

P R E S S E - I N F O R M A T I O N

Wirtschaft / Technik / IT / 06-12-19

KI: Zu Risiken und Nebenwirkungen fragen Sie Ihren zertifizierten Übersetzungsdienst- leister

Open-Source-Tools gefährden Datenschutz / Wie erkennt man zertifizierte Übersetzungsunternehmen? / Lohnt sich die Implementierung von eigenen Übersetzungsmaschinen? / PwC: Geschickte Kombination von Mensch und Maschine verspricht größtmöglichen Erfolg

Offenbach am Main.- Die gesamte Belegschaft erwartete das Ergebnis mit Hochspannung: Wer würde gewinnen? Ihre eigens entwickelte Maschine oder Google Translate? Die Fashion-Webseite, die es zu übersetzen galt, bestand zu gleichen Teilen aus Marketingsprache und juristischen Fachbegriffen, wie z.B. Versandbedingungen oder Datenschutz. Die Auflösung ergab, dass beide Maschinen im Vergleich zu menschlichen Übersetzer*innen deutliche Schwächen zeigten. Die firmeneigene Maschine bearbeitete jedoch präziser die juristischen Termini. Für den internationalen Sprachdienstleister EVS Translations, der das Experiment in seiner bulgarischen Niederlassung durchführte, bestätigte das Ergebnis eine lang gehegte Vermutung: Das Translation-Memory-System macht den Unterschied. Jeder Übersetzungsauftrag füttert seit Jahrzehnten eine interne Datenbank mit Trainingsdaten, die nicht nur verwaltet und gepflegt, sondern auch – im Gegensatz zu herkömmlichen Systemen – zusätzlich unterschiedlich gewichtet werden können. Dies fehlt bei Google Translate.

Woran es der öffentlich zugänglichen Maschine mangelt, sind Qualitätskontrolle und Datenschutz. Denn welcher CEO eines DAX-Konzerns käme auf die Idee, Arbeitsverträge, Compliance-Richtlinien oder Unterlagen eines Rechtsstreits einem Netzwerk anzuvertrauen, das für seinen fragwürdigen Umgang mit sensiblen Daten

bekannt ist? Ein Verstoß gegen Geheimhaltungspflichten kann schnell einen lang eingefädelt Deal zunichtemachen, wenn Informationen unkontrolliert verloren gehen. Was folgt, ist ein Imageverlust für das ganze Unternehmen, der auch den Markenwert beeinträchtigen kann.

Doch wie lassen sich die Vorteile einer qualitativ hochwertigen, aber zugleich kostengünstigen Übersetzung mittels maschineller Hilfe und einer möglichst sicheren Handhabung von „Datengold“ realisieren? Die gute Nachricht: Es gibt Lösungen. Eine davon trägt den Namen ISO 18587.

I. Datengold gehört in den Tresor

Automatisierte Übersetzungsmethoden erscheinen verlockend einfach. Doch der Preis ist, dass Millionen von Nutzern die online verfügbaren Übersetzungsmaschinen 24 Stunden am Tag mit Unmengen ihrer eigenen Daten füttern. Vertrauliches kann versehentlich gespeichert und öffentlich zugänglich gemacht werden. Demzufolge liegt das Augenmerk für Übersetzungsunternehmen im diskreten Handling von Kundeninformationen, deren gewissenhafter Speicherung sowie Schutz vor Eingriffen von außen.

Dies bestätigt auch die KI-Studie 2019 der Beratungsfirma Deloitte, die weltweit KI-Verantwortliche befragt hat, darunter 100 Entscheider aus deutschen Unternehmen. Gerade hinsichtlich Cyber-Risiken bestehen Bedenken, etwa was den Diebstahl sensibler Daten oder Algorithmen anbelangt. Dieser Aspekt der Studie deckt sich mit den Ergebnissen für die ausländischen Vergleichsmärkte. Als gravierendste Herausforderungen im Rahmen von KI-Initiativen nennen die befragten deutschen Firmen an erster Stelle datenbezogene Probleme (z.B. bezüglich Datenschutz, Datenqualität, Bereinigung, Integration, maschinelle Interpretation etc.).

