-mt-lecture2014012>>, 28.11.2014
- [Luckhardt 2010] H. Luckhardt. Virtuelles Handbuch Informationswissenschaft. Einführung in die Informationswissenschaft. Computergest ü tzte und maschinelle Ü bersetzung für Fachinformation. In: Vituelles Handbuch, Universität des Saarlandes, 02.06.2010. Online-Quelle: <http://is.uni-sb.de/studium/handbuch/exkurs_mt.html>, 28.11.2014
- [Lucy 2014] Lucy LT KWIK Translator. Kostenlose Ü bersetzung. Online-Quelle: <<http://www.lucysoftware.com/deutsch/maschinelle-uebersetzung/kwik-translator/kostenlose-uebersetzung-lucy-lt-kwik-translator.html>>, 28.11.2014

[Moses 2013A] MBOT Moses Translation System in vorläufiger Version von 5. August 2013. In: IMS, Institut für Maschinelle Sprachverarbeitung an der Universität Stuttgart. Online-Quelle:

<<http://www.ims.uni-stuttgart.de/forschung/ressourcen/werkzeuge/mbotmoses.html>>, 28.11.2014

[Moses 2013B] Das phrasenbasierte SMÜ-System Moses. Online-Quelle:

<<http://www.statmt.org/moses/>>, 28.11.2014

[PROMT 2014] Übersetzungssysteme von PROMT: PROMT 10. Online-Quelle:

<http://www.promt.de/translation_software/desktop/>, 28.11.2014

[Ramlow 2008] M. Ramlow; M. Krenz; Hrg. U. Seewald-Heeg. Maschinelle Übersetzung und XML im Übersetzungsprozess. Prozesse der Translation und Lokalisierung im Wandel. Zwei Beiträge. Hg. Frank & Timme GmbH, Verlag für wissenschaftliche Literatur, Berlin, 2008

[Reinke 2005] U. Reinke. Einführung in die Maschinelle Übersetzung (MÜ). Vorlesung „Einführung in die Sprachdatenverarbeitung“. Universität des Saarlandes. Saarland, 2005.

Online-Quelle:

<http://www.f03.fh-koeln.de/imperia/md/content/personen/reinke_uwe/sdv_cl_6gross.pdf>, 28.11.2014

[Reverso 2013A] Reverso Online. Kostenloser Online-Übersetzer in deutscher Version.

Online-Quelle: <http://www.reverso.net/text_translation.aspx?lang=DE>, 28.11.2014

[Reverso 2013B] Reverso Localize. Website-Übersetzung und –Indizierung in englischer

Version. Online-Quelle: <<http://localize.reverso.net/Default.aspx?lang=en>>, 28.11.2014

[Schmitz 2002] K. Schmitz. Softwarelokalisierung – ein neues Arbeitsfeld für Übersetzerinnen und Übersetzer. In: Linguistica Antverpiensia, New Series 1/2002: 376.

Online-Quelle: <<https://lans-tts.ua.ac.be/index.php/LANS-TTS/article/viewFile/26/25>>, 28.11.2014

[SDL 2013] SDL Grundlagen. Regelbasierte und statistische Ansätze der maschinellen Übersetzung. Lösungsübersicht in deutscher Sprache. In: SDL Download-Center. 2013.

Online-Quelle:

<<http://www.sdl.com/de/download/approaches-machine-translation/25416/>>, 28.11.2014

