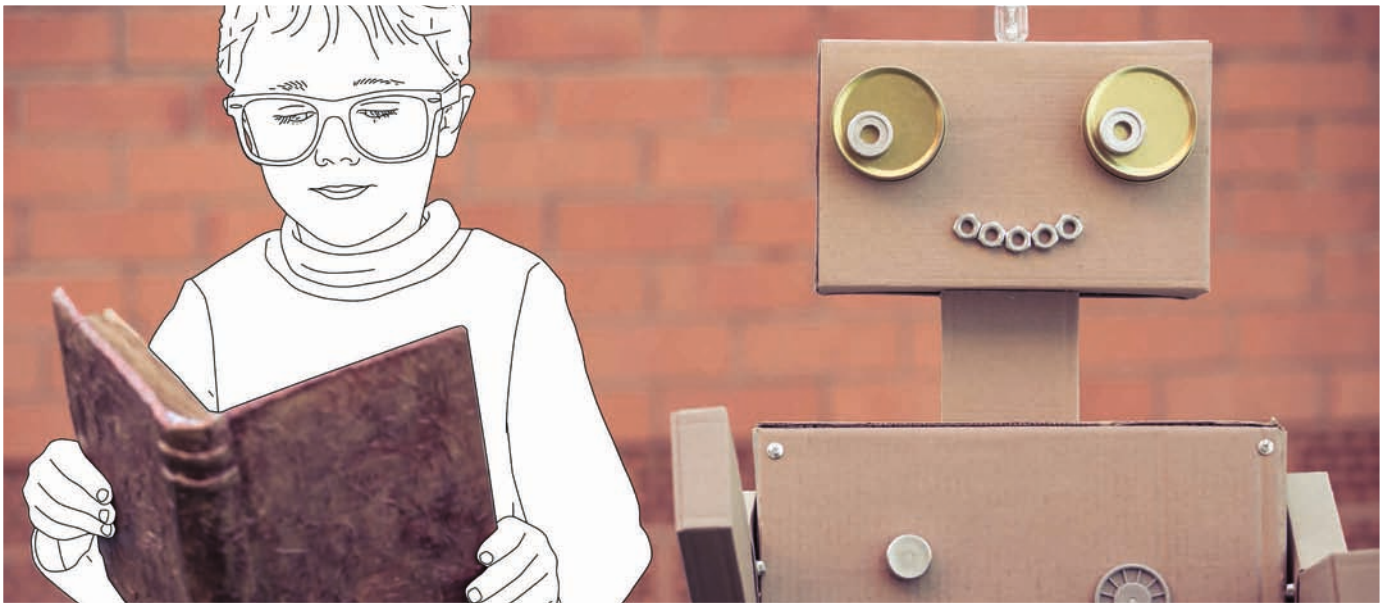

ILLUSTRATION *CSH / Adobe Stock*

Übersetzen mit Mensch und Maschine

Dank Künstlicher Intelligenz hat die Maschinelle Übersetzung einen enormen Produktivitätssprung getan. Wie kann es jetzt weitergehen, etwa mit Hilfe von Pre-Editing oder dem Trainieren mit eigenen Daten? Gefragt ist das Fachwissen der Technischen Redaktion.

TEXT *Rachel Herwartz*

Bei der neuronalen Maschinellen Übersetzung (NMÜ) wird für jedes Wort im Ausgangssatz („Eingabewort") eine Repräsentation in Form eines Zahlenvektors erzeugt. In der Maschinellen Übersetzung reichert der Encoder jedes Eingabewort mit Kontextinformationen aus dem ganzen Eingabesatz an. Der Decoder generiert die Ausgabewörter und baut Wort für Wort den Ausgabesatz auf. Die Wahrscheinlichkeit für das jeweils nächste Wort ergibt sich dabei aus den Wörtern des Eingabesatzes und allen bisher generierten Ausgabewörtern (ABB. 01, S. 46).

Der Erfolg dieser Methode liegt in der Nachbildung eines neuronalen Netzes. Im Gehirn werden die Nervenzellen (Neuronen) durch Synapsen miteinander verbunden und spannen dadurch „neuronale Netze" auf. Künstliche neuronale Netze (KNN) sind Programme, die im Training durch die Verknüpfung von Daten lernen, was die erwarteten Ausgaben für bestimmte Eingangswerte sind. Auf diese Weise können KI-Systeme eigenständig Verkehrsschilder auf einer Straße erkennen und Fußgänger von Autos unterscheiden. Sie können aber versagen, wenn ein LKW umkippt und quer auf der Straße liegt. So eine Situation entspricht nicht dem „gelernten" Idealbild. >

