

Übersetzung durch Mensch und Maschine

Geschichte der Maschinellen Übersetzung

- Einleitung
- Entwicklungsgeschichte der MÜ

Übersetzung durch Mensch und Maschine

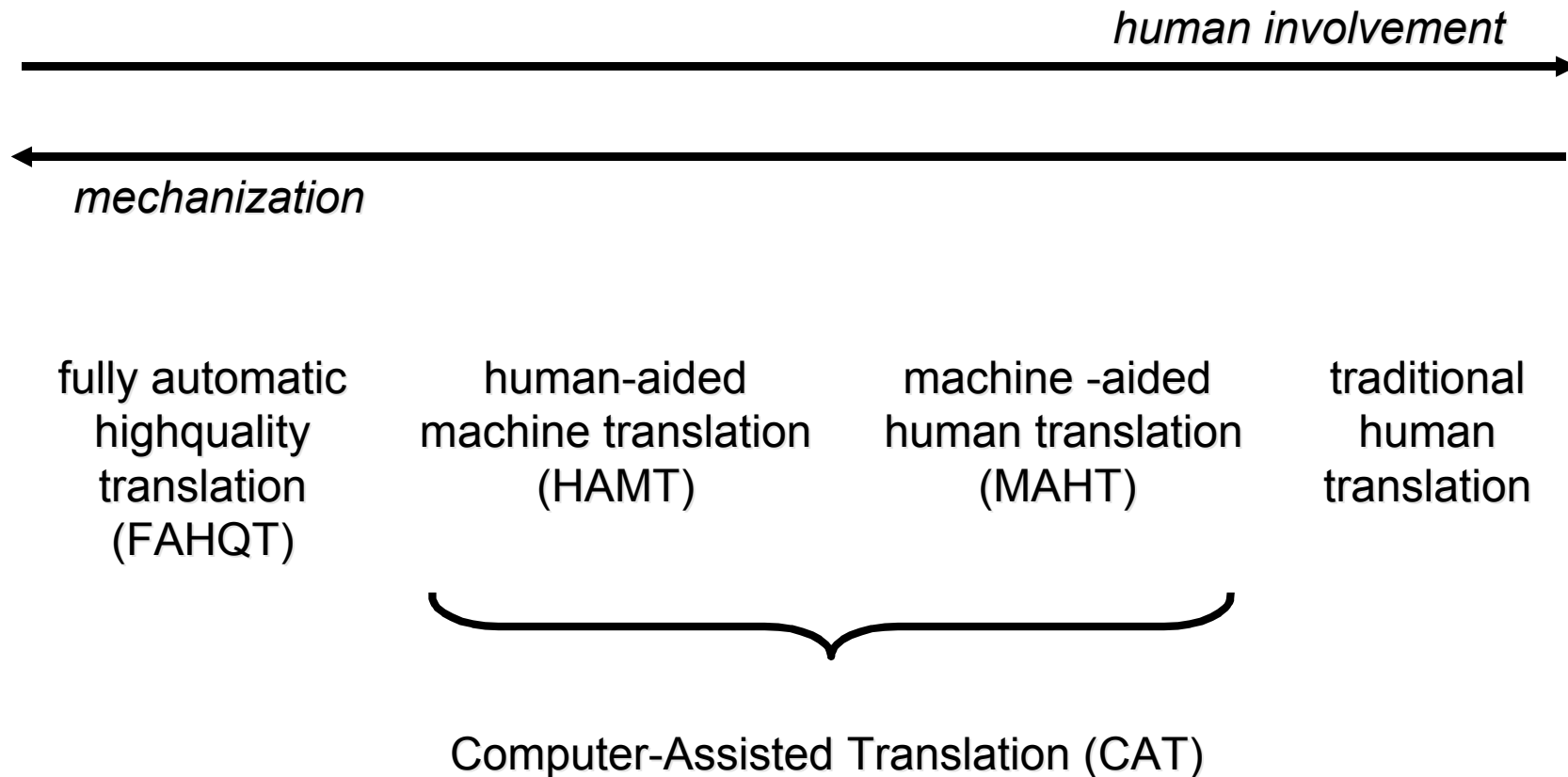
Einleitung

- MÜ als Modellierung des Übersetzungsprozesse
- MÜ als Mechanisierung menschlicher Tätigkeit
 - ▶ Hochqualitative vollautomatische Übersetzung
 - ▶ Maschinelle Übersetzung mit menschlicher Intervention
 - ▶ Maschinenegestützte konventionelle Übersetzung
 - ▶ Konventionelle Übersetzung durch Menschen



