

10. Maschinelle Übersetzung

Antonia Schmalz

Das Internet ist dazu geeignet, Wissen zu demokratisieren und Informationen allen zugänglich zu machen sowie Zusammenarbeit zu fördern. Grundvoraussetzung dafür ist jedoch, dass man sich gegenseitig versteht. Ein automatisches Übersetzungssystem könnte helfen, Sprachbarrieren zu überwinden, doch die menschliche Sprache ist komplex und mehrdeutig. Übersetzungen können vom Kontext abhängen oder zusätzliches Wissen erfordern. Versuche der maschinellen Übersetzung lieferten daher bisher unzuverlässige und unnatürliche Ergebnisse. KI-Methoden könnten endlich Abhilfe schaffen: Künstliche neuronale Netze (KNN) sind in der Lage, selbstständig aus großen Datensätzen zusätzliches Wissen zu extrahieren, z. B. darüber, welche Wortkombinationen in einem bestimmten Kontext sinnvoll oder welche Formulierungen üblich sind. So kann es gelingen, der Vision, menschliche Sprache möglichst natürlich und korrekt wiederzugeben, tatsächlich einen Schritt näher zu kommen.

Laut Überlieferung der Bibel scheiterte bereits das Projektmanagement des Turmbaus zu Babel daran, dass die Strafe für dieses vermessene Vorhaben eine effiziente Kommunikation durch eine fehlende gemeinsame Sprache und damit die Errichtung des Turms selbst unmöglich machte. Zu den Menschheitsträumen zählt seither so etwas wie ein Universalübersetzungsgerät, allerdings gibt es das bislang nur in der Science-Fiction, etwa in Douglas Adams Buch „Per Anhalter durch die Galaxis“. Darin tritt der Universalübersetzer durch einen „Babelfisch“ in Erscheinung, dessen Namensgebung wieder auf die erste bekannte Problemsituation dieser Art verweist. Der Babelfisch ermöglicht demjenigen, der ihn im Ohr trägt, direkt in einer ihm fremden Sprache zu kommunizieren.

Als nun in den vergangenen Jahrzehnten die Welt durch das aufkommende und explosionsartig wachsende Internet näher zusammenrücken konnte und Informationsaustausch und Zusammenarbeit globaler wurden, spürte man erneut schmerzlich das Fehlen eines funktionstüchtigen Übersetzungswerkzeugs. So wurde im Jahr 1997 auch der erste Babelfisch vermeintlich Realität. Der Online-Dienst Babel Fish der Suchmaschine AltaVista, später Yahoo, übersetzte Texte auf Knopfdruck zwischen 36 Sprachenpaaren aus 13 Sprachen. Die zugrundeliegende Software der Firma Systran nutzte ein regelbasiertes Übersetzungssystem, das manuell vordefi-

nierte Syntax- und Grammatikregeln sowie Wörterbücher der betrachteten Sprache beherrschte und Sätze Wort für Wort abarbeitete.

Schnell zeigten sich jedoch die engen Grenzen dieser maschinellen Übersetzung. Ohne jegliche Alltagserfahrung und ohne ein zumindest rudimentäres „Verständnis“ für die Zusammenhänge im Satz oder gar im gesamten Text ist es dem Computer nicht möglich, Mehrdeutigkeiten aufzulösen. Handelt es sich nicht gerade um einen Horrroman, ist jedem menschlichen Leser mit einer gewissen Lebenserfahrung ohne weiteren Zusammenhang klar, dass für die Aussage „Ich öffnete das Einmachglas mit meiner rechten Hand“ kein Blut fließen musste. Zur zweifelsfreien Interpretation des Berichts „Ich traf den Sohn des Nachbarn mit einem blauen Ball“ ist hingegen auch für den Menschen mehr Kontext nötig. Ein Computer ohne Lebenserfahrung, der nur einzelne oder wenige Wörter im Zusammenhang betrachtet, muss hier scheitern.

Die Babel-Fish-Webseite wurde daher weithin berühmt für die Kreation absurder Aussagen – im deutschen Sprachraum insbesondere für die maschinelle Übersetzung der Starr-Reporte zur Lewinsky-Affäre des damaligen US-Präsidenten Clinton, die AltaVista 1998 zuerst stolz direkt auf der Startseite verlinkte. Als die Internetreaktionen klarmachten, dass Übersetzungen wie „Im Verlauf des Flirtings mit ihm hob sie ihre Jacke in der Rückenseite und zeigte ihm die Brücken ihrer Zapfenunterwäsche.“ (z. B. Strassmann 1998) keine Werbung für den Babel Fish waren, wurde der Text wieder entfernt. Der Babel Fish konnte dennoch im Jahr 2001 ca. 1,3 Mio. Übersetzungen pro Tag verzeichnen (Yang 2003). Sein Nutzen war, den Lesenden schnell und unkompliziert die wesentliche Aussage zu vermitteln, sogenanntes „gisting“ – nicht einen publizierbaren Text zu generieren.

Bis ins Jahr 2018 hat sich die Qualität der maschinellen Übersetzung durch den Übergang zu selbstlernenden Systemen (siehe Einleitung zu Kapitel Technologie „Entwicklungswege zur KI“) zwar deutlich verbessert, die grundlegenden Probleme und Herausforderungen sind aber die gleichen geblieben. Die Tatsache, dass einem Computer die für die kontextsensitive Übersetzung nötige Alltagserfahrung fehlt, legt sogar die Vermutung nahe, dass eine zuverlässig fehlerfreie, maschinelle Übersetzung nie möglich sein wird (Winther Madsen 2009).

Abbildung 10.1 zeigt links eine maschinelle Übersetzung nach dem aktuellen Stand der Technik eines beliebig ausgewählten Textes aus der New York Times. Der Sinn des ersten Satzes ist nicht zu entnehmen, da die Doppeldeutigkeit des englischen Wortes „skate“ (hier eigentlich „Rochen“, nicht „Schlittschuh“) falsch aufgelöst wurde und „land vertebrates“ (eigentlich Landwirbeltiere, nicht Landwirt) falsch übersetzt wurde. Die nachfolgenden Sätze sind jedoch gut verständlich.

ENGLISCH ▼

You can trace the heritage of little skates and land vertebrates back to a common ancestor that lived about 420 million years ago.

