
Auf lange Sicht

Sie ist hübsch. Er ist stark. Er ist Lehrer. Sie ist Kindergärtnerin

Google Translate übersetzte Texte gerne so, dass Klischees zementiert werden. Der Konzern gelobte Besserung. Die Republik hat den Test gemacht.

Von Marie-José Kolly und Simon Schmid, 19.04.2021

Tippt man bei Google Translate ins erste Feld *the nurse*, so erscheint im zweiten: die Krankenschwester. Und nicht: der Krankenpfleger. Obwohl das englische Wort *nurse* beides bezeichnen kann, eine Frau oder einen Mann.

Klischiert, aber gerade noch so erträglich? Schauen wir noch ein Beispiel an.

Stellt man das erste Feld auf Deutsch um und tippt dort «der Krankenpfleger», so erscheint im zweiten auf Französisch: *l'infirmière*. Die Krankenpflegerin.

Das ist nun erstens wirklich falsch und zweitens sexistisch – denken Menschen. Anders funktionieren die Algorithmen, die mit Daten – vielen Millionen Texten aus dem Internet – trainiert wurden. Sie reproduzieren die Muster und damit auch die Verzerrungen, *biases*, die sie in den Daten vorfinden. Und ein solches Muster ist eben: Pflegepersonal, das sind häufiger Frauen. So entsteht der Gender-Bias.

Google Translate und andere Algorithmen übersetzen immer schneller und besser. Wie gut sind sie aber mittlerweile darin, diesen Bias loszuwerden?

Der Algorithmus tut, als wüsste er es

Besonders leicht lässt sich das an Übersetzungen aus dem Finnischen nachzeichnen. Denn diese Sprache kennt nur ein einziges Pronomen, um in der Einzahl über Drittpersonen zu sprechen: *hän*. Das Wort steht für er, sie, es, kann also Frauen, Männer, Buben, Mädchen bezeichnen.

Finninnen unterläuft deshalb oft ein typischer Fehler: Sie sagen auf Englisch etwa *she* und zeigen dabei auf einen Mann. Oder sprechen von *him*, meinen dabei aber eine Freundin, weil sie es nicht gewohnt sind, dass sie im Personalpronomen ein Geschlecht ausdrücken müssen.

Auch automatische Übersetzer müssen *hän* ein Geschlecht zuweisen, wenn sie das Wort in eine Sprache wie Deutsch übertragen. Und auch sie treffen oft seltsame Entscheidungen.

Eine Kostprobe aus Google Translate:

Hän on opettaja. Hän on lastentarhanopettaja. Er ist ein Lehrer. Sie ist Kindergärtnerin.

Dabei funktionieren *opettaja* und *lastentarhanopettaja* ähnlich wie *nurse* auf Englisch: Weder das Wort selbst noch seine grammatischen Eigenschaften weisen auf ein bestimmtes Geschlecht hin. Wird es nicht aus dem Kontext klar, weiss man schlicht nicht, ob es einen Mann oder eine Frau bezeichnet.

Aber der Übersetzungsalgorithmus tut so, als wüsste man es.

Und er tut auch so – das zeigen weitere Übersetzungsversuche –, als wäre stets klar, dass er Athlet ist, sie Tänzerin. Er Tierarzt, sie Tiersitterin. Er Zahnarzt, sie Dentalhygienikerin. Er Busfahrer, sie Psychologin.

Wie umfassend ist dieser Gender-Bias?

Um dies herauszufinden, haben wir hundert finnische Sätze nach dem obigen Schema – *Hän on opettaja* – bei Google Translate eingegeben. Die 100 Sätze umfassen in der Schweiz häufig ausgeübte Berufe und ein paar weitere Beschäftigungen. Das Ergebnis: Nur bei 20 Berufen spuckte der Algorithmus auf Deutsch ein weibliches Personalpronomen aus. 80 Berufe und damit der Grossteil wurden in der Übersetzung dagegen als männlich eingestuft.

