

Harald H. Zimmermann, Saarbrücken

Textverarbeitung und Maschinelle Übersetzung

(im Original: Wordprocessing and Machine Translation)

An den Anfang der Überlegungen zu diesem Thema soll eine Frage gestellt werden, die - gleichgültig, welcher Religionsgemeinschaft jemand angehört oder welche Weltanschauung er vertritt - eigentlich jedermann bewegt, vielleicht ausgenommen die Sprachgenies à la Schliemann: Viele entscheidende Barrieren zwischen Personen, Institutionen oder auch Völkern werden durch die Verschiedenheit des wesentlichen Kommunikationsinstruments, die natürliche Sprache, hervorgerufen. Dabei ist erwiesen, dass es im Grunde keine "besseren" oder "schlechteren" Sprachen, vielleicht nicht einmal verschieden "schwere" Sprachen gibt: alle erfüllen seit Jahrhunderten ihren Zweck, Wissen zu vermitteln, soziale Handlungen zu unterstützen usw.. Jedes Sprachsystem ist für sich betrachtet also - trotz mancher historisch bedingter Erker - sei es durch die frühe Verfestigung eines älteren Lautstandes in der Schriftsprache, sei es durch die natürliche Konventionalisierung (anstelle einer theoretisch deutbaren "logischen" Systementwicklung: man vgl. die Ansätze zu Kunstsprachen wie ESPERANTO - also voll entwickelt und nutzbar, zudem repräsentiert sich die Kultur eines Volkes oft in eben dieser seiner Sprache am besten.

Es soll im folgenden auch nicht in Frage gestellt werden, daß manches Denken so sehr mit den Sprachmöglichkeiten verbunden ist, dass eine "Übersetzung" in eine andere natürliche Sprache (Fremdsprache) - modern ausgedrückt - zu Informationsverlusten führt. Da (bislang wenigstens) natürlichsprachliche Wissensvermittlung (oder allgemeiner: Kommunikation) immer - also auch innerhalb eines Sprachsystems mit Vagheiten verbunden ist, erwartet man eigentlich auch nirgendwo eine 1:1 Übersetzung zwischen natürlichen Sprachen. Damit muss der Übersetzer und Dolmetscher "leben". Die zunehmende Internationalisierung im sozialen wie Wissenschaftsbereich lässt umgekehrt - vorausgesetzt, Mittel sind dafür ausreichend vorhanden - den Übersetzer- und Dolmetscherberuf als gesellschaftlich wichtig und attraktiv erscheinen.

Wenn man sich nun - wie es scheint - nicht auf eine Kunstsprache wie ESPERANTO bei der internationalen Kommunikation abstützen möchte, so bleibt eigentlich als Alternative - will man den Aufwand zum Erlernen einer Fremdsprache möglichst gering halten - nur die Verständigung auf eine natürliche Sprache als internationale Kommunikationssprache. Dies ist sicherlich ebenso längst eine Binsenweisheit wie auch nach aller Erfahrung wenig realistisch. Selbst eine Entwicklung wie die Europäische Gemeinschaft ist bislang nicht in der Lage (oder auch Willens), hier sozusagen auf "technischem" Wege eine interne Lösung zu realisieren, die z.B. mittelfristig den Haushalt der EG-Kommission drastisch entlasten könnte.

Wenn es nach der Sprach-Politik ginge, so hätte daher - auch angesichts der leeren Kassen - das Computer-Übersetzungssystem Hochkonjunktur, das in der Lage wäre, die Sprachbarrierenproblematik zumindest in der wissenschaftlich-technischen Kommunikation (auf Textebene) befriedigend zu lösen. Dieses Problem der Überwindung von Sprachbarrieren durch MT-Systeme, z.B. in der Fachinformation, ist jedoch bis heute nicht gelöst. Es geht dabei nicht einmal um hochfliegende Wunschträume der Menschheit, nach einer idealen, schranken- und barrierenlosen Kommunikationsgesellschaft. Wenn man realistisch ist, wird man sogar die Vorstellung zumindest zurückstellen, dass wir uns einmal dank einer (Mikro-)Computersoftware, die (gesprochene) Sprache so weit übersetzt, wie dies heute ein menschlicher Übersetzer oder Dolmetscher tut, dabei "brauchbar" verständigen können, auch wenn die Gesprächspartner unterschiedliche (natürliche) Sprachen verwenden.

