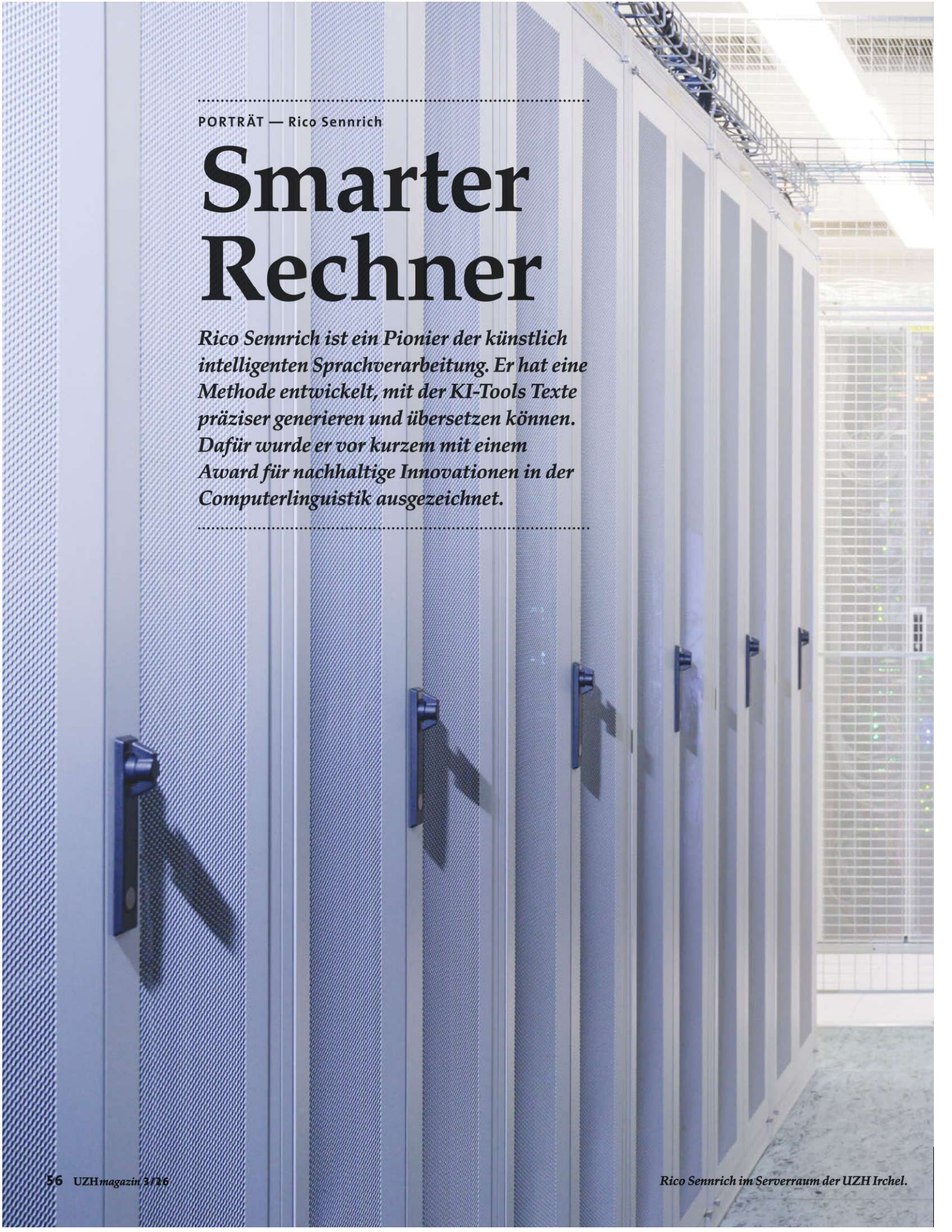




.....