- [SDL 2014A] SDL Download-Center. Ressourcen zur MÜ. 2014. Online-Quelle: <<http://www.sdl.com/de/resources/?Download+Type=&Area+of+interest=Maschinelle+%C3%9Cbersetzung&Product=&Service=>>, 28.11.2014
- [SDL 2014B] SDL FreeTranslation.com. Online-Übersetzer für Dokumente, Websites und Texten. Online-Quelle: <<http://www.freetranslation.com/>>, 28.11.2014
- [Seewald-Heeg 2009] Hrg. U. Seewald-Heeg; D. Stein. Maschinelle Übersetzung - Von der Theorie zur Praxis. In: JLCL 24/3, 2009. Online-Quelle: <http://media.dwds.de/jlcl/2009_Heft3/ldv-forum-24-3.pdf>, 28.11.2014
- [Seewald-Heeg 2011] U. Seewald-Heeg. Historischer Überblick über die Entwicklung der MÜ. Maschinelle Übersetzung. Vorlesung „Lokalisierung Grundlagen“. Hochschule Anhalt. 2011
- [Stein 2009] D. Stein. Hrg. Seewald-Heeg. Maschinelle Übersetzung - Von der Theorie zur Praxis. Maschinelle Übersetzung – ein Überblick. In: JLCL 24/3, 2009:5-18. Online-Quelle: <http://media.dwds.de/jlcl/2009_Heft3/daniel_stein.pdf>, 28.11.2014
- [Systran 2014] Desktop-Übersetzungssoftware von Systran. Online-Quelle: <<http://store.systran.de/>>, 28.11.2014
- [TAUS 2006] Putting Machine Translation to Work. Report on the TAUS Executive Forum on March 23-24, 2006. Washington DC, April 2006. Online-Quelle: <<https://www.taus.net/downloads/finish/56-public-reports/16-taus-putting-machine-translation-to-work-april-2006>>, 28.11.2014
- [TAUS 2014A] Machine Translation Post-Editing Guidelines. In: TAUS, Translation Automaton User Society. November, 2010. Online-Quelle: <<https://evaluation.taus.net/resources/guidelines/post-editing/machine-translation-post-editing-guidelines>>, 28.11.2014
- [TAUS 2014B] Free TAUS Reports. In: TAUS, Translation Automaton User Society. 2014. Online-Quelle: <<https://www.taus.net/free-reports/free-taus-reports>>, 28.11.2014

[Temnikova 2009] I. Temnikova; C. Orasan. Post-editing Experiments with MT for a Controlled Language. Research Institute in Information and Language Processing, University of Wolverhampton, UK. In: Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Language (ISMTCL), Besancon, France, July 2009. Online-Quelle:

<<http://clg.wlv.ac.uk/papers/Temnikova-Orasan-ISMTCL.pdf>>, 28.11.2014

<http://www.academia.edu/2718763/Postediting_Experiments_with_MT_for_a_Controlled_Language>, 28.11.2014

[WorldLingo 2014] WorldLingo. Free Translation Online. Online-Quelle:

<http://www.worldlingo.com/en/products_services/worldlingo_translator.html>, 28.11.2014

[Yu 2014] Jingsong. S. Yu. Principles and Practice of Computer Aided Translation. Grundkonzept computerunterstützter Übersetzungstechnologie und Verwendung computerunterstützter Übersetzungswerkzeuge. MOOC in chinesischer Sprache mit englischen Untertiteln an der Peking Universität. In: MOOC 学院, das größte chinesische Studienforum. Studienbeginn am 22.09.2014. Online-Quelle:

<<http://mooc.guokr.com/course/875/Principles-and-Practice-of-Computer-Aided-Translation/>>, 28.11.2014

Abbildungsverzeichnis

Abb. 1: Ins Deutsche maschinell übersetzte Homepage von Microsoft Translator..	2
Abb. 2: Entwicklungsgeschichte der MÜ von 1950 bis 1990.....	7
Abb. 3: Historischer Überblick über die Entwicklung der MÜ	9
Abb. 4: Regelbasierte MÜ vs. statistikbasierte MÜ im Überblick.....	10
Abb. 5: Das Vauquoi'sche Dreieck der Transferebenen.....	11
Abb. 6: Lehrfilm zum Abstract auf der Website von „MOOC 学院“	20
Abb. 7: Zertifikat bei „MOOC 学院“	21
Abb. 8: Filteroptionen des SDL Download-Centers	22
Abb. 9: Pflichtfelder zum Erhalten von SDL-Ressourcen	23
Abb. 10: Ressourcen auf dem TAUS-Portal	24
Abb. 11: Homepage von Reverso Online in deutscher Version	27
Abb. 12: Lucy LT KWIK Translator	29
Abb. 13: GUI von SDL FreeTranslation.com	29