Wenn man unter diesem Vorzeichen beispielsweise in die Schubladen der Experten für Künstliche Intelligenz (KI) schaut, so findet sich eine Reihe von Theorien und Modellen, hier und da auch so genannte "Mini-Welten", in denen exemplarisch und ansatzweise verschiedene (z.T. frappierende) Lösungen zum sog. "Sprachverstehen" vorgestellt bzw. angeboten werden. (Ein interessantes Beispiel ist das in Hamburg entwickelte Modell HAM-ANS.) Doch trotz z.T. jahrzehntelanger Forschungen von Sprachwissenschaftlern, Computerspezialisten, Psychologen usw. ist ein Durchbruch zu praktikablen, hoch entwickelten Verfahren, die das Problem zumindest äquivalent zu den möglichen Leistungen menschlicher Übersetzer, ohne eine entsprechende menschliche Interaktion

lösen, nicht in Sicht. Die Fachwelt hofft nun auf die Ergebnisse der Forschungen der "5th Generation Computer" in Japan; im Rahmen einer "westeuropäischen" Antwort auf dieses japanische Forschungsprogramm, dem europäischen ESPRIT-Konzept, soll wie in Japan das Problem des Sprachverstehens durch Computer mitbehandelt werden. Doch was an praktikablen Lösungen, (oder besser: Lösungsansätzen) herauskommen wird, ist heute noch völlig ungewiss. Oder umgekehrt: nach aller bisherigen Erfahrung wird man auf die 6. oder 7. oder n-te Computergeneration warten müssen, ehe (vielleicht) derartige Wunschträume Wirklichkeit werden.

Woran liegt das? Nun: einmal an der Komplexität und der ungeheuren Vielfalt der natürlichen Sprache. Zu (fast) jeder sprachlichen Regel gibt es eine Ausnahme, die ein maschinelles Sprachanalyse- oder -verstehenssystem aufschwemmt, es unübersichtlich und unhandlich macht. Dies gilt analog für die Evaluierung der Ergebnisse, da die Bewertung fast mehr Aufwand verursacht als die Erstellung der Verfahren selbst. Zum anderen stellt sich ein "internes" Transfer- und ErgebnisSicherungsproblem: die Forschungs- und Arbeitsgruppen, die sich mit derartigen Themen befassen, sind (zu) klein, die damit beschäftigten Wissenschaftler eher an prinzipiellen Lösungen interessiert als an "Knochenarbeit". Die verwendeten Technologien wandeln sich zudem so schnell, dass Anpassungen an neue Rahmen (Soft- und Hardware) häufig mehr Zeit verschlingen als die Entwicklungsarbeiten) selbst. Dem Financier grundlegender problemspezifischer Entwicklungen (sei es nun ein Ministerium oder auch die Industrie) geht - meist unter einem ökonomischen Erfolgszwang stehend - frühzeitig der Atem aus.

Arbeitsgruppen zerfallen - man denke in Deutschland an das LIMAS-Projekt, an die PLIDIS-Entwicklungen, an die CONDOR-Gruppe - und nur mühevoll können erreichte Teilergebnisse halbwegs gesichert werden.

Lange Zeit bedeutete die (mangelnde) Geschwindigkeit des Computers und die relativ geringe Speicherkapazität ein Handicap für die diesbezügliche Forschung und Entwicklung. Dies scheint sich jetzt langsam zu bessern, jedenfalls wird den technischen Rahmenbedingungen zunehmend weniger Bedeutung zukommen. Umgekehrt wächst der Bedarf an derartigen Verfahren oder zumindest an Teillösungen mit dem Ansteigen der Nutzung von Tele- und Bürokommunikation, aber auch mit der zunehmenden Internationalisierung der Fachinformation.

Es bringt nämlich wenig, wenn man heute aus Japan auf deutschsprachige Informationsbanken zugreifen kann (und umgekehrt), ohne dass man dabei den Titel oder das Abstract (und später - beim On-Demand-Publishing - auch die Textfassung) lesen bzw. verstehen kann. Natürlich könnte sich - wie erwähnt - die Wissenschaft - und tut es ja zum Teil - auf eine 'Wissenschaftssprache' (z. B. Englisch) verständigen, aber jedermann leuchtet ein, dass die Verfügbarkeit textueller Information in der jeweiligen Muttersprache eine weitaus interessantere Lösung darstellen würde - vorausgesetzt, sie lässt sich ökonomisch und auch qualitativ ausreichend gut herstellen.