PORTRÄT — Rico Sennrich

Smarter Rechner

Rico Sennrich ist ein Pionier der künstlich intelligenten Sprachverarbeitung. Er hat eine Methode entwickelt, mit der KI-Tools Texte präziser generieren und übersetzen können. Dafür wurde er vor kurzem mit einem Award für nachhaltige Innovationen in der Computerlinguistik ausgezeichnet.

.....

A man with short brown hair and glasses, wearing a dark blue button-down shirt and dark trousers, stands in a server room. He is holding a black laptop with both hands and looking towards the camera with a slight smile. The server room has rows of white server racks with glass doors, and blue cables are visible running along the top of the racks. The lighting is bright and even.

Wenn wir von KI-Helfern Texte übersetzen und generieren lassen, rechnen im Hintergrund die Algorithmen von Rico Sennrich.

Text: Simona Ryser
Bilder: Marc Latzel

Seine Schwiegermutter schickt Grüsse auf WeChat in chinesischen Schriftzeichen. Rico Sennrich berührt den Translate-Button auf dem Smartphone und bekommt die Nachricht übersetzt. Der Computerlinguist spricht auch etwas Chinesisch. Sprachen interessieren ihn. Ursprünglich ist er Anglist, und einst am Gymnasium in St. Gallen hatte er das altsprachliche Profil gewählt. Vor allem fasziniert ihn, wie Sprachen funktionieren – menschliche und maschinelle. Rico Sennrich ist Professor für Computerlinguistik und hat vor kurzem den Test-of-Time-Award der Association for Computational Linguistics erhalten. Der Preis wird für Innovationen aus der Grundlagenforschung, die auch nach zehn Jahren noch Bestand haben, verliehen.

Und das hat seine Erfindung. Sie erleichtert uns täglich das Leben mit Smartphones, Laptops und Computern. Während wir ganz selbstverständlich mit ChatGPT, Claude und all den zahlreichen KI-Helfern Texte übersetzen und generieren lassen, rechnen im Hintergrund die Algorithmen von Rico Sennrich im Hölletempo Bits und Bytes rauf und runter. Vor exakt zehn Jahren, 2026, präsentierten Sennrich und sein Team eine neue Methode, die prägend war für die meisten kommerziellen neuronalen Sprachmodelle. Mit der so genannten Tokenisierung, der Zerteilung der Sprache in kleinste Textbausteine, gelang es, ein Verfahren einzuführen, das es GPT-Modellen ermöglicht, sämtliche Wörter einer Sprache zu verarbeiten und sich nicht mehr wie davor auf die häufigsten zu beschränken. Ursprünglich für das maschinelle Übersetzen entwickelt, ist die Tokenisierung heute Grundlage praktisch aller grossen Sprachmodelle.

Durchschlagende Idee

Rico Sennrich lächelt verschmitzt. 2013 bis 2015 war er als SNF-Postdoctoral-Fellow Gast an der University of Edinburgh. Als er sich anschliessend für eine ausgeschriebene Postdoc-Stelle bewarb, wurde er beim Job-Interview gefragt, was für ein Projekt er verfolgen würde,

wenn er sich mit den damals noch neuen künstlichen neuronalen Methoden beschäftigen könnte. Er habe da so eine Idee, habe er geantwortet, wie man von dem maschinellen Übersetzungssystem mit begrenztem Wortschatz zu einem System mit unbegrenzter Übersetzungsmöglichkeit kommen könne. Damals bekam Sennrich den Job, in diesem Jahr nun den Award für die Idee, die sich als durchschlagend herausstellen sollte.

Sennrich springt vom Besprechungstisch auf, setzt sich an den Computer und schreibt ein langes Wort ins blinkende Feld: Coronapandemiemassnahmengegner:innen. Der Computerlinguist zeigt mir, wie ChatGPT es in Token schneidet. Das Wort leuchtet zerstückelt in verschiedenen bunten Abschnitten. Früher waren zusammengesetzte Wörter, wie sie gerade im Deutschen besonders häufig vorkommen, für maschinelle Übersetzungssysteme eine grosse Herausforderung. Auch Sätze

«KI erlaubt es uns, in Fremdsprachen, die wir nicht kennen, zu kommunizieren, das ist eine grosse Chance, um Barrieren abzubauen.»