Prof. Dr. Rachel Herwartz berät bei der Gestaltung von Übersetzungs- und Terminologieprozessen. Sie ist Geschäftsführerin der TermSolutions GmbH, Anbieter von Terminologie-Software und Consulting. An der Internationalen Hochschule SDI München leitet sie den Studiengang „Translation Management" und ist für die Online-Zertifikatskurse „Post-Editing" und „Maschinelle Übersetzung" verantwortlich.
rachel.herwartz@termsolutions.de
www.term-solutions.com

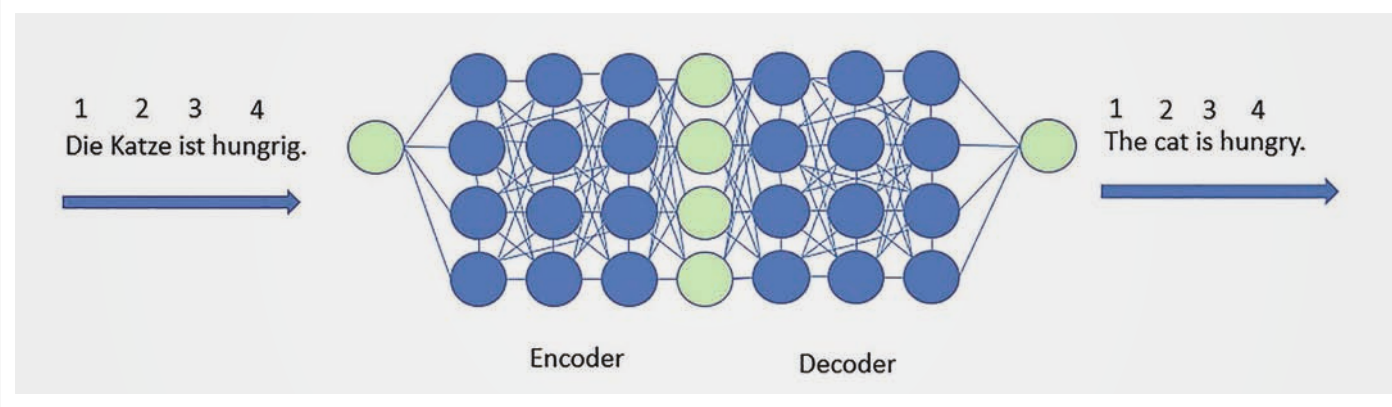

ABB. 01 Encoder-Decoder-Architektur.
QUELLE *Rachel Herwartz*

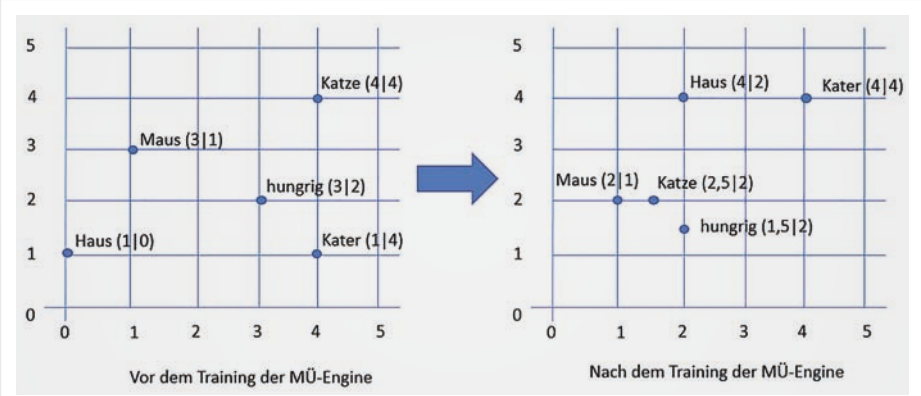

ABB. 02 Beispiel für Zahlenvektoren vor und nach dem Training.
QUELLE *Rachel Herwartz*

> Diese Funktionsweise liegt ebenfalls der Neuronalen Maschinellen Übersetzung zugrunde. Insofern haben auch Übersetzungen mittlerweile „etwas mit KI zu tun".

Training eines neuronalen Systems

Für das Training der „NMÜ-Engine", des KI-Systems, wird für jedes Wort zunächst willkürlich ein Zahlenvektor vergeben (ABB. 02, LINKS). Das Trainingsmaterial besteht in diesem Fall aus bilingualen Satzpaaren. Generische Systeme wie die von Google, Microsoft oder DeepL wurden dabei mit Millionen von Daten aus EU-Beständen trainiert. Mit jedem neuen Satzpaar, das in das Trainingsmaterial aufgenommen wird, „lernt" die Engine, welche Wörter immer wieder in ähnlichen Kontexten auftauchen. Diese Wörter nähern sich im Laufe des Trainings einander an und erhalten somit auch ähnliche Koordinaten (ABB. 02, RECHTS). Die Vorhersage des Systems basiert daher nicht auf grammatikalischen Regeln, sondern auf Häufigkeiten der Verteilung von Wörtern im Trainingsmaterial.