This primitive fish gave rise to the vast majority of vertebrate species, including all fish with paired fins. Typically, these pectoral and pelvic fins — which correspond to forelimbs and hindlimbs in land vertebrates — are small and largely used for balance and steering. To really propel themselves, most fish wriggle their powerful tails.

DEUTSCH ▼

Sie können das Erbe der kleinen Schlittschuhe und Landwirte auf einen gemeinsamen Vorfahren zurückverfolgen, der vor etwa 420 Millionen Jahren gelebt hat.

Aus diesem primitiven Fisch entstanden die meisten Wirbeltierarten, einschließlich aller Fische mit gepaarten Flossen. Typischerweise sind diese Brust- und Beckenflossen - die Vordergliedmaßen und Hintergliedmaßen in Landvertebraten entsprechen - klein und weitgehend für Gleichgewicht und Steuerung verwendet. Um sich selbst voranzutreiben, winden die meisten Fische ihre mächtigen Schwänze.



Abbildung 10.1: Textausschnitt aus einem zufällig ausgewählten Artikel der New York Times sowie das Ergebnis der maschinellen Übersetzung mit Google Translate (eigene Darstellung).

Der heilige Gral der maschinellen Übersetzung wird jedoch auch für die nahe Zukunft unerreichbar bleiben, nämlich eine zuverlässig korrekte, aber vor allem auch eine natürlich klingende Übersetzung zu erzeugen.

Anwendungen

Was ein Übersetzungsprogramm wirklich leisten muss, hängt stark vom Zweck ab. Wird eine automatische Übersetzung im kommerziellen, politischen oder juristischen Umfeld genutzt, ist es für die Außendarstellung eines Unternehmens wünschenswert oder rechtlich sogar nötig, dass die Übersetzung vollständig korrekt ist. Beispiele hierfür sind die Pflege der Webseite eines internationalen Unternehmens oder Protokolle von internationalen, politischen Tagungen. Hier müssen menschliche Übersetzerinnen und Übersetzer den Text in jedem Fall nachbearbeiten, und das Übersetzungsprogramm dient im Wesentlichen der Zeitersparnis. Die Einsatzdomänen sind meist sehr spezifisch, a priori bekannt, und oftmals existieren schon individuell gepflegte Datenbanken mit verifizierten Übersetzungen von Phrasen, Sätzen und Textabschnitten, auf die ein Programm zurückgreifen kann und die wiederverwendet werden können (Übersetzungsspeicher). Die Übersetzungsprogramme können sehr speziell auf die jeweiligen Anwendungsdomänen ausgerichtet und für diese vorkon-

figuriert werden. Eine Ausnahme sind zum Beispiel Filmuntertitel, die verschiedenste Bereiche berühren können. Anbieter von Übersetzungsprogrammen, die primär diesen Markt adressieren, sind z. B. Omniscien Technology, SDL, Yandex oder Systran. Im Angebot sind sowohl Serverlösungen, die der Kunde – z. B. aus Sicherheits- und Geheimhaltungsgründen – direkt bei sich betreiben kann, als auch Cloud-Services. Die maschinelle Übersetzung hat für diese Anwendungen eine überwiegend wirtschaftliche Bedeutung, in dem Sinn, dass primär Kosten und Zeit von sonst nötigen professionellen Übersetzern eingespart werden können.

Auf der anderen Seite gibt es Anbieter wie Google, Microsoft, DeepL oder auch Facebook und Amazon, die eine Generalübersetzung zur Verfügung stellen wollen – als Service für den Internetnutzer, den Social-Media-Nutzer, Reisende, die die Online-Bewertung von Hotels verstehen möchten, oder den Hausmann, der das leckere Bibimbap aus dem letzten Asienurlaub möglichst original nachkochen möchte. Solche Übersetzungsprogramme müssen alle möglichen Domänen abdecken können, es ist aber meistens ausreichend, wenn sie den Sinn des Ursprungstextes wiedergeben können. In diesem Anwendungsfeld ist neben einem wirtschaftlichen Nutzen vor allem auch ein gesellschaftlicher Vorteil zu sehen, da die Kommunikation über Sprachgrenzen erleichtert wird, und insbesondere auch mehr Menschen Zugang zu Wissen und Informationen bekommen, die etwa im Internet im Wesentlichen auf Englisch verfügbar sind (siehe auch Sprachenerweiterung und „Low-Resource Languages“).

Nicht zu vernachlässigende Treiber der Entwicklung von Übersetzungsprogrammen sind auch Polizei und Militär. Im Rahmen der Terrorismusbekämpfung soll gegebenenfalls automatisiert fremdsprachige Kommunikation nach Hinweisen auf verdächtiges Verhalten durchsucht werden, die über einzelne Schlagwörter hinausgehen. Auch Soldaten in den Einsatzgebieten sollen in der Kommunikation mit der Bevölkerung vor Ort unterstützt werden.

Durch zunehmend international agierende Unternehmen und sprachübergreifende Social-Media-Anwendungen entsteht ein starker „Application Pull“, auf den verstärkt und verstärkend wirkend ein „Technology Push“ durch die drastische Weiterentwicklung von Deep-Learning-Algorithmen und neuronalen Netzwerken trifft.

Die Übersetzungsalgorithmen, die derzeit weiterentwickelt werden, sind selbstlernende Systeme ohne vorgegebene Sprachregeln, die bei der Übersetzung mehr als ein einzelnes Wort im Zusammenhang betrachten können. Hierfür werten sie eine möglichst große Datenbank von Texten aus, die parallel in den betrachteten Sprachen vorliegen (Prinzip des Rosettasteins). Üblicherweise ist ein zentraler Bestandteil dieser Datenbank das Archiv der Vereinten Nationen mit mehreren Milliarden Wörtern in zahlreichen Sprachen.

Statistische Maschinelle Übersetzung (SMT)

Im Jahr 2007 führte Google als einer der ersten Anbieter ein selbstlernendes Übersetzungssystem ein, das Datenbanken statistisch auswertete (Statistical Machine Translation, SMT) und so die wahrscheinlichsten Übersetzungen eruierte. Aktuell werden bei der statistischen Übersetzung jeweils eine Reihe von aufeinanderfolgenden Wörtern gleichzeitig betrachtet (Phrase-based Machine Translation, PBMT) und damit auch die Häufigkeit von Wortkombinationen evaluiert, wodurch implizit auch der Kontext zu einem geringen Grad berücksichtigt wird, wenn die sinnrelevanten Wörter eng genug beisammenstehen.