Diese Entwicklung berücksichtigen seriöse Sprachdienstleister oft seit Jahrzehnten und hüten das „Datengold“ ihrer Kunden wie einen Schatz im Tresor. Unabhängige Institute prüfen regelmäßig interne Verfahrensprozesse nach objektiven Qualitätsstandards und vergeben Gütesiegel an die Übersetzer. Doch meist finden sich die Kunden im Dickicht der diversen Zertifizierungen nicht zurecht.

Die ISO-Norm 27001 bedeutet für die Übersetzungsindustrie, dass alle Daten besonders gesichert werden müssen. Der weltweit tätige Sprachdienstleister EVS Translations z.B. sichert durch sein externes Rechenzentrum ein Servicelevel von 24/7/365 zu, vollständige Videoüberwachung sowie einen bewaffneten Sicherheitsdienst, eine dreistufige Authentifizierung, eine umfassende Zertifizierung nach ISO 27001 und durch TÜV Rheinland CAT III+ (99,982% Verfügbarkeit) sowie Verfügbarkeitsklasse 3 nach DIN EN 50600. Für Unternehmenskunden gibt es separate Mailserver mit individuellen Verschlüsselungsoptionen für Sender und Empfänger – maßgeschneiderte Lösungen inbegriffen.

II. Outsourcen oder selbst in KI investieren?

In der KI-Studie von Deloitte heißt es: „Nur 15 Prozent der Studienteilnehmer implementieren KI primär mit betriebseigenen Kräften“. Künstliche Intelligenz wird mehrheitlich eingekauft: Der Anteil der Unternehmen, die KI „as a Service“ einsetzen, liegt laut Studie bei 65 Prozent.

Dieser Aussage können Übersetzungsunternehmen nur zustimmen: Sie haben vielfach beobachtet, dass sich die Bedürfnisse der Kunden geändert haben. Gefragt sind Rundum-Sorglos-Pakete aus einer Hand, weniger einzelne Services. Doch wer oft externe Dienstleister bezahlt, fragt sich unweigerlich irgendwann, ob nicht eine eigene Investition in neuronale Netze (Neural Machine Translation – NMT) Sinn macht. Edward Vick, Geschäftsführer von EVS Translations, gibt zu bedenken: „Es kann sein, dass eine interne Übersetzungsabteilung z.B. eine Mio. Euro Kosten verursacht, das Outsourcing jedoch nur mit 750.000 Euro zu Buche schlägt. Jeder einzelne Übersetzungsfehler kostet in seiner Bearbeitung Geld, Zeit und Ressourcen – Einsatz, den eine vierköpfige Abteilung möglicherweise nicht so gut bewältigen kann wie ein Expertenteam von 200 Menschen. Effektiver ist es da, sich an eine bestehende Übersetzungsplattform anzudocken, die einen höheren Durchsatz und damit einhergehend eine höhere Qualität bietet. Die Frage ist immer: Was genau sucht man?“

Bevor ein Unternehmen jedoch den Schritt eines eigenen KI-Systems in Betracht zieht, ist es wichtig, drei bestimmte Aspekte zu berücksichtigen:

- Ressourcen
- Lernkurve
- Spezialisierung

#Ressourcen

Die Implementierung erfordert eine erhebliche Anfangsinvestition in Hardware und neuronale Netze, ebenso in nachfolgende Schulungs- und Wartungskosten, die für den Betrieb dieser Technologie benötigt werden. Neben den physischen Anforderungen, wie z. B. Raum- und Grafik-Mikroprozessoren, wird NMT auch erhebliche Investitionen in Zeit, Daten und Training der künstlichen Intelligenz erfordern, bevor Ergebnisse erzielt werden können. Auch wenn das System dann Ergebnisse liefert, bedeutet dies nicht, dass es die richtigen für das Unternehmen sind. Wie bei allem anderen und insbesondere bei NMT gibt es eine Lernkurve. KI kann nur so gut sein, wie sie programmiert wurde.

#Lernkurve

Der Fokus liegt auf angewandter KI, in der intelligente Maschinen eine spezielle Aufgabe ausführen können und nur nach gründlichem Lernen und Training mit großen Datenmengen mithilfe von Algorithmen simulieren, was bei jeder Kombination von Eingaben passieren würde. Werte berechnen die effektivste Ausgabe basierend auf den verfügbaren Daten. Mit anderen Worten, die Intelligenz von KI hängt von der Menge der Daten ab, die sie verarbeitet. D.h. dass jedes Mal, wenn ein Unternehmen eine neue Sprache hinzufügen muss, ein neues, spezialisiertes Training erforderlich ist. Schwierig wird es, wenn je nach Sprache mal mehr, mal weniger öffentlich zugängliche Trainingsdaten zur Verfügung stehen.