Umgang mit Homonymen

Im Beispiel in Abbildung 02, auf der rechten Seite, sind die drei Wörter „Maus", „Katze" und „hungrig" durch das Training bereits zusammengerückt. Allerdings steht das Wort „Kater" weit abseits, was daran liegt, dass „Kater" ein Homonym ist. Homonyme, das heißt Wörter mit mehreren Bedeutungen, von denen eine der Bedeutungen im Trainingsmaterial vorherrscht, werden von der NMÜ-Engine auch nach der häufig vorkommenden Bedeutung übersetzt. In diesem Fall scheint der „Kater" als „männliche Katze" seltener im Trainingsmaterial vorhanden zu sein als mit der häufig vorkommenden Bedeutung „Kater" im Sinne einer „schlechten körperlichen und seelischen Verfassung nach [über]reichlichem Alkoholgenuss" (Duden). Somit sind auch seine Koordinaten denen anderer Wörter wie „Bier", „Wein" und „trinken" ähnlicher.

Tatsächlich ist die zweite Bedeutung von „Kater" in den generischen NMÜ-Systemen von Google und Microsoft die vorherrschende. Daher wird es für das Englische mit „hangover", im Französischen mit „gueule du bois" übersetzt und bei Google sogar mit einem Häkchen hervorgehoben, dem Symbol für „trusted". DeepL hingegen schlägt für „Kater" mehrere Übersetzungen vor: cat, tomcat und hangover.

Wie schon gezeigt, wird in der Encoder-Decoder-Architektur in Abbildung 01 während der Übersetzung der Kontext beachtet. Hierfür werden in der Generierung des Zielsatzes alle Wörter des Ausgangssatzes und alle schon übersetzten Wörter des Zielsatzes berücksichtigt. So können die Sätze „Der Kater fängt die Maus" und „Der Kater war hungrig" im Laufe der Übersetzung entsprechend korrigiert werden (ABB. 03).

Das alles hat jedoch weniger mit Grammatik oder Semantik zu tun, sondern mehr mit Mathematik und Statistik. In der automatischen Vorschlagsfunktion von Google Übersetzer scheint dieser Mechanismus allerdings nicht zu greifen (ABB. 04). Man sollte sich daher nicht blind auf die Vorschläge (unten links) verlassen, sondern den Text tatsächlich komplett eingeben.

Missverständliche Personalpronomen

Die NMÜ scheitert auch an Personalpronomen, deren Zuordnung nur durch den Kontext klar wird. Ein Beispiel, das die Maschine im Gegensatz zum Menschen nicht auflösen kann, lautet: „Immer wenn Marie ihre Katze füttert, beißt sie sie". [1] Dass die Katze Marie beißt und nicht umgekehrt, basiert auf Weltwissen und wird vom Humanübersetzer in die Zielsprache übertragen.

Dass dieser Satz in der Maschinellen Übersetzung jedoch Probleme macht, zeigt sich schon bei einer Übersetzung ins Englische. Google und Microsoft übersetzen hier „Whenever Marie feeds her cat, she bites her". DeepL hingegen übersetzt in ein falsches „Whenever Marie feeds her cat, she bites it".

In den romanischen Sprachen erzeugen die Systeme sehr unterschiedliche Ergebnis-

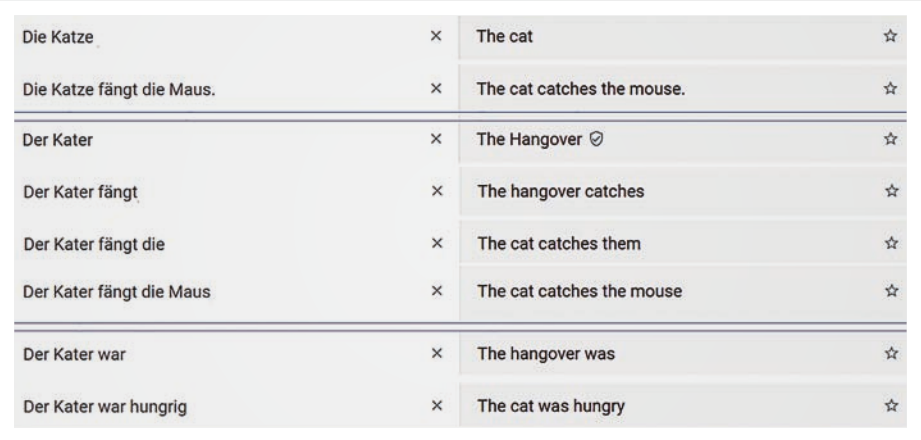

ABB. 03 Homonyme wie „der Kater" werden erst im Lauf der Übersetzung im Kontext erkannt.
QUELLE *Rachel Herwartz*