Berufsleute sind meistens Männer

Anzahl der Übersetzungen nach Geschlecht



Quelle: Eigene Analyse, [Google Translate](#).

Zur vollständigen Liste: Wie die Berufe übersetzt wurden

Er ist Sekretär.
Er ist Verkäufer.
Er ist Krankenpfleger.
Sie ist eine Putzfrau.
Sie ist Kellnerin.
Er ist ein Techniker.
Er ist ein Lehrer.
Sie ist Kindergärtnerin.
Er ist ein Bauer.
Sie ist Sozialarbeiterin.
Er ist Koch.
Er ist ein Zimmermann.
Er ist Programmierer.
Er ist Buchhalter.
Er ist ein Mechaniker.
Er ist Elektriker.
Er ist Zeichner.
Er ist ein Zinnschmied. (Wortwörtliche Übersetzung von «Spengler»)
Er ist ein Architekt.
Sie ist eine Friseurin.
Er ist Wirtschaftsprüfer.
Er ist Berater.
Sie ist eine Ärztin.

Sie ist Arzthelferin.
Er ist Gärtner.
Er ist zehn. (*Kymppi* heisst einerseits «Geschäftsführer», andererseits ist es ein umgangssprachliches Wort für *kymmenen*, «zehn». Der Übersetzungsalgorithmus hat sich, trotz dem Kontext im Satz, für Letzteres entschieden.)
Er ist Maurer.
Er ist LKW-Fahrer.
Sie ist Babysitterin.
Er ist ein Maler.
Er ist ein Ingenieur.
Er ist ein Polizeibeamter.
Er ist ein Anführer.
Er ist ein Immobilienmakler.
Er ist Anlageberater.
Er ist ein Trainer.
Sie ist Physiotherapeutin.
Er ist Professor.
Er ist ein Anwalt.
Er ist Bäcker.
Sie ist Konditorin.
Sie ist Zahnhygienikerin.
Er ist Uhrmacher.
Sie ist eine Kosmetikerin.
Sie ist Psychologin.
Er ist ein Busfahrer.
Er ist Installateur.
Er ist ein Journalist.
Er ist Grafiker.
Er ist Rezeptionist.
Er ist Maschinenbediener.
Er ist Kassierer.
Er ist Chemiker.
Er ist ein Musiker.
Er ist ein Anwalt.
Er ist Wirtschaftsanalytiker.
Er ist ein Reiseführer.
Er ist ein Postbote.
Er ist Metzger.
Er ist Innenarchitekt.
Er ist ein Rapper.
Er ist Statistiker.
Er ist Elektronikinstallateur.
Er ist Polizeikommandant.
Er ist Medizintechniker.
Er ist Lokführer.
Er ist wissenschaftlicher Mitarbeiter.
Er ist Apotheker.
Er ist Biologe.
Er ist ein Handwerker.
Er ist ein Kaplan.
Er ist ein Zahnarzt.
Er ist Übersetzer.
Er ist Dolmetscher.
Er ist Designer.
Er ist Forscher.
Er ist Barkeeper.
Sie ist Bibliothekarin.
Er ist Optiker.
Er ist Filmregisseur.
Er ist ein Fotograf.
Er ist Dachdecker.
Er ist Zollbeamter.

Er ist Förster.
Er ist Schneider.
Er ist Banker.
Sie ist Schauspielerin.
Er ist Unternehmer.
Sie ist Hebamme.
Sie ist eine Tänzerin.
Er ist ein Athlet.
Er ist Administrator.
Sie ist Logopädin.
Sie ist Sanitäterin.
Er ist ein Künstler.
Er ist ein Schriftsteller.
Er ist ein Pilot.
Er ist ein Zimmermann.
Sie ist eine Tiersitterin.
Er ist Tierarzt.

Wie entsteht der Gender-Bias bei Google Translate? Streng genommen ist nicht der Algorithmus das Problem – sondern das Material, mit dem er gefüttert wird. Also die Daten, mit denen das Übersetzungsprogramm «trainiert» wird. Dieses Material besteht aus Texten, die eben nicht gender-neutral sind, sondern Stereotype, Klischees, Verzerrungen enthalten.

