

**tecteam**

EVA TENSCHERT  
TECTEAM GMBH

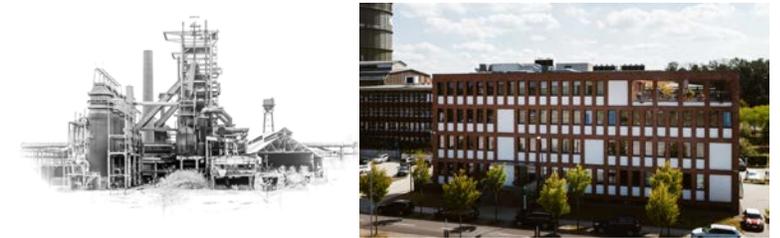
**PRE-EDITING**  
VORAUSSCHAUEN  
STATT  
NACHSEHEN

JUNI 2021



WWW.TECTEAM.DE

**tecteam**



## Dienstleistung

**Konzeption, Technische Redaktion, Technische Illustration, Übersetzungs-Management**

tecteam erstellt Ihre Technische Dokumentation.

Wir schreiben Bedienungsanleitungen, Betriebsanleitungen, Montageanleitungen und vieles mehr für Kunden aus unterschiedlichen Branchen.

Mit Technischer Redaktion und Technischer Illustration begleiten wir Ihre Redaktions- und Übersetzungsprozesse, führen Schwachstellenanalysen durch und erstellen Redaktionsleitfäden.



## Beratung

**CE-Prozess und EU-Richtlinien, Redaktionssysteme (CCMS), Texterstellung und Terminologearbeit, Automatisierung, Content Delivery, Redaktionsprozesse**

Bei tecteam erhalten Sie kompetente Beratung rund um die Kernthemen zeitgemäßer Technischer Redaktion.

Unsere Berater sind erfahrene Spezialisten und Mitglieder verschiedener Fachgremien.

Durch die Betreuung unterschiedlicher Kundenprojekte und die kontinuierliche eigene Fortbildung gewährleisten sie unabhängige Beratung auf dem aktuellen Stand von Wissenschaft und Technik.



## Bildung

**Studium Master of Science, Weiterbildung, tekam-Zertifikat, Technisches Volontariat, E-Learning, Seminare**

Das tecteam Bildungsinstitut ist Marktführer für die Aus- und Weiterbildung Technischer Redakteurinnen und Redakteure.

Mit unserem modularen Weiterbildungskonzept lassen sich Ausbildungs- und Qualifizierungswege individuell auf die Erfordernisse von Berufseinsteigern, Mitarbeitern und Unternehmen zuschneiden.

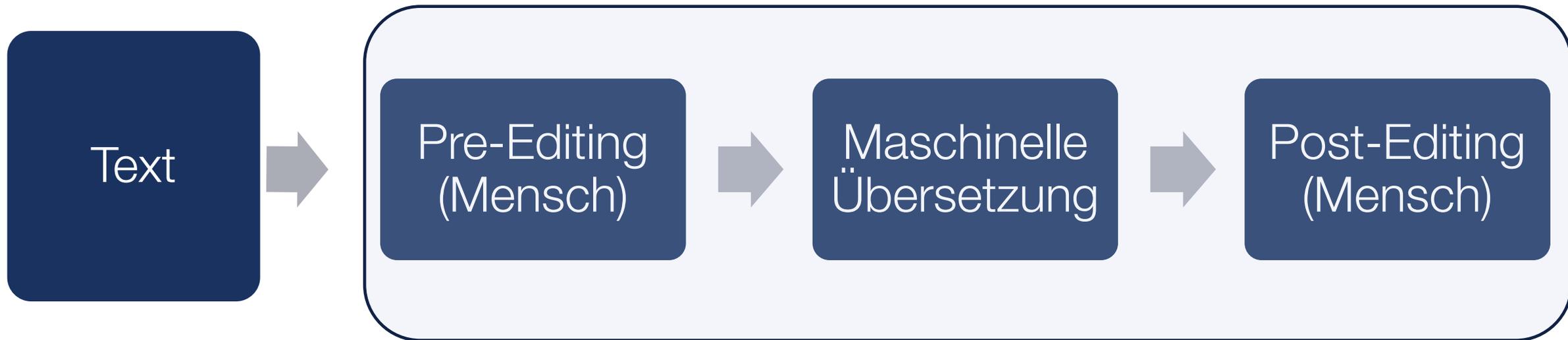
In der Personalentwicklung bieten wir damit in Deutschland einzigartige Möglichkeiten.

# ONLINE-REGIONALGRUPPENVERANSTALTUNG RHEIN-MAIN

## 10. JUNI 2021

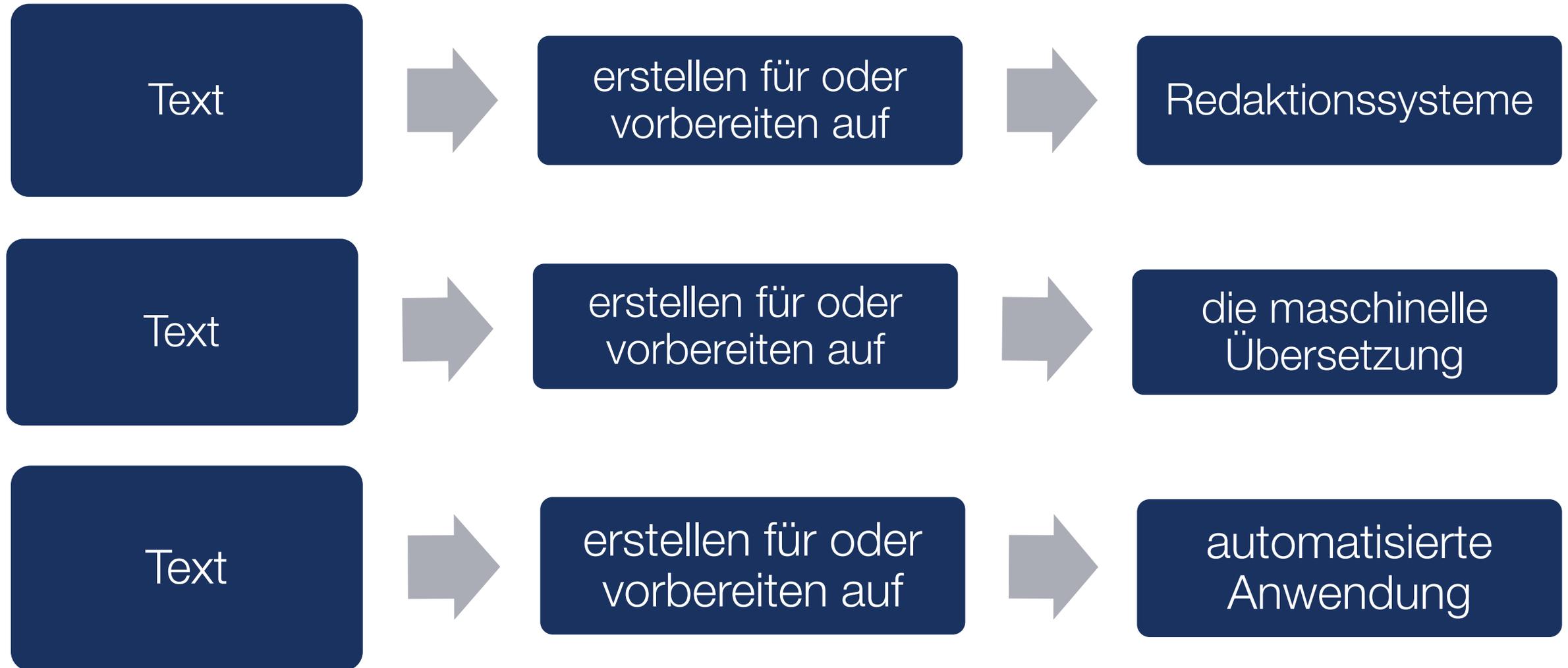
- Pre-Editing
  - Was bedeutet das Wort, was ist das und wozu dient es?
- Mensch – Text – Maschine
  - menschliche Leistungen
  - Textqualität, Merkmale, geeignet und nicht geeignet
  - maschinelle Leistungen (Wie „tickt“ und wie „lernt“ eine Maschine?)

