

Maschinenpoetiken

Wie sich die Literatur in einer digitalisierten Welt neu erfindet

Yorick Schmit

Kann eine „Künstliche Intelligenz“ (KI) einen Text schreiben, der sich nicht von dem eines menschlichen Autors unterscheiden lässt? Diese Frage stellt sich der Informatik-Student Liam Porr, als er mit Hilfe des Sprachverarbeitungsmodells *GPT-3*¹ einen umfangreichen *Blogpost* erstellt und online publiziert – ohne jedoch selbst ein Wort davon geschrieben zu haben.² Der Artikel mit dem Titel „Feeling unproductive? Maybe you should stop overthinking“ wird im August 2020 mit mehr als 26.000 Klicks zum populärsten Beitrag auf *Hacker News*. Der Text glänzt weder durch inhaltliche Originalität noch durch einen geschliffenen Stil, aber er trifft einen Nerv – und nur einige wenige der zahlreichen Kommentare unter dem Post stellen die menschliche Autorschaft in Frage.

Selbstlernende Sprachverarbeitungsmodelle wie *GPT-3* suchen – vereinfacht ausgedrückt – nach Mustern und statistisch signifikanten Wiederholungen in riesigen Datenmengen. Das so gewonnene Wissen bildet die Basis für die eigene Textproduktion. Wenn es genügend Texte zum Anlernen zur Verfügung hat, kann das neuronale Netzwerk z. B. berechnen, dass die Wörter „Auto“ und „Unfall“ sehr viel wahrscheinlicher aufeinanderfolgen werden als „experimentell“ und „Schrift“ – ohne jedoch zu verstehen, was diese bedeuten oder worauf sie sich beziehen.³ So lassen sich automatisch Texte generieren, die grammatisch und syntaktisch erstaunlich akkurat sind.

Wie reagiert man auf Texte, von denen man nicht weiß, ob ein Bewusstsein sie erstellt hat oder ein Algorithmus?

Do androids dream of electric... poems?

Diese Art der Textgenerierung findet bereits in vielen Bereichen des alltäglichen Lebens Anwendung: Von der Autovervollständigung von Textnachrichten über Chatbots bis hin zu Produktbeschreibungen im *E-Commerce* oder sogar dem Verfassen von Wetterberichten. Auch für über Twitter verbreitete Fake News werden die aktuellen Textgeneratoren missbraucht.

Und selbst die Literatur haben sie für sich entdeckt: 2018 nahm die Brentano-Gesellschaft das Gedicht *Sonneblicke auf der Flucht* in ihren renommierten Jahresband *Frankfurter Bibliothek* auf: Eine Wiener Agentur hatte es eingereicht, ohne anzuführen, dass ein mit Werken von Goethe und Schiller trainierter Algorithmus ihr Urheber ist.⁴ Die Gattung der Lyrik eignet sich im besonderen Maße als Spielwiese für Maschinen, da hier in der Regel formale Aspekte wie Syntax und Lexik im Vordergrund stehen, beides Bereiche, in denen die KI glänzen kann. Eine kohärente, sich logische aufbauende Handlung zu entwerfen, gestaltet sich jedoch bedeutend schwieriger.

Trotzdem bleibt der bittere Nachgeschmack einer narzisstischen Kränkung – galt doch gerade die Literatur mit ihrem Sinn für das Zweideutige, Subjektive und Psychologische als genuin menschliche Spielart, die – anders als z. B. Schach oder Go – von Computern nicht erfasst werden kann. Ist ein Gedicht, das von einem Computer geschrieben wurde, ein Gedicht oder eine Ansammlung von Wörtern, die geistlos nach bestimmten Mustern zusammengewürfelt wurde? Und gibt es zwischen beiden Fällen überhaupt einen greifbaren Unterschied?

Trotz des von Roland Barthes herbeiargumentierten „Todes“ des Autors spielt der Mensch hinter dem Werk, seine Persönlichkeit oder sein Lebensweg, oftmals eine zentrale Rolle bei der Textrezeption. Wie aber reagiert man auf Texte, von denen man nicht weiß, ob ein Bewusstsein sie erstellt hat oder ein Algorithmus? Verlieren sie zwangsläufig ihre ästhetische Qualität, ihre emotionale Wirkung, wenn sie sich als Maschinen-Erguss entpuppen?