Übersetzung durch Mensch und Maschine

Übersetzen durch Mensch und Maschine



Übersetzung durch Mensch und Maschine

Entwicklungsgeschichte der MÜ

Ahnen

- ▶ Descartes, Leibniz (Wörterbücher auf der Grundlage eines universellen Zahlenkodes)
- ▶ Joh. J. Becher 1661
- ▶ John Wilkins 1668 (*Essay towards a Real Character and a Philosophical Language*)

Vorläufer

- ▶ George Artstruni 1933
- ▶ Petr P. Smirnov-Trojanskij 1933

Pioniere

- ▶ Warren Weaver, Mathematiker, *Translation* (1949)
- ▶ Yehoshua Bar-Hillel, Logiker



Übersetzung durch Mensch und Maschine

JOH. J. BECHERI,
Spirensis
CHARACTER,
Pro
NOTITIA LIN-
guarum Universalī.

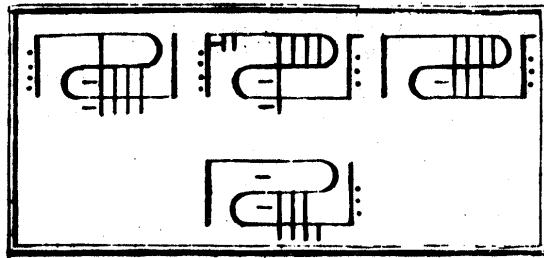
INVENTUM STEGANO-
GRAPHICUM HACTENUS IN-
auditum quo quilibet suam Legendo vernaculam
diversas imò omnes Linguas, unius etiam diei
informatione, explicare ac intelli-
gere potest.



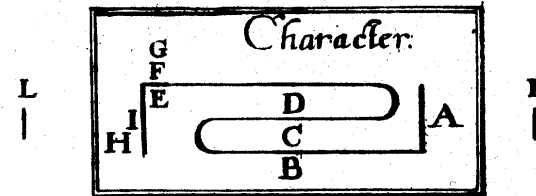
FRANCOFURTI,
Sumpt. Johannis Wilh. Ammonii
& Wilhelmi Serlini,
Typis JOHANNIS GEORGII SPÖRLIN.
ANNO M. DC. LXI.



Übersetzung durch Mensch und Maschine



*Huc spectant Exempla pro
Numeri notitia et post haec
Tabula pro Variationibus Sensuum.*



Exempla pro Numeri notitia.

5.	3.	3.	2.	1.
10	9.	8.	7.	6.
15.	14.	13.	12.	11.
20.	19.	18.	17.	16.
60.	50.	40.	30.	25.
200.	100.	90.	80.	70.
700.	600.	500.	400.	300.
3000.	2000.	1000.	900.	800.
8000.	7000.	6000.	5000.	4000.
1003.	1580.	11000.	10000.	9000.
			1111.	7327.

Übersetzung durch Mensch und Maschine



LEXICON

Pro

Resolutione primæ Characteris

partis

A. B. C. D.

L. S.

Ineas Characteris secundum informationem in tuæ Vernaculæ numerum transfer quem in eisdem LEXICO evolve, ita designatum vocabulum pro conditione punctationis flectendum invenies.

A			
Ab	1	Abbatissa	10
Abactus	2	Abdere	11
Abacus	3	Abdicare	12
Abaculus	4	Abditus	13
Abavus	5	Abdomen	14
Abax	6	Aberrare	15
Abbas	7	Abesse	16
Abbatia	8	Abhinc	17
	9	Abhorrere	18

Abies

CHARACTERIS.

