

# Automatische Handschriftenerkennung

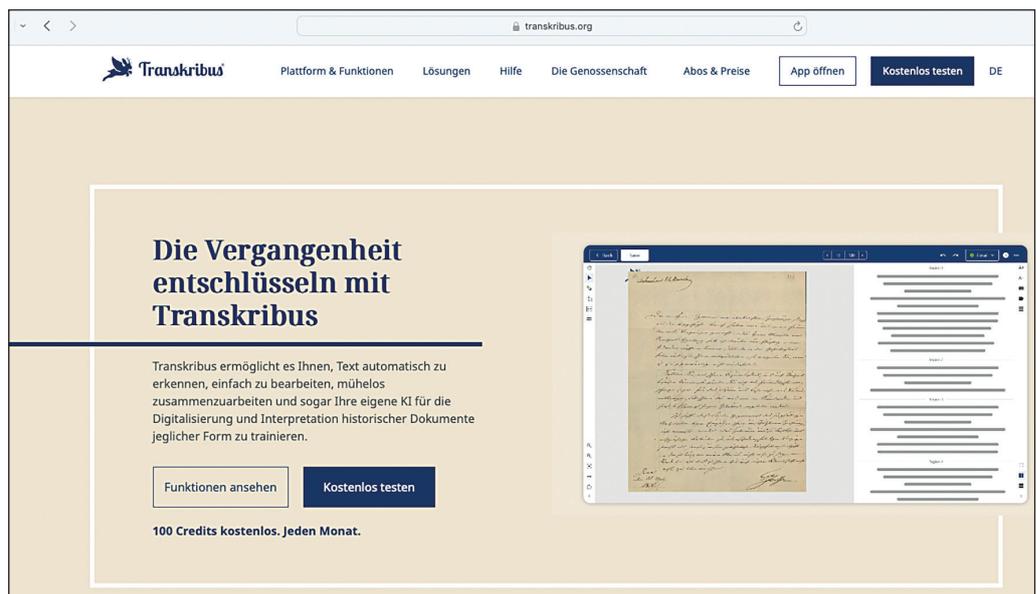
## Transkribus unter der Lupe

Traut man dem euphorischen Tenor von Fachleuten und Technikbegeisterten, steht die Menschheit mit der Entwicklung von sogenannter künstlicher Intelligenz (KI) vor einem Paradigmenwechsel. Eine neue Ära sei eingeläutet worden, die den zunehmend steiler verlaufenden Bogen exponentiellen technischen Fortschritts der letzten Jahrzehnte fortführt. Mit KI-Modellen, welche mit der menschlichen Geisteskraft im Wettkampf stehen und diese in Teilen schon überflügelt haben sollen, ist ein vorläufiger Gipfel erreicht. Selbst augenscheinliche Kreativleistungen – die letzte sicher geglaubte Bastion des Menschen gegenüber der sich ausweitenden digitalen Revolution – wie das Schreiben von Texten, das Erstellen von Bildern und sogar von Videos werden mittels sogenannter „generativer KI“ ermöglicht. Ein früher und naheliegenderer Einsatzbereich aber, das Analysieren von visuellen Informationen, führt uns zu einer Anwendung, die für uns, als Freunde der deutschen Schrift, wohl am interessantesten erscheint: das Auslesen und Transkribieren alter Handschriften.

Diesem Vorhaben widmete sich von 2016 bis 2019 das mit 8,2 Mio. Euro geförderte EU-Projekt „READ“ (Recognition and Enrichment of Archival Documents, zu Deutsch: Erkennung und Anreicherung von archivalischen Dokumenten) unter Schirmherrschaft der Univer-

sität Innsbruck. READ startete mit der Ambition, das kulturelle europäische Erbe, welches in kilometerlangen Bänden in Archiven vor sich hinschlummert, zu bergen und die Niederschriften der letzten Jahrhunderte der Forschung und Öffentlichkeit zugänglich zu machen. So entstand in disziplinübergreifender Zusammenarbeit von Archivaren, Computer- und Geisteswissenschaftlern als auch Freiwilligen die Plattform „Transkribus“, mit der die automatische Texterkennung und Transkription (handschriftlicher) Texte möglich gemacht werden sollte.

Werfen wir zunächst einen Blick auf die dem Programm zugrunde liegenden Mechanismen. Zwei Begriffe sind hierbei zentral: Künstliche Intelligenz und ihr Teilbereich Maschinelles Lernen. Abgeleitet von der KI-