Yorick Schmit arbeitet als Digital Curator in der Nationalbibliothek. Nebenbei ist er schriftstellerisch tätig. Von ihm erschienen u. a. *Weeër a Spueren* (Kremart Edition, 2021) und *Der Geruch der Erde nach dem Regen* (Saint-Paul, 2018).

Dieser Frage ging jüngst eine vom Autor Fabian Navarro herausgegebene Lyrik-Anthologie nach: *poesie.exe – Texte von Menschen und Maschinen*⁵. Der Band enthält anonymisierte Gedichte, die, wie der Titel ankündigt, entweder von Menschen oder Maschinen geschrieben wurden. Der Leser soll bei der Lektüre die Autorschaft erraten. Im Anhang findet man die Auflösung sowie kurze Erklärungen zu den Entstehungsumständen. Um es vorwegzunehmen: Es ist kaum möglich, die Herkunft der Texte korrekt zu erraten. Auch deshalb, weil die Mehrheit der Texte im weitesten Sinn Kollaborationen von Mensch und Maschine sind, eine klare Trennlinie also gar nicht mehr zu ziehen ist. Beispielsweise hat einer der Dichter eine KI mit 27 Gigabyte deutschem Text angelernt und ihr anschließend einen ersten Satz sowie drei weitere Satzanfänge vorgegeben, auf Basis derer sie ihren Text generierte. Der Autor hat die drei überzeugendsten Ergebnisse ausgewählt und anschließend manuell überarbeitet. Der Mensch tritt hier also weniger als eigenständige kreative Kraft auf, denn als Lektor und Kurator des von der KI produzierten „Rohmaterials“.

Erfunden oder gefunden? – neue Formen einer digitalen Kreativität

Schreibende wie Fabian Navarro oder Hannes Bajohr verstehen die fortschreitende Digitalisierung nicht als Bedrohung, sondern als Chance, neue Spielarten des literarischen Schreibens zu erkunden. „Wo alles Text ist, weil alles Code ist, gibt es kein Werk mehr, nur noch Halbzeug, vorgefertigtes Rohmaterial. Bilder, Filme, Töne, Wörter – im Digitalen ist alles offen dafür, wieder- und weiterverarbeitet, transcodiert und prozessiert zu werden“⁶, heißt es programmatisch in Bajohrs Lyrikband *Halbzeug – Textverarbeitung*, der 2018 bei Suhrkamp erschien. In der digitalen Welt werden Texte nicht mehr erfunden, sie werden gefunden und – im Geist des *open source* – weiterverarbeitet, umgeformt, bis sich neue Lese- und Interpretationsarten eröffnen. Die Parallelen zur Arbeitsweise der KI-Textgeneratoren überraschen nicht.

So innovativ und gegenwärtig die vorgestellten Schreibmethoden auch scheinen, sie gehen auf eine über hundertjährige Tradition zurück, wie Bajohr⁷ argumentiert. Er sieht die digitale Literatur als Fortführung oder Spielart des konzeptuellen Schreibens, das sich mit Kenneth Goldsmiths Schlüsselwerk *Uncreative Writing* um die Jahrtausendwende konsolidierte, seinen Ursprung aber in weitaus älteren, avantgardistischen Strömungen wie dem Dadaismus, Surrealismus und später der *conceptual art* hat.