Probleme treten hier vor allem bei Sprachpaaren auf, bei denen sich die Wortreihenfolge und/oder die grammatikalische Struktur stark unterscheiden. Durch die Einbindung zusätzlicher Sprachmodelle, „Reordering“- (Umsortierungs-) Modellen und weiterer unterstützender Algorithmen wird diesen Defiziten entgegengewirkt. Durch den statistischen Ansatz entsteht außerdem eine starke Abhängigkeit vom Trainingsmaterial. Trainiert man etwa als Extrembeispiel einen Algorithmus ausschließlich mit Texten aus der Zoologie, wird in der anschließenden Übersetzung eines Berichts über ein Baseballspiel das englische Wort „bat“ trotzdem immer als „Fledermaus“, nicht aber als „Schläger“ wiedergegeben werden. Im ersten Hype um das damals neue Verfahren zitiert die Zeitschrift Computerbild (Hülsbörner, 2007) Philipp Köhn, einen der Mitentwickler der SMT: „Technische und politische Texte können wir völlig problemlos übersetzen lassen – von Sportberichten und Kochrezepten lassen wir dagegen lieber die Finger“. Ursache dieser Einschätzung war die thematische Beschränkung der Themen, die in den frei verfügbaren und zum Training nutzbaren Texten der Vereinten Nationen oder Europäischen Union existiert.

Neuronale Maschinelle Übersetzung (NMT)

Im November 2016 führte Google dann mit großer Medienresonanz das erste Übersetzungssystem ein, das sich eines neuronalen Netzes (Neural Machine Translation, NMT) bediente und versprach, mit natürlicheren Sätzen die Lücke zwischen menschlicher und maschineller Übersetzung zu schließen („Bridging the Gap between Human and Machine Translation“; (Yonghui Wu et al., 2016)). Bis dato gab es zwar vielversprechende Arbeiten zur Nutzung neuronaler Netze, die NMT-Systeme schnitten aber im Praxistest schlechter ab als ausgereifte SMT-Übersetzer. Google konnte erstmals verschiedene Schwächen, wie etwa die langsamere Trainingsgeschwindigkeit, die ineffiziente Behandlung von seltenen Wörtern oder das Problem, dass manchmal nicht alle Wörter des Ursprungssatzes übersetzt wurden, erfolgreich ausgleichen.

Anders als bei der SMT wird bei der NMT immer ein vollständiger Satz gleichzeitig betrachtet. Mathematisch wird von einer linearen Abbildung zwischen Eingangs-

und Ausgangssprache in der SMT übergegangen zu einer nicht-linearen Abbildung in der NMT, basierend auf einer Vektordarstellung des Ein- und Ausgangssatzes über mehrere Zwischenstufen und mehrere Abstraktionsgrade. Dadurch kann jedes einzelne Wort jeweils in der Abhängigkeit von allen anderen Wörtern im Satz betrachtet werden. Zusätzliche Sprachmodelle sind hier nicht nötig. Durch die gleichzeitige Betrachtung eines vollständigen Satzes mit allen Abhängigkeiten können NMT-Algorithmen insbesondere besser mit sogenannten „long-distance dependencies“ umgehen, d. h. getrennten Satzteilen, die grammatikalisch und/oder inhaltlich voneinander abhängen und die Aussage stark beeinflussen können. Besonders im Deutschen sind solche „long-distance dependencies“ häufig: „Für die Geburtstagsfeier der dreijährigen Lise brachte der Vater ein Kaninchen mit einer blauen Schleife mit/um.“ Auch Sprachpaare, bei denen die Wortreihenfolge stark unterschiedlich ist, beherrscht NMT besser. Die starke Abhängigkeit von Umfang, Qualität und Ausrichtung des Trainingsmaterials besteht jedoch auch hier. Allgemein erzeugen NMT-Systeme deutlich natürlicher klingende, für den Menschen verständlichere Sätze als SMT-Systeme, die auch den Inhalt besser wiedergeben. Die Satzstruktur ist besser, es gibt weniger Syntaxfehler und weniger Fehler in der Wortreihenfolge.

Auf der Webseite des Übersetzungsprogramms Microsoft Translator (Microsoft) kann der interessierte Nutzer selbst vergleichende Tests durchführen und ein Gefühl für die jeweilige Übersetzungsqualität bekommen. Die beiden Beispiele in Abbildung 10.2 demonstrieren einerseits das bessere Niveau der NMT-Übersetzung, andererseits aber auch die immer noch vorhandene Unzulänglichkeit bei Übersetzungen, die ein Verständnis der Aussage erfordern.

Auch wenn nach Google im Verlauf des Jahres 2017 fast alle großen Anbieter von Übersetzungswerkzeugen, wie Microsoft/Bing, Yandex, Systran, SDL oder Omniscien Technologies, die NMT-Technologie in ihre Systeme integriert haben, wurde dadurch der SMT-Algorithmus nicht automatisch ersetzt. Oftmals werden hybride Verfahren angeboten, die die Vorteile beider Ansätze nutzen. So lassen sich bei SMT-Algorithmen etwa leichter problem- bzw. anwendungsspezifische Datenbanken einbinden, die eine bestimmte fachspezifische Terminologie oder regelmäßig wiederkehrende Sätze und Formulierungen mit verifizierter Übersetzung (Übersetzungsspeicher) enthalten. Anders als NMT kann dies dadurch auch sicherstellen, dass feste Bezeichnungen über einen Text hinweg konsistent übersetzt werden. Abhängig vom Anwendungszweck, der Anwendungsdomäne, dem verfügbaren Trainingsmaterial und auch dem betrachteten Sprachenpaar kann SMT auch die bessere Übersetzungsqualität liefern (siehe z. B. M. Farajian et al.).

Im August 2017 trat überraschend ein neuer Player auf den Markt. Das deutsche Unternehmen DeepL behauptet auf seiner Webseite (DeepL) unbescheiden, dass „im Blindtest [...] DeepLs Resultate etwa drei Mal so häufig als beste Übersetzung gewählt

DEUTSCH ▼

Zum 15. Jahrestag von 9/11 erklärt Expertin Amanda Ripley, wieso manche Menschen im World Trade Center auf die Anschläge vorbereitet waren und sich retten konnten – und andere nicht.

