



Maschinelle Übersetzung bei Volkswagen

Sprache als betriebswirtschaftlicher Faktor

Jörg Porsiel

Über das Volkswagen Sprachenportal können die Mitarbeiter des Volkswagen Konzerns E-Mails und Berichte maschinell übersetzen lassen. Um diese Funktion richtig nutzen zu können, ist Aufklärungsarbeit erforderlich. Dieser Artikel erscheint im Dezember 2008 außerdem in englischer Sprache in der Zeitschrift Multilingual.

Seit Sommer 2002 bietet das Intranet der Volkswagen AG das innovative „Volkswagen Sprachenportal“ an. Es bündelt zentral und für alle Konzernmitarbeiter zugänglich eine Vielzahl von Informationen zum Thema Sprache im Allgemeinen (u. a. bzgl. der neuen deutschen Rechtschreibung mit Links zu entsprechenden Websites) sowie Übersetzen und Dolmetschen im Besonderen. Darüber hinaus bietet es Zugriff auf eine umfangreiche, mehrsprachige, unternehmensspezifische Terminologiedatenbank mit zurzeit ca. 15 000 Einträgen und auf maschinelle Übersetzung (nachfolgend auch als MT bezeichnet) in den Sprachen Deutsch, Englisch, Französisch, Russisch und Spanisch. Das Portal und seine Inhalte sind in all diesen Sprachen verfügbar.

Ausgangssituation und Rahmenbedingungen

Der Volkswagen Konzern, der nahezu 330 000 Mitarbeiter zählt (davon etwas über 50 % in Deutschland), liefert seine Autos in mehr als 150 Länder weltweit. An 48 Produktionsstandorten werden Fahrzeuge der Marken Volkswagen, Volkswagen Nutzfahrzeuge, Audi, Seat, Škoda, Bentley, Lamborghini und Bugatti gefertigt. 2007 wurden knapp 6,2 Mil-

lionen Fahrzeuge verkauft und ein Umsatz von 108,9 Milliarden Euro generiert. Das VW-interne Fremdsprachenmanagement in Wolfsburg wickelt jährlich ca. 30 000 Übersetzungsaufträge in bis zu 40 Sprachen und ca. 1 200 Dolmetschereinsätze ab.

Ausschlaggebend für die Bereitstellung des Sprachenportals (siehe Abb. 1), der Terminologiedatenbank und der maschinellen Übersetzung waren mehrere Faktoren: Zum einen die wachsende Notwendigkeit im Zeitalter der Internationalisierung und Globalisierung, Texterstellungs- und Übersetzungsprozesse qualitativ zu optimieren und zu beschleunigen, indem u. a. das Bewusstsein für die (betriebswirtschaftliche) Bedeutung einer einheitlichen Unternehmenssprache geschärft sowie umfassendes Terminologie-management im Bereich der technischen Dokumentation weiter vorangetrieben wurde. Diese Ziele sollten durch die Be-



reitstellung einer möglichst unkomplizierten Anwendung zur Bewältigung des in den vergangenen Jahren extrem gestiegenen E-Mail-Aufkommens sowohl inner- als auch außerhalb des Konzerns erreicht werden. Darüber hinaus musste ebenfalls der starke Anstieg des Textvolumens der technischen Dokumentation aufgrund der Erweiterung der Fahrzeugmodellpalette sowie der Erschließung neuer Märkte bewältigt werden. Weltweit werden täglich mehrere Milliarden E-Mails versandt (Spam-Mails nicht mitgerechnet). Im Durchschnitt erhält ein Mitarbeiter eines in den For-