In der Folge ist auch die Übersetzung *biased*: Ihr Output orientiert sich an dem, was die Trainingsdaten als «Standard-Mensch» kennen (etwa: weiss, männlich, heterosexuell). Und nicht zwingend an dem, was man als Nutzerin eingegeben hatte oder als wünschbares Ergebnis erachten würde.

Deshalb kommt «Sie ist ...» in den übersetzten Sätzen nicht nur viel seltener vor als «Er ist ...». Die weibliche Form steht auch neben ganz bestimmten Berufsbezeichnungen. Nicht beim Ingenieur, dem Schriftsteller, dem Professor. Auch nicht neben dem Metzger, Maurer oder Polizeibeamten. Sondern neben der Logopädin, der Bibliothekarin, der Friseurin.

Nur sporadisch weichen die Übersetzungen von den Berufstereotypen ab. Zum Beispiel übersetzt Google Translate aus dem Finnischen: Sie ist Ärztin. Er ist Krankenpfleger. Und nicht: Er ist Arzt. Sie ist Krankenpflegerin.

Wie dies zustande kommt, ist unklar. Interessanterweise kehren sich die Geschlechter um, wenn man den Ausgangssatz nicht in der dritten Person, sondern in der ersten Person formuliert: mit *minä* (ich) statt mit *hän* (er, sie).

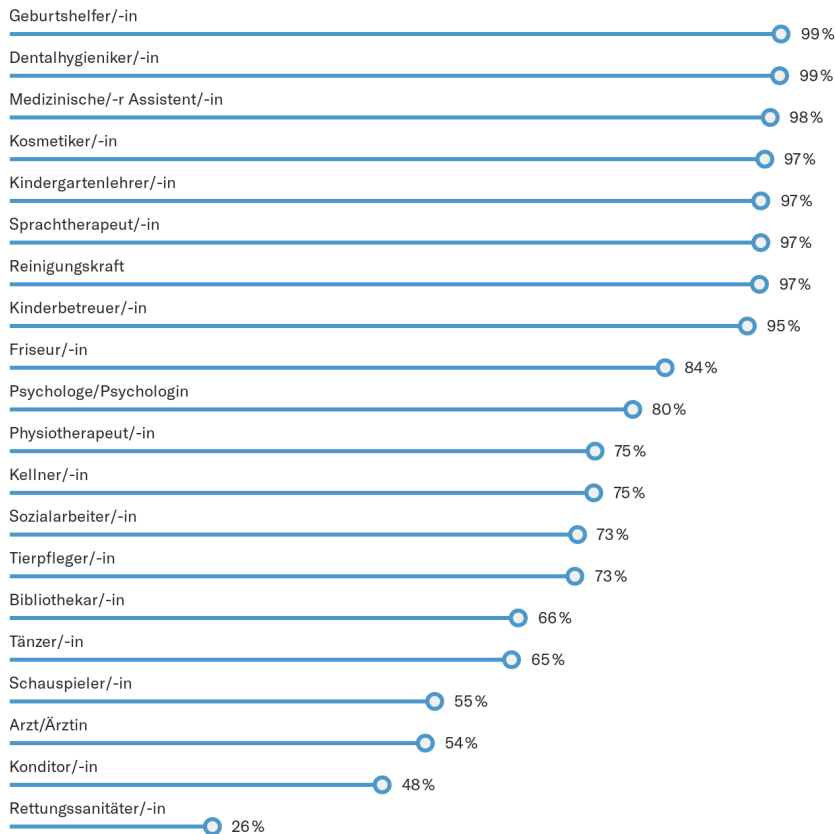
Minä olen lääkäri. Minä olen sairaanhoitaja. Ich bin ein Arzt. Ich bin eine Krankenschwester.