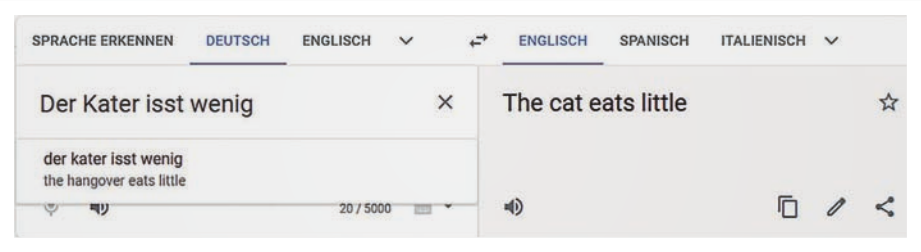

ABB. 04 Falsche Übersetzungsvorschläge in Google Übersetzer (links unten im Bild).
QUELLE *Rachel Herwartz*

se (ABB. 05, S. 48). In vielen Fällen wird tatsächlich eine Bedeutung „Marie beißt die Katze" erzeugt. Es ist jedoch kaum vorhersehbar, wann der Fehler auftritt, da diese drei Sprachen eigentlich ähnliche Ergebnisse erhalten sollten.

Wird im Satz „die Katze" durch „der Kater" ersetzt, werden die Bezüge auch für die Maschine klarer. Jetzt ist der deutsche Satz durch das unterschiedliche Geschlecht eindeutiger und wird korrekt übersetzt (ABB. 06, S. 48).

Allerdings zeigt sich in der französischen Übersetzung von Microsoft Übersetzer ein neuer Fehler: Der Kater wurde trotz des vorhandenen Kontextes „füttern" mit „gueule de bois" übersetzt. Das zeigt, dass eine Terminologiekontrolle gerade in der NMÜ unerlässlich ist. [2] Denn der Decoder kann bereits während der Übersetzung zwar Eingabewörter einbeziehen, die er für sinnvoll erachtet. Wie das Beispiel für Französisch zeigt, gibt es aber keine Garantie dafür, dass der Decoder bei Erreichen des Satzendes alle Encoder-Ausgaben berücksichtigt hat. Dadurch können sich ungewollte Hinzufügungen („overtranslation") oder Auslassungen („undertranslation") in den Zieltexten entwickeln. Es können auch „Halluzinationen" entstehen, bei denen Wörter mehrfach wiederholt werden. [3]

Vorteile von NMÜ

Texte, die mit NMÜ-Systemen übersetzt werden, sind flüssiger lesbar und von wesentlich höherer Qualität als solche der bisher vorherrschenden regelbasierten (RMÜ) oder statistischen Übersetzungssysteme (SMÜ). Die Übersetzungsqualität kann dabei nicht nur von einem Humanübersetzer gemessen werden, sondern auch mit Hilfe automatisch berechneter Scores wie BLEU dargestellt wird.

Bei BLEU wird eine maschinell erzeugte Übersetzung mit der Referenzübersetzung eines Humanübersetzers verglichen. Je weiter die Maschinelle Übersetzung von der Originalübersetzung entfernt ist, umso niedriger ist der dafür ausgegebene BLEU-Score. Allerdings ist zu bedenken, dass auch Humanübersetzungen nie tatsächlich zu 100 Prozent an diese erste Referenzübersetzung herankommen: Selbst ein Übersetzer, der seinen eigenen Text noch einmal Monate später übersetzt, wird keine hundertprozentige Übereinstimmung mit dem Ausgangstext mehr erreichen.

Verschiedene Studien zeigen anhand des BLEU-Scores eine signifikante Qualitätssteigerung der neuronalen gegenüber anderen Arten der Maschinellen Übersetzung. [4] Das trifft zum Beispiel für die Sprachkombinationen Deutsch/Englisch und Englisch/ >

Deutscher Satz *„Immer wenn Marie ihre Katze füttert, beißt sie sie."*	
Bedeutung A „Die Katze beißt Marie"	**Bedeutung B „Marie beißt die Katze"**
Google und DeepL Italienisch *Ogni volta che Marie nutre il suo gatto, la morde.*	Microsoft Italienisch *Ogni volta che Marie dà da mangiare al suo gatto, lo morde.*
Google und Microsoft Spanisch *Siempre/Cada vez que Marie alimenta a su gato, la muerde.*	DeepL Spanisch *Cada vez que Marie da de comer a su gato, lo muerde.*
	Google, Microsoft, DeepL Französisch *Chaque fois que Marie nourrit son chat, elle le mord.*

ABB. 05 „Immer wenn Marie ihre Katze füttert, beißt sie sie." (IT/SP/FR)
QUELLE *Rachel Herwartz*

