

Häufige Fehler maschineller Übersetzungssysteme

Hauptkategorie	Fehlertyp	Beschreibung	Beispiel
Syntax	Grammatikfehler	zum Beispiel sprachabhängige Fehler aufgrund des Kontextes	DE: „Er hat es ihr gegeben“ FR: „Il le lui a donné“ anstatt „Il la lui a donné“ (männlich „es = le“ statt weiblich „es = la“)
	Falsche oder vergessene Zeichensetzung	In manchen Satzkonstruktionen verursachen sie schwere Fehler. Siehe dieses klassische Beispiel:	DE: Wir essen, Opa EN: We eat grandpa
Wortschatz	Falsches Synonym	Einsatz eines Wortes mit anderer Bedeutung im Satz	EN: 1. Open the lid of the <u>battery box</u> . 2. Remove the old battery from the <u>box</u> DE: 1. Öffnen Sie den Deckel des <u>Batteriekastens</u> . 2. Nehmen Sie die alte Batterie aus dem <u>Karton</u> .
	Fachterminologie	Falsche oder inkonsistente Verwendung der Terminologie	EN: How expensive is it to train an <u>MT engine</u> ? DE: Wie teuer ist die Ausbildung eines <u>MÜ-Triebwerks</u> ?
	Präzision	Die MÜ ist genauer bzw. allgemeiner als die Originalaussage.	DE: In Elektroautos werden verschiedene Arten von Lithium-Ionen-Batterien verwendet. EN: Various types of lithium-ion batteries are used in electric vehicles.
	Jargon	Oft werden umgangssprachliche Benennungen aus der Technik von MÜ nicht verstanden	DE: Er fährt einen <u>Ami</u> mit OD (er fährt einen <u>amerikanischen Wagen</u> mit Overdrive) EN: He drives a <u>Yank</u> with OD
	Neues Wort	Für Wörter, die nicht trainiert wurden oder neu sind, versucht das System, manchmal vergeblich, eine Übersetzung zu generieren, da auch Tokens unterhalb der Worteinheit gelernt wurden.	DE: Verbrannt, verkohlt und ausgemerkelt (Buchtitel über Kanzlerschaften mit Bezug auf Brandt, Kohl und Merkel) EN: Burnt, charred and pitted
	Rechtschreibfehler	Meistens werden die von der MÜ-Engine erkannt und korrigiert, aber ...	DE: Zwei Themen wurden in die engere <u>Wal</u> gezogen. EN: Two topics were dragged into the narrower <u>whale</u> .
Stil / Sprachebene	Rollen- und Geschlechtszuweisung	Klassisches und aktuelles Thema, das stark von den Trainingsdaten (kuratierte Daten oder nicht) abhängt.	EN: The <u>manager</u> informed the <u>nurses</u> about the duty schedule via email. DE: Der <u>Manager</u> informierte die <u>Krankenschwestern</u> per E-Mail über den Dienstplan.
	Falsches/unangemessenes Sprachregister	Duzen, Gender, Regionalismus	EN: He said to the 50-year-old: Your father is there. DE: Er sagte zu dem 50-Jährigen: <u>Dein Vater</u> ist dort. (bei <u>your son</u> wird gesiezt)
	Übernahme des Satzbaus der Ausgangssprache	Weniger geläufige Konstruktionen in der Zielsprache	DE: Rückwirkend soll die [Preis]Bremsen dann auch für Januar und Februar gelten. EN: Retroactively, the brake will then also apply to January and February.
Verständnis	Bezugswort	MÜ erkennt Bezüge aus vorherigen Aussagen nicht immer korrekt. Die Netzwerkkonstruktion (RNN, CNN, Transformer) beeinflusst hier das Verhalten der Engine.	EN: Human language is very complex. <u>It</u> is often ambiguous. DE: Die menschliche Sprache ist sehr komplex. <u>Es</u> ist oft mehrdeutig.
	Weltwissen	Das kommt zum Beispiel bei Nachrichtentexten vor.	DE: Die Ampel möchte Mehrstaatlichkeit künftig für alle erlauben. EN: The traffic light would like to allow multiple states for everyone in the future.
	Kontext	Übersetzung ist für den Kontext nicht geeignet. Das kommt zum Beispiel bei Überschriften vor.	EN: Clean training data (Überschrift) DE: Saubere Trainingsdaten (anstatt: Trainingsdaten bereinigen)
	Bedeutung	Bedeutung nicht erkannt	DE: Vonseiten der Union hagelt es Kritik. EN: The CDU/CSU has been criticized.
	Redewendungen falsch umgesetzt	Eine Redewendung wird wörtlich übersetzt	DE: Er hat ein Brett vor dem Kopf EN: He has a board in front of his head (he is a blockhead)

Hauptkategorie	Fehlertyp	Beschreibung	Beispiel
Textänderung	Hinzufügen von Inhalten	Erfolgt aufgrund der Wahrscheinlichkeitsberechnung des Modells für das nächste Wort.	DE: Wenn die Maschine ausreichend abgekühlt ist, erlischt die gelbe Leuchtdiode, und die Maschine kann wieder eingesetzt werden. EN: When the machine has cooled down sufficiently, the yellow LED goes out and the new machine can be used.
	Auslassen von Inhalten	ebd.	EN: Suitable for monitoring of the bearing <u>temperature</u> of electrical motors. DE: Besonders geeignet zur Lagerüberwachung elektrischer Motoren.
	Zahlenfehler	MÜ-Systeme können bspw. Zahlen vergessen oder inkonsistent umrechnen. Auch mit dem Datum haben sie manchmal Probleme.	EN: Battery capacities of current EVs range from a mere 17.6 kWh in the Smart EQ ForTwo with a range of just 58 miles, up to 100 kWh in the Tesla Model S and Model X that can run for over 300 miles before needing a charge. DE: Die Batteriekapazität aktueller Elektrofahrzeuge reicht von nur 17,6 kWh im Smart EQ ForTwo mit einer Reichweite von nur 94 km bis zu 100 kWh im Tesla Model S und Model X, die über 300 Meilen fahren können, bevor sie aufgeladen werden müssen.
	Formatierungsfehler	Es kann vorkommen, dass Tags oder Formatierungsinformationen während der Übersetzung ausgelassen werden.	

TAB. 01 QUELLE François Massion