Accipiter	83	Acidus	115
Acclivis	84	Acies	116
Accola	85	Acinaces	117
Accommodare	86	Acinus	118
Accumbere	87	Acinofus	119
Accumulare	88	Acipenser	120
Accurate	89	Aclis	121
Accusare	90	Aconitum	122
Accusavi	91	Acorus	123
Accusabo	92	Acquiescere	124
Accusa	93	Acquirere	125
Accusum	94	Acredula	126
Accusaverim	95	Acris, cre	127
Accusavero	96	Acriter	128
Accusavisse	97	Acta, orum	129
Accusatum ire	98	Actio	130
Accusans	99	Actor	131
Accusatio	100	Actum est	132
Acedia	101	Actuarius	133
Acer	102	Actuosus	134
Acer, acris	103	Actus	135
Acerbus	104	Actutum	136
Aceria	105	Acuere	137
Acervus	106	Aculeus	138
Acerzare	107	Acumen	139
Acetabulum	108	Acupictus	140
Acetarium	109	Acus, eris	141
Acetosa	110	Acus	142
Acetofella	111	Acufacere	143
Acetum	112	Acupingere	144
Achates	113	Acutus	145
Acicula	114	Ad	146

C Ad

Übersetzung durch Mensch und Maschine

Georges Artsouni (1933)

- Der Franco-Armenier Georges Artstruni entwarf 1933 eine Speichervorrichtung auf Papierband, mit der für ein beliebiges Wort die Entsprechung in einer anderen Sprache aufgefunden werden sollte. Sie wurde als Patent angemeldet und anscheinend wurde 1937 ein Prototyp dieser Vorrichtung vorgeführt.