Glossar

Cloud bezeichnet ein Netzwerk mehrerer verteilter Rechner.¹⁴³

Cloud Computing bedeutet die „Nutzung von IT-Infrastrukturen und –Dienstleistungen, die nicht vor Ort auf lokalen Rechnern vorgehalten, sondern als Dienst gemietet werden und auf die über ein Netzwerk (z.B. das Internet) zugegriffen wird“. ¹⁴⁴ IT-Infrastruktur bezeichnet alle notwendigen materiellen und immateriellen Güter, die den Betrieb von Anwendungssoftware ermöglichen.¹⁴⁵

Globalisierung umfasst „alle Aktivitäten eines Unternehmens im Hinblick auf die Vermarktung eines Produktes oder einer Dienstleistung außerhalb des nationalen Marktes“. Ziel der Globalisierung ist „ein erfolgreiches Marketing in vielen unterschiedlichen regionalen Märkten unter besonderer Berücksichtigung“ auf technische, wirtschaftliche und gesetzliche Aspekte.¹⁴⁶

GUI ist die englische Abkürzung von „graphical user interface“. Auf Deutsch heißt das Wort „grafische Benutzeroberfläche“ und wird als Benutzerschnittstelle eines Computers über grafische Symbole bezeichnet.¹⁴⁷

Internationalisierung bezeichnet die technische Entwicklung von Produkten und Dienstleistungen im Hinblick auf eine leichte Anpassung an andere Märkte und bedeutet für die Lokalisierung, dass Programme so entwickelt werden, dass für neue Länderversionen keine Neuprogrammierung der Software erforderlich ist.¹⁴⁸

siehe: **Lokalisierung**

Kontrollierte Sprache bezeichnet eine [natürliche Sprache](#) wie Deutsch, die nach bestimmten Regeln hinsichtlich Fachwörter, Sprachstil, Grammatik und Wortschatz eingeschränkt wird. Ziel ist es, technische Dokumentationen und Handbücher verständlicher zu machen, potentielle Mehrdeutigkeit zu minimieren und Übersetzungen zu erleichtern.

¹⁴³ Vgl. [Duden 2013]

¹⁴⁴ Vgl. [Duden 2013]

¹⁴⁵ Siehe auch, <http://de.wikipedia.org/wiki/IT-Infrastruktur> [28.11.2014]

¹⁴⁶ [Schmitz 2002:376]

¹⁴⁷ Siehe auch, http://de.wikipedia.org/wiki/Grafische_Benutzeroberfl%C3%A4che [28.11.2014]

¹⁴⁸ Vgl. [Schmitz 2002:376]

Lokalisierung, auch Softwarelokalisierung genannt, bezeichnet die Anpassung oder Übersetzung eines Software-Produkts oder einer Internet-Präsenz an einen lokalen Markt mit sprachlichen und kulturellen Besonderheiten und mit Hilfe spezieller technologischer Methoden und Werkzeuge.¹⁴⁹

Maschinelle Übersetzung bedeutet im engeren Sinne, dass einzelne Wörter, Sätze oder ganze Texte von einer Sprache A in eine Sprache B mit Hilfe von Computerprogrammen automatisch übersetzt werden.¹⁵⁰

Post-Edition ist das Verfahren zur Qualitätsverbesserung eines maschinell übersetzten Textes mittels menschlicher Nachbearbeitung.

Regelbasierte maschinelle Übersetzung bedeutet, dass der Ausgangstext mittels eines Satzes von linguistischen Regeln und einem Wörterbuch in Strukturen der Zielsprache analysiert und übersetzt wird, aus denen der Zieltext generiert wird.¹⁵¹

Statistikbasierte maschinelle Übersetzung, auch als statistische maschinelle Übersetzung bezeichnet, bedeutet, dass die Übersetzung mittels statistischer Wahrscheinlichkeitsmodelle aus einem zweisprachigen Korpus erzeugt wird.¹⁵²

¹⁴⁹ Vgl. [Seewald-Heeg 2011:141], [Schmitz 2002:376]

¹⁵⁰ Vgl. [Luckhardt 2010]

¹⁵¹ Vgl. [Seewald-Heeg 2009:20]