# Was bedeutet das Wort „Pre-Editing“?

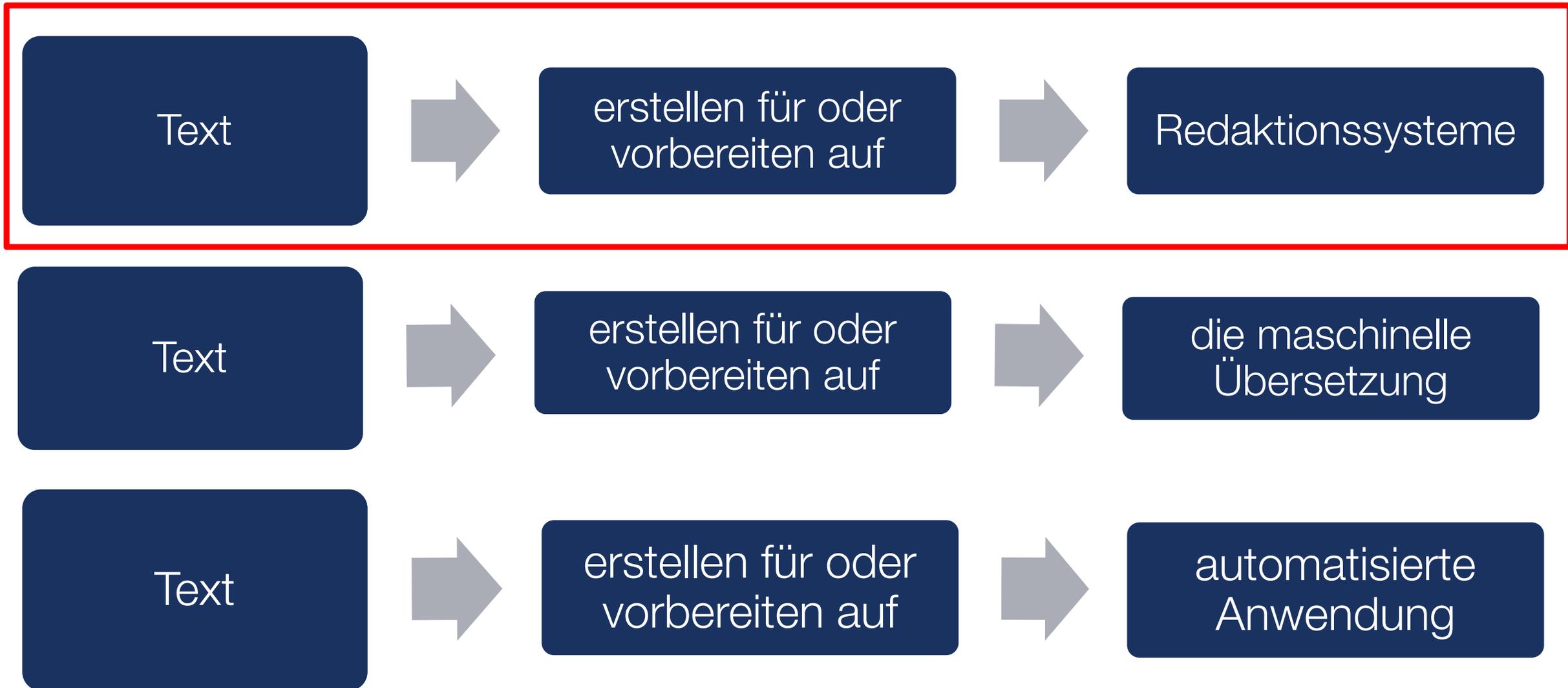


---

Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten



Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten



## Beispiel „Pre-Editing, Texte aufbereiten für ... ein Redaktionssystem“

Eine 20 Jahre alte Bedienungsanleitung im Stil „Genre Bleiwüste“ dient einem Redakteur als (eine!) Recherchequelle für die Informationseinheiten („Bausteine“, „information objects“, „Knoten“ ...), die in einem Redaktionssystem erfasst werden sollen.



„monolithische  
Quelle“



... modularisieren  
... u. v. m.



Redaktionssystem

---

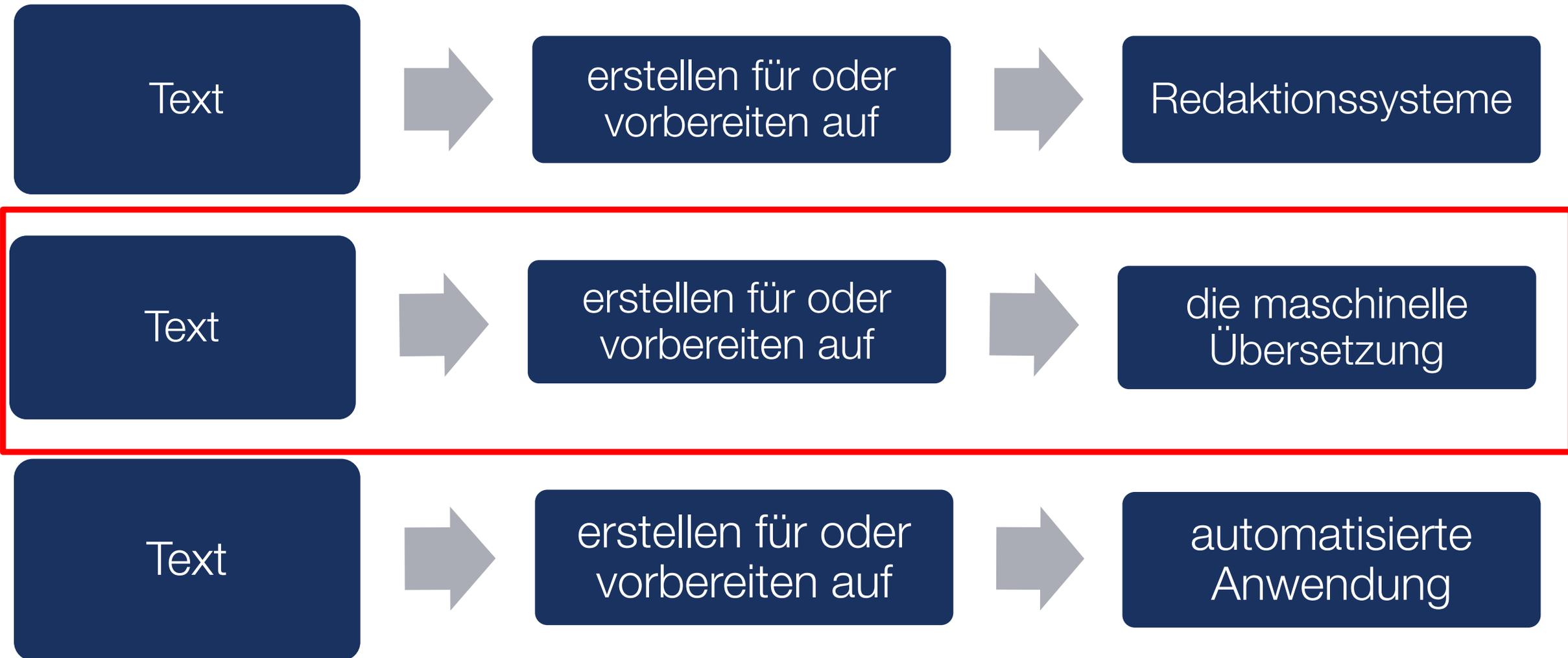
Editing und Pre-Editing für ein Redaktionssystem  
= Texte zielgerichtet erstellen/aufbereiten  
= typische Arbeiten eines ausgebildeten TR