Übersetzung durch Mensch und Maschine

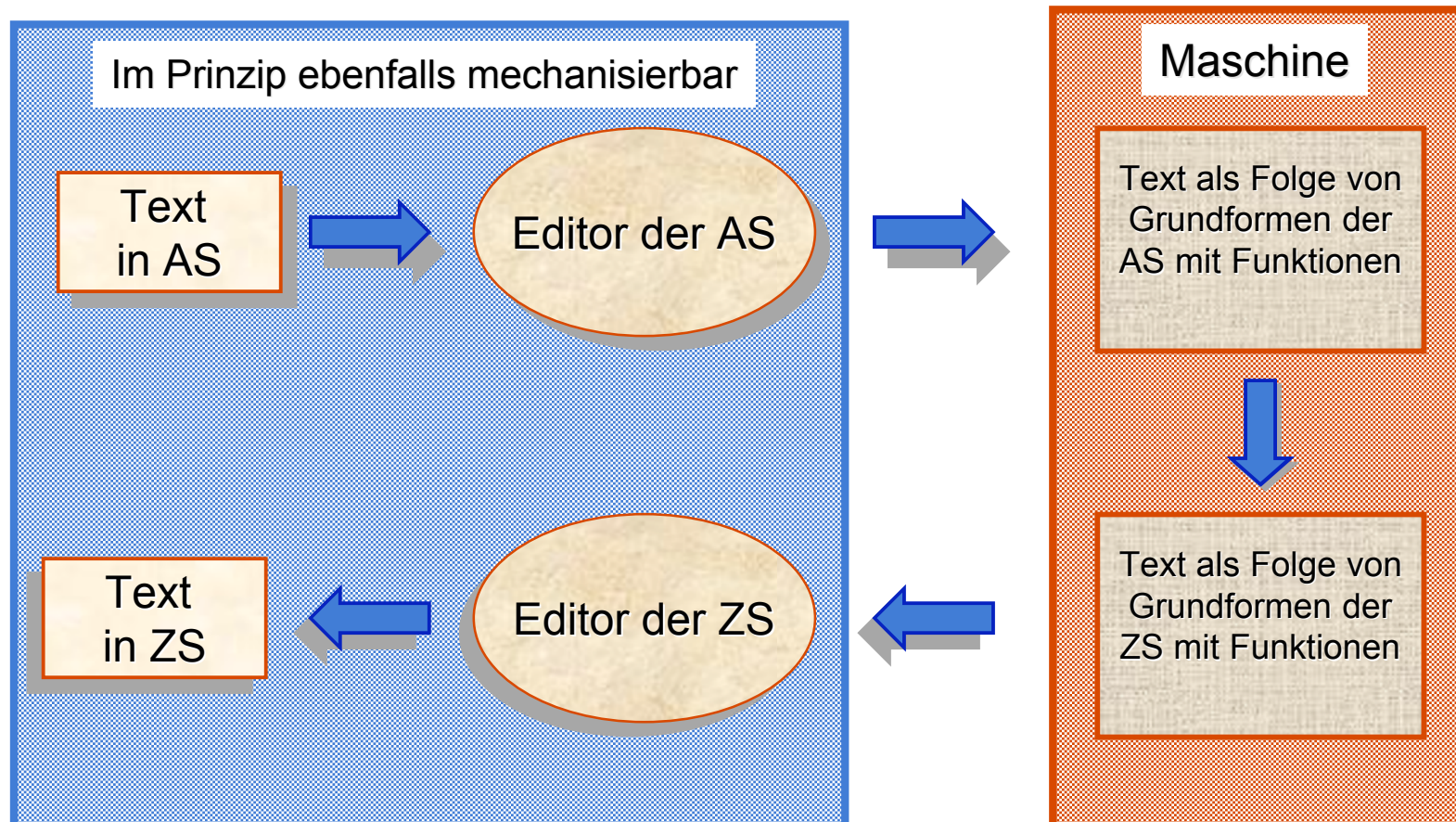
Petr Smirnov-Troyanskii

- Die Vorschläge von Smirnov_Troyanskij, ebenfalls als Patent angemeldet, waren aus heutiger Sicht weit bedeutsamer. Er sah drei Stadien der mechanischen Übersetzung vor
 1. Ein Editor, der nur die Ausgangssprache (AS) beherrschte sollte die 'logische' Analyse der Wörter im Ausgangstext (AT) in ihre Grundformen und syntaktischen Funktionen vornehmen.
 2. Die Maschine sollte Folgen von Grundformen und deren Funktionen in äquivalente Folgen in der Zielsprache (ZS) transformieren.
 3. Ein weiterer Editor, der nur die Zielsprache kannte sollte dann diesen Output in entsprechende normale Wortformen seiner eignen Sprache konvertieren.
- Troyanskij sah sowohl eine bilinguale als auch eine multilinguale Übersetzung vor. Obwohl sein Patent sich nur auf die Maschine bezog, die das zweite Stadium ausführte, glaubte Troyanskii, dass der Prozess der logischen Analyse selbst ebenfalls mechanisiert werden könne.