Deutscher Satz *„Immer wenn Marie ihren Kater füttert, beißt er sie."*
Bedeutung A „Der Kater beißt Marie"
Google, Microsoft und DeepL Italienisch *Ogni volta che Marie dà da mangiare al suo gatto/nutre il suo gatto (lui) la morde.*
Google, Microsoft und DeepL Spanisch *Siempre/Cada vez que Marie alimenta a su gato, él/éste la muerde.*
Google und DeepL Französisch *Chaque fois que Marie nourrit son chat, il la mord.* Microsoft Französisch *Chaque fois que Marie nourrit sa gueule du bois, il la mord.*

ABB. 06 „Immer wenn Marie ihren Kater füttert, beißt er sie." (IT/SP/FR)
QUELLE *Rachel Herwartz*

Beispiele für Formulierungsmuster

DE: Bitte Motor ausbauen. EN: Please remove engine.
DE: Bitte bauen Sie den Motor aus. EN: Please remove the engine.
DE: Motor ausbauen. EN: Remove engine.

INF. 01 QUELLE *Rachel Herwartz*

> Deutsch sowie Übersetzungen aus dem Chinesischen ins Englische zu.

In einer Bewertung durch Humanübersetzer ließen sich für NMÜ in der Sprachrichtung Englisch-Deutsch 50 Prozent weniger Wortstellungsfehler, 17 Prozent weniger lexikalische Fehler, 19 Prozent weniger morphologische Fehler und insgesamt 26 Prozent weniger Post-Editing-Aufwand als für RMÜ feststellen. [5]

Auch Sprachdienstleister, die schon seit längerem Maschinelle Übersetzung verwenden, berichten ebenfalls von einer enormen Qualitätssteigerung beim Umstieg von SMÜ auf NMÜ. So können NMÜ-Elemente häufiger komplett übernommen werden, und es gibt insgesamt weniger Nachbearbeitungsbedarf. Gerade die Übersetzungen vom Englischen ins Deutsche und umgekehrt haben für einige Experten mittlerweile eine so hohe Qualität erreicht, dass es für die Ausbildung von Übersetzern „sinnvoll sein könnte, einen Großteil der reinen Übersetzungsübungen durch Post-Editing-Kurse zu ersetzen". [6]

Die Qualität der MÜ kann aber meist noch weiter verbessert werden. Das gelingt durch Training mit eigenem Material und durch ein vorangehendes Pre-Editing.

Prüfen und anpassen

Parallel zum „Post-Editing" (PE), in DIN ISO 18578 definiert als Nachbearbeitung eines ausschließlich maschinell übersetzten Textes, bezeichnet „Pre-Editing" eine Überarbeitung des Ausgangstextes mit dem Ziel, ein besseres Ergebnis in der Maschinellen Übersetzung zu erzielen. Laut Norm fällt dies in den Aufgabenbereich des Übersetzers bzw. Post-Editors, obwohl er sich häufig nicht zur Überarbeitung des Ausgangstextes berufen fühlt. Wenn hingegen Technische Redakteure ihre Texte unter Beachtung der geltenden Rechtschreibung, mit korrekter Formatierung und nach den Regeln für übersetzungsgerechtes Schreiben verfassen, sind diese Arbeitsschritte vom Pre-Editing nicht mehr weit entfernt.

Bereits seit langem wird darüber geforscht, wie sich Kontrollierte Sprache auf die Qualität Maschineller Übersetzung auswirkt. [7] Für die regelbasierte MÜ wurden bereits 1996 große Erfolge erzielt: „Die Integration des Pre-Editing auf der Basis des SDD [Siemens Dokumentationsdeutsch, R.H.] in den Übersetzungsprozess mit dem firmeninternen MÜ-System ermöglichte es Siemens, sich weitgehend das Post-Editing zu ersparen." [8]

In einer Studie zum Pre-Editing in der statistischen MÜ konnten 2017 über 90 Pro-

ABB. 07 Arbeitsschritte in (Re-)Training und Customization. QUELLE *Rachel Herwartz*

zent des Ausgangstextes so bearbeitet werden, dass der Text mit Hilfe eines statistischen Maschinellen Übersetzungssystems in einer ausreichend hohen Qualität maschinell übersetzt werden konnte. [9] Diese Erfolge konnten für die NMÜ leider bisher nicht eindeutig nachgewiesen werden.

Marzouk [10] hat auf der Basis von neun ausgewählten Regeln aus der tekom-Leitlinie 2013 [11] und einem Korpus von zehn deutschen Benutzerhandbüchern festgestellt, dass der BLEU-Score für die NMÜ mit 83 Prozent doppelt so hoch ist wie der von SMÜ und RMÜ. Dieses gute Ergebnis trifft allerdings auf die präeditierten sowie die nicht präeditierten Texte zu.