Angesichts der generellen Problematik, aber auch angesichts der mangelnden Verfügbarkeit von Mitteln (die - um einen drastischen Vergleich zu wagen - letztlich in ihrer Gesamtheit den Investitionen für ein kleines Raumfahrtprogramm entsprechen dürften) muss zur Überwindung der Sprachbarrieren eine Politik der "kleinen Schritte" entwickelt werden. Für eine derartige Strategie sind unter Kosten/Nutzen-Gesichtspunkten vor allem zwei Kriterien wichtig:

- die jeweiligen Zwischenschritte müssen praxisrelevante Ergebnisse bringen;
- die Zwischenschritte müssen nach Möglichkeit auf den vorausgehenden Etappen (Verfahren, Daten) aufbauen.

Vielleicht könnte man noch ein drittes Element hinzufügen:

- vorhandene Ressourcen sollten so weit wie möglich mitgenutzt werden.

Zu diesen Ressourcen für eine maschinelle Sprachdatenverarbeitung in der Fachinformation gehören v. a. bestehende Lexika, Thesauri und Enzyklopädien. Bei aller Problematik einer unmittelbaren Übernahme bestehender "gedruckter" Sprachdatensammlungen - sie sind meist wenig formalisiert, z.T. ad hoc auf den "menschlichen" Benutzer zugeschnitten: derartige Inventare enthalten so viel "kondensiertes" sprachliches Wissen, dass sie zumindest als Steinbruch, vielfach auch mit ihren sprachlichen Kodierungen für den Aufbau eines systematischen Wortinventars herangezogen werden können.

Was sind das aber für Zwischenschritte? Hier lassen sich verschiedene Wege gehen. Einmal kann man die Software-Entwicklungen zunächst auf bestimmte Fachgebiete/Branchen und innerhalb dieser Bereiche auf bestimmte Textsorten konzentrieren. Zugleich kann man - vorausgesetzt, dies lässt sich praktisch verwirklichen - nur Teilstrukturen der natürlichen Sprache behandeln. Ein derartiger Weg wird z.B. in der Textinformation mit TITUS beschrieben. Hierbei gelangt man frühzeitig zu maschinellen Übersetzungen (die z.T. noch einer knappen Nachredaktion bedürfen), v.a. verfügt man über ein interessantes terminologisches Kontrollinstrument. Analog zu TITUS kann dann anschließend versucht werden, die verfügbaren Sprachstrukturen schrittweise zu erweitern.

In Kanada hat man vor einigen Jahren den Versuch gemacht, eine sehr stark restringierte Textsorte zu bearbeiten. Es ging darum, Wettermeldungen maschinell zu übersetzen (System METEO). Ein scheinbar trivialer Fall, doch der Teufel steckt hier im Detail - heute ist dieses zunächst viel versprechende Projekt bereits Makulatur.

Ein weiterer Ansatz, der bei den Systemen SYSTRAN (auf Großrechnerebene) und LOGOS (auf Mini rechner-Basis) unternommen wird, ist die Unterstützung des menschlichen Übersetzers bei der Übersetzungsarbeit. Hier ist bereits das Vorhandensein eines umfassenden Fachwörterbuchs eine wichtige Grundlage; zudem muss der Übersetzer über eine Schnittstelle verfügen, die einerseits ein bequemes Ergänzen von Wörterbuchdaten ermöglicht, andererseits zugleich die Edition der maschinellen Roh-Übersetzungen (mehr schafft der Computer bei "normalen" Texten nicht) über Textsystem-Funktionen zulässt. Hier heißt die Gleichung ganz einfach: ist ein derartiges maschinelles Verfahren preiswerter (und sichert es daneben eine größere Konsistenz?). Diese Frage ist heute vielleicht noch nicht ganz so eindeutig beantwortet, zeigt aber eine richtige Entwicklungsrichtung an.

Wenn man kurz- und mittelfristig das Mengenproblem im Hinblick auf die Sprachbarrieren in der Fachinformation überwinden will, so kann die Nachredaktion oder Post-Edi tion (sie setzt einen qualifizierten Übersetzer voraus) auf Dauer nicht die Lösung sein. Daher wird an der Universität des Saarlandes - mit Unterstützung des BMFT - ein anderer Weg versucht: die Entwicklung eines so genannten maschinellen "Informativ-Textübersetzungs-Systems" (ITS). Voraussetzung - wie bei allen anderen Verfahren - sind umfangreiche maschinelle Lexika,

die fachgebietsspezifische und auch textsortenrelevante Kennungen haben, um dem Problem der Mehrdeutigkeit von Benennungen zu begegnen. Nützlich sind auch thesaurusartige Begriffsvernetzungen, die zusammen mit Teilwortrelationen und Wortableitungen ein lexikalisches Netzwerk bilden, das bezogen ist auf algorithmisch umgesetzte sprachliche Regeln. Damit stellt es ein praktikables - wenn auch nicht perfektes - Instrumentarium für begriffliche Vereindeutigungen dar. Der Aufbau eines derartigen Lexikons bildet im Grunde die wesentliche Investition.