#Spezialisierung

Aufgrund der Fehlerhaftigkeit maschinengestützter Übersetzungen steigt mit zunehmendem Einsatz von NMT unweigerlich die Nachfrage nach Posteditoren, Korrekturlesern und Übersetzern. Formatierung, Stil und Genauigkeit sowie alle marken- und branchenbezogenen Anforderungen müssen präzise erfüllt sein, ebenso spezifische Terminologie und Stilrichtlinien. Bedeutende organisatorische Änderungen müssen dafür eingeführt und umgesetzt werden, damit intelligente Ma-

schinen und Menschen effektiv zusammenarbeiten. Zusätzlich gibt es zahlreiche ungeklärte Gesetze, wenn es um ethische Entscheidungen geht, die von Maschinen getroffen werden.

III. Der beste Trainer heißt „Big Data“

Da der Zugriff auf Daten der Hauptfaktor für die effektive Einführung der KI-Technologie ist, sind das Datenmanagement und die Technologie mit hohen Investitionskosten in Sachen Datenschutz die größten Herausforderungen. Der Geschäftsführer von EVS Translations, Edward Vick, erläutert: „Die Rolle von Daten in der angewandten KI ist weitaus wichtiger als in Big Data- und Analyse-Applikationen. Um Geschäftswert zu generieren, müssen KI-Anwendungen auf große Mengen unternehmensspezifischer Daten geschult werden. Seit fast 30 Jahren übersetzen wir täglich über 50.000 Sätze in 163 Sprachkombinationen. Wir haben Big Data im Haus.“ Dies ist nur möglich, weil alle Unternehmensprozesse und -abläufe in einem hochmodernen Datenmanagementsystem automatisiert sind und in terminologiespezifischen Clustern strukturiert.

Damit die Qualität der maschinell erzeugten Übersetzung nicht leidet, kam 2017 die ISO-Norm 18587 neu hinzu. Sie legt die Anforderungen fest, die Unternehmen befolgen müssen, wenn sie maschinelle Übersetzungen und deren Nachbearbeitung anbieten. Hierbei wird unterschieden zwischen einer Full-Service-Nachbearbeitung, die einen mit einer menschlichen Übersetzung vergleichbaren Text erstellen soll, und einer leichten Nachbearbeitung, die lediglich darauf abzielt, den maschinell übersetzten Text verständlich zu machen. Alle Vor-Ort-Prozesse werden von einem externen, unabhängigen Auditor der docConsult GmbH auf die Einhaltung der ISO-Norm 18587 geprüft und die Erfüllung ihrer Anforderungen bestätigt.

IV. Gemeinsam geht's besser

Um den Übersetzungs- und Lokalisierungsbedarf internationaler Unternehmen zu decken und dringende, großvolumige Übersetzungsaufträge entgegenzunehmen und abzuwickeln, ist unter Umständen eine maschinelle Übersetzung erforderlich. Zum einen können Dokumente, für die es bisher weder Zeit noch Budget gab, auf diese Weise übersetzt werden. Zum anderen bleiben Sprachdienstleister wettbewerbsfähig, wenn sie

ihre Dienstleistungen erweitern und sie an die wachsende Kundennachfrage nach maschinellen Übersetzungslösungen anpassen.

Kein statistisches oder neuronales maschinelles Übersetzungssystem kann so hochwertige Fachübersetzungen liefern wie eine Person. Daher kombiniert ein durchdachtes hybrides Modell beide positiven Effekte: qualitativ hochwertige (Mensch) und kostengünstige (Maschine) Übersetzungen. Laut der Wirtschaftsprüfungsgesellschaft PricewaterhouseCoopers AG bringt erst die intelligente Verknüpfung entscheidende Wettbewerbsvorteile und damit Erfolg: Ökonomische Effizienz, menschlicher Faktor, maschinelle Unterstützung. Wenn das keine gute Nachricht für Übersetzer ist!