Folgende Regeln der tekom-Leitlinie wurden für die Analyse ausgewählt:

- Oberflächentexte mit Anführungszeichen auszeichnen
- Funktionsverbgefüge vermeiden
- Konditionalsätze mit „Wenn" einleiten
- eindeutige pronominale Bezüge herstellen
- Partizipialkonstruktionen und Passiv vermeiden
- Konstruktionen mit „sein" und „zu" vermeiden
- überflüssige Präfixe streichen
- keine Wortteile weglassen

Bei einer Bewertung der Übersetzungen durch Humanübersetzer wurde oft kritisiert, dass der durch kontrollierte Sprache optimierte Text in der Ausgangssprache in der Übersetzung dann „nicht mehr so schön klingt". [12] Diese Bewertung trifft nicht nur auf den übersetzten Text, sondern auch bereits auf die kontrollierte Sprache im Ausgangstext zu.

Der Übersetzungsdienstleister Dr. François Massion betont, dass zur Erkennung von Schwierigkeiten, die sich beim Einsatz von MÜ ergeben könnten, Merkmale wie die Satzlänge in Wörtern, die Wortlänge pro Satz, die Interpunktion, die Anzahl der Verben im Satz sowie die Komplexität der Satzstruktur und Mehrdeutigkeit zählen. [13] Dies führt beim Übersetzer dann nicht zum Pre-Editing des Ausgangstextes, sondern zu einer Bewertung des Übersetzungsauftrags in die drei Kategorien:

- → für MÜ nicht geeignet
- → für PE geeignet
- → braucht keine MÜ

Die Experten Miyata und Fujita haben 2021 ihre Studie zum Pre-Editing von 2017 noch einmal wiederholt, diesmal für neuronale Maschinelle Übersetzung. Galt vorher immer der Grundsatz, dass für Maschinelle Übersetzung die Sätze möglichst kurz und einfach sein sollten, so stellten sie jetzt jedoch für die NMÜ fest: „Rather, it is more important to make the content, syntactic relations, and word senses clearer and more explicit, even if the ST becomes longer." [14]

Domäne oder generisches Training

Einen zweiten Ansatzpunkt für die Verbesserung des MÜ-Ergebnisses bietet die Verwendung von spezifischerem Trainingsmaterial. Trotz der „Blackbox" NMÜ kann bereits im Training versucht werden, die Verbindungen zwischen den Wörtern so zu beeinflussen, dass bestimmte Verbindungswege gegenüber anderen bevorzugt werden. Dies gilt zum Beispiel für Beeinflussungen im Sinn von „bevorzuge nur Informationen aus dem Sachgebiet X", „aus der Textsorte Y" oder „für die Zielgruppe Z". Abhilfe versprechen „domänenspezifisch" trainierte Engines, etwa im Übersetzungssystem Matecat, wo für jeden Übersetzungsauftrag derzeit zwischen 36 verschiedenen „Subjects" gewählt werden kann.

Mit eigenen Daten

Andere Anbieter stellen es ihren Kunden frei, die generische Engine zu verwenden oder ein mit eigenen Daten erstelltes bzw. angereichertes System zu nutzen. Bei Google AutoML wird für jede Übersetzung das Ergebnis der generischen Engine und das der selbsttrainierten Engine angezeigt. Eine aktuelle Übersicht über die gängigen Systeme bietet eine Studie, die jährlich aktualisiert wird. [15]

Dabei ist zu unterscheiden zwischen einem Training mit einem komplett leeren System und einem „Re-Training" bzw. der „Customization" von generischen Systemen. Da man für Ersteres mindestens eine Million Satzpaare pro Sprachrichtung und pro Fachgebiet benötigen würde, ist – außer für Behörden und Großkonzerne – meist nur die Anpassung durch das zusätzliche Einspeisen eigener Daten realistisch. [16] So kann man schon mit 5.000 bis 15.000 Satzpaaren pro Sprachrichtung gute Ergebnisse erzielen. Mehr eigene Daten können bei den Systemen manchmal die MÜ-Qualität verschlechtern. Auch eine Priorisierung von Trainingspaaren ist möglich, zum Beispiel durch eine Vorauswahl von „guten" Segmenten, die dann höher gewichtet werden als die anderen.

Gutes Trainingsmaterial im Sinn von bilingualen Satzpaaren ist vor allem in Orthografie und Interpunktion fehlerfrei, ohne auseinandergerissene Satzfragmente und enthält keine überflüssigen Tags oder Formatierungen. Dies ist in den Translation Memories der Unternehmen nicht immer gegeben und muss daher vor dem Training bereinigt werden (ABB. 07).