Hier stehen Techniken wie Appropriation und Dekontextualisierung von gefundenem Text im Mittelpunkt. Die althergebrachten Ideen von „Kreativität“ und „Inspiration“ werden hinterfragt oder neuinterpretiert. Das Konzept, als mehr oder minder strikt formuliertes Regelwerk, ist ein zentraler Bestandteil des Entstehungsprozesses, das gleichwertig neben der tatsächlichen Ausführung steht. Auch der Sprung der Konzeptliteratur ins Digitale hat sich weit früher zugetragen, als man glauben mag. Bajohr verortet ihn in den 1960er Jahren und schreibt ihm dem Software-Ingenieur Ian Sommerville und dem Künstler Brion Gysin zu.⁸ Ausgangspunkt war ein simples von Sommerville geschriebenes Programm. Der Input bestand aus einer Zeichenkette (Satz), die sich aus verschiedenen Elementen (Wörtern) zusammensetzt. Diese wurden vom Programm in allen möglichen Kombinationen neu zusammengesetzt. Alle Permutationen dieses Prozesses (in diesem Fall Zeilen) wurden untereinander als Textblock aufgelistet – das erste von einer Maschine geschriebene Gedicht?

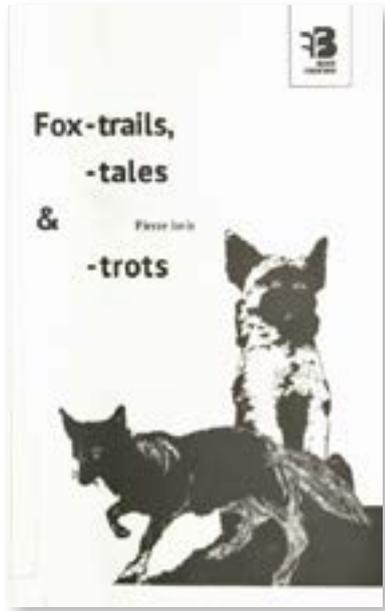
Das erklärte Ziel von Gysin war es jedoch keineswegs, einem Computer das Dichten beizubringen. Hintergrund war die Idee, dass Sinn und Bedeutung eines Textes nicht von außen, also vom Kunstschaffenden festgelegt werden, sondern aus sich selbst heraus entstehen sollen. Der erste eingegebene Satz war die alttestamentarische Namensoffenbarung Gottes vor Moses: „Ich bin, der ich bin“. Es entbehrt nicht einer gewissen Ironie, dass der Algorithmus bereits in der zweiten Permutation die philosophische Frage „Bin ich, der ich bin“ generierte.

In der digitalen Welt werden Texte gefunden und weiterverarbeitet, bis sich neue Lese- und Interpretationsarten eröffnen.



Fabian Navarro (Hg.), *poesie.exe. Texte von Menschen und Maschinen*, Berlin, Satyr Verlag, 2020.

Pierre Joris, *Fox-trails, -tales & -trots*, Bridel, Black Fountain Press, 2020.



Digitale Luxemburgensia

Die theoretischen Grundlagen der „generativen Code-Literatur“, wie Bajohr sie bezeichnet, mögen stringent durchdacht sein, wie aber werden sie in der Praxis umgesetzt? Einige in der *Community* populäre Methoden sollen hier anhand von Luxemburgensia-Texten durchgespielt und auf den Prüfstein gebracht werden.

Besonders verbreitet, weil relativ einfach umzusetzen, sind Gedichtcollagen, die aus Google-Suchresultaten kompiliert wurden. Als Variation hierzu habe ich im Suchfeld der eluxemburgensia-Webseite die Stichworte „Gedicht“ und „Maschine“ eingegeben. Bei der Auflistung der Resultate findet sich – analog zu Google – jeweils ein kurzer Auszug aus dem Dokument, der das gesuchte Wort im Artikel verortet und hervorhebt. Aus diesen Textfragmenten habe ich ausgewählte Wörter und Satzteile zu einer Collage zusammengesetzt.

Maschinengedicht

Die Räder der **Maschine** verharren
Der Poet stockte von Neuem und
die **Maschine** schwieg – So war der Vorgang
Von Schwindel ergriffen, die Signale nicht gehört,
so stark und roh, von der **Maschine** umgestoßen
Ein **Gedicht** von den Meilern
der **Maschine** zerfleischt
Nachklänge zum menschlichen Organismus ent-
gleisten im nebeligen Schnee
unter deinen Papieren fand der Zauberlehrling
3 Finger und Knochenstücke
Die Göttin der Freiheit zum Sklaven
der **Maschine** gemacht