V.i.S.d.P.: EVS Translations GmbH
Luisenstr. 3
63067 Offenbach am Main
% Annika Krauth-Wentritt
Tel. 0160/944 63 813
E-Mail: annika.krauth@web.de
Internet: www.evs-translations.com/de

Maschinenübersetzung: Das 1x1 der Schwachstellen

Auch bei weiterentwickelten Übersetzungstechniken wie der Nutzung neuronaler Netze (Neural Machine Translation - NMT) spielt der Blick auf das Qualitätsniveau des „Endproduktes“ eine wichtige Rolle. Neuronale Netzwerke imitieren die Arbeitsweise des menschlichen Gehirns und dessen Fähigkeit, Informationen zu verstehen und daraus zu lernen. Google oder Microsoft verfügen mittlerweile über attraktive Modelle, Dokumente ohne menschliche Leistung zu übersetzen und entstandene Fehler als weitere Trainingsquelle für NMT zu nutzen. Klarer Vorteil: Übersetzungen sind innerhalb kürzester Zeit für wenig Geld zu haben. War die Sprachbranche seit je her von starkem Preisdruck geprägt, hat sich dieser Aspekt durch den Einsatz von NMT noch verschärft.

600.000 Linguisten und 21.000 Sprachdienstleister weltweit sehen diese Entwicklung mit gemischten Gefühlen: Verlieren sie nun ihren Arbeitsplatz an Roboter? Der Geschäftsführer von EVS Translations, Edward

Vick, steht neuen Technologien grundsätzlich aufgeschlossen gegenüber, stellt aber zugleich klar: „Im Zentrum der digitalen Strategie steht nach wie vor der Mensch. Er bedient und programmiert Computer und steuert Prozesse. Auch wenn wir neuronale Netzwerke in verschiedenen Stadien einsetzen, prüfen wir immer genau, ob der Kunde einen konkreten Nutzen davon hat, z.B. wieviel Zeit oder Geld er sparen kann.“ Daher hält es Vick für verfrüht, vor der Übermacht der Maschine zu kapitulieren. Auch zeichnen sich bislang noch keine Lösungen für unausgegorene „Kinderkrankheiten“ ab.

1 – Zwischen den Zeilen lesen

Einer Übersetzungsmaschine fehlt Textverständnis. Während ein Humanübersetzer alle rhetorischen Feinheiten und sinnhaften Unterschiede eines Textes gekonnt aufspürt und überträgt, kann eine Übersetzungsmaschine nur stupide das übersetzen, was konkret geschrieben steht. Unklare Aussagen, schwammige Wort-hülsen oder ausschweifende Formulierungen werden von maschinellen Übersetzungssystemen meist schlecht oder sinnfrei übersetzt.

2 – Uneinheitliche Begriffswahl

Zwar übersetzen maschinelle Systeme zunehmend konsistenter, verwenden Fachterminologie also einheitlicher und mit aufwändigen menschlichen Eingriffen. Im Vorfeld kann auch bei Maschinenübersetzungen ein weitestgehend konsistenter Sprachgebrauch gewährleistet werden. Eine 100%ige Sicherheit hat man jedoch nur mit einem Humanübersetzer, der Konsistenz auch dann erzeugt, wenn der Ausgangstext selbst nicht einheitlich ist.

3 – Schwerwiegende Sinnfehler

Das größte Problem bei der neuronalen maschinellen Übersetzung (NMÜ) ist, dass der Text überraschend gut klingt – zu gut. Früher konnte man fast jeden maschinell übersetzten Satz daran erkennen, dass 1) die Formulierungen holprig oder unvollständig waren und 2) der Text im Ganzen nicht fließend lesbar war. Dieses Problem ist mittlerweile kaum noch zu finden. Stattdessen lassen sich maschinell übersetzte Texte heutzutage wunderbar lesen und nur bei genauem Hinschauen oder im Vergleich mit dem Ausgangstext lässt sich feststellen, dass der Zieltext zwar angenehm klingt, aber gefährlich weit vom Ausgangstext entfernt ist. Die Konsequenzen können weitreichend sein: Im schlimmsten Fall bedrohen

fehlerhafte Übersetzungen z.B. im Medizin- und
Pharma-Bereich Leben.