Standardisierung im Trainingsmaterial

Häufig werden fehlende Metadata-Attribute wie „User Manual" und „Marketing Material" in den Translation Memories beanstandet. Manchmal handelt es sich aber auch um unstrukturierte Translation Memories, die sich über die Jahre angesammelt haben: „A linguistic data management is still missing at both client and vendor side." [17]

Allerdings kann man bei den Anforderungen noch einen Schritt weitergehen. Selbst wenn anhand von Metadaten eine Textsorte wie „Technische Dokumentationen" identifiziert und damit eine Engine trainiert wird, ist zu beachten, dass die Sätze einheitliche Formulierungsmuster zum Beispiel in der Anrede der Kunden aufweisen (INF. 01). Man sollte auch keine Translation Memories verwenden, in denen britisches und amerikanisches Englisch vermischt ist. Denn sonst „lernt" die Engine im Training diese Varianten und gibt dann in der Übersetzung willkürlich die eine oder die andere Variante aus.

Anreicherung des Trainingsmaterials

Wenn nicht genügend eigenes bilinguales Material vorhanden ist, ist es möglich, durch >

Kontrolliertes Vokabular in den Prozessschritten

Im Trainingsmaterial der MÜ-Engine	Als Pre-Editing im Ausgangstext	Während des laufenden MÜ-Prozesses	Im Post-Editing	In der QS
Ausgangssprache: keine Homonyme, Synonyme vermeiden[1]	Keine Homonyme verwenden, Synonyme vermeiden	Meist nur bilinguale Glossare mit zwei Spalten einbindbar, daher zielgruppenspezifischer Export aus der Terminologiedatenbank	Termerkennung aus einer Termdatenbank	Voraussetzung: Präskriptive Terminologiearbeit
Zielsprache: Homonyme möglich, Synonyme vermeiden[1]	Einsatz einer Terminologieprüfung im Ausgangstext	Einfaches nachträgliches Suchen und Ersetzen führt oft zu Grammatikfehlern	Argumentation für oder gegen Benennungen in TDB ablegen	Einsatz einer Terminologieprüfung im TMS oder der MÜ
Idealerweise Abkürzungen und Eigennamen kennzeichnen	In Zukunft: Eigennamen und Abkürzungen kennzeichnen	Einbindungsversuche wie Provokation werden nicht von allen Systemen unterstützt	Gründe für Terminologiefehler als Feedback an Sprachtechnologen	Feedback an die Sprachtechnologen im TMS oder im MÜ-System

[1] Eventuell sind Inkonsistenzen im Ausgangstext gut zur Identifikation von Synonymen, wenn die Äquivalente im Zieltext dann eindeutig sind. Allerdings kann das Material so nicht für das Training im Ausgangstext in beiden Sprachrichtungen verwendet werden. [18]

TAB. 01 QUELLE *Rachel Herwartz*

> eine maschinelle Rückübersetzung bzw. die Generierung weiterer Satzpaare durch Platzhalter und Duplizierung von Sätzen, zum Beispiel mit Zahlen oder Datumsangaben, das Trainingsmaterial anzureichern. Allerdings kann man auch domänenspezifische Sprachpaare käuflich erwerben oder Open-Source-Sammlungen frei für das Training verwenden.

Für seltene Sprachkombinationen empfiehlt es sich in manchen Fällen, über eine Pivot-Übersetzung bzw. Relaissprache zunächst ins Englische und dann vom Englischen in die Zielsprache zu übersetzen. Bei ähnlich aufgebauten Sprachen kann man auch mit mehrsprachigen Satzpaaren arbeiten. Zunächst wird nur ein Sprachpaar wie zum Beispiel Deutsch – Spanisch trainiert. Daran „hängen" weitere romanische Sprachen wie Italienisch, Französisch, Portugiesisch und Rumänisch.

Terminologie einbinden

Die Einbindung der Terminologie muss in allen Prozessschritten einer Maschinellen Übersetzung erfolgen: im Trainingsmaterial der MÜ-Engine, im Pre-Editing des Ausgangstextes, während des laufenden MÜ-Prozesses, im Post-Editing durch den Humanübersetzer und zuletzt in der abschließenden Qualitätssicherung (TAB. 01).

Neue Aufgabengebiete

Zur Wahl eines Tools für die Maschinelle Übersetzung gibt es bis heute keine „One-size-fits-all"-Anwendung. Vielmehr handelt es sich um Lösungen, die für jede Sprachrichtung, jede Textsorte, jedes Sachgebiet (domain) und ggf. auch für jede Zielgruppe passen. [19] Die Anreicherung mit sinnvollen Metadaten, die Standardisierung von Formulierungsmustern und die Terminologiekontrolle fällt dabei für die Ausgangssprache in das Aufgabengebiet der Technischen Redaktion.