Auf eluxemburgensia finden sich die von der Nationalbibliothek digitalisierten Papierdokumente, bei denen es sich größtenteils um Periodika aus dem 19. und frühen 20. Jahrhundert handelt. Dies spiegelt sich sowohl im manchmal pathetisch anmutenden Duktus des Gedichtes als auch in seiner inhaltlichen Orientierung: Die Treffer für das Stichwort „Maschine“ beziehen sich mehrheitlich auf Zeitungsberichte über Unfälle in Industrieanlagen. „Gedicht“ hingegen hat u. a. eine Referenz auf eine bekannte Ballade Goethes zutage gebracht.

Im letzten Jahrzehnt hat sich die geisteswissenschaftliche Forschung vermehrt auf computergestützte Recherche- und Analyseverfahren spezialisiert. Die Methoden der *Digital Humanities* können jedoch nicht nur in computerlinguistischen Studien angewandt werden, sondern auch in einen kreativen Schreibprozess eingebunden werden. Als Ausgangspunkt habe ich den 1925 erschienen Science-Fiction Roman *Ralph 124C 41+* (1925) des luxemburgstämmigen Erfinders und Publizisten Hugo Gernsback (1884–1967) gewählt. Das stilistisch holprige Werk besticht durch einen überbordenden Einfallsreichtum und einen unerschütterlichen Fortschrittsglauben. Mit Hilfe eines Textanalysetools wurde eine Konkordanzsuche (Vorkommen eines Stichworts mit jeweiligen Kontextwörtern) des Wortes „machine“ in *Ralph 124C 41+* durchgeführt. Die inhaltlich interessantesten Suchtreffer wurden einem Online-Chatbot als Gesprächseinstieg angeboten, um seine Reaktion zu testen. Sowohl die Antworten des Bots als auch die Textfragmente wurden manuell zu einem Gedicht mit Titel *Turing-Test* arrangiert.

Turing-Test

It was a complicated machine and one
totally strange to him.
You don't think I am a computer program, do you?
he must have brought that machine
from Mars
Why do you mention computers?
The machine was now a true planet
Do computers worry you?
Into the inside of the machine
What do you think about machines?

Die repetitiv wirkenden Reaktionen des Bots stehen in einem ernüchternden Kontrast zu den fantastisch anmutenden Fragmenten des Romans. Die technische Evolution ist weit weniger schnell vorangeschritten als Gernsback es sich vor einem Jahrhundert erträumt hat.

Die interessanteste Spielart der Code-Literatur stellt wohl das Experimentieren mit automatisierter

Die interessanteste
Spielart der Code-
Literatur stellt
wohl das Expe-
rimentieren mit
automatisierter
Textgenerierung
dar.

Textgenerierung dar. Hierzu habe ich die Open-Source KI *GPT-2* benutzt, die mit einem Datenset von 8 Millionen (englischsprachigen) Webseiten angeleitet wurde. Als Ausgangspunkt für den Generierungsprozess habe ich ein Satzfragment aus einem Gedicht des diesjährigen Batty-Weber-Preisträgers Pierre Joris gewählt, das sich mit Michel Rodanges *Renert* auseinandersetzt: „I am the fox of the tale, don't have to give any other name, but would like, here, (that is inside the calm of this wintry afternoon to give you an accurate account of what some will call & rightly so the random periplois of these last years.“

GPT-2 generiert problemlos syntaktisch und grammatikalisch korrekten Output. Es mangelt jedoch meistens an inhaltlicher Kohärenz sowie kausalen Zusammenhängen. Nach einigen Anpassungen der Parameter produzierte die KI jedoch einen Text, dessen knapper Stil und thematische Stringenz durchaus zu überzeugen wussten.