- Troyanskij war seiner Zeit weit voraus und war ausserhalb Russlands unbekannt.



Übersetzung durch Mensch und Maschine

Petr Smirnov-Troyanskij



Übersetzung durch Mensch und Maschine

Pioniere

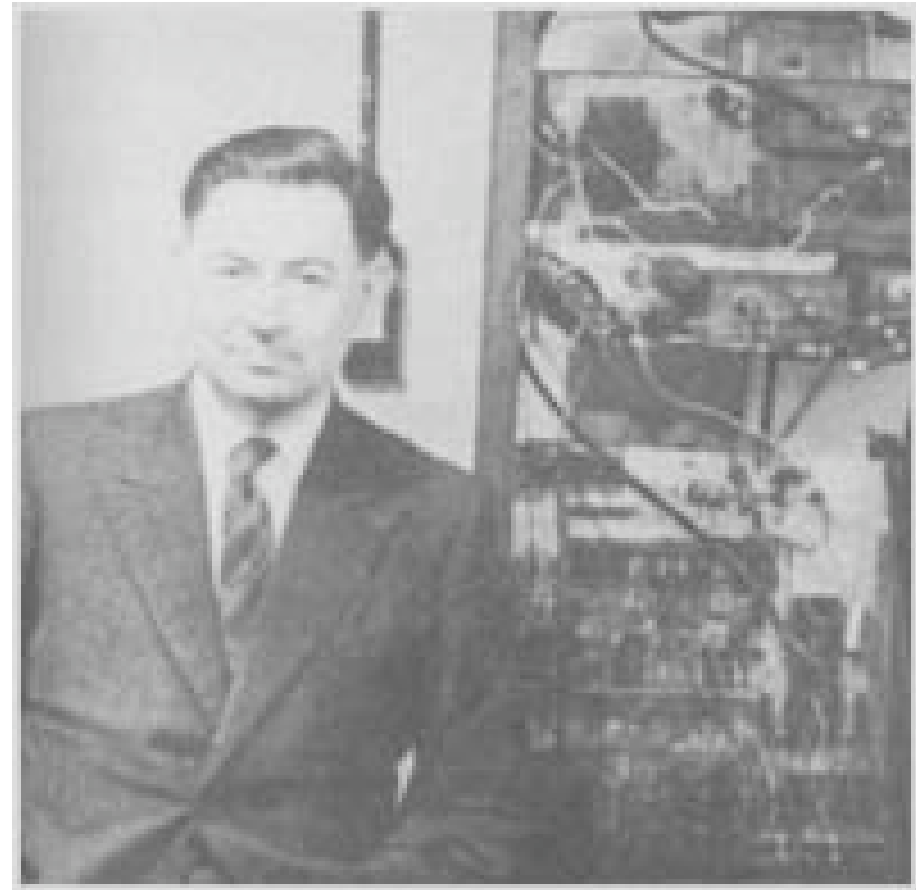
- Nur wenige Jahre nach der Erfindung des Computers, wurde die Möglichkeit diskutiert, Computer für die automatische Übersetzung einzusetzen, und zwar zuerst 1947 von Warren Weaver von der Rockefeller Foundation und von Andrew D. Booth, einem britischen Kristallographen.
- Nach seiner Rückkehr an das Birkbeck College (London), erforschte Booth die Mechanisierung eines zweisprachigen Wörterbuchs und initiierte eine Zusammenarbeit mit Richard H. Richens (Cambridge), der unabhängig davon Lochkarten für die Erstellung von groben Wort-für-Wort-Übersetzungen von wissenschaftlichen Texten verwendet hatte.



