

Vom Wörterbuch zur modern

en KI: Übersetzung im Wandel



© CoxinhaFotos auf Pixabay

Die Arbeit des Übersetzers wurde in den letzten Jahrzehnten maßgeblich durch den Einsatz elektronischer Datenverarbeitung (EDV) verändert. Von den ersten regelbasierten maschinellen Übersetzungssystemen bis hin zu modernen KI-gestützten Lösungen hat sich die Technologie kontinuierlich weiterentwickelt. Doch trotz dieser Fortschritte hat der menschliche Übersetzer aufgrund der nach wie vor bestehenden Einschränkungen der Maschinen weiterhin wichtige Aufgaben.

Die Geschichte des Übersetzens ist genauso alt wie die Menschheit selbst. Seit Menschen aus verschiedenen Kulturen miteinander in Kontakt treten, besteht die Notwendigkeit, Gedanken, Ideen und Konzepte in andere Sprachen zu übertragen. Über Jahrtausende hinweg war dies eine Tätigkeit, die allein auf dem Wissen, der Erfahrung und der Sprachbeherrschung von Übersetzern basierte. Doch in der zweiten Hälfte des 20. Jahrhunderts, nach dem Aufkommen von Computern und elektronischer Datenverarbeitung (EDV), begann sich auch die Welt des

Übersetzens zu verändern. Dieser Wandel verlief alles andere als geradlinig, er war eher durch zahlreiche Herausforderungen und auch durch Rückschläge geprägt.

Übersetzen ohne EDV-Unterstützung: Die Zeit vor 1960

Bis in die 1960er Jahre hinein war Übersetzen eine Aufgabe, die vollständig auf menschlichen Fähigkeiten beruhte. Übersetzer saßen an ihren Schreibtischen, umge-

ben von Wörterbüchern, Fachliteratur und Notizbüchern, und arbeiteten sich Zeile für Zeile durch die Texte.

Besonders in Bereichen wie Recht, Technik oder Medizin, in denen Präzision und spezifische Terminologie entscheidend sind, waren Übersetzer darauf angewiesen, umfangreiche eigene Recherchen durchzuführen, um sicherzustellen, dass die von ihnen verwendeten Begriffe korrekt und verständlich waren.

In dieser Ära erschien die Idee, Maschinen könnten eines Tages Übersetzungsarbeit unterstützen, als reine Utopie.

Zwar wurden erste Übersetzungsprogramme entwickelt, doch sie waren in ihrer Kapazität und Leistungsfähigkeit sehr begrenzt. Es schien vor allem den Fachübersetzern unmöglich, dass die Computer eines Tages in der Lage sein würden, mit der Komplexität menschlicher Sprache umzugehen.

Die Anfänge der maschinellen Übersetzung: Regelbasierte Systeme und ihre Grenzen [1]

Trotz der technischen und sprachlichen Limitierungen begannen Wissenschaftler in den 1950er und 1960er Jahren, sich mit der Idee der maschinellen Übersetzung auseinanderzusetzen. [2] Diese frühen Systeme basierten auf der Idee, dass Sprache durch festgelegte grammatische und lexikalische Regeln in eine andere Sprache übertragen werden könne. Diese „regelbasierten Systeme“ (Rule-Based Machine Translation, RBMT) versuchten, menschliche Übersetzungen nach festen Mustern zu replizieren. [3] Man glaubte, dass es möglich sei, eine Sprache in einer Art „syntaktischen Struktur“ zu fassen, die dann einfach in die syntaktische Struktur einer anderen Sprache zu übertragen sei.

Das berühmte Georgetown-IBM-Experiment [4] im Januar 1954 ist eines der bekanntesten Beispiele für die frühe Begeisterung für maschinelle Übersetzung. Der Wortschatz war jedoch auf 250 Lexeme beschränkt und es konnten nur einfache Sätze übersetzt werden. Allerdings gingen einige Wissenschaftler davon aus, dass Maschinen innerhalb weniger Jahre in der Lage sein würden, jede Sprache präzise zu übersetzen.

Diese Hoffnung erfüllte sich nicht. Die regelbasierten Systeme waren sehr limitiert, da sie kaum in der Lage waren, die Nuancen, Ambiguitäten und die Vieldeutigkeit natürlicher Sprache zu verstehen. Die Ergebnisse dieser frühen MT-Systeme waren daher oft kaum brauchbar.