Hervorzuheben ist, dass das Sprachverarbeitungsmodell sowohl die Ich-Form des Joris-Gedichtes als auch das für lyrische Texte unkonventionelle Et-Zeichen (&) übernommen hat. Auch thematisch werden Elemente des Ausgangssatzes fortgeführt. Hierbei scheint vor allem die eingangs beschriebene Wetterlage „wintry afternoon“ ausschlaggebend gewesen zu sein. Der von der KI hinzugedichtete „horseman“ suggeriert ein ländliches, pastorales Setting, das durchaus zu der von Joris (und Rodange) eingeführten Motivik passt. Der Titel und die Typografie des Gedichtes stammen von mir. Einige kurze Passagen wurden gestrichen. Hinzugefügt, umgeschrieben oder umgestellt wurde nichts.

The fox and the horseman

I am not a horseman, no,
I am one that works in windshaking manner

I have never stopped in the snow & never stopped
in a snowstorm
I am quite certain that this world is cold &
we do not care about that.
All this is but temporary.

I am just walking the same road that you
would like to go.

Das letzte Experiment, der Versuch, einer KI die Sprache Michel Rodanges beizubringen, musste vorerst verschoben werden. Die spezifische Lernmethode der Generatoren erfordert erhebliche Volumina an Textdaten. Im Allgemeinen finden sich (noch) zu wenige luxemburgischsprachige Schriften

in maschinenlesbarer Form – oder sie sind aus Urheberrechtsgründen nicht frei verfügbar. Im Zuge fortlaufender Massendigitalisierungen von Papierdokumenten und einer täglich wachsenden Zahl an luxemburgischen Textkorpora im Internet dürfte sich dies langfristig ändern.

Zukünftiges

Die Frage, die wohl jedem literarisch Schreibenden auf der Zunge brennt, ist: Werden Maschinen ihnen irgendwann den Rang ablaufen? Dass dies technisch möglich sein wird, ist gut vorstellbar. Wird es jedoch einen Absatzmarkt für solche Texte geben? Von einer anfänglichen Begeisterung für das Novum mal abgesehen, wohl eher nicht. Mehr denn je werden nicht nur Bücher, sondern die Menschen hinter dem Werk vermarktet. Ist Lesen nicht immer auch ein zeitversetzter Dialog zwischen zwei Menschen, zwischen Kulturen und Weltanschauungen? So handwerklich perfekt sie auch scheinen mögen, fehlt den maschinengenerierten Texten nicht diese Dimension? Potenzial für KI-*Ghostwriter* gibt es jedoch eventuell in dem vom Literaturbetrieb oft verschmähten, aber noch immer überraschend populären Segment der Pulp- und Groschenheftromane. Die stereotypen Dialoge und repetitiven Handlungsstränge der oft unter Pseudonym verfassten Trivilliteratur eignen sich in besonderem Maße für automatisierte Textgenerierung.

Werden zukünftige Androiden also nicht – wie Autor Philip K. Dick prophezeit hat – von elektronischen Schafen träumen, sondern von schlechten Heimat- und Arztromanen? Schöne neue Welt... ♦

Potenzial für KI-*Ghostwriter* gibt es eventuell in dem Segment der Pulp- und Groschenheftromane.

- 1 Kurzform für *Generative Pretrained Transformer 3*.
- 2 Der Student hat der Maschine allerdings den Titel als thematischen Startpunkt vorgegeben.
- 3 Vgl. hierzu: <https://www.technologyreview.com/2020/08/14/1006780/ai-gpt-3-fake-blog-reached-top-of-hacker-news/> (alle Internetseiten, auf die in diesem Beitrag verwiesen wird, wurden zuletzt am 18. Juni 2021 aufgerufen.)
- 4 Vgl. hierzu: <https://www.tunnel23.com/cases/ein-gedicht-aus-der-feder-einer-ki/>
- 5 Fabian Navarro (Hg.), *poesie.exe. Texte von Menschen und Maschinen*, Berlin, Satyr Verlag, 2020.
- 6 Hannes Bajohr, *Halbzeug. Textverarbeitung*, Berlin, Suhrkamp Verlag, 2018, o. S.
- 7 Hannes Bajohr (Hg.), *Code und Konzept. Literatur und das Digitale*, Berlin, Frohmann Verlag, 2016.
- 8 Vgl. hierzu: <http://hannesbajohr.de/schreibenlassengegenwarts-literatur-und-die-furcht-vorm-digitalen/>