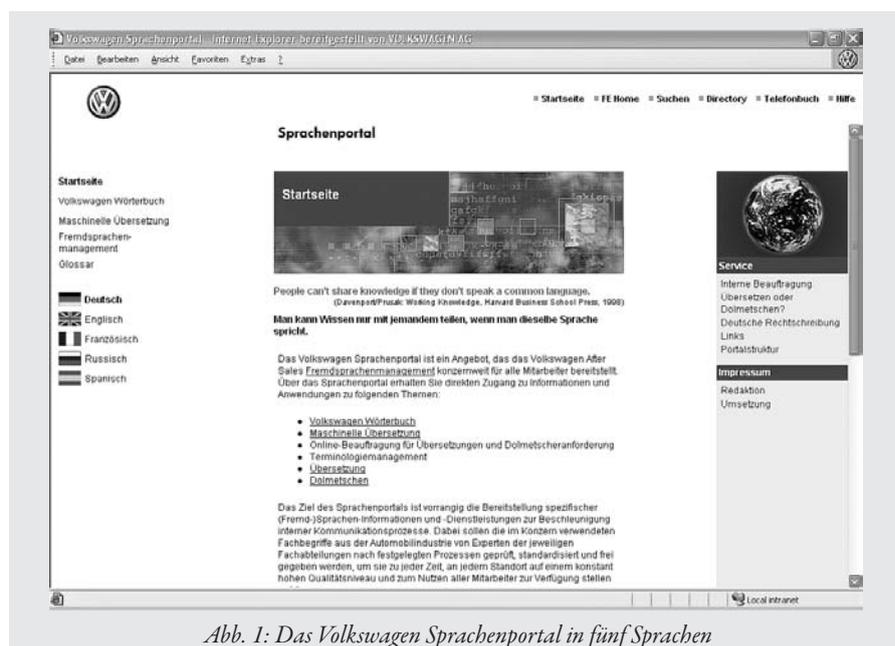


Abb. 1: Das Volkswagen Sprachenportal in fünf Sprachen



tune 500 gelisteten Unternehmens, wozu auch Volkswagen gehört, pro Tag 80 E-Mails. Schätzungen zufolge können 70 % der Mitarbeiter eines Großunternehmens in Deutschland täglich 60 Minuten ihrer Arbeitszeit allein durch unverständliche E-Mails (dazu gehören auch fremdsprachige) verlieren. Zudem führt die steigende Zahl der Mails nicht zwangsläufig zu einer Verbesserung der Kommunikation bzw. des Informationsaustausches. Ineffizienter Informationsaustausch durch unverständliche E-Mails kann auch darin bestehen, dass fremdsprachige E-Mails aufgrund unzureichender Sprachkenntnisse nicht bzw. missverstanden oder wegen der Wartezeit für eine notwendige Humanübersetzung (zu) spät beantwortet werden. Hier sorgt MT für erhebliche Prozessbeschleunigung: In Sekundenschnelle können Textmengen von einer Zeile bis hin zu mehreren Hundert Seiten übersetzt werden – wobei die Übersetzungsqualität zunächst von untergeordneter Bedeutung ist. Zeitgewinn ist das vorrangige Ziel beim Einsatz maschineller Übersetzung!

Ziele des Volkswagen Sprachenportals

Das Sprachenportal verfolgt mit seinem Angebot mehrere Ziele unterschiedlicher Tragweite. Neben der weiteren Verbesserung von Sprachqualität und Geschwindigkeit in der Kommunikation des Unternehmens sollen die im Portal verfügbaren Inhalte nachhaltig bewussteinbildend in Bezug auf das Thema Sprache im Allgemeinen und der Bedeutung einer einheitlichen und eindeutigen Unternehmenssprache im Besonderen wirken.

Betriebswirtschaftlicher Nutzen

Das Sprachenportal und dort vor allem MT wird von den Mitarbeitern gegenwärtig zur beschleunigten Abarbeitung täglicher Routineaufgaben, wie z. B. der Bearbeitung/Übersetzung von E-Mails, Sitzungsprotokollen oder Berichten eingesetzt. Dabei geht es um die Bereitstellung so genannter *gist* oder *indicative translations*, also von Rohübersetzungen. Obwohl die Qualität dieser Rohübersetzungen in Abhängigkeit von der Qualität der Ausgangstexte (GIGO-Prinzip, s. dazu unten: Warum maschinelle Über-

setzung „so schlecht“ ist) zum Teil stark schwanken kann, reicht den Anwendern das Übersetzungsergebnis in der Regel aus, um innerhalb von Sekunden und in Verbindung mit dem eigenen Vor- und Fachwissen eine Entscheidung treffen zu können, wie weiter zu verfahren ist. Aus der Konstellation von Informationsbedarf, -beschaffung und -verarbeitung – in Verbindung mit dem enormen Zeitgewinn durch MT – resultiert der primäre Nutzen für das Unternehmen. In Zeiten immer kürzer werdender Produktzyklen (time-to-market) und fortschreitender Globalisierung ist die beschleunigte Informationsbeschaffung, -aufbereitung und „Veredelung“ zu Wissen von wesentlicher Bedeutung für die Sicherung von Wettbewerbsvorteilen. Insbesondere bei weltweit über alle Zeitzonen hinweg operierenden Unternehmen ist der Faktor Zeit entscheidend. Der Einsatz von MT bei Routineaufgaben, wie dem Lesen, Beantworten oder Weiterleiten von E-Mails o. Ä., führt zu einer bisher nicht gekannten Beschleunigung fremdsprachiger Kommunikations- und Entscheidungsfindungsprozesse, gleichzeitig verbunden mit einer erheblichen und nachhaltigen Senkung von Bearbeitungs- und Übersetzungskosten.