standardisierte,  
strukturierte,  
modularisierte ...  
bereinigte Texte



in ein Redaktionssystem  
integrieren für Redakteure und  
für Anwender

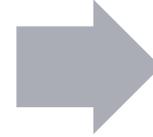
Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten



# Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten = Profession des Technischen Redakteurs

Ziele bei der Erstellung und Bearbeitung

- inhaltlich und formal korrekt (Normen berücksichtigt)
- mit Review-Prozessen optimiert
- qualitativ hochwertig und fehlerfrei
- übersetzungsgerecht
- terminologisch bereinigt
- strukturiert, standardisiert, gem. regelbasierter Sprache
- konkret statt vage und/oder abstrakt
- eindeutig im Ausdruck



Engines der MÜ  
Sprachspeicher für die MÜ

... letztlich, um verständliche  
und sinnvolle Ausgangstexte  
und ebenso gute  
Übersetzungen für den  
Menschen zu erreichen!

# Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten = Profession des Technischen Redakteurs\*

Ziele bei der Erstellung und Bearbeitung

- inhaltlich und formal korrekt (Normen berücksichtigt)
- mit Review-Prozessen optimiert
- qualitativ hochwertig und fehlerfrei
- übersetzungsgerecht
- terminologisch bereinigt
- strukturiert, standardisiert, gem. regelbasierter Sprache
- konkret statt vage und/oder abstrakt
- eindeutig im Ausdruck



Engines der MÜ  
Sprachspeicher für die MÜ

... letztlich, um verständliche  
und sinnvolle Ausgangstexte  
und ebenso gute  
Übersetzungen für den  
Menschen zu erreichen!

\*andere Bezeichnungen („Wildwuchs“): Informationsmanager, information experts, information professionals ...

# Studium, Aus- u. Weiterbildung zum technischen Redakteur

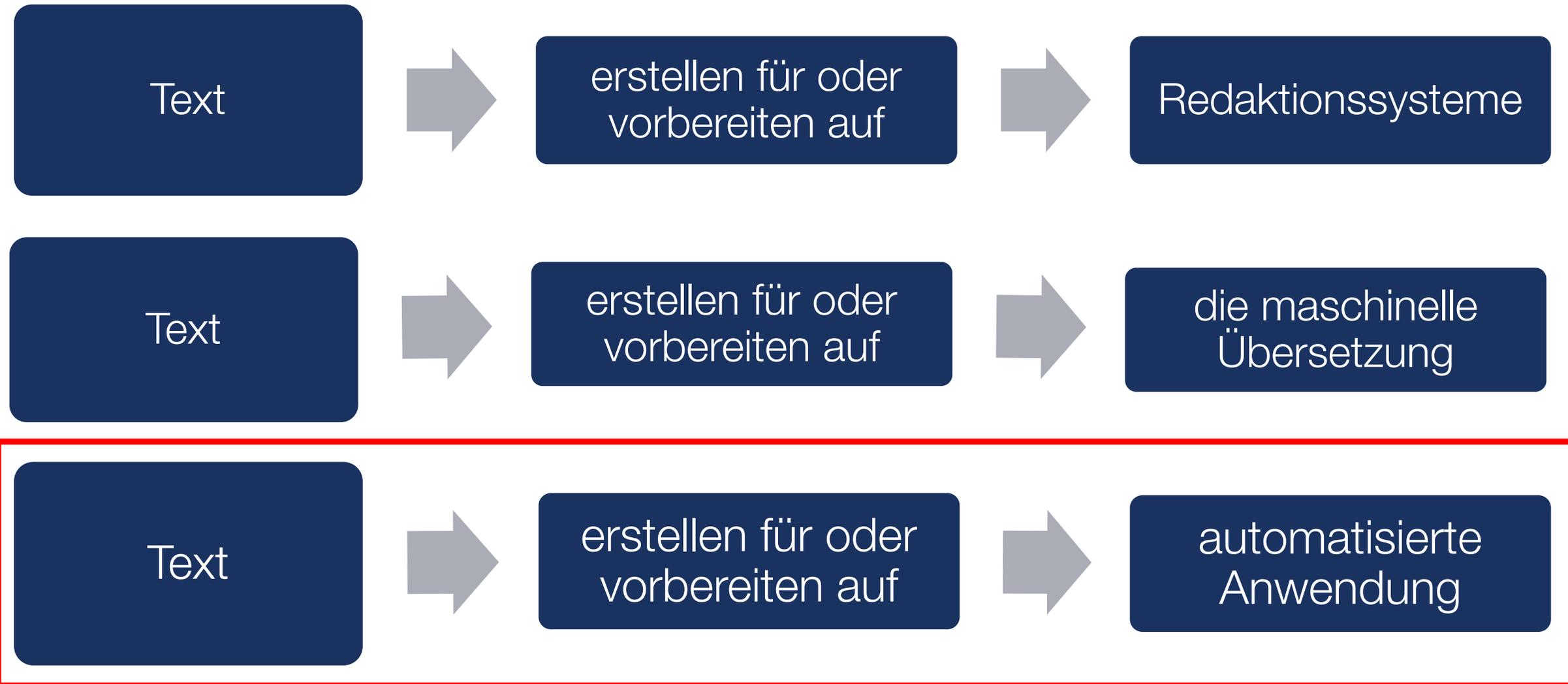
## Inhalte Ausbildung

- normative Anforderungen
- Redaktionsprozesse
- Erstellungsprozess einer Anleitung
  - Informationsarten, Funktionsdesign  
Grammatik, Stilistik
  - Schreibregeln
- Terminologie-Grundlagenwissen
- Methoden, Basis ...