Frauen betreuen, assistieren und putzen

Es gibt Statistiken, die den Übersetzungen von Google-Translate eine gewisse Plausibilität verleihen. So zeigt sich etwa: Die meisten Berufe, denen der Algorithmus ein weibliches Pronomen zuteilt, werden in der Schweiz effektiv häufiger von Frauen ausgeübt. Beispielsweise sind 99 Prozent aller Dentalhygienikerinnen tatsächlich Frauen, und auch der Beruf Bibliothekarin weist einen Frauenanteil von 66 Prozent auf.

Typische Frauenjobs

Frauenanteil in den weiblich übersetzten Berufen



Basierend auf der Strukturerhebung 2016–2018, siehe «Zu Daten und Methode» unten. Quelle: [Bundesamt für Statistik](#).

Das ist ein Erklärungsansatz – aber keine Ausrede für grosse Tech-Firmen, solchen Verzerrungen nicht aktiv gegenzusteuern. Dies sollten sie aus zwei Gründen tun.

Einerseits um der Qualität der Übersetzung selbst willen: Wenn der deutsche «Krankenpfleger» zur französischen *infirmière* wird, ist das schlicht falsch. Und dass geschlechtsneutrale Begriffe – *opettaja*, *nurse*, Pflegepersonal – in der Zielsprache ein Geschlecht tragen, ist ebenfalls nicht korrekt.

Ein Beispiel dazu: Übergab man Google Translate vor zwei Jahren «Pflegepersonal», so spuckte es auf Französisch aus: *infirmières*. Der Algorithmus wählte dabei die weibliche Form nicht nur aus dem Repertoire einer männlichen und einer weiblichen aus (*infirmiers*, *infirmières*), sondern übersah auch, dass es eigentlich genderneutrale Entsprechungen gegeben hätte: *personnel de soin* oder *personnel soignant*. Dies wäre auch Wort für Wort viel naheliegender gewesen.

Zur Übersetzung gehört auch die Übertragung des Geschlechts, ob weiblich, männlich, nicht binär oder genderneutral. Statistische Rollenverteilungen in der Bevölkerung haben dabei selbstverständlich keine Rolle zu spielen.

Andererseits ist es wichtig, solchen Biases entgegenzusteuern, um stereotype, häufig veraltete Rollenbilder nicht zu zementieren.

Textkorpora, die zum Training von Übersetzungsalgorithmen verwendet werden, widerspiegeln nicht unbedingt reale oder aktuelle Verteilungen zwischen Männern und Frauen in bestimmten Rollen, Berufen oder Ämtern. (Zum Beispiel erbringen Frauen auf dem Schweizer Arbeitsmarkt

knapp 40 Prozent der geleisteten, bezahlten Arbeitsstunden – nicht einen Fünftel, wie unser kleiner Übersetzungsdatensatz suggerieren könnte.) Zudem verschärfen Algorithmen typischerweise die verzerrten Verhältnisse aus Trainingsdaten. Klischees werden also nicht nur wiedergegeben, sie werden geradezu potenziert.

Das zeigt auch Teil zwei unseres Übersetzungsexperiments.

Frauen sind schön – und depressiv

Hier haben wir den Algorithmus nicht mit Berufsbezeichnungen, sondern mit persönlichen Eigenschaften, also mit Adjektiven gefüttert.

Eine kleine Kostprobe dazu:

Hän on sievä. Hän on älykäs. Sie ist hübsch. Er ist intelligent.

Ähnlich wie bei den Berufen zeigt sich, dass Google Translate viel häufiger männliche als weibliche Übersetzungen wählt. Von den 50 Adjektiven, die wir getestet haben, wurden 45 in Kombination mit «Er ist ...» übersetzt. Nur bei 5 Adjektiven kam die weibliche Form zum Zug: «Sie ist ...».

Eigenschaften haben meistens Männer

Anzahl der Übersetzungen nach Geschlecht



Quelle: Eigene Analyse; [Google Translate](#).