Fehler in Grammatik, Wortwahl und Syntax führten häufig zu unverständlichen oder ungenauen Übersetzungen

Die 1970er Jahre: Terminologie-Datenbanken als Unterstützung für Übersetzer

Trotz der enttäuschenden Ergebnisse der ersten MT-Versuche schritten die Entwicklungen im Bereich der EDV und der Sprachverarbeitung weiter voran. In den 1970er Jahren begann der Vormarsch der Terminologie-Datenbanken. [5] Diese spezialisierten Datenbanken wurden entwickelt, um Übersetzern bei der Arbeit mit technischen und spezialisierten Texten zu helfen.

Die Idee dahinter war, dass in vielen Fachbereichen bestimmte Begriffe immer wieder verwendet werden und dass es sinnvoll wäre, diese Begriffe in einer Datenbank zu speichern, auf die der Übersetzer schnell und einfach zugreifen konnte. Diese frühen Terminologie-Datenbanken erleichterten die Arbeit von Übersetzern erheblich, da sie ihnen ermöglichten, präzise und konsistente Übersetzungen von Fachbegriffen zu verwenden, ohne jedes Mal nach der richtigen Übersetzung suchen zu müssen. Obwohl diese Systeme noch weit davon entfernt waren, automatische Übersetzungen zu realisieren, stellten sie einen wichtigen Schritt in Richtung eines EDV-unterstützten Übersetzungsprozesses dar.

Die frühen 1990er Jahre: Translation-Memory-Systeme und die Revolution der Übersetzungsbranche

Einer der bedeutendsten Fortschritte in der Geschichte des EDV-Einsatzes beim Übersetzen erfolgte in den frühen 1990er Jahren mit der Einführung von Translation-Memory-Programmen. [6] Diese Programme revolutionierten die Übersetzungsbranche, indem sie es ermöglichten, bereits übersetzte Sätze oder Textsegmente zu speichern und bei Bedarf wiederzuverwenden.

Das Prinzip der Translation Memories (TM) ist einfach: Wenn ein Übersetzer einen Text übersetzt, wird jeder Satz oder jedes Textsegment in einer Datenbank gespeichert. Sobald der Übersetzer später auf ein gleiches oder ähnliches Segment stößt, schlägt das Programm das zuvor Gespeicherte vor. Diese Technologie hat viele Vorteile:

■ Erstens ermöglicht sie eine erhebliche Zeitersparnis, da Übersetzer nicht jedes Mal denselben Satz neu übersetzen müssen.

Zweitens sorgt sie für Konsistenz in der Übersetzung, da gleiche oder ähnliche Sätze immer auf dieselbe Weise übersetzt werden.

Diese Programme erwiesen sich als besonders nützlich in Bereichen wie Recht und Technik, in denen viele Sätze und Phrasen standardisiert sind und immer wieder verwendet werden. Übersetzer konnten mit TM-Programmen ihre Effizienz steigern und gleichzeitig die Qualität ihrer Übersetzungen verbessern. Es war jedoch wichtig, dass menschliche Übersetzer die vorgeschlagenen Übersetzungen überprüften und gegebenenfalls anpassten, da die Systeme noch nicht in der Lage waren, die komplexen Kontexte und Bedeutungsnuancen vollständig zu verstehen und es oft nötig war, bestimmte Sätze in einem Ausgangstext aufgrund des Kontextes unterschiedlich zu übersetzen. So kann der Satz: „Er hat einen Vogel“, je nach Kontext übersetzt werden als: „He has a bird.“, oder als: „He is crazy.“ Auch die fortschrittlichsten TM-Systeme waren also kein Ersatz für Humanübersetzer. Sie waren jedoch zweifelsohne ein wichtiges Werkzeug für die Verbesserung der Produktivität von Humanübersetzern. [7]

Die 2010er-Jahre: Statistische maschinelle Übersetzung und der Aufstieg der KI

Während die regelbasierten MT-Systeme in den 1960er und 1970er Jahren gescheitert waren, gab es in den 2010er Jahren einen neuen Ansatz, der die maschinelle Übersetzung erheblich verbesserte: die statistische maschinelle Übersetzung (Statistical Machine Translation, SMT) [8] Anstatt sich auf festgelegte Regeln zu verlassen, analysierten diese Systeme riesige Mengen an zweisprachigen Texten und nutzten statistische Modelle, um Muster und Wahrscheinlichkeiten zu erkennen. [9]