Rico Sennrich, Computerlinguist

Berg oder Strand?

Wimmelbuch und Wirbelwind

Wo sind Sie am kreativsten?

Die meisten Ideen kommen mir bei Gesprächen mit Mitarbeitenden und Studierenden, aber grössere Texte oder Projekte entwerfe unterwegs in der Natur.

Was machen Sie, um den Kopf auszulüften und auf neue Gedanken zu kommen?

Mein zweijähriger Sohn ist ein Wirbelwind und lüftet meinen Kopf am Abend mühelos aus.

Mit welcher berühmten Persönlichkeit würden Sie gerne zu Abend essen und weshalb?

Der Komiker Jon Stewart verbindet Humor mit politischem Biss und wäre sicher ein interessanter Gesprächspartner.

Drei Bücher, die Sie auf die einsame Insel mitnehmen würden?

Auf meinen letzten Inselferien war «Midnight's Children» von Salman Rushdie dabei. Dazu vielleicht noch ein Buch, das ich gerne mal lesen wollte, aber nie dazu gekommen bin: «Hundert Jahre Einsamkeit» von Gabriel García Márquez. Und hoffentlich bin ich auf der Insel nicht ganz allein, sondern mit meiner Familie, und dann kommt noch ein Wimmelbuch für meinen Sohn mit.

Kugelschreiber oder Laptop?

Mein Laptop ist bei Weitem mein wichtigstes Arbeitsinstrument, aber Kugelschreiber und Papier schätze ich weiterhin fürs Brainstorming und die Arbeit mit Textentwürfen.

Berg oder Strand?

Beides. Die nächsten Ferien verbringe ich mit meiner Familie in den Bergen.

wie «Er hängte die Wäsche auf» oder «Sie legte den Mantel ab» waren problematisch, weil die abgetrennten Verbpartikel «ab» beziehungsweise «auf» nicht erkannt werden konnten. Alles, was unbekannt oder selten ist, stellte für frühere maschinelle Übersetzungssysteme ein Problem dar. «Allerdings steckt oft gerade in solchen Elementen die entscheidende Information eines Satzes», erklärt Sennrich. Ein Beispiel: «Ich wohne in Hindelbank». Für einen Menschen ist klar, dass es sich beim letzten Wort um einen Ortsnamen handelt. Für frühere Übersetzungssysteme hingegen war alles lesbar, ausser das unbekannte Wort am Satzende.

Sennrich sitzt noch immer am Bildschirm. «Bei der Tokenisierung, die mit Hilfe des sogenannten Byte-Pair-Encoding geschieht, werden häufig vorkommende Zeichenfolgen zu einer Einheit zusammengefügt, während seltene Folgen aufgeteilt werden», erklärt er. Dadurch ist es für neuronale Netze möglich, auch seltene Wörter zu verarbeiten. Zuvor musste der Computer gewissermassen jedes Wort einer Sprache einzeln auswendig lernen. Durch die Zerlegung in kleinste Teile, kann das System nun auch gewisse unbekannte Wörter präzise übersetzen, indem es diese wie bei einem Puzzle aus bekannten Einzelteilen zusammensetzt.

Digitale Welt erkunden

Ein kleiner Roboter steht neben dem Computer, im Bauch trägt er eine tickende Uhr. Der Roboter sieht nett und etwas veraltet aus. Rico Sennrich wuchs in den 1990er-Jahren in St. Gallen auf. Seine Eltern betrieben ein Coiffeurgeschäft. Als kleiner Bub interessierte er sich für Science-Fiction-Literatur. Man kann ihn sich gut vorstellen, wie er im Friseursalon der Eltern zwischen fliegenden Haaren und Föhngeöstöber in Büchern las. Als er etwa zehn Jahre alt war, schenkte ihm der Götti seinen



alten Computer. Sennrich überlegt kurz. Vermutlich war damals Windows 3.1 installiert. Der Computer faszinierte ihn. Er spielte auch die damals üblichen Spiele, etwa Minesweeper oder Solitaire. Vor allem aber machte es ihm Spass, die digitale Welt zu erkunden und mit Zeichenprogrammen zu spielen. Und bald schrieb er erste kleine Programme, um zu verstehen, wie ein Computerprogramm funktioniert.