¹⁵² Vgl. [Seewald-Heeg 2009:20]

Anhang***Bibliographie online verfügbarer Ressourcen für MÜ***

[ALPAC Report 1966]

Language and Machines. Computers in Translation and Linguistics. A Report by the Automatic Language Processing Advisory Committee. Division of Behavioral Sciences, National Academy of Sciences, National Research Council, Publication 1416 Washington, D. C. 1966. Online Quelle:

<http://books.nap.edu/html/alpac_lm/ARC000005.pdf>, 28.11.2014

[Arnold 1996]

D.J. Arnold; L. Balkan; S. Meijer; R. Lee. Humphreys; L. Sadler. Machine Translation: an Introductory Guide, ISBN: 1855542-17x. Original Hg. NCC Blackwell Ltd, London, 1994. Kopie im PDF-Format, 1996. Online-Quelle:

<<http://promethee.philo.ulg.ac.be/engdep1/download/bacIII/Arnold%20et%20al%20Machine%20Translation.pdf>>, 28.11.2014

[BabelFsh 2014]

BabelFish. Free Online Translator. Online-Quelle: <<http://www.babelfish.com/>>, 28.11.2014

[Baidu 2014]

Baidu fanyi. Freie chinesische online Übersetzung. Online-Quelle:

<<http://fanyi.baidu.com/#auto/zh/>>, 28.11.2014

[Bing 2014]

Bing Translator. Online-Quelle: <<http://www.bing.com/translator/>>, 28.11.2014

[Collins 2013]

M. Collins. Natural language processing (NLP). Columbia University. In: Coursera, online course. Studienfrist: 24.02.2013 - 05.05.2013. Online-Quelle:

<<https://www.coursera.org/course/nlangp>>, 28.11.2014

[D.O.G. 2006]

Die Übersetzungstechnologien und ihre Pioniere. In: D.O.G. GmbH Dokumentation ohne Grenzen. D.O.G. news, 01/2006. Online-Quelle:

<<http://www.dog-gmbh.de/uebersetzung-lokalisierung/uebersetzung-lokalisierung-fachartikel/die-uebersetzungstechnologien-und-ihre-pioniere.html>>, 28.11.2014

[EAGLES 1996]

EAGLES MT Evaluation Working Group. EAGLES Evaluation of Natural Language Processing System, Final Report. EAGLES Document EAG-EWG-PR.2, ISBN 87-90708-00-8. Center for Sprogteknologi, Copenhagen, 1996. In: ISSCO, Institute for Semantic and Cognitive Studies. Inhaltsübersicht unter: <<http://www.issco.unige.ch/en/research/projects/ewg96/ewg96.html>>, 28.11.2014
Online-Quelle: <<http://www.issco.unige.ch/downloads/ewg96.pdf>>, 28.11.2014

[EAMT 2014]

EAMT: European Association for Machine Translation. Machine Translation Archive. Electronic repository and bibliography of articles, books and papers on topics in machine translation, computer translation systems, and computer-based translation tools. Letztes Update: 7. November 2014. Online-Quelle: <<http://www.mt-archive.info/>>, 28.11.2014

[Google 2014]

Google Translate. Online-Quelle: <<https://translate.google.com/>>, 28.11.2014

[Hutchins 1992A]

W. John Hutchins; Harold L. Somers. An introduction to machine translation. METAL. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-15.pdf>>, 28.11.2014

[Hutchins 1992B]

W. John Hutchins; Harold L. Somers. An introduction to machine translation. M é t ó. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-12.pdf>>, 28.11.2014

[Hutchins 1992C]

W. John Hutchins; Harold L. Somers: An introduction to machine translation. Hg. Academic Press, London, 1992. Inhaltsübersicht unter: <<http://www.hutchinsweb.me.uk/IntroMT-TOC.htm>>, 28.11.2014

[Hutchins 1992D]

W. John Hutchins; Harold L. Somers: An introduction to machine translation. General introduction and brief history. Hg. Academic Press, London, 1992. Online-Quelle: <<http://www.hutchinsweb.me.uk/IntroMT-1.pdf>>, 28.11.2014