Auch die übersetzungsgerechte Erstellung eines Ausgangstextes ist für die Technische Redaktion nichts Neues. Da alle Ausgangstexte und ihre Übersetzungen als Trainingsmaterial für maschinelle Übersetzungssysteme dienen können, wirkt sich die Qualität der Ausgangstexte auch auf die Qualität der Maschinellen Übersetzungen und den Post-Editing-Aufwand aus.

Wie bereits bei Einführung der mittlerweile etablierten Translation-Memory-Systeme gibt es durch die Maschinelle Übersetzung in der Abstimmung zwischen Technischer Redaktion und Übersetzung zahlreiche neue Herausforderungen. Sie lassen sich nur gemeinsam meistern.

LINKS UND LITERATUR ZUM BEITRAG

[1] Burchard, Aljoscha/Porsiel, Jörg: Vorwort (2017): *Was kann maschinelle Übersetzung und was nicht. In: Porsiel, Jörg (Hrsg.), S. 16.*

[2] Herwartz, Rachel (2021): *Neuerung mit Folgen. In: technische kommunikation. H. 4, S. 42-47.*

[3] Van Genabith, Josef: (2020): *Neural Machine Translation. In: Porsiel, Jörg (Hrsg.): Maschinelle Übersetzung für Übersetzungsprofis. Sammelband. S. 109.*

[4] Vashee; Kirti (2017): *Neural Machine Translation; A Practicioner's Viewpoint. In: Porsiel, Jörg (Hrsg.): Maschinelle Übersetzung. Grundlagen für den professionellen Einsatz, Sammelband. S. 44–58.*

[5] Bentivogli, Luisa et al. (2016): *Neural versus Phrase-Based Machine Translation Quality: A Case Study. In: Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing.*

[6] Hansen-Schirra, Silvia/Schaeffer, Moritz/Nitzke, Jean (2017): *Post-Editing: Strategien, Qualität, Effizienz. In: Porsiel, Jörg (Hrsg.), S. 188.*

[7] Lehrndorfer, Anne (1996): *Kontrollierte Sprache für Technische Dokumentation – Ein Ansatz für das Deutsch. In: Krings, Hans-Peter (Hrsg.): Wissenschaftliche Grundlagen der technischen Kommunikation;* und Lehrndorfer, Anne (2001): *Kontrolliertes Deutsch. Linguistische und sprachpsychologische Leitlinien für eine (maschinell) kontrollierte Sprache in der Technischen Dokumentation.*

[8] Marzouk, Shaimaa/Hansen-Schirra, Silvia (2020): *Kontrollierte Sprache im Zeitalter der neuronalen Maschinellen Übersetzung. In: Porsiel, Jörg (Hrsg.), S. 237.*

[9] Miyata, Rei/Fujita, Atsushi (2017): *Dissecting human pre-editing toward better use of off-the-shelf machine translation systems, in: Proceedings of the 20th Annual Conference of the European Association for Machine Translation (EAMT), S. 54–59.*

[10] Marzouk, Shaimaa/Hansen-Schirra, Silvia (2020): *Kontrollierte Sprache im Zeitalter der neuronalen Maschinellen Übersetzung. In: Porsiel, Jörg (Hrsg.), S. 234–256.*

[11] tekom-Leitlinie (2013): *Regelbasiertes Schreiben. Deutsch für die Technische Kommunikation. 2. Auflage;* und tekom-Leitlinie (2014): *Regelbasiertes Schreiben – Englisch für deutschsprachige Autoren. bzw. Rule-Based Writing – English for Non-Native Writers.*

[12] Marzouk/Hansen-Schirra (2020): *S. 248.*

[13] Massion, Francois (2020): *NMT im Einsatz bei einem Dienstleister: von der Systemauswahl bis zum fertigen MÜ-Workflow. In: Porsiel, Jörg (Hrsg.), S. 267.*

[14] Miyata, Rei/Fujita, Atsushi (2021): *Understanding Pre-Editing for Black-Box Neural Machine Translation. In: Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics, S. 1539–1550, Zitat S. 1547.*

[15] Nimdzi (2021): *Nimdzi Language Technology Atlas: the Definitive Guide to the Language Technology Landscape https://www.nimdzi.com/language-technology-atlas/ [zuletzt aufgerufen am 23.8.2021]*

[16] Winter, Tom/Zielinski, Daniel (2020): *Terminologie in der neuronalen maschinellen Übersetzung. In: Porsiel, Jörg (Hrsg., 2020), S. 221f.*

[17] Lieske, Christian (2020): *Metadata and Machine Translation. In: Porsiel, Jörg (Hrsg.). S. 152.*

[18] Winter/Zielinski (2020): *S. 225.*

[19] Burchard/Porsiel (2020): *S. 9.*