Maschinelle Übersetzung im Einsatz

Die im Sprachenportal eingesetzte maschinelle Übersetzung wurde nach umfangreichen Tests verschiedener Produkte

ausgewählt. Ausdrücklich war (und ist) dabei nicht beabsichtigt, Humanübersetzer durch MT zu ersetzen oder MT in einem produktiven Umfeld zu nutzen, d. h. für die Übersetzung von Texten einzusetzen, die bisher von Humanübersetzern bearbeitet wurden. Vielmehr wird MT fast ausschließlich für die Übersetzung von Texten verwendet, die der Anwender bisher selbst, z. B. mit Hilfe eines handelsüblichen Wörterbuchs oder ähnlicher Hilfsmittel, und recht zeitaufwändig übersetzen musste. Für alle angebotenen Sprachpaare wird ein so genanntes „regelbasiertes System“ (rule-based machine translation [RBMT]) der Firma Lucy Software and Services GmbH eingesetzt. In den vergangenen Jahren verstärkt auf den Markt drängende statistische (statistical machine translation [SMT]) oder Hybrid-Systeme, also eine Mischung aus SMT und RBMT, waren Mitte 2002, dem Testzeitpunkt bei Volkswagen, in den Sprachen Deutsch, Englisch und Spanisch entweder noch nicht marktreif oder überhaupt nicht verfügbar, konnten folglich auch nicht geprüft werden.

Kennzahlen

Das Sprachenportal und damit die maschinelle Übersetzung sind konzern-, d. h. weltweit (auch auf Dienstreisen), erreichbar. Nach einer forcierten Erweiterung des Sprachenangebotes um Russisch (Februar 2008) und Französisch (Juli 2008) sowie



Abb. 2: Maschinelle Übersetzung im Sprachenportal



der Integration von über 40 000 VW-spezifischen Begriffen (d. h. hauptsächlich aus Kfz-Technik und Maschinenbau) werden derzeit durchschnittlich 7 000 Übersetzungen pro Arbeitstag maschinell verarbeitet, wobei „eine Übersetzung“ sowohl lediglich ein einziges Wort (s. dazu unten unter: Warum maschinelle Übersetzung „so schlecht“ ist) als auch Texte mit Umfängen von mehr als 100 DIN-A4-Seiten umfassen kann. Insgesamt bedeutet dies umgerechnet ein Übersetzungsvolumen von ca. 1 500 Standardseiten pro Arbeitstag, wobei – das sei nochmals betont – die Qualität zunächst keine Rolle spielt. Hier geht es vorrangig und ausschließlich um Geschwindigkeit, also Zeitgewinn. In den vergangenen zwei Jahren wurde das Sprachenportal inhaltlich stark erweitert, u. a. durch den Ausbau der dort zugänglichen Terminologiedatenbank, die nun in 20 Sprachen (mit ca. 15 000 deutschen Benennungen) zur Verfügung steht und ausschließlich konzerntypische Begriffe enthält. Darüber hinaus wurde die maschinelle Übersetzung um die Sprachen Französisch und Russisch ergänzt sowie die VW-Terminologie in den nunmehr fünf einsetzbaren Sprachen integriert, um die Übersetzungsergebnisse zu verbessern. Die Inhalte des Sprachenportals sind vollständig in den derzeit fünf Portalsprachen lokalisiert.

„Die Grenzen meiner Sprache sind die Grenzen meiner [Arbeits-]Welt.“
(frei nach Ludwig Wittgenstein:
Tractatus Logico-Philosophicus)

Was maschinelle Übersetzung kann

Soll maschinelle Übersetzung (siehe Abb. 2) sinnvoll eingesetzt werden, muss man sich vorher darüber im Klaren sein, wo deren Verwendung nützlich ist und was MT nicht zu leisten vermag und vor allem warum. Das bedeutet für den Portalbetreiber, in diesem Sinne bewusstseinsbildend tätig werden zu müssen. Der Einsatz von MT lohnt unter folgenden Voraussetzungen: Es muss spezifische Unternehmensterminologie in möglichst großem Umfang und von möglichst guter Qualität sowohl in der jeweiligen Ausgangs- wie auch in der Zielsprache vorhanden sein. Jeder terminologische Begriff sollte eindeutig und

darüber hinaus einem Sachgebiet zugeordnet sein. Idealerweise sollte es eine Schnittstelle zwischen Terminologiedatenbank und maschineller Übersetzung geben, oder es sollten doch wenigstens regelmäßige Updates neuer Terminologie durchgeführt werden, damit die Qualität der MT-Ergebnisse kontinuierlich und spürbar besser wird. Neben der erheblichen Zeitersparnis führt dies zu einer Steigerung der Akzeptanz seitens der Nutzer, die maschineller Übersetzung im Allgemeinen anfänglich eher skeptisch gegenüberstehen.