Übersetzung durch Mensch und Maschine

Pioniere: Warren Weaver und Andrew Booth

Warren Weaver



Übersetzung durch Mensch und Maschine

Pioniere: Andrew Booth mit Frau



Übersetzung durch Mensch und Maschine

Weaver Memorandum

- Es war jedoch ein Memorandum von Warren Weaver vom Juli 1949, das die Idee einer MÜ allgemein bekannt machte (Weaver 1949). Er skizzierte die Erfolgsaussichten und schlug verschiedene Methoden vor: die Verwendung von im Krieg entwickelten Dechiffrier-Techniken, statistische Methoden, Shannon's Informationstheorie, und die Erforschung der der Sprache zugrunde liegenden Logik und ihrer universellen Eigenschaften, 'the common base of human communication.'



Übersetzung durch Mensch und Machine

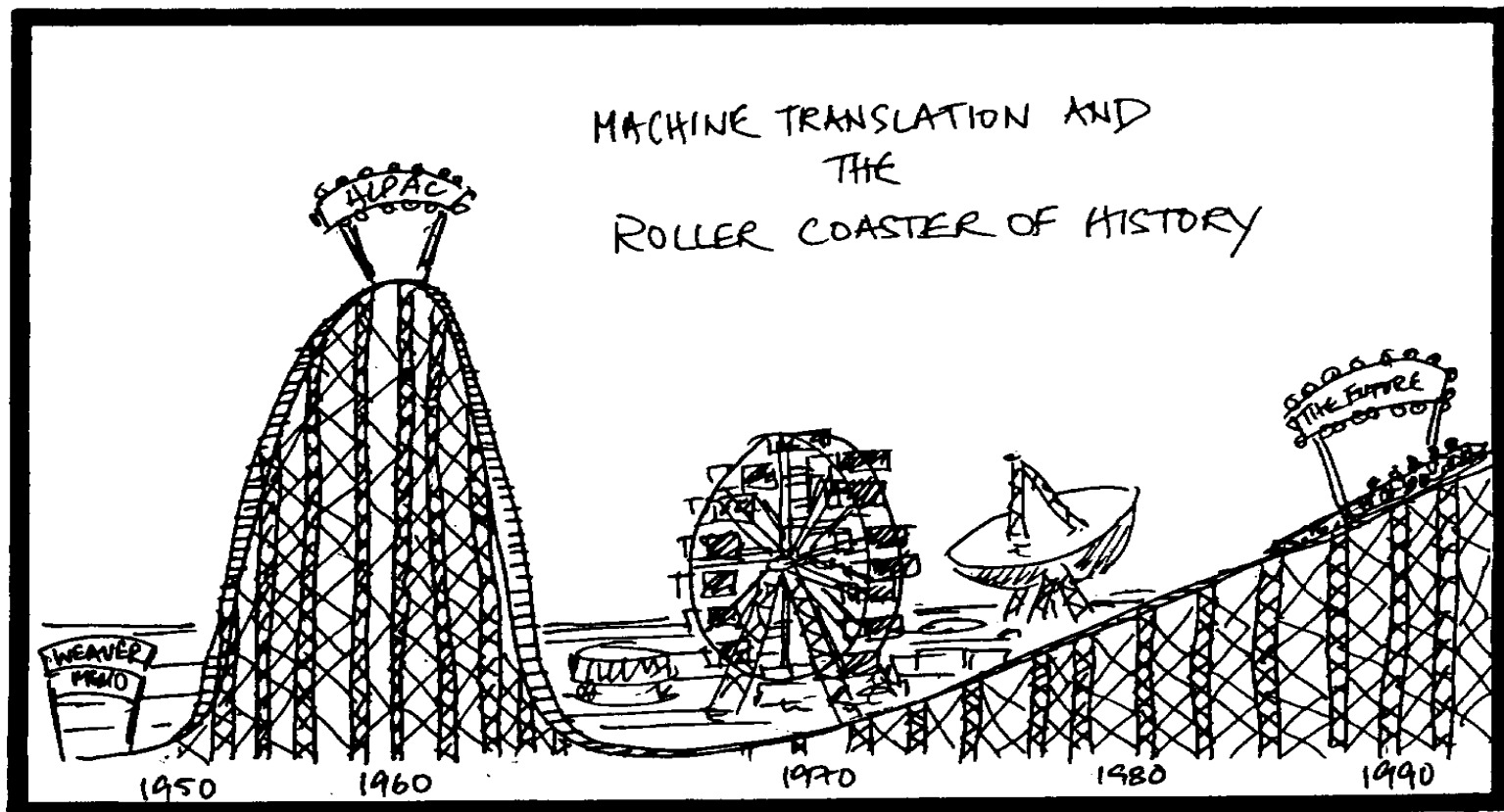
Weaver Memorandum

- Innerhalb weniger Jahre hatten an der Washington University (Seattle), der University of California (Los Angeles) und am Massachusetts Institute of Technology Forschungsarbeiten zur MÜ begonnen.
- Am MIT wurde auch im Jahre 1951 der erste Forscher ernannt, der sich hauptamtlich mit MÜ beschäftigte: Yehoshua Bar_Hillel.
- Ein Jahr später rief er die erste MÜ-Konferenz zusammen, wo die Grundzüge der zukünftigen Forschungsaktivitäten herauskristallisierten. Es gab Vorschläge zur Behandlung der Syntax von Oswald und Bar_Hillel (seine Kategorialgrammatik), Vorschläge, den Ausgangstext in einer MÜ_orientierten restringierten Sprache abzufassen bzw. zu bearbeiten, und Argumente für die Konstruktion von Subsprachensystemen.
- Es war offensichtlich, dass eine vollautomatische Übersetzung ohne langfristig angelegte Grundlagenforschung nicht möglich ist, und dass man in der Zwischenzeit ohne menschliche Intervention nicht auskommen konnte. Der Einsatz von Menschen war erforderlich, um entweder die Ausgangstexte für die Übersetzung aufzubereiten oder um den Output zu redigieren (damals schon bekannt als pre_ und postediting).



Übersetzung durch Mensch und Maschine

MÜ und die Achterbahn der Geschichte



Machine Translation and the Roller Coaster of History



Übersetzung durch Mensch und Maschine

Entwicklungsgeschichte der MÜ

Schneller Aufstieg und Ernüchterung

- ▶ Erste Konferenz über MÜ 1952 am MIT
- ▶ Georgetown University Experiment 1954
- ▶ Übersetzungssysteme der ersten Generation