Zur vollständigen Liste: Wie die Eigenschaften übersetzt wurden

Er ist irrelevant.
Er ist wichtig.
Er ist lebhaft.
Er ist ruhig.
Er ist gross.
Er ist klein.
Er ist gut.
Er ist böse.
Sie ist schön.
Er ist hässlich.
Er ist stark.
Er ist schwach.
Er ist freundlich.
Er ist glücklich.
Sie ist hübsch.
Sie ist attraktiv.
Er ist hart. (Synonym zu «laut».)
Er ist ruhig.
Er ist alt.
Er ist jung.
Er ist berühmt.
Er ist lustig.
Er ist beliebt.
Er ist interessant.
Er ist gelangweilt.
Er ist intelligent.

Er ist dumm.
Er ist unfähig.
Er ist qualifiziert.
Er ist mutig.
Sie ist depressiv.
Er ist glücklich.
Er ist traurig.
Er ist emotional.
Er ist gesund.
Er ist krank.
Er ist traditionell.
Er ist progressiv.
Er ist eifersüchtig.
Er ist grosszügig.
Er hat es geschafft. (Freie Übersetzung von «Er ist erfolgreich.»)
Er ist reich.
Er ist arm.
Er ist hilfreich. (Synonym zu «hilfsbereit».)
Er ist abweisend.
Er ist wettbewerbsfähig. (Synonym zu «kompetitiv».)
Er ist kritisch.
Sie ist locker.
Er ist gefährlich.
Er ist besorgt.

Die Geschichte der weiblichen Eigenschaften ist schnell erzählt: «Sie» ist schön (wahlweise auch hübsch oder attraktiv), locker oder depressiv.

«Er» hingegen ist gut, böse, wichtig, irrelevant, stark, schwach, intelligent, dumm, erfolgreich oder mutig – und vieles mehr. Und auch bei komplexeren Sätzen produziert der Algorithmus ein Klischee nach dem anderen:

Hän tilaa oluen baarissa. Hän keittää kotona. Er bestellt Bier an der Bar. Sie kocht zu Hause.

Das sind nun alles, gelinde ausgedrückt, ziemlich krasse Verzerrungen.

Sie kommen dadurch zustande, dass Algorithmen statistische Prinzipien auswerten: Sie ermitteln in Trainingstexten, wie häufig ein Wort in der Nähe anderer Wörter auftritt – etwa «Kindergarten» oder «Hochschule» neben männlichen Vornamen oder Formen wie «Vater», «er», «Mann».

Sie kommen auch durch methodische Entscheide, etwa die Verwendung sogenannter Pivot-Sprachen, zustande. Für manche Sprachenpaare wie Englisch und Deutsch gibt es viele Texte, die im Netz in beiden Sprachen vorliegen und sich als Trainingsmaterial verwenden lassen. Für andere Sprachenpaare – etwa Ungarisch und Swahili – gibt es viel weniger solcher Parallelkorpora. Hier kommen die Pivot-Sprachen ins Spiel: Für eine Übersetzung von Ungarisch zu Swahili übersetzt man zuerst das Ungarische in eine Pivot-Sprache und anschliessend die Pivot-Sprache ins Swahili.

Als Pivot verwendet man häufig: Englisch. Und da Englisch in grossen Teilen des Wortschatzes nicht zwischen Geschlechtern unterscheidet, gelangt so vermutlich zumindest ein Teil des Gender-Bias in Übersetzungen: An die Stelle des Geschlechts, das in der Ausgangssprache verwendet wurde, tritt das stereotype Geschlecht, das ausgehend vom Englischen gewählt wird.

Auf diese Weise gelangt Gender-Bias via sehr, sehr viele Augenpaare in die Köpfe der Menschen, die Google Translate benutzen, um sich zu informie-

ren, zu unterhalten, sich mit anderen auszutauschen. Nicolas Kayser-Bril von Algorithm Watch nennt solche Methoden «digitalen Kolonialismus».