[Koehn 2014A]

P. Koehn. Ressourcen-Sammlung zur maschinellen Übersetzung. Software und Ressourcen, Publikationen, Videos und Materialien. Online-Quelle: <<http://homepages.inf.ed.ac.uk/pkoehn/>>, 28.11.2014

[Koehn 2014B]

P. Koehn. Ressourcen-Sammlung zur statistischen maschinellen Übersetzung. Einführung, Software, parallele Korpora. Online-Quelle: <http://www.statmt.org/>, 28.11.2014

[Kolbs 2012]

P. Kolbs. Ressourcen-Sammlung für statistische maschinelle Übersetzung. Letzt geändert am 28.12.2012. Online-Quelle: <http://www.ling.uni-potsdam.de/~kolb/smt.html>, 28.11.2014

[Lingenio 2012]

Übersetzungssysteme von Lingenio: translate version 12. Online-Quelle: <http://static.lingenio.de/Deutsch/Produkte/Uebersetzungssysteme.htm>, 28.11.2014

[Linguetec 2014]

Personal Translator. Das preisgekrönte Übersetzungsprogramm von Linguetec. Online-Quelle: <http://www.linguetec.de/products/tr/pt>, 28.11.2014

[Lopez 2010]

A. Lopez. Machine Translation: Models, Search, and Evaluation. In: videolec- tures.net. exchange ideas & share knowledge. Online-Quelle: http://videolec- tures.net/clsps09_lopez_mt/, 28.11.2014

[Lopez 2014A]

A. Lopez; M. Post; Y. Cao. Machine Translation. Introduction to Machine Translation. 29.01.2014. Online-Quelle: <http://www.slideshare.net/alopezfoo/jhu-mt-lecture2014012>, 28.11.2014

[Lopez 2014B]

A. Lopez. Referate zur MÜ. Online-Quelle: <https://alopez.github.io/papers.html>, 28.11.2014

[Lucy 2014]

Lucy LT KWIK Translator. Kostenlose Übersetzung. Online-Quelle: <http://www.lucysoftware.com/deutsch/maschinelle-uebersetzung/kwik-translator/kostenlose-uebersetzung-lucy-lt-kwik-translator.html>, 28.11.2014

[Mitamura 1999]

T. Mitamura. Controlled Language for Multilingual Machine Translation. Proceedings of Machine Translation Summit VII, Singapore, September 13-17, 1999. Online-Quelle: <http://www.lti.cs.cmu.edu/Research/Kant/PDF/MTSummit99.pdf>, 28.11.2014

[Moses 2013A]

MBOT Moses Translation System in vorläufiger Version von 5. August 2013. In: IMS, Institut für Maschinelle Sprachverarbeitung an der Universität Stuttgart.

Online-Quelle:

<<http://www.ims.uni-stuttgart.de/forschung/ressourcen/werkzeuge/mbotmoses.html>>, 28.11.2014

[Moses 2013B]

Das phrasenbasierte SMÜ-System Moses. Online-Quelle:

<<http://www.statmt.org/moses/>>, 28.11.2014

[PROMT 2014]

Übersetzungssysteme von PROMT: PROMT 10. Online-Quelle:

<http://www.promt.de/translation_software/desktop/>, 28.11.2014

[Reverso 2013A]

Reverso Online. Kostenloser Online-Übersetzer in deutscher Version. Online-Quelle:

<http://www.reverso.net/text_translation.aspx?lang=DE>, 28.11.2014

[Reverso 2013B]

Reverso Localize. Website-Übersetzung und -Indizierung in englischer Version.