## Befähigung zu / Training für

- inhaltlich und formal korrekt
- mit Review-Prozessen optimiert
- qualitativ hochwertig und fehlerfrei
- übersetzungsgerecht
- terminologisch bereinigt
- strukturiert, standardisiert, regelbasierte Sprache
- konkret statt vage und/oder abstrakt
- eindeutig im Ausdruck

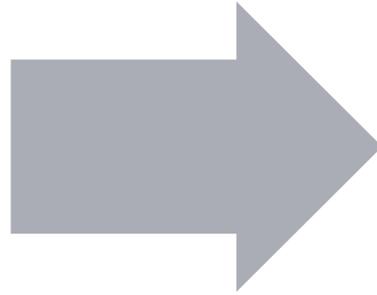
Pre-Editing = Texte zielgerichtet erstellen, vor- und aufbereiten



---

## Pre-Editing: Texte zielgerichtet erstellen, vor- und aufbereiten

alle bereits genannten  
Eigenschaften sowie  
umfangreiche  
Textgrundlagen  
(Korpus,  
Beispieltexte)



„automatisierte Texte“, automatisierte  
Anwendungen

## automatisierte Texte

Automaten-Texte, Display-Texte, Vorlagen für  
gesprochene Ansagen, Textvorlage für den  
Sprachspeicher zahlreicher  
Alltagsanwendungen

Wettervorhersagen-Texte,  
„Roboter-Journalismus“,  
Hilfe-Bots oder  
„KI-Assistenten“

eigene  
Texte, z.B.  
AB

**Technische  
Redakteure sind  
Profis für alle Texte,  
die funktional und leicht  
verständlich sein  
sollen!**



Texte für Systeme,  
Texte für die Automatisierung ...

Automaten-Texte, Display-Texte,  
Vorlagen für gesprochene Ansagen,  
Textvorlage für den Sprachspeicher  
zahlreicher Alltagsanwendungen

Wettervorhersagen-Texte,  
„Roboter-Journalismus“,  
Hilfe-Bots oder  
„KI-Assistenten“

eigene  
Texte, z.B.  
AB

**Alle Texte werden zunehmend in vielen Sprachen benötigt!**

**Autoren dieser Texte sind**

- **idealerweise mehrfach qualifiziert**
- **kooperieren mit qualifizierten Personen aus den Bereichen**
  - **Technische Kommunikation**
  - **Terminologie**
  - **Sprachwissenschaften, Translatorik**

**Idealerweise wissen diese Autoren auch, was eine KI bzw. eine maschinelle Übersetzung macht und was nicht!**



# MENSCH

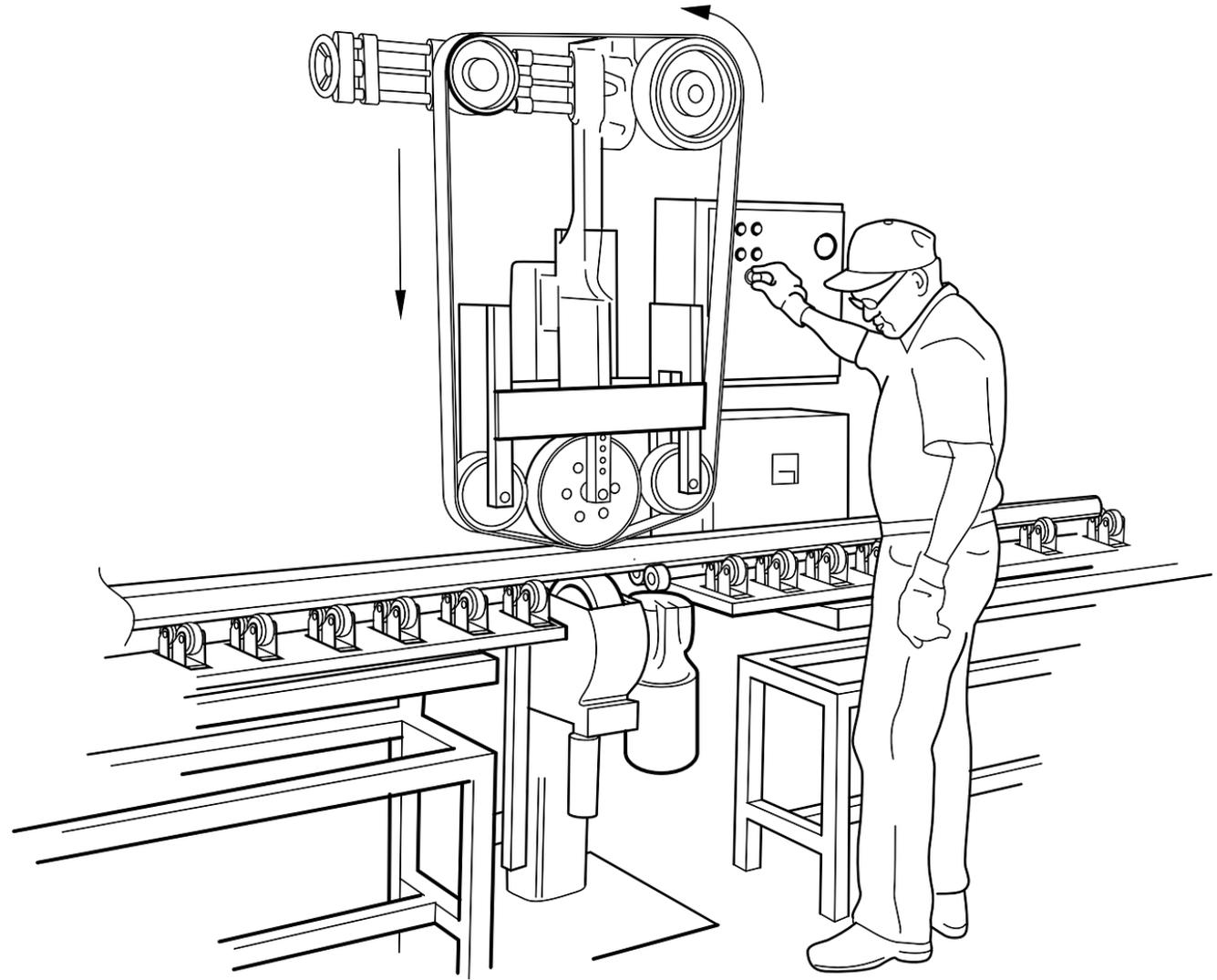
Bewusstsein:  
Inhalt, Sinn, Kohärenz  
Verstehen, „kreieren“,  
handeln, sinnvolle,  
menschliche  
Lösungen finden,  
kann ganzheitliche  
Entscheidungen  
treffen