ENGLISCH ▼

SMT

To the 15th anniversary of 9/11, an expert does not explain Amanda Ripley, why some people in the World Trade Center to the attacks were prepared and could – save yourself and others.others not)

NMT

On the 15th anniversary of 9/11, expert Amanda Ripley explains why some people in the World Trade Center were prepared for the attacks and saved themselves – and not others. (korrekt: - and others not)

DEUTSCH ▼

Zwei Tage später veränderte sich alles in seinem Leben, als er nach der Physiotherapie mit dem Auto zum Trainingszentrum fahren wollte – dieses aber nicht mehr erreichte.

ENGLISCH ▼

SMT

Two days later everything in his life changed but no longer reached when he wanted to – go after the physiotherapy training centre by car this.

NMT

Two days later, everything changed in his life when he wanted to drive to the training center by car after physiotherapy – but this was no longer achieved. (korrekt: reached)



Abbildung 10.2: Auf der Webseite des Microsoft Translators (Microsoft) können eigene Texte oder Beispielsätze aus einer Datenbank mit den beiden verfügbaren Algorithmen (SMT, NMT) übersetzt werden (eigene Darstellung der Übersetzungsergebnisse).

werden, wie die der anderen.“ Die zugrundeliegenden neuronalen Netze sind bei DeepL erstmals keine sogenannten rekurrenten neuronalen Netze, sondern „Convolutional Networks“. Solche Architekturen sind für die Bilderkennung üblich, für die Übersetzung gab es bisher Forschungsarbeiten – auch z. B. von Google und Facebook –, aber keine kommerziellen Implementierungen. Der Vorteil von Convolutional Networks ist, dass alle Wörter parallel übersetzt werden können und bereits optimierte Bibliotheken für die Berechnung existieren (Merkert, 2017). Darüber hinaus profitiert die Qualität des DeepL-Systems von einem weiteren Faktor. DeepL bot vorher unter dem Firmennamen Linguee eine Suchmaschine für Übersetzungen an und konnte damit extrem umfangreiche, qualitativ sehr hochwertige Trainingsdaten sammeln.

Die zentrale Technologie der Zukunft wird sicherlich ein selbstlernendes System jenseits der rein statistischen Methoden sein. Das ist neben der reinen Qualitätsdebatte unter anderem auch begründet durch die Möglichkeit von Zero-Shot-Übersetzungen, das heißt direkte Übersetzungen zwischen Sprachenpaaren, für die keine parallelen Texte zum Training vorliegen (siehe auch Textfundus). Welche Ausprägung sich durchsetzt und welche Versprechen und Visionen damit tatsächlich erfüllt werden können, wird sich in den kommenden Jahren zeigen. Die technologischen Fortschritte überholen sich gerade, die Nachfrage aus möglichen Anwendungen ist enorm. Die Themen Machine Learning (ML) und Deep Learning (DL) standen 2017 in Gartners Hype Cycle (Columbus 2017) auf dem höchsten Punkt, der für aufgeblasene Erwartungen an diese Technologie steht. Der Antwort auf die Frage, wie nahe eine Übersetzung durch eine Künstliche Intelligenz der natürlichen Sprache kommen kann, werden uns die nächsten Jahre also zumindest etwas näher bringen.

Evaluation

Um eine definitive Aussage darüber treffen zu können, welcher Algorithmus die besseren Ergebnisse liefert, muss die Übersetzungsqualität objektiv gemessen werden. Das ist insbesondere wichtig, um die Parametereinstellungen vorhandener Algorithmen optimieren beziehungsweise weiterentwickeln zu können. Wird in anderem Kontext ein KI-System z. B. trainiert, Straßenschilder in Videoströmen zu finden, lässt sich leicht überprüfen, wie viele Schilder richtig als solche identifiziert und wie viele übersehen wurden. Die quantitative Bewertung der Qualität einer automatischen Sprachübersetzung ist jedoch selbst Gegenstand wissenschaftlicher Debatten. Wie die „beste“ Übersetzung eines Satzes lautet, ist subjektiv, es gibt oft mehrere inhaltlich und grammatikalisch korrekte Versionen. Ein automatisierter, objektiver Vergleich mit einer subjektiven Referenzübersetzung wird nie eine absolute Aussagekraft haben.

Eine weit verbreitete Metrik ist das BLEU-System (Papineni et al. 2002). Grundidee hier ist zu vergleichen, wie viele Wörter aus der automatischen Übersetzung auch in einer humanen Referenzübersetzung nahe zusammen vorkommen. Grammatik und Syntax werden nicht evaluiert. Schwächen dieser Metrik lassen sich am einfachsten an einem Extrembeispiel demonstrieren. Den englischen Satz „The food in prison was horrible.“ übersetzt Google Translate korrekt mit „Das Essen im Gefängnis war schrecklich“. Lautet der deutsche Referenzsatz, mit dem die Übersetzung für die Bewertung verglichen wird, aber „Die Nahrung in der Haftanstalt war fürchterlich.“, wird der BLEU-Wert miserabel ausfallen, da einzig das Wort „war“ identisch ist. Die Metrik kann wegen solcher Unzulänglichkeiten nicht zur Evaluierung eines einzelnen Satzes genutzt werden, sondern muss über längere Texte gemittelt werden.