ALPAC und seine Folgen

- ▶ Automatic Language Processing Advisory Committee (ALPAC) Bericht 1966

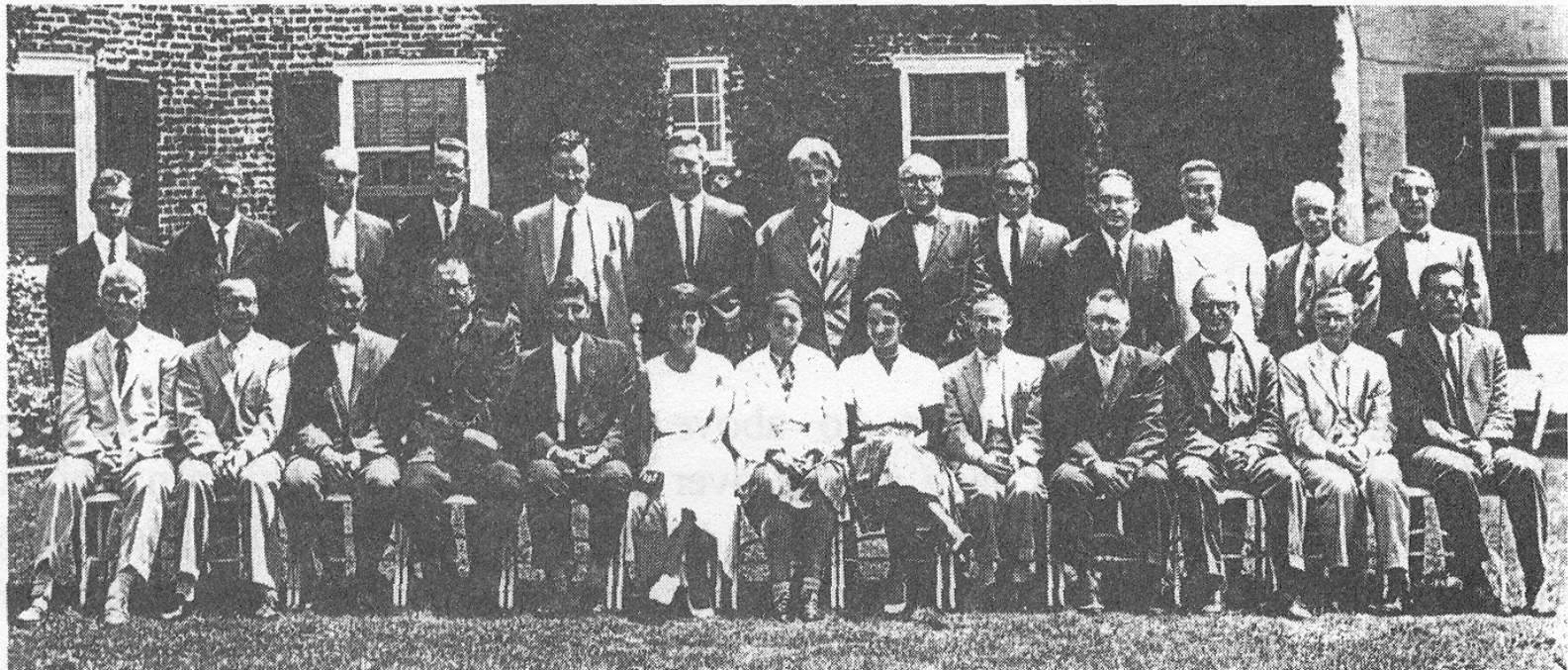
Konsolidierung

- ▶ Übersetzungssysteme der 2. Generation



Übersetzung durch Mensch und Maschine

Princeton 1960



Machine Translation Conference, Princeton Inn, July 1960



Übersetzung durch Mensch und Maschine



Association for Machine Translation and Computational Linguistics,
Denver, Colorado, August 1963



Übersetzung durch Mensch und Maschine

Pioniere: Victor H. Yngve und Anthony G. Oettinger



Übersetzung durch Mensch und Maschine

Peter Toma: Erster Entwickler von Systran



Übersetzung durch Mensch und Maschine

Entwicklungsphasen

- Computerwissenschaft
- Computerlinguistik
- Künstliche Intelligenz
- Corpuslinguistik



Übersetzung durch Mensch und Machine

Entwicklungsphasen: Computerwissenschaft

- MÜ als ingenieurwissenschaftliches Problem
- lokal begrenzter empirischer Ansatz
- linguistische Naivität verknüpft mit trickreicher Programmierung
- fehlende Modularität
- keine Trennung zwischen linguistischen Daten (Lexikon & Grammatik) und Verarbeitungsalgorithmen



Übersetzung durch Mensch und Machine

Entwicklungsphasen: Computerlinguistik

- Zunehmende "Linguistisierung"
- MÜ primär als computerlinguistisches Problem
- Maximale Unabhängigkeit von Analyse und Synthese (Modularität)
- Maximale Trennung von linguistischen Daten und Algorithmen
- Indirekte Übersetzungsverfahren durch Stratifizierung und Transfer, evtl. Interlingua



Übersetzung durch Mensch und Machine

Entwicklungsphasen: Künstliche Intelligenz

- Berücksichtigung von Hintergrundwissen
- Kontext
- kognitive Schemata
- Weltwissen



Übersetzung durch Mensch und Maschine

Typische Architekturen von MÜ Systemen

- Direkte Übersetzung
- Transfer-basierte Übersetzung
- Interlingua-basierte Übersetzung