Im algorithmischen/bearbeitungsstrategischen Bereich verzichtet man demgegenüber zunächst auf komplizierte Regeln, die jeden möglichen 'Schlenker' der natürlichen Sprache mitmachen, in der Hoffnung, dass zwar nicht unbedingt ein "gutes" Deutsch oder Englisch, aber ein verstehbarer, informativer Text entsteht. "Versteht" der Computer einen Satz nicht, so zeigt er dies in einem Statusbericht an: so hat man ggf. Gelegenheit, in einer Art "Interedition" einen allzu komplexen Ausgangssatz an die System-Grammatik anzupassen. Da in der Regel sowieso bestimmte Prä-Editionen erforderlich sind, z. B. um Rechtschreibfehler zu korrigieren, die das System aufgrund eines Wörterbuchabgleichs feststellt, kann ein System-Experte die eine oder andere Problematik schon "voraussehen" und den Text - möglichst ohne Sinnänderung - an das System formal leicht anpassen.

Das System ITS (Informativ-Textübersetzungs-System) stellt eine anwendungsorientierte Systemvariante des an der Universität des Saarlandes entwickelten Basis-Systems SUSY dar, so dass alle im Grundsystem realisierten Funktionen (v. a. die Analyse- und Syntheserichtungen zu den Sprachen Deutsch, Englisch, Französisch und Russisch) nebst den dazugehörigen Basis-Wörterbüchern Verwendung finden können.

Und noch ein weiterer Vorteil kann hier genannt werden: ITS ist ein modulares System, so dass wesentliche Teile identisch sind mit dem ebenfalls in Saarbrücken verfügbaren System CTX (Computergestütztes Texterschließungssystem). Automatische Indexierung und maschinelle Informativübersetzung greifen also ineinander über, so dass sowohl die einsprachige Texterschließung als auch die Textübersetzung lexikalisch, strukturell und auch rechenökonomisch voneinander profitieren.

Das Indexierungssystem CTX hat in mehreren, z.T. umfangreichen Anwendungen die ersten praktischen Proben bereits bestanden. So werden beispielsweise deutschsprachige Dokumente einer Reihe von Mitgliedern der Arbeitsgemeinschaft Fachinformation (AFI), z.B. Daten des Deutschen Patentamts, des Fachinformationszentrums Werkstoffe (FIZ 5) und aus dem Rechtsbereich zum Datenschutz verarbeitet). ITS ist dagegen erst in der ersten Entwicklungsphase. Aber auch bei der Informativübersetzung wird bereits mit handfesten Daten (wiederum des Deutschen Patentamts, aber auch des Fachinformationszentrums Technik, FIZ 16) gearbeitet.

Die Einbeziehung von Textsystemen in die maschinelle Sprachübersetzung muss unter zwei Aspekten gesehen werden.

Einmal ist ein Textsystem der ideale Rahmen für die Manipulation von Zeichenketten oder auch "höherwertiger" Elementen (Wörter, Zeichen, Sätze, Abschnitte ...). Professionale spezialisierte Systeme - zunehmend aber auch angepasste allgemein verwendbare Mikrocomputer (PC-Computer mit Textverarbeitungs- und Editionshilfen wie WORD oder WORDSTAR) unterstützen den Übersetzer bei der Prä-, Inter- oder Postedition: Unter Präedition wird dabei die Manipulation des Inputs (Quellsprachen-Texte) zur Anpassung an die (technischen wie inhaltlichen) "Grenzen" des MT-Verfahrens verstanden. Bei der Interedition wird während des maschinellen Übersetzungsprozesses durch Systemrückfragen die Möglichkeit gegeben, den zu übersetzenden Text zu modifizieren. Bei der Postedition wird (i.d.R. durch einen kompetenten Sprecher der Ausgangs- und Zielsprache) eine Verbesserung der Übersetzung selbst angestrebt. Daneben wird es über die externen Textsystem-Schnittstellen möglich, Fremdtex te in den Prozess zu integrieren bzw. den Output ohne Neufassung an weiterverarbeitende Systeme (z.B. zum Photosatz) zu übermitteln.