Mit der Veröffentlichung des NMT-Systems gab Google an, für die Übersetzung eines häufig verwendeten Referenztexts (WMF'14) vom Englischen ins Französische einen

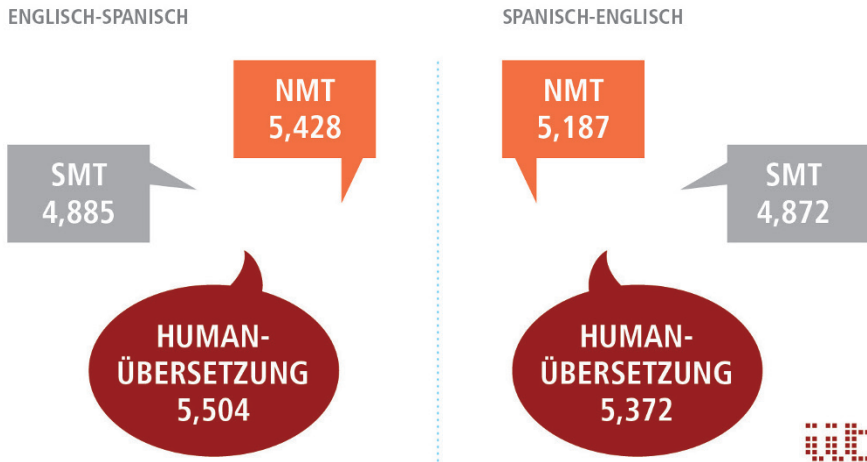


Abbildung 10.3: Die betrachtete SMT ist der phrasenbasierte Algorithmus „Google Translate“ und die NMT der neuronale Algorithmus „Google Neural Machine Translation (GNMT)“. Die Humanübersetzung wurde durch bilinguale Menschen erstellt. Die Zahlenwerte geben die Bewertung der jeweiligen Übersetzung durch menschliche Prüfer an (0 = schlecht, 6 = perfekt).

BLEU-Wert von 38.95 (aus 0 bis 100) zu erreichen und für die Übersetzung vom Englischen ins Deutsche 24,17 BLEU (Wu et al. 2016). DeepL gibt an, auf den gleichen Testdaten 44,7 für Englisch-Französisch und 31,1 BLEU für Englisch-Deutsch zu erreichen. Kritiker der BLEU-Metrik bemerken jedoch, dass durch die Schwächen des Messverfahrens Änderungen um wenige Punkte nicht aussagekräftig sind (Pan 2016).

Zusätzlich zur BLEU-Evaluierung ließ Google im Rahmen der gleichen Veröffentlichung menschliche Prüfer drei Versionen einer Übersetzung bewerten: SMT, NMT und Human. Die Prüfer vergaben 0 bis sechs Punkte für die Qualität der Übersetzung. Abbildung 10.3 zeigt das Ergebnis und die Verbesserung der Qualität durch das neuronale Netz. Doch allein die Tatsache, dass selbst die Humanübersetzung nicht die volle Punktzahl erreichte, zeigt, wie subjektiv eine Qualitätsbewertung ist.

Der Hersteller von Übersetzungssoftware Omniscien Technologies weist in einem White Paper außerdem darauf hin, dass die Qualität der Übersetzung sehr stark von verschiedenen Randbedingungen abhängt. Neben dem gewählten Algorithmus spielen auch die Trainingsdaten (Qualität, Umfang, spezifische Domänen), das konkrete Sprachenpaar und die Art des Textes (technisch, formal, umgangssprachlich) wie auch die Zieldomäne und konkrete Anwendung eine wichtige Rolle.

Textfundus und Zero Shot Translation

Ein kritischer Faktor für die Qualität einer Übersetzung ist der Textfundus, mit dem der selbstlernende Algorithmus trainiert wird. Traditionell wurden hier Dokumente aus dem Umfeld der UN oder der EU herangezogen, die professionelle Übersetzer in zahlreichen Sprachen parallel erstellen und die frei verfügbar sind. Bücher, die in mehrere Sprachen übersetzt wurden, sind typischerweise nicht im großen Umfang frei zugänglich. Der Anbieter DeepL gründet seinen Erfolg unter anderem darauf, dass er aus seiner Unternehmensvorgeschichte Zugriff auf Milliarden qualitativ hochwertiger Übersetzungen hat. Die großen Anbieter, die bestimmte kommerzielle Domänen adressieren, heben sich vor allem auch durch das jeweilige domänenspezifische Trainingsmaterial voneinander ab. Unter den Generalübersetzern hat z. B. Google durch seinen Zugriff auf riesige Datenmengen in verschiedenen Sprachen einen Vorteil gegenüber Wettbewerbern. Da diese Daten jedoch nicht zwingend verifiziert und von guter Qualität sind, kann ein geringerer Datenumfang die gleiche oder sogar bessere Übersetzungsqualität liefern, solange die Güte der Trainingsdaten zuverlässig hoch ist.

Einen von vielen interessanten, im Zusammenhang mit der Auswahl des Textfundus auftretenden Effekten beschrieb Nataly Kelly, Vice President bei Smartling: „Given that male pronouns have been over-represented throughout history in most languages and cultures, machine translation tends to reflect this historical gender bias.“ (Errens) (Übersetzung mit DeepL: „Da männliche Pronomen in der Geschichte in den meisten Sprachen und Kulturen überrepräsentiert waren, spiegelt die maschinelle Übersetzung diese historische geschlechtsspezifische Ausrichtung wider.“). So kann es etwa passieren, dass das englische „engineer“ unabhängig vom Kontext eher mit Ingenieur als mit Ingenieurin übersetzt wird.

Wenn nur beschränkt viel Textmaterial verfügbar ist, zeigt sich eine weitere Stärke der NMT. So können z. B. prinzipiell auch einsprachige Texte zum Lernfortschritt beitragen, da hieraus auch Sprachstruktur einer einzelnen Sprache und begünstigte Wortkombinationen entnommen werden können. Die Verbesserung der Lernprozesse, insbesondere unter Einbeziehung von monolinguaalem Trainingsmaterial ist derzeit denn auch ein zentraler Entwicklungspunkt. Das ist besonders für diejenigen Sprachenpaare relevant, für die keine oder nur wenige gemeinsame, bilinguale Texte vorliegen.

In genau diesen Fällen erweist sich die NMT der SMT durch ein weiteres Potenzial als überlegen: Sie bietet die Möglichkeit zur sogenannten Zero-Shot-Translation. Gibt es etwa für das Sprachenpaar Finnisch-Afrikaans nicht genügend zweisprachiges Trainingsmaterial, muss mit dem SMT-Verfahren eine Zwischensprache genutzt werden, mit der es jeweils bilinguale Texte gibt, die parallel statistisch ausgewertet werden können („Pivot-Translation“). So muss erst vom Finnischen z. B. ins Englische und im

Anschluss vom Englischen in Afrikaans übersetzt werden. In diesem Prozess akkumulieren sich Übersetzungsfehler drastisch und die Ausgabe in der Zielsprache ist zum Teil kaum mehr verständlich. Ein neuronales Netz hingegen kann prinzipiell gleichzeitig mit allen Trainingsdaten mehrerer Sprachen gefüttert werden (Johnson et al. 2017) und direkt Verbindungen über mehrere Sprachen hinweg aufbauen. Wenn keine direkte Verbindung zwischen zwei Sprachen besteht, kann über Logikketten eine Übersetzung von anderen Sprachenpaaren abgeleitet werden, die relevante Informationen beinhalten (Abbildung 10.4). In der Realität werden derzeit jedoch viele neuronale Systeme für konkrete Sprachenpaare konfiguriert und trainiert. Die Zero-Shot Translation steht erst am Anfang der Entwicklungen und wird zu weiteren Verbesserungen in der Übersetzungsqualität bei allen Sprachpaaren beitragen.