Den Anwendern wiederum muss das Vorhandensein einer standardisierten Unternehmensterminologie vermittelt werden, wobei der übergeordnete betriebswirtschaftliche, aber auch der persönliche Nutzen dazustellen sind. Des Weiteren müssen die Nutzer dazu angehalten werden, diese Terminologie auch tatsächlich für ihre tägliche Arbeit zu verwenden. In Bereichen wie der technischen Dokumentation wäre es darüber hinaus überaus hilfreich, maschinelle Lektoratsysteme einzusetzen, um die Effizienz zu erhöhen, (Übersetzungs-)Kosten zu senken und das Sprachbewusstsein nachhaltig zu fördern.

Die Qualität der Übersetzungsergebnisse hängt bei MT – weitaus mehr noch als bei der Humanübersetzung – von der Qualität der Ausgangstexte ab. Hier schlägt das „garbage-in-garbage-out“-Prinzip (GIGO) voll durch.

Warum maschinelle Übersetzung „so schlecht“ ist

Gelegentlich beschwerten sich Anwender, dass die Ergebnisse der maschinellen Übersetzung „so schlecht“ seien. Besonders zwei „Beweise“ werden dafür angeführt: 1. Die Übersetzung einzelner Begriffe „funktioniert nicht“. „Ich habe xy eingegeben, aber keine Übersetzung erhalten.“ Variante: „... aber eine völlig falsche Übersetzung erhalten.“ 2. Einige Nutzer sind der Meinung, das ideale Messmittel für Übersetzungsqualität gefunden zu haben: nämlich die „Rückübersetzung“.

Sie unterliegen aber hier mangels Sprachbewusstsein dem Trugschluss, Sprache sei in Aufbau und Funktionsweise mit Mathematik gleichzusetzen: Analog zur Gegenprobe in der Mathematik wird der ge-



Jörg Porsiel

Diplom-Übersetzer. Studium in Brüssel, Edinburgh, Metz und Heidelberg (Abschluss). Seit 1990 im Bereich Übersetzung, Terminologiemanagement und (fremdsprachige) Unternehmenskommunikation tätig. Beim Fremdsprachenmanagement der Volkswagen AG seit 2002 für das Sprachenportal und maschinelle Übersetzung verantwortlich.

*FS-ZM für VST-1/3
Fremdsprachenmanagement,
Volkswagen AG, Brieffach 1943,
38436 Wolfsburg*

extern.joerg.porsiel@volkswagen.de

rade übersetzte Text rückübersetzt. Ist das Ergebnis nicht vollkommen oder wenigstens doch zu 90 % identisch mit dem Ursprungstext, handelt es sich nach Ansicht dieser Nutzer automatisch um eine „schlechte Übersetzung“. Beide „Beweise“ führen drastisch vor Augen, wo die Defizite auf Anwenderseite liegen bzw. wie viel Arbeit noch in Bezug auf das Verständnis von Sprache im Allgemeinen und (maschineller) Übersetzung im Besonderen weiterhin zu leisten ist.

Es ist festzustellen, dass eine der Hauptfehlerquellen bei der maschinellen Übersetzung (neben unzureichender Terminologie, Softwaremängeln, sonstigen technischen Problemen) der Anwender selbst ist. Einige Beispiele: Aufgrund fehlenden Sprachgefühls und meist nicht vorhandener Kenntnis dessen, was maschinelle Übersetzung leisten kann und was nicht, versuchen Nutzer häufig Einzelbegriffe (wie mit einem Wörterbuch oder einer Terminologiedatenbank) zu übersetzen. Das funktioniert manchmal, meistens aber nicht (richtig). Anwender geben häufig auch unvollständige Sätze ein, solche mit fehlender oder falscher Interpunktion (vor allem fehlendem Satzendezeichen), fehlerhafter, um nicht zu sagen beliebiger, Rechtschreibung oder mit unbekanntem bzw. selbst erdachten