# MASCHINE „KI“

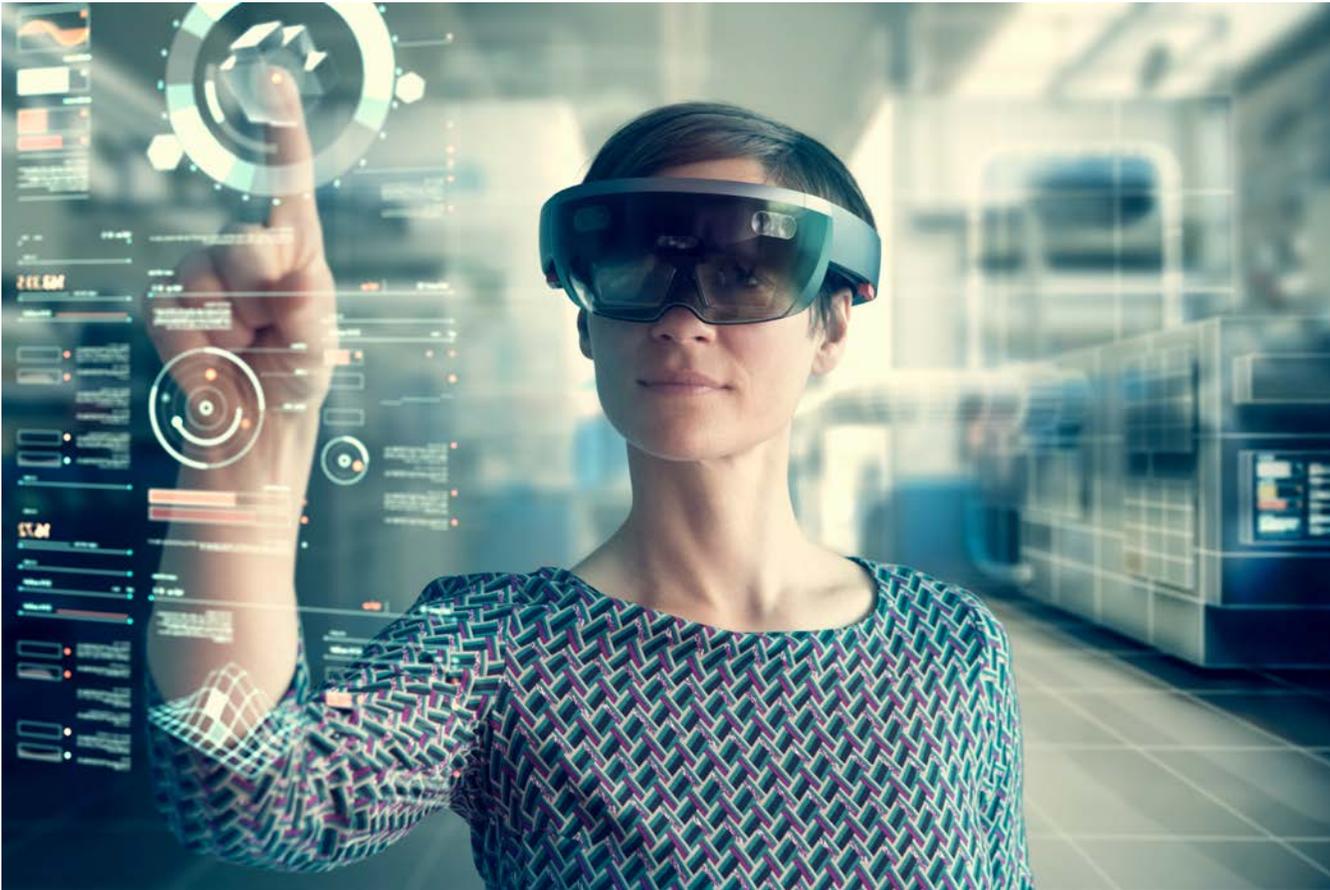
verarbeitet ohne  
Bewusstsein  
vorgefertigte Inhalte



# MENSCH UND MASCHINE

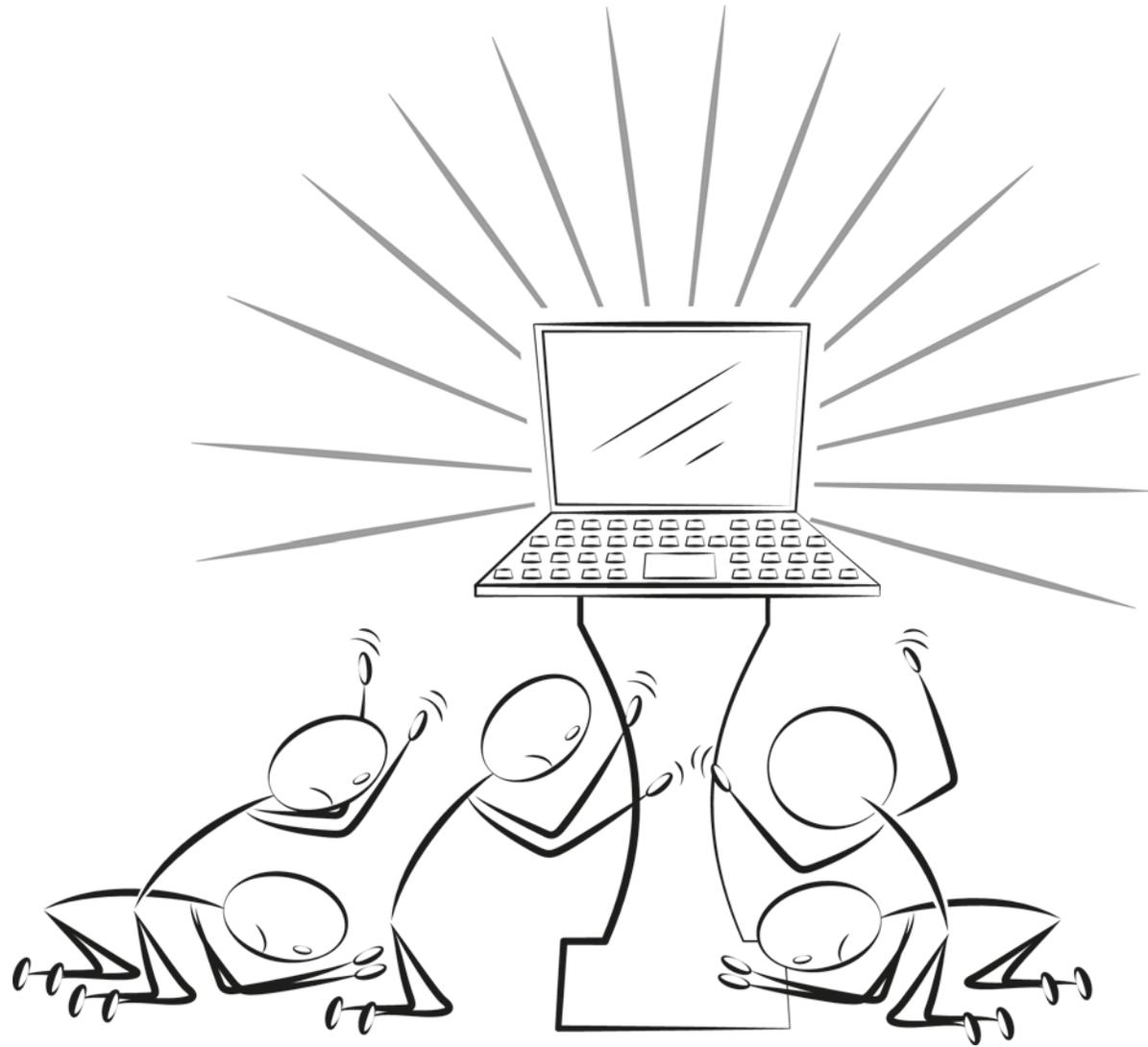


Texte sind von Menschen für Menschen.  
Technik, Programme, Tools ... sind nur Hilfsmittel.



- Erstellung
- Überarbeitung
- Übersetzung
- Überprüfung
- Automatisierung

Technik, Programme und  
Tools sind nur Hilfsmittel  
(und keine Götter)  
(oder andere Mythen).