Dank statistischer Modelle konnten Systeme Übersetzungen erzeugen, die die Leistung regelbasierter Systeme deutlich übertrafen, weil erstmals auch der Kontext der Ausgangstexte berücksichtigt werden konnte. [10]

Der Übergang von der regelbasierten zur statistischen maschinellen Übersetzung bedeutete, dass Maschinen nun in der Lage waren, komplexere Sprachmuster zu erkennen und Übersetzungen zu erzeugen, die flüssiger und natürlicher klangen. Dennoch waren diese Systeme nicht fehlerfrei. Da sie auf Wahrscheinlichkeiten basierten, konnte es vorkommen, dass die Maschinen falsche oder unpassende Übersetzungen lieferten, insbesondere wenn die Ausgangstexte mehrdeutig oder gar fehlerhaft waren.

Diese Fortschritte führten jedoch zu einer zunehmenden Integration von MT-Systemen in den Übersetzungsprozess. Professionelle Übersetzer begannen, MT-Systeme als ergänzendes Werkzeug zu nutzen, um ihre Arbeit zu beschleunigen. Sie überprüften die maschinellen

Übersetzungen und korrigierten Fehler, was zu einem hybriden Modell führte, in dem Mensch und Maschine zusammenarbeiteten. [11]

Die 2020er Jahre: Künstliche Intelligenz und Large Language Models

In den zurückliegenden Jahren hat sich die maschinelle Übersetzung erneut revolutioniert – diesmal durch die Einführung von Künstlicher Intelligenz (KI) und so genannten Large Language Models (LLMs). [12] Diese Modelle basieren auf tiefen neuronalen Netzen und sind in der Lage, riesige Mengen an Texten zu analysieren und auf dieser Grundlage Übersetzungen zu generieren.

Im Gegensatz zu früheren MT-Systemen, die auf Regeln oder statistischen Modellen basierten, nutzen LLMs eine Form des maschinellen Lernens, bei dem sie durch die Analyse großer Mengen an Texten lernen, Muster in der Sprache zu erkennen. Diese Systeme sind in der Lage, sehr flüssige und natürlich klingende Übersetzungen zu erzeugen und können sogar in gewissem Maße den Kontext eines Textes verstehen.

Trotz dieser beeindruckenden Fortschritte haben auch die modernsten KI-Systeme ihre Grenzen. Während LLMs in der Lage sind, Wörter und Phrasen auf eine sehr komplexe Weise zu kombinieren, können sie nicht wirklich „denken“ oder den tiefen Sinn eines Textes erfassen. Sie basieren auf mathematischen Modellen, die Muster erkennen, aber keine logischen Schlussfolgerungen ziehen können. Dies führt dazu, dass auch in modernen maschinellen Übersetzungen immer noch inhaltliche Fehler auftreten.

Besonders problematisch sind dabei immer noch Situationen, in denen der Ausgangstext Fehler enthält. Auch LLMs sind (noch) nicht in der Lage, diese Fehler zu erkennen oder zu korrigieren. Sie übernehmen die Fehler oft in die Übersetzung, was zu ungenauen oder falschen Ergebnissen führen kann. Fachübersetzer hingegen haben die Fähigkeit, Fehler im Ausgangstext zu identifizieren und den Ausgangstext zu hinterfragen.

Die Rolle des menschlichen Übersetzers in einer KI-getriebenen Welt

Obwohl die Technologie im Bereich der maschinellen Übersetzung beeindruckende Fortschritte gemacht hat, bleibt der menschliche Übersetzer unverzichtbar. KI-Systeme und LLMs sind hervorragende Werkzeuge, um den Übersetzungsprozess zu beschleunigen und zu unterstützen, doch sie sind nicht in der Lage, die Nuancen, die logischen Zusammenhänge und die kulturellen Feinheiten eines Textes vollständig zu verstehen.

Der Fachübersetzer spielt daher nach wie vor eine zentrale Rolle, insbesondere bei der Übersetzung komple-

xer Fachtexte. Er überprüft die maschinellen Übersetzungen, identifiziert und korrigiert Fehler und sorgt dafür, dass die Übersetzung nicht nur sprachlich korrekt, sondern auch inhaltlich stimmig ist.