Nach diesem Modell arbeitet z.B. die EG-Anwendung des SYSTRAN-Systems. Dabei ist ein IBM-Textsystem mit einem zentralen (IBM- oder SIEMENS-)Großrechner gekoppelt. Nach Fertigstellung eines zu übersetzenden Textes wird diese als Datei mit dem Auftrag zur Übersetzung an den Großrechner weitergegeben. Im Prinzip ist nunmehr eine "getrennte" SYSTRAN-Weiterverarbeitung möglich, z.B. mit der Möglichkeit eines vorherigen Lexikon-Update durch das SYSTRAN-Servicezentrum. Es kann aber auch eine Quick-and-dirty-Übersetzung erfolgen.

Nach Abschluß des Übersetzungsprozesses wird die Zielsprache-Datei mit Bezug zur Originaldatei wieder über eine Software-Brücke an das betreffende Textsystem zurückgegeben. Dem Übersetzer, der an dieser Stelle evtl. zum ersten Mal eingreift, steht nunmehr der Originaltext nebst Rohübersetzung zur Verfügung, der nun je nach Qualitätsanforderungen nachgebessert wird.

Die heutige Konzeption der EG-Kommission zum Einsatz von SYSTRAN ist voll auf die drei genannten Kriterien abgestimmt: Die Gesamtkosten des spezifischen Übersetzungsprozesses müssen niedriger sein als bei "reiner" Humanübersetzung, wobei auch die bei der Humanübersetzung erforderliche typographische Erfassung des Zielsprachigen Textes mit einzurechnen ist, umgekehrt muß der Kostenanteil für die Textverarbeitungshard- und Software mit eingerechnet werden. Es hat sich bislang offenbar gezeigt, daß trotz eines höheren Anteils einer Nachredaktion gegenüber der Humanübersetzung (die bei der EG ebenfalls intellektuell überprüft wird) insgesamt - die Investitionskosten zur System- und Lexikontwicklung nicht eingerechnet - durchaus kostengünstig(er) ist bzw. angesichts der allgemeinen technologischen Entwicklung in naher Zukunft sein wird. Bei geeigneter Organisation - auch dies wurde erprobt - können die Verweilzeiten (d.h. die Zeit, die ab dem Auftrag zur Übersetzung bis zur Realisierung vergeht) ebenfalls gesenkt werden. Die Integration in einen Textverarbeitungsrahmen ließe zudem auch eine Electronic-Mail-Komponente zu, über die dem Auftraggeber z.B. auch die Rohübersetzung (sofern einigermaßen brauchbar) als Informationsquelle bereitgestellt werden könnte. Dies kann als ein Beispiel für die mögliche Einbindung in das Informationsumfeld betrachtet werden. Bei der Entwicklung des LOGOS-Systems, wird auch für den Übersetzungsvorgang ein Microprozessor-System zugrundgelegt. Während bei SYSTRAN der Übersetzer selbst nur in der Vor- oder Nachbereitungsphase einbezogen wird, der Übersetzungsprozess jedoch von ihm nicht beeinflusst werden kann, sieht das LOGOS-System - z.Z. in begrenzter Form - auch eine Interaktion vor: der Übersetzer hat die Möglichkeit, das Lexikon um fehlende Begriffe und diesbezügliche Kodierungen zu ergänzen. Die LOGOS-Software ist gleichsam als "Branchensoftware" auf einem WANG-Computer realisiert, wobei die Textverarbeitungsfunktionen voll genutzt werden können. Nach Angaben des Systementwicklers können am Tag technisch rund 20.000 laufende Wörter übersetzt werden.

Auch hierbei geht es im Grunde um die Reduktion der Kosten, hier bezogen auf einen integrierten Übersetzerarbeitsplatz. Ein Humanübersetzer, der - dies gilt eigentlich für alle irgendwo heute eingesetzten MT-Systeme - zum ersten Male mit den Rohübersetzungen eines derartigen Systems konfrontiert wird, ist zunächst einmal schockiert über die vermeintlich unzureichende Qualität. In vielen Fällen - es gibt natürlich auch Übersetzungen, die unverständlich sind - reicht aber eine Wortumstellung bzw. ein leichter stilistischer Eingriff (bestimmter anstelle des unbestimmten Artikels) aus, um den zielsprachigen Text lesbar zu machen bzw. zu berichtigen.