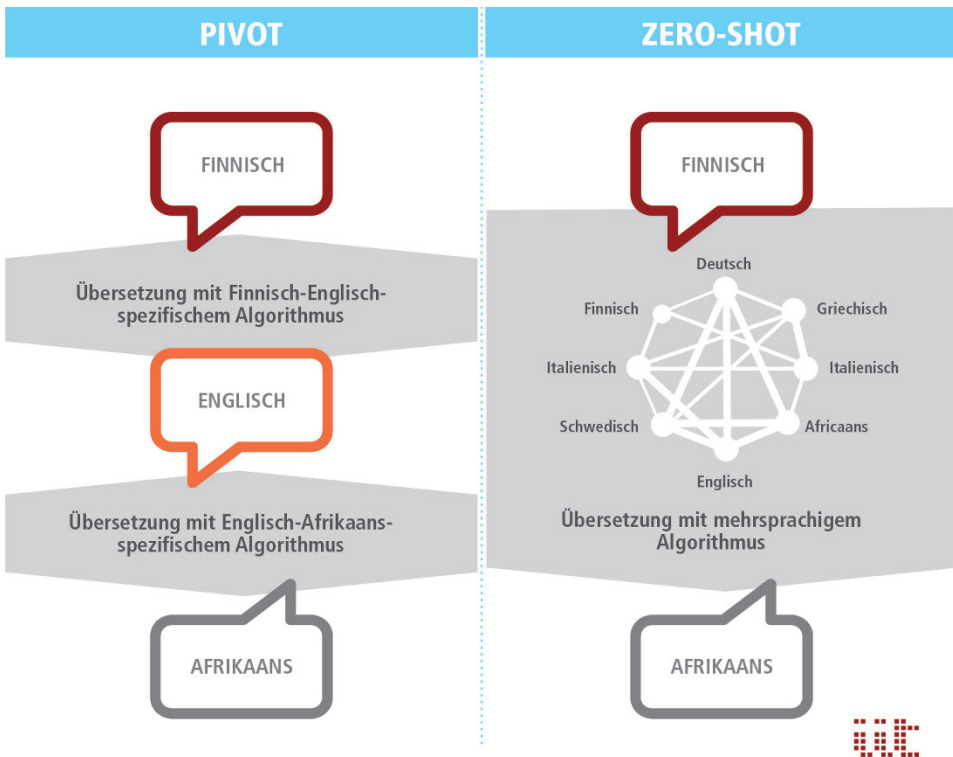


Abbildung 10.4: Bisher musste für Sprachpaare, die nicht über ausreichend viele gemeinsame Textquellen verfügen, eine sogenannte „Pivot-Translation“ über eine gemeinsame Zwischensprache durchgeführt werden. Ein NMT-Algorithmus hingegen kann prinzipiell gleichzeitig mit allen Trainingsdaten mehrerer Sprachen gefüttert werden und direkt Verbindungen über mehrere Sprachen hinweg aufbauen – per „Zero-Shot-Translation“ (eigene Darstellung angelehnt an Lommel 2017).

Insgesamt ist der Einfluss des Trainingsmaterials und des Trainingsprozesses auf die Übersetzungsqualität enorm. Die Verfügbarkeit von geeignetem, hochwertigem Trainingsmaterial wird zukünftig entscheidend sein, um die Vision einer natürlichen, zuverlässigen Sprachübersetzung zu erreichen.

Sprachenerweiterung und „Low-Resource languages“

Google bot im Februar 2018 die Möglichkeit, 103 Sprachen zu übersetzen, wobei noch nicht alle Sprachenpaare die NMT nutzen können, sondern mit dem SMT-Algorithmus auskommen müssen. DeepL unterstützte sieben Sprachen. Auf der Welt gibt es etwa 6.900 verschiedene Sprachen. Google deckt davon also weniger als 2 Prozent ab. Fast das gesamte Wissen liegt in ca. 1 Prozent der Sprachen vor (Carbonell 2016).

Um das volle Potenzial des Internets für einen gerechten, barrierefreien Zugang zu Wissen und Informationen ausschöpfen zu können, muss sichergestellt werden, dass auch die Sprachbarriere fällt. Und zwar nicht nur für die 77 Sprachen, die mehr als zehn Millionen Menschen jeweils sprechen, sondern auch und gerade für diejenigen Sprachen, in denen sich nur kleine Gruppen verständigen. Für diese Sprachen steht oft nur wenig Trainingsmaterial für die maschinelle Übersetzung zur Verfügung („Low-resource languages“). Allein in Indien gibt es bis zu 400 linguistisch unterschiedliche Sprachen und selbst dort, in der ehemaligen britischen Kolonie, die Englisch als eine offizielle Amtssprache besitzt, ist es bei weitem keine Selbstverständlichkeit, dass Englisch – die Sprache des Internets – verstanden wird. Auch Soldaten oder internationale Hilfsorganisationen in Krisengebieten treffen häufig auf Vertreter der „Low-resource-languages“ und könnten von maschinellen Übersetzern profitieren. „Low-resource-languages“ und das Auffinden von nutzbarem Textmaterial für diese Sprachen stehen daher auch regelmäßig im Fokus der Ideenschmiede des US-Verteidigungsministeriums DARPA.