Online-Quelle: <<http://localize.reverso.net/Default.aspx?lang=en>>, 28.11.2014

[SDL 2013]

SDL Grundlagen. Regelbasierte und statistische Ansätze der maschinellen Übersetzung. Lösungsübersicht in deutscher Sprache. In: SDL Download-Center. 2013. Online-Quelle:

<<http://www.sdl.com/de/download/approaches-machine-translation/25416/>>, 28.11.2014

[SDL 2014A]

SDL Download-Center. Ressourcen zur MÜ. 2014. Online-Quelle:

<<http://www.sdl.com/de/resources/?Download+Type=&Area+of+interest=Maschinelle+%C3%9Cbersetzung&Product=&Service=>>>, 28.11.2014

[SDL 2014B]

SDL FreeTranslation.com. Online-Übersetzer für Dokumente, Websites und

Texten. Online-Quelle: <<http://www.freetranslation.com/>>, 28.11.2014

[SDL 2014C]

SDL FreeTranslation.com. Apps zur maschinellen Übersetzung. Online-Quelle:

<<http://www.translationzone.com/de/openexchange/automated-translation.html>>, 28.11.2014

[Seewald-Heeg 2009]

Hrg. U. Seewald-Heeg; D. Stein. Maschinelle Übersetzung - Von der Theorie zur Praxis. In: JLCL 24/3, 2009. Online-Quelle:
<http://media.dwds.de/jlcl/2009_Heft3/ldv-forum-24-3.pdf>, 28.11.2014

[Stein 2009]

D. Stein. Hrg. Seewald-Heeg. Maschinelle Übersetzung - Von der Theorie zur Praxis. Maschinelle Übersetzung – ein Überblick. In: JLCL 24/3, 2009:5-18. Online-Quelle:
<http://media.dwds.de/jlcl/2009_Heft3/daniel_stein.pdf>, 28.11.2014

[Systran 2014]

Desktop-Übersetzungssoftware von Systran. Online-Quelle:
<<http://store.systran.de/>>, 28.11.2014

[TAUS 2006]

Putting Machine Translation to Work. Report on the TAUS Executive Forum on March 23-24, 2006. Washington DC, April 2006. Online-Quelle:
<<https://www.taus.net/downloads/finish/56-public-reports/16-taus-putting-machine-translation-to-work-april-2006>>, 28.11.2014

[TAUS 2014A]

Machine Translation Post-Editing Guidelines. In: TAUS, Translation Automaton User Society. November, 2010. Online-Quelle:
<<https://evaluation.taus.net/resources/guidelines/post-editing/machine-translation-post-editing-guidelines>>, 28.11.2014

[TAUS 2014B]

Free TAUS Reports. In: TAUS, Translation Automaton User Society. 2014. Online-Quelle:
<<https://www.taus.net/free-reports/free-taus-reports>>, 28.11.2014

[TAUS 2014C]

A translation automation timeline. In: TAUS, Translation Automaton User Society. 2014. Online-Quelle:
<<https://www.taus.net/timeline/a-translation-automation-timeline>>, 28.11.2014

[Temnikova 2009]

I. Temnikova; C. Orasan. Post-editing Experiments with MT for a Controlled Language. Research Institute in Information and Language Processing, University of Wolverhampton, UK. In: Proceedings of the International Symposium on Data and Sense Mining, Machine Translation and Controlled Language (ISMTCL), Besancon, France, July 2009. Online-Quelle:
<<http://clg.wlv.ac.uk/papers/Temnikova-Orasan-ISMTCL.pdf>>, 28.11.2014
<http://www.academia.edu/2718763/Postediting_Experiments_with_MT_for_a_Controlled_Language>, 28.11.2014

[TMpro 2011]

TMpro. TMpro's translation engine. Online trademark Translator in Demo.
Online-Quelle: <<http://www.tmproapp.com/>>, 28.11.2014

[WorldLingo 2014]

WorldLingo. Free Translation Online. Online-Quelle:
<http://www.worldlingo.com/en/products_services/worldlingo_translator.html>,
28.11.2014

[Yu 2014]

Jingsong. S. Yu. Principles and Practice of Computer Aided Translation.
Grundkonzept computerunterstützter Übersetzungstechnologie und Verwendung
computerunterstützter Übersetzungswerkzeuge. MOOC in chinesischer Sprache
mit englischen Untertiteln an der Peking Universität. In: MOOC 学院, das größte
chinesische Studienforum. Studienbeginn am 22.09.2014. Online-Quelle:
<<http://mooc.guokr.com/course/875/Principles-and-Practice-of-Computer-Aided-Translation/>>, 28.11.2014