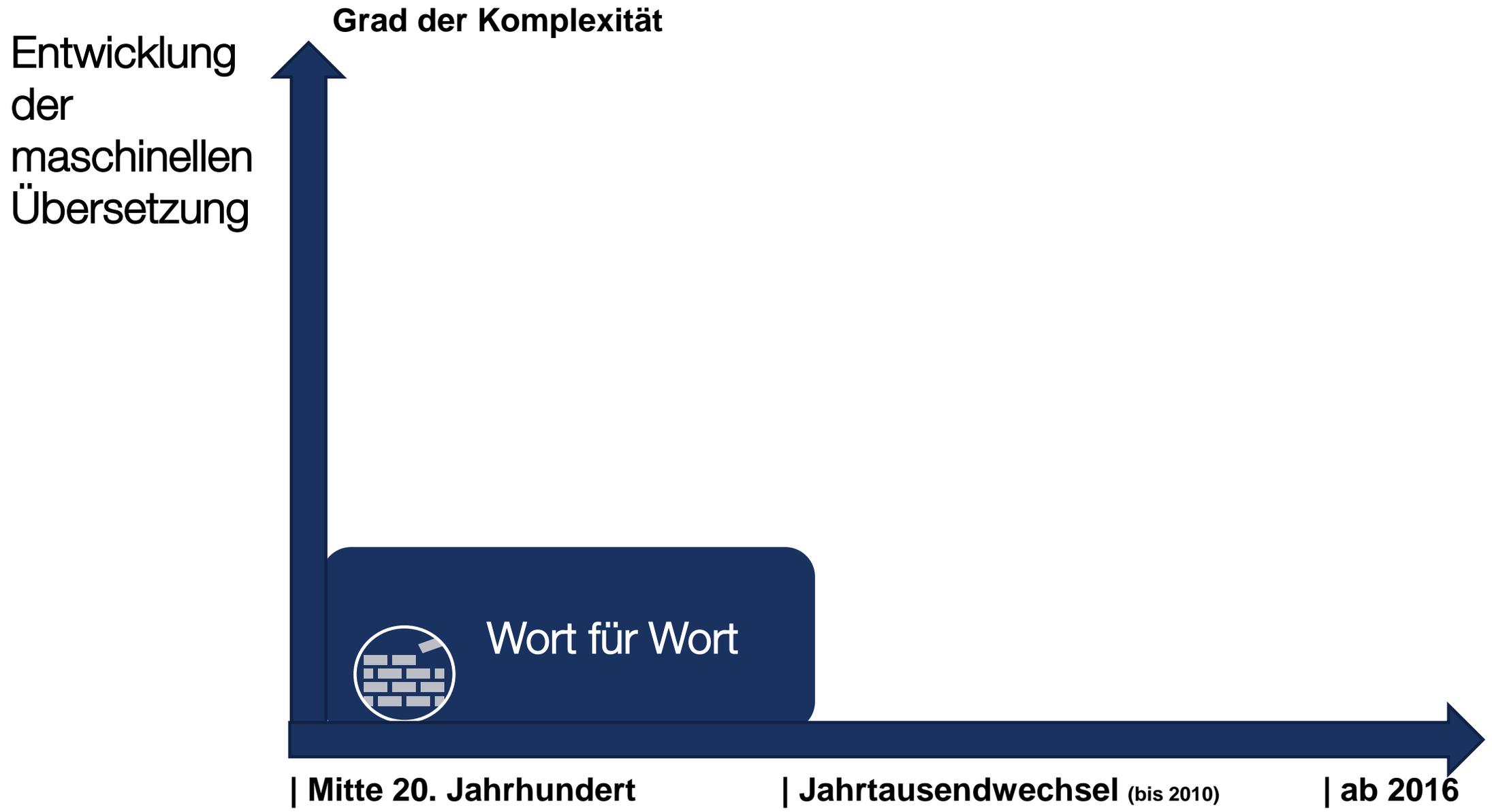


**WIE  
„TICKT“  
UND  
WIE  
„LERNT“  
EINE  
MASCHINE?**



# ÜBERSETZUNG





http://www.nuvaring.de



Ermächtigen Sie bitte Plätzchen für Loggen in zu dieser Stelle und erfrischen Sie diese Seite.

OK

Entwicklung  
der  
maschinellen  
Übersetzung

Grad der Komplexität



Wort für Wort



Segmente, Wort-Kombinationen



statistische Verfahren



Kombinationen, Strukturen,  
Segmente mit Metadaten  
(Kontextualisierung),  
Google: Interaktion mit Eingaben

Grundlage:  
Maschinelles Lernen

| Mitte 20. Jahrhundert

| Jahrtausendwechsel (bis 2010)

| ab 2016



Was ist die  
klassische  
Programmierung?



Was ist  
maschinelles  
Lernen?



Quelle:  
Machine Learning Zero to Hero Google I/O 19  
Laurence Moroney and Karmel Allison

Trainingsphase  
für das  
maschinelle  
Lernen

ERGEBNIS,  
ANTWORTEN

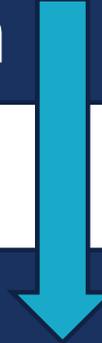


DATEN



maschinelles  
Lernen

REGELN



*Neuronales* Netzwerk

Quelle:  
Machine Learning Zero to Hero  
Google I/O 19  
Laurence Moroney and Karmel Allison

Trainingsphase  
für das  
maschinelle  
Lernen

ERGEBNIS,  
ANTWORTEN



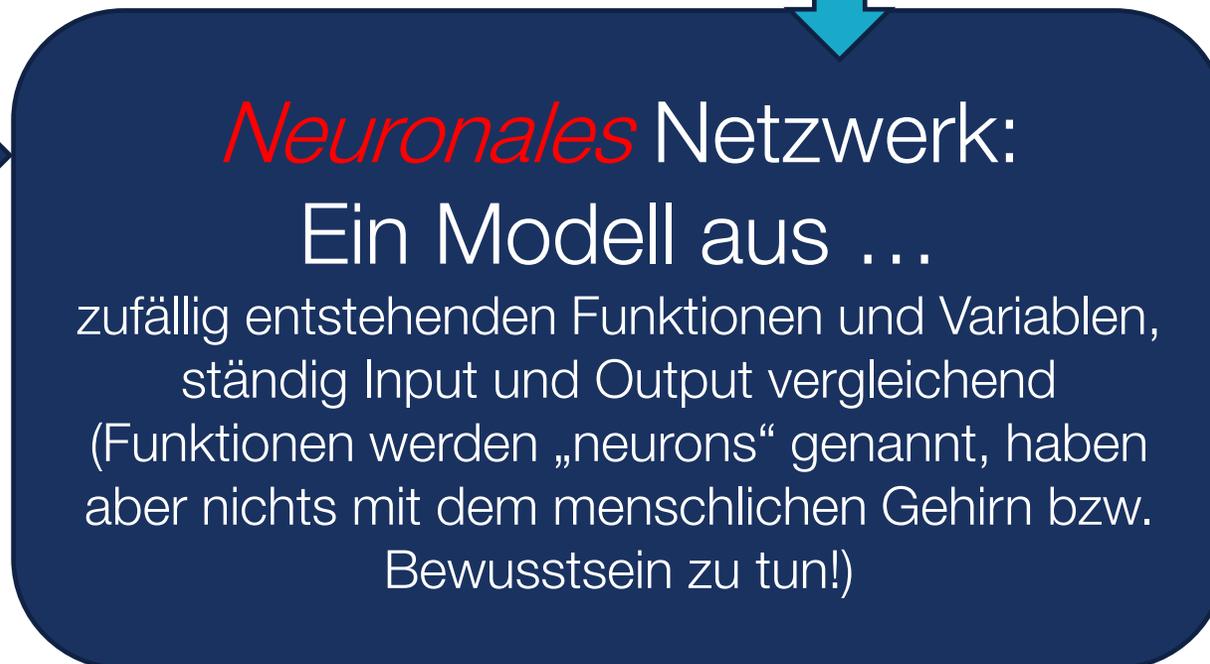
DATEN



REGELN



DATEN



*Neuronales* Netzwerk:

Ein Modell aus ...

zufällig entstehenden Funktionen und Variablen,  
ständig Input und Output vergleichend  
(Funktionen werden „neurons“ genannt, haben  
aber nichts mit dem menschlichen Gehirn bzw.  
Bewusstsein zu tun!)

VORHERSAGEN



Quelle:  
Machine Learning Zero to Hero  
Google I/O 19  
Laurence Moroney and Karmel Allison

# DAMALS UND HEUTE

## ■ Artikel Maschinelle Sprachübersetzung in Europa

<http://www.flc.kyushu-u.ac.jp/~de/ddj/Archiv/genko.html>

Zeit der ersten maschinellen Übersetzungen

- Übersetzung für „Armbanduhr“ lautete damals noch „poor tape clock“

## ■ (Mindest)segmentzahl

- bedeutet: Mindest-Input an Daten, Ergebnissen, Antworten für eine Engine der MÜ
  - bedeutet auch: bei seltenen Sprachpaaren oder sehr wenig Material keine MÜ!
- Minimum an Segmenten? Einige zehntausend? Hunderttausend? Je nach Inhalt!
- Google: verfügt über Milliarden Segmente, die „in Echtzeit aufgefüllt werden“, und zwar durch jede Anfrage und Eingabe (heutige Übersetzung für „Armbanduhr“: „[wrist-]watch“)
  - Stimmt die Gleichung: „hat viel Textmaterial = weiß viel, ist klüger oder besser“?

DATEN



Texte  
Wortlisten  
Korpus

maschinelle Übersetzung, MÜ  
machine translation, MT  
neuronal MT, NMT



VORHERSAGEN

„geratene“  
Übersetzung, die für  
eine Maschine gemäß  
formaler Regeln richtig  
erscheint, die  
allerdings hinsichtlich  
des Sinns und Inhalts  
nicht von der  
Maschine geprüft  
werden kann



Daten  
(nur Text,  
kein „Sinn“)



VORHERSAGEN  
„geratene“  
Übersetzung  
ohne Sinnprüfung



# TERMINOLOGIEARBEIT – ZUERST DAS WISSENSMANAGEMENT, DANN DER TEXT

- Zitat Prof. Kageura:  
„Terms are attributed to terminology and thus the system of knowledge of the domain, not to the texts produced in the domain.“

Quelle: Prof. Kyo Kageura, Universität Tokyo

Vortrag: The status of terms and concepts in the learned use of language

Invoking the Wüsterian spirit in the era of machine learning

DTT Symposion 2021

# TERMINOLOGIEARBEIT – ZUERST DAS WISSENSMANAGEMENT, DANN DER TEXT

- Zitat Prof. Kageura:  
„Terms are attributed to terminology and thus the **system of knowledge of the domain**, not to the texts produced in the domain.“

Quelle: Prof. Kyo Kageura, Universität Tokyo

Vortrag: The status of terms and concepts in the learned use of language

Invoking the Wüsterian spirit in the era of machine learning

DTT Symposion 2021



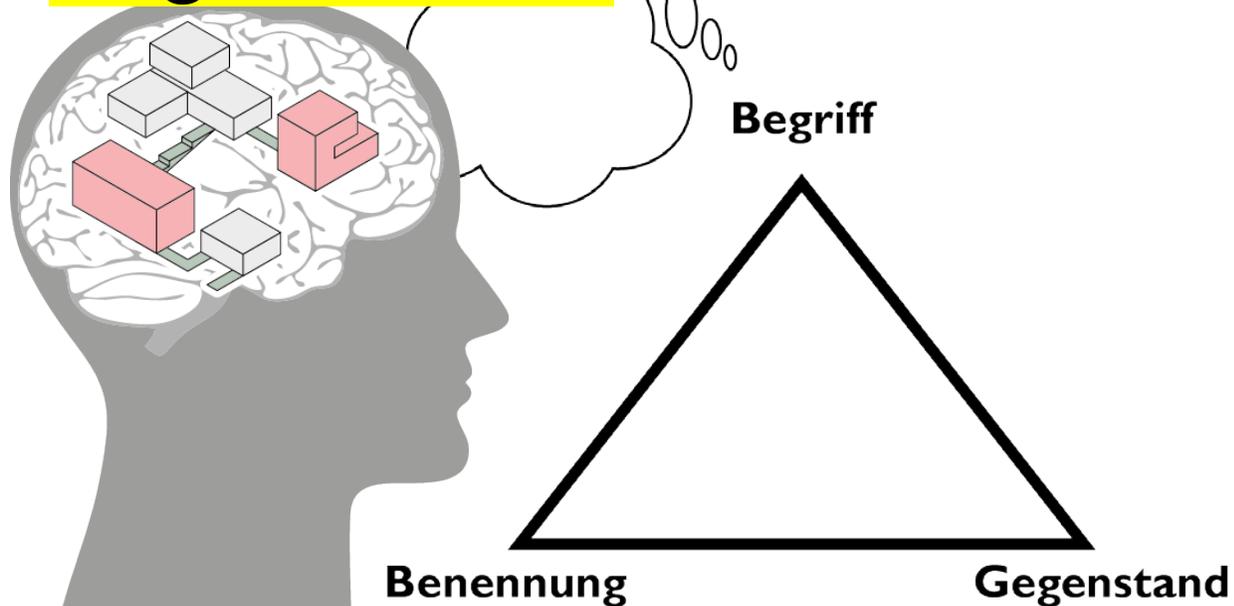
## WISSENSSYSTEM (KNOWLEDGE SYSTEM)

ÜBEREINKUNFT, ANGLEICHUNG, ABSTIMMUNG ...  
VON VORSTELLUNGEN UND BEGRIFFEN  
(HARMONIZATION OF CONCEPTS)