Grundsätzlich übersetzen derzeit SMT-Algorithmen „Low-resource languages“ besser, da diese mit weniger Trainingsmaterial bessere Ergebnisse liefern als datenhungrige neuronale Netze. Gleichzeitig sind gerade für diese Sprachen erst recht keine zweisprachigen Texte in beliebigen Zielsprachen vorhanden. Ein sprachenunspezifischer NMT-Algorithmus, der mit allen verfügbaren Sprachen gleichzeitig trainiert wird, Zero-Shot-Translation ermöglicht und gegebenenfalls durch einsprachiges Trainingsmaterial unterstützt werden kann, birgt hier großes Leistungspotenzial.

Integration

Die Vision, die Douglas Adams' Babelfisch verkörpert, geht deutlich über das reine Übersetzen von Texten hinaus. Nachdem der Reisende den Babelfisch ins Ohr gesteckt

hat, kann er die Problematik „Sprache“ komplett ausblenden und sich auf seine Handlungen konzentrieren. Als ein englischsprachiger Google-Manager und seine schwedische Kollegin sich im Herbst 2017 bei einer Presseveranstaltung Googles neue In-Ear-Kopfhörer ins Ohr steckten und begannen, sich miteinander in ihrer jeweiligen Sprache zu unterhalten, lag in den Medien der Vergleich mit dem Babel-fisch nahe. Was tatsächlich passierte, war, dass die Bluetooth-Kopfhörer mit einem Smartphone verbunden waren, auf dem die App Google Translate lief. Die App übertrug über den Cloud-Service die gesprochene englische Sprache in englischen Text, übersetzte diesen ins Schwedische und verwandelte den schwedischen Text wieder in eine schwedische Sprachausgabe. Weil das drahtlose Netzwerk bei der Presskonferenz offensichtlich eine hohe Datenrate unterstützte, konnte das Ganze fast in Echtzeit geschehen.

Damit wir die Sprachbarriere tatsächlich ignorieren können, muss die Übersetzung vollständig in unsere Umgebung integriert werden. Im ausschließlich textlich digitalen Umfeld ist das bereits gut umgesetzt. Google Translate kann mittlerweile in jeder anderen App genutzt werden, Webseiten können aus der Suchergebnisliste direkt in der automatisch übersetzten Version aufgerufen werden, Facebook zeigt Neuigkeiten auf Wunsch direkt zusammen mit der Übersetzung an. Sprache-zu-Sprache-Übersetzungsdienste hingegen gibt es zwar – neben Google bietet dies zum Beispiel auch Microsoft im Skype Translator an – die derzeitige Qualität lässt jedoch zu wünschen übrig. Das maschinelle Verstehen von gesprochener Sprache ist eine immense Herausforderung und auch nur durch das exzessive Training einer Künstlichen Intelligenz mit einer riesigen Datenmenge zu erreichen. Ob individuelle Stimmlagen, Dialekte, Akzente, Sprachfehler, undeutliche Aussprache oder Hintergrundgeräusche: die vielfältigen akustischen Variationen machen das Training eines Spracherkennungsalgorithmus zu einer Herkulesaufgabe. In der Verkettung der Spracherkennung mit einem Übersetzungsprogramm, das mit unsauberen Formulierungen, unvollständigen Sätzen, Slang und Umgangsvokabular noch einmal besonders herausgefordert wird, addieren sich die Fehler.

Trotz dieser nachvollziehbaren Herausforderungen funktionieren die existierenden Apps in überschaubaren Alltagssituation überraschend gut und lassen gespannt auf die Entwicklungen der nächsten Jahre blicken. Allein das kontinuierliche Anwachsen der Datenbasis, z. B. von Google, trainiert die Künstliche Intelligenz ständig weiter, sodass sogar ohne sonstiges Zutun immer mehr Sprachnuancen unterschieden und erkannt werden können.

Im Jahr 2006 urteilte die britische Wochenzeitung Economist mit Recht: „Translation systems are of limited use if they cannot be used by people on the move“ (Übersetzung mit DeepL: „Übersetzungssysteme sind nur begrenzt nutzbar, wenn sie von Personen, die unterwegs sind, nicht genutzt werden können.“; The Economist 2006).

Dem sind wir mit Apps wie Google Translate, für die im Übrigen auch ein Offline-Modus zur Verfügung steht und mit denen auch Text aus Bildern übersetzt werden kann, schon deutlich näher gekommen, ein Babelfish ist es aber noch nicht.

Ausblick

Will man sich aktuell zum Thema maschinelle Übersetzung schlau machen, muss man penibel das Erstellungsdatum der Informationsquelle berücksichtigen. Aussagen, die vor weniger als zwei Jahren richtig waren, sind mit Einführung der NMT ab Ende 2016 bereits überholt, und die Entwicklungsgeschwindigkeit scheint nicht abzunehmen. Durch die massiven Investitionen großer Konzerne wie Google, Amazon, Microsoft in die Forschung zum Deep Learning und Maschinenlernen werden kontinuierlich Fortschritte erzielt (Columbus 2017). Bis zum Jahr 2016 war die gesamte Übersetzungsbranche zu einer 40-Milliarden-USD-Industrie angewachsen mit einem jährlichen Wachstum von 7 Prozent. Der Unterbereich der maschinellen Übersetzung ist hingegen sogar um 20 bis 25 Prozent gewachsen (Vogel 2016).

Parallel zu den Entwicklungen in der reinen Sprachübersetzung wachsen die Ausprägungen und Anwendungsfelder von KI stetig an. Die Abhängigkeit von Qualität und Umfang der Trainingsdaten, das mangelnde Verständnis für den Kontext und die fehlende Alltagserfahrung sind Probleme, die über die verschiedenen Anwendungen hinweg immer wieder auftauchen. Lösungen können sich gegenseitig befruchten, wie auch DeepL erfolgreich die Ansätze aus der Bilderkennung auf die Sprachübersetzung übertragen hat.

Für eine zuverlässig fehlerfreie Übersetzung benötigen Computer semantisches Wissen: Gebäude haben Dächer, Türen können sich öffnen und schließen, Autos können fahren, Flugzeuge fliegen in der Luft.... Das gleiche Wissen, das eine Übersetzung erleichtert und prinzipiell in Datenbanken vorgehalten werden kann, vermag auch einen Algorithmus bei der Bilderkennung oder beim Interpretieren gesprochener Sprache zu unterstützen. Für die richtige Interpretation einer Aussage und damit eine zweifelsfreie Übersetzung ist aber über solche einfachen Zusammenhänge hinaus noch zusätzliches empirisches Wissen nötig. In den folgenden beiden Beispielsätzen bezieht sich „sie“ einmal auf die Stadtverwaltung und einmal auf die Demonstranten. Ohne ein komplexeres Hintergrundwissen über Stadtverwaltungen, Demonstranten und Politik kann aber mit keinem Regelsatz dieser Zusammenhang hergestellt werden (angelehnt an Hobbs 1976).