Das alles gilt erst recht für Texte, die einen starken Kulturbefug haben. Übersetzungen der Harry-Potter-Bücher haben jedenfalls, unabhängig von der Zielsprache und unabhängig vom Einsatz von MT-Systemen, zu intensiven Diskussionen über das erlaubte Maß der übersetzerischen Freiheiten geführt. [13]

Fazit: Die Zukunft der Übersetzungstechnologie

Die Geschichte des EDV-Einsatzes beim Übersetzen ist eine Geschichte des ständigen Fortschritts und der Innovation. Von den frühen regelbasierten MT-Systemen

über Translation-Memory-Programme bis hin zu KI-gestützten Übersetzungstools hat sich die Technologie stark entwickelt. Dennoch bleibt die maschinelle Übersetzung eine Ergänzung, kein Ersatz für den menschlichen Übersetzer. MT wird zweifelsohne dazu beitragen, dass die Kosten für Übersetzungen noch weiter sinken werden, bei einer gleichzeitigen Verbesserung der Qualität.

In den kommenden Jahren wird sich die Technologie weiterentwickeln, und es ist wahrscheinlich, dass die maschinellen Übersetzungen noch besser und präziser werden. Übersetzer werden diese Werkzeuge zu nutzen wissen und sie werden sie nutzen müssen, um Sprachdienstleistungen weiter wettbewerbsfähig anbieten zu können.

Das tiefgehende Verständnis von Sprache und Kultur bleibt aber auch künftig ein Bereich, in dem der Mensch den Maschinen überlegen ist.

Literatur:

- [1] Weaver, Warren. (1955). Translation. In Machine Translation of Languages* (pp. 15-23). MIT Press. • [2] Hutchins, W. John, & Somers, Harold L. (1992). An Introduction to Machine Translation* Academic Press. • [3] Melby, Alan K. (1982). The Possibility of Language: A Discussion of the Nature of Language, with Implications for Human and Machine Translation. John Benjamins Publishing Company. • [4] Hutchins, W.J. (2004). The Georgetown-IBM Experiment Demonstrated in January 1954. In: Frederking, R.E., Taylor, K.B. (eds) Machine Translation: From Real Users to Research. AMTA 2004. Lecture Notes in Computer Science(), vol 3265. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30194-3_12 • [4] Sager, Juan C. (1990). A Practical Course in Terminology Processing. John Benjamins Publishing Company. • [6] Baker, Mona. (1992). In Other Words: A Coursebook on Translation. Routledge. • [7] Garcia, Ignacio. (2009). Beyond Translation Memory: Computers and the Professional Translator. Journal of Specialized Translation, (12), 199-214. • [8] Hutchins, W. John. (1995). The History of Machine Translation in a Nutshell. In Computers and Translation: A Translator's Guide (pp. 431-445). John Benjamins Publishing Company. • [9] Koehn, Philipp. (2010). Statistical Machine Translation. Cambridge University Press. • [10] Och, Franz Josef, & Ney, Hermann. (2004). The Alignment Template Approach to Statistical Machine Translation. Computational Linguistics, 30(4), 417-449. • [11] Schäfer, Björn, & Schulz, Anne-Kathrin. (2010). Statistical Machine Translation: Challenges, Opportunities, and Applications. John Benjamins Publishing Company. • [12] Vaswani, Ashish et al. (2017). Attention is All You Need* Advances in Neural Information Processing Systems. • [13] <https://read-eat.de/harry-potter-und-der-stein-des-anstosses-ist-harry-potter-schlecht-uebersetzt/>



Der Autor:



Dr.-Ing. Wolfgang Sturz hat sich als Ingenieur und Gründer des Sprachdienstleisters Transline einen Ruf als Querdenker der Übersetzungsbranche erworben. Übersetzen – ein humanistisches Kunstwerk? Oder ein prozessgesteuerter, IT-gestützter Produktionsablauf? Letzteres war immer sein persönliches Steckbrief: maximale Effizienz durch den gezielten Einsatz von Informationstechnologien – stets unter den wachsamen Augen eines qualifizierten Teams aus Fachübersetzern und Projektmanagern. Es war daher nur folgerichtig, dass er seit den Anfängen seines Unternehmens einen starken Fokus auf das Wissensmanagement gelegt hat. Eine Leidenschaft, aus der vor nun genau 25 Jahren das Magazin „wissensmanagement“ entstanden ist.

✉ sturz@wissensmanagement.net