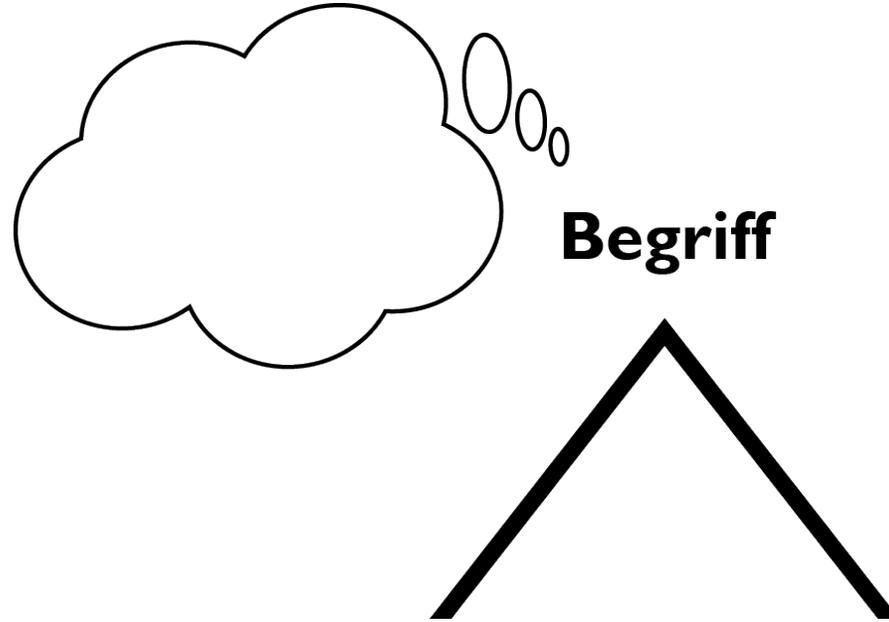
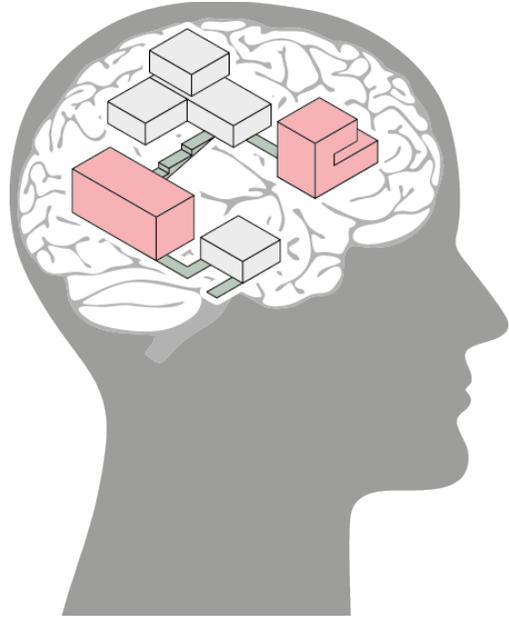
# **Begriff** (englisch: concept)

## **Begriffsebene**



Vorstellungen, Ideen,  
im Kopf organisiertes  
**Wissen**

sich über Ideen und  
Vorstellungen  
austauschen



----- Semiotik -----

----- Semantik -----

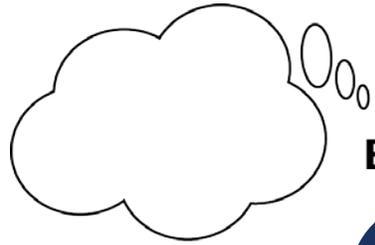
Eine „künstliche Intelligenz“ ist  
auf einer formalen Ebene;  
Wortebene, Satzebene, strukturelle Ebene ...



**Wortebene**  
**Satzebene,**  
**formale Ebene**

└─  
**Benennung**

Zusammenarbeit von menschlichen  
„Wissensorganisatoren“ aus verschiedenen Bereichen



**Begriff**

Kenntnisse  
und Praxis  
in der  
Redaktion

Redakteure  
Redaktion, Taxonomien,  
Terminologie, Ontologien

Ontologen

Terminologen

Fachexperten  
und  
Sprachexperten

→ Strukturen,  
die für eine KI  
„verträglich“  
sind und „Sinn“  
transportieren?



## Literaturtipp zum Vertiefen

Zeitschrift „edition“ des Deutschen Terminologie-Tag e. V.  
online erhältlich, siehe [www.dttev.org](http://www.dttev.org)

edition, Ausgabe 2/17, Terminologie und KI

edition, Ausgabe 1/18, Terminologie und strukturierte Informationen

edition, Ausgabe 2/18, Terminologie in der maschinellen Übersetzung

edition, Ausgabe 2/20, Terminologie bei der Live-Untertitelung

edition, Ausgabe 1/21, Neuronale Maschinelle Übersetzung

**MASCHINEN  
UND  
KI  
ANGEMESSEN  
NÜTZEN**



# MERKMALE EINES GEEIGNETEN TEXTES

- inhaltlich korrekt, sachlich-logisch und formal-logisch korrekt, das heißt kohärent und konsistent
- mithilfe regelbasierter Sprache erstellt und nachweislich möglichst fehlerfrei (Reviews, Lektorate)
- nach funktionalen Aspekten strukturiert und je nach Erfordernis modularisiert
- konkret, eindeutig, einfach und prägnant hinsichtlich Grammatik und Wortwahl
- übersetzungsgerecht formuliert
- mit sprachneutralen Bildern ergänzt
- terminologisch bereinigt = zuerst das Wissensmanagement und Terminologearbeit, erst dann die Texterstellung

# MERKMALE EINES GEEIGNETEN SATZES

- kurz – ist also nicht in mehrere hypotaktische Segmente aufgeteilt
- idealerweise nur mit einer oder zwei Aussagen „befüllt“  
(Grundwissen Grammatik: Satzgegenstand ist das Subjekt, Satzaussage ist das Prädikat)
- im Aktiv (Passiv vermeiden)
- im Indikativ (Möglichkeitsform vermeiden)
- positiv formuliert (Negationen vermeiden)
- einer Informationsart zuordenbar (Instruktion und Sicherheit z. B. nicht vermischen)
- falls möglich und zielführend: in leichter Sprache verfasst

## BEISPIEL

### GEEIGNETE SÄTZE

VERSUS

### NICHT GEEIGNETER SATZ



- (...)
- (...)
- Füllen Sie Wasser in den Wassertank.
- Schalten Sie die Kaffeemaschine ein.

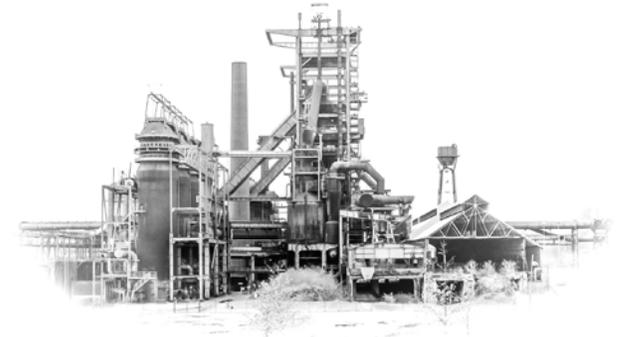


- Es kann die Befüllung des Behälters mit Wasser vor oder nach der Befüllung mit Kaffeepulver erfolgen, sofern stets der abgeschaltete Zustand vor oder während der Wasserzufuhr eingehalten wurde wegen der Verbrühungsgefahr.

VIELEN DANK FÜR  
IHRE  
AUFMERKSAMKEIT!



TECTEAM GMBH  
EVA TENSCHERT  
ANTONIO-SEGNI-STR. 4  
44263 DORTMUND



TELEFON: 0231 9206 184  
E.TENSCHERT@TECTEAM.DE  
WWW.TECTEAM.DE