Die Stadträte verweigerten den Demonstranten eine Erlaubnis, weil sie Gewalt fürchteten.

Die Stadträte verweigerten den Demonstranten eine Erlaubnis, weil sie Gewalt befürworteten.

Ob eine solche sprachliche Ambiguität in ferner Zukunft automatisiert aufgelöst werden kann, vielleicht durch einen immer größeren Datenschatz und die Kombination unterschiedlicher KI-Anwendungen, ist abzuwarten. Vielleicht kann das aber auch erst durch den aktiv agierenden und interagierenden Roboter geschehen, der dann wirklich direkt im Alltag lernen kann. Bis dahin werden die Entwicklungen aber einem Babelfisch zunehmend nahe kommen und Sprachbarrieren können weiter abnehmen.

Literatur

- Carbonell, Jaime (2016): *Massively Multilingual Language Technologies*. Anniversary Symposium "Building Bridges, Breaking Barriers", Baden-Baden. Online verfügbar unter http://videlectures.net/interACT2016_carbonell_multilingual_technologies/, zuletzt geprüft am 22.06.2018.
- Columbus, Louis (Forbes, Hrsg., 2017): *Gartner's Hype Cycle for Emerging Technologies, 2017 Adds 5G And Deep Learning For First Time*. Online verfügbar unter <https://www.forbes.com/sites/louiscolumbus/2017/08/15/gartners-hype-cycle-for-emerging-technologies-2017-adds-5g-and-deep-learning-for-first-time/#3472122b5043>, zuletzt geprüft am 22.06.2018.
- DeepL. Online verfügbar unter <https://www.deepl.com/press.html>, zuletzt geprüft am 22.06.2018.
- Errens, Julie: *The past, present and future of machine translation*. Online verfügbar unter <http://www.monotype.com/resources/articles/the-past-present-and-future-of-machine-translation/>, zuletzt geprüft am 22.06.2018.
- Farajian, M. Amin et al.: Neural vs. Phrase-Based Machine Translation in a Multi-Domain Scenario. In Association for Computational Linguistics (Hrsg.), *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers* (S. 280–284). Online verfügbar unter <http://www.aclweb.org/anthology/E17-2045>, zuletzt geprüft am 22.06.2018.
- Hobbs, J.R. (1976): *Pronoun resolution*. Research Report, City University of New York. New York. Verfügbar unter <http://www.isi.edu/~hobbs/PronounResolution.pdf>, zuletzt geprüft am 22.06.2018.
- Hülsbörner, Simon (Computerwoche, Hrsg., 2007, 27. April): *Elektronische Dolmetscher holen auf*. Online verfügbar unter <https://www.computerwoche.de/a/elektronische-dolmetscher-holen-auf,591945>, zuletzt geprüft am 22.06.2018.
- Johnson, Melvin et al. (2017): Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. *Transactions of the Association for Computational Linguistics* (5), 339–35. Online verfügbar unter <http://aclweb.org/anthology/Q17-1024>, zuletzt geprüft am 22.06.2018.
- Lommel, Arle (Common Sense Advisory Blog, Hrsg.): *Zero-Shot Translation Is Both More and Less Important Than You Think*. Online verfügbar unter <http://www.commonsenseadvisory.com/default.aspx?Contenttype=ArticleDetAD&tabID=63&Aid=37905&modul=390>, zuletzt geprüft am 22.06.2018.
- Merkert, Johannes (Heise online, Hrsg., 2017): *Maschinelle Übersetzer: DeepL macht Google Translate Konkurrenz*. Online verfügbar unter <https://www.heise.de/newsticker/meldung/Maschinelle-Uebersetzer-DeepL-macht-Google-Translate-Konkurrenz-3813882.html>, zuletzt geprüft am 22.06.2018.
- Microsoft: *Microsoft Translator - Try & Compare*. Online verfügbar unter <https://translator.microsoft.com/neural/>, zuletzt geprüft am 22.06.2018.

- Pan, Hazel Mae (2016): *How BLEU Measures Translation and Why It Matters*. Online verfügbar unter <https://sator.com/technology/how-bleu-measures-translation-and-why-it-matters/>, zuletzt geprüft am 22.06.2018.
- Papineni, Kishor et al. (2002, Juli): BLEU: a Method for Automatic Evaluation of Machine Translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)* (S. 311–318). Online verfügbar unter <http://aclweb.org/anthology/P/P02/P02-1040.pdf>, zuletzt geprüft am 22.06.2018.
- Strassmann, Burkhard (1998, 17. September): Mundgeschlecht im ovalen Büro. OHNE MENSCHLICHEN EINGRIFF NUN AUCH AUF DEUTSCH: DER STERRBERICHT. *taz. die tageszeitung*, S. 20. Online verfügbar unter <http://www.taz.de/!1324949/>, zuletzt geprüft am 22.06.2018.
- The Economist* (10. Juni 2006): How to build a Babel fish. Technology Quarterly Section. Online verfügbar unter <http://www.economist.com/node/7001819>, zuletzt geprüft am 22.06.2018.
- Vogel, Stephan (2016): *Machine Translation - Winds of Change*. Anniversary Symposium "Building Bridges, Breaking Barriers", Baden-Baden. Online verfügbar unter http://videolectures.net/interACT2016_vogel_machine_translation/, zuletzt geprüft am 22.06.2018.
- Winther Madsen, Matthias (2009, 23. Dezember): *The Limits of Machine Translation*. Masterarbeit, Universität Kopenhagen: Kopenhagen.
- Wu, Yonghui et al. (2016): *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation* (Computing Research Repository (CoRR), Hrsg.). Online verfügbar unter <http://arxiv.org/abs/1609.08144>, zuletzt geprüft am 22.06.2018.
- Yang, Jin E. L. (2003): Going live on the Internet. In Harold Somers (Hrsg.), *Computers and Translation. A translator's guide* (S. 194). John Benjamins Publishing.



Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.