Wie man es richten könnte

Welche Daten man verwendet, um Algorithmen zu trainieren, ist also bedeutsam:

- Man kann relativ wahllos im Internet Texte zusammensuchen (vielleicht geraten dabei als Beifang auch automatisch übersetzte und damit extraverzerrte Texte mit ins Textkorpus).
- Man kann Textdaten sorgfältig danach auswählen, von wem sie wie übersetzt wurden.
- Oder man kann Texte selber erstellen, ergänzen, anpassen, um dem Algorithmus mehrere Optionen zu gleichen Anteilen zu verfüttern: Er ist Ingenieur. Sie ist Ingenieurin. Sie/er ist Ingenieur/-in.

Einer der Lösungsansätze für Debiasing zielt darauf ab, stereotype statistische Verknüpfungen, etwa zwischen «Kindergarten» und «Frau», maschinell zu entfernen und Trainingsdaten damit ausgeglichener zu gestalten.

Ein weiterer ist die neue Methode, die der Internetgigant vor einem Jahr als «dramatisch andersartiges Paradigma» beschrieb. Für manche Sprachenpaare:

1. produziert Google Translate eine defaultmässige Übersetzung;
2. überprüft, ob die Ausgangssprachliche Form genderneutral war und ob die Zielsprachliche Übersetzung ein spezifisches Geschlecht trägt;
3. und stellt in diesem Fall eine zweite Übersetzung mit dem jeweils anderen Geschlecht dazu.

Ein Beispiel dafür (mit der Zielsprache Deutsch funktioniert es nicht):

Hän on opettaja.She is a teacher / He is a teacher.

Google Research versprach im April 2020, Gender-Bias in englischen Sätzen werde nun bei Übersetzungen aus dem Ungarischen, Finnischen, Türkischen und Persischen um 90 Prozent oder mehr niedriger ausfallen.

Lässt man Google Translate heute «Pflegepersonal» auf Französisch übersetzen, erhält man tatsächlich: *personnel infirmier*. Unser Experiment mit 150 übersetzten Sätzen zeigt aber: Im Zweifelsfall ist der Mensch ein Mann. Wenn er eine Frau ist, dann meist dort, wo es zum althergebrachten Klischee passt. Ein Google-Mediensprecher wollte sich zu unserer Frage nach Debiasing für Zielprachen wie Deutsch und Französisch nicht öffentlich äussern. Auch wie es seine Übersetzungen künftig weiter *debiasen* will, beantwortet das Unternehmen auf Anfrage nicht.

Klar ist: Die Technik allein wird es nicht richten – es wird Menschen mit dem entsprechenden Augenmerk brauchen, die mit der Maschine zusammenarbeiten. Sprachwissenschaftlerinnen werden die Big-Data-Modelle vermutlich wieder vermehrt an die Hand nehmen müssen, zusammen mit Genderforschern, Entwicklerinnen und mit Menschen verschiedenen Geschlechts, Alters sowie verschiedener Hautfarbe, geografischer und kultureller Herkunft.

Zu Daten und Methode

Zahlreiche Berufsbezeichnungen im Finnischen enden mit *-mies*: Mann. *Kirvesmies* («Axt-Mann») etwa steht für den Beruf des Zimmermanns beziehungsweise der Zimmerleute. Wann immer ein genderneutrales Synonym existierte, haben wir diesen Begriff verwendet. Statt *postimies* («Post-Mann») haben wir zum Beispiel den Begriff *postinkantaja* («Post-Träger») gewählt, um den Algorithmus nicht schon durch die Begriffe der Ausgangssprache in eine bestimmte Richtung zu beeinflussen. Nur für eine von 100 Berufsbezeichnungen, beim Zimmermann, haben wir kein brauchbares Synonym gefunden.

Die Daten, die wir zum Anteil der Frauen in verschiedenen Berufsgattungen zeigen, basieren auf der ISCO-19-Kategorisierung. Sie kamen durch Strukturerhebungen vom Bundesamt für Statistik zustande, wobei eine Stichprobe von 200'000 Personen schriftlich befragt wurde. Die angegebenen Werte sind Hochrechnungen aufgrund der Erhebungen. In der Grafik zeigen wir Frauenanteile für die naheliegendsten Berufe der ISCO-19-Kategorisierung.