Als er sich dann nach Zürich an die Universität zum Studieren aufmachte, schwankte er zwischen Englischer Sprachwissenschaft und Informatik. Er entschied sich für die Sprachwissenschaft und entdeckte an der UZH die Computerlinguistik, die er schliesslich als Nebenfach wählte. «An der Computerlinguistik fand ich interessant, dass man eigene Theorien einfacher testen kann», sagt Rico Sennrich. So schrieb er denn als erstes Projekt in der Computerlinguistik ein Programm, das die grammatischen Beziehungen zwischen den Wörtern eines Satzes bestimmt und in einem Syntaxbaum visualisiert. Sennrich spezialisierte sich auf «Natural Language Processing», die Verarbeitung von natürlichen Sprachen. In seiner Dissertation war er dann beim maschinellen Übersetzen gelandet.

Kosten fairer abrechnen

Die Vielsprachigkeit lebt Sennrich auch im persönlichen Alltag. Mit seiner chinesischen Frau – sie kennen sich aus Edinburgh, auch sie ist Sprachwissenschaftlerin – spricht er Englisch, mit dem zweijährigen Sohn Schweizerdeutsch. Mit seinen Schwiegereltern kommuniziert er, unterstützt von diversen KI-Tools, chinesisch und englisch. Er lacht. So manche Touristen zücken eine Übersetzungsass, wenn sie mit gekrauter Stirn rätselnd eine fremdsprachige Menükarte studieren. «KI erlaubt es uns auch in Fremdsprachen, die wir nicht kennen, zu

kommunizieren», sagt Sennrich, «das ist eine grosse Chance, um Barrieren abzubauen und sich über Sprach- und Kulturgrenzen hinweg zu verständigen.»

Und was steht als Nächstes an? Bei kleineren Sprachen gibt es Verbesserungspotenzial. Gerade hat Sennrichs Forschungsgruppe einen Weg gefunden, wie es die KI schafft, von Rumantsch Grischun in die unterschiedlichen rätoromanischen Dialekte zu übersetzen. Auch die Übertragung von längeren wissenschaftlichen Dokumenten in verschiedene Sprachen bereitet den neuronalen Sprachmodellen noch Schwierigkeiten, gilt es doch, den anspruchsvollen Kontext korrekt zu erkennen. «Wir trainieren sie gezielt auf Kontextebene, damit die Texte noch kohärenter werden und nichts ausgelassen wird», sagt Sennrich. Auch strukturelle Textgegebenheiten sollen berücksichtigt werden. Diese elaborierten Algorithmen wird man später auch auf andere Texte anwenden können.

Und noch ein Problem gibt es, das der smarte Computerlinguist angehen will. Die Tokenisierung der Sprachen produziert ungleiche Kosten. Englische Wörter bestehen aus lauter einteiligen Token ganz im Gegensatz zu anderen Sprachen wie zum Beispiel dem Rätoromanischen, bei dem sich ein Wort in vier bis fünf Token teilt. Abgerechnet wird in harter Währung pro Token. Das bedeutet, dass Englisch kostengünstiger ist als manch andere Sprache. Sennrich will Fairness zwischen den Sprachen und hat sich darangemacht, ein Verfahren zu entwickeln, das auch komplexe Sprachen mit längeren Tokenfolgen pro Wort komprimiert.

Nun, die Association for Computational Linguistics ACL – das Gremium, das ihm den Test-of-Time-Award verliehen hat –, wird das Auge auf Rico Sennrichs smarte Algorithmen halten. Wer weiss, vielleicht meldet sie sich in zehn Jahren wieder.