Wenn man so will, und dies sei ein erstes Fazit, ist eine maschinelle Übersetzung heute ohne die Integration in Anbindung an die Textverarbeitung überhaupt nicht praktikabel. Es wird jedoch für die Zukunft interessant sein, nicht nur die Primitiv-Funktionen (z. B. Zeichen, Wort, Satz löschen, umstellen, einfügen) zu nutzen, sondern MT-spezifische Funktionen zu integrieren, die einerseits die Umgebung, in der ein Text steht bzw. entsteht, als Information für den MT-Prozess nutzbar machen. In einer Büroumgebung beispielsweise treten Texte strukturiert (als Bericht, Brief, Notiz, Protokoll, Termin) auf, ähnliches gilt für fach- und branchenspezifische Texte/Daten: Angebote, Bilanzen, Rechnungen, Produktbeschreibungen, Kataloge usw. Umgekehrt können Fachgebietsskennungen, Strukturhinweise (auf "Textsorten") die Qualität einer maschinellen (Roh-)Übersetzung erhöhen, sofern sie systematisch ausgenutzt werden. Dies sowie eine stärkere Interaktionskomponente wird es erst möglich machen, MT-Systeme im Büro z. B. zur Übersetzung von Korrespondenz aus der Muttersprache in eine (dem Schreiber weitgehend unbekannte) Zielsprache umzusetzen.

Derartige "Entwicklungen" stellen dabei im Grunde zugleich "Spezialisierungen" und Fortentwicklungen der Textverarbeitung dar: Silbentrennung, Rechtschreib- und Stilhilfen sind "lexikalisch" orientiert. Insofern könnte die maschinelle Übersetzung, mit entsprechenden Teilkomponenten verknüpft, durchaus einen größeren Marktanteil erhalten, als dies bei einer Spezialentwicklung für den Humanübersetzer jemals der Fall sein wird.

Die Telekommunikation wird daneben eine Reihe von Alternativen eröffnen: so ist es vorstellbar, dass MT-gestützte Übersetzungszentren entstehen, die über ein Mailbox-System elektronisch Aufträge zur Übersetzung eines Textes erhalten (also analog zum SYSTRAN-Verfahren) und diese im Service bearbeiten, d. h. die Ergebnisse in der Mailbox des Auftraggebers wieder ablegen. Mit Bild-

schirmtext-Kopplungen zu Bürosystemen bzw. Personal Computern ist der technische Grundstein von Rechnern hierzu bereits gelegt.

Umgekehrt ist es vorstellbar, dass Systeme zur Informativübersetzung (analog zu ITS) bei einem Informationsvorgang (z. B. bei einem Abruf fremdsprachiger Datenbanken) "zwischen geschaltet" werden: das Ergebnis der Recherche (d. h. die Rohübersetzung nebst Original) kann ggf. je nach Bedarf über ein sog. "Downloading", d. h. z. B. durch Übernahme in eine Textdatei des Textsystems bzw. Personal Computers aufbereitet und abgelegt werden.

Es sollte deutlich geworden sein, dass auch im Vorfeld der Realisierung des Wunschtraumes einer perfekten maschinellen Übersetzung eine Reihe realistischer Zwischenformen zu finden sind, die alle einen angemessenen Einsatzbereich finden könnten. Wenn auch die Integration in die moderne Textverarbeitung bzw. Büro- und Telekommunikation erst den Boden für diese praktisch einsetzbaren Systeme bereitet, so darf die Basisentwicklung, d. h. die Realisierung höherwertiger Systeme zur MT, nicht stehenbleiben. Fortentwicklungen werden v. a. durch die Forschungen zur Wissensrepräsentation und -verarbeitung im Bereich der sog. "Künstlichen Intelligenz" erhofft, doch werden die diesbezüglichen Investitionen sehr, sehr hoch sein. Die von den Staaten der EG geplante Entwicklung eines eigenen Systems zur maschinellen Übersetzung (EUROTRA) stellt auf der Grundlage relativ gesicherter moderner Erkenntnisse v. a. der Sprachwissenschaft und Informatik einen Versuch dar, ein brauchbar-praktisches System zu entwickeln, das den Ansprüchen der nächsten zehn Jahre gerecht wird. Auch EUROTRA wird konzeptionell in das Tätigkeitsfeld der Textbe- und -verarbeitung eingebettet sein. Unter diesen Vorzeichen der Selbstbeschränkung auf machbare Lösungen hat inzwischen also die eigentlich interessante Phase einer anwendungsorientierten maschinellen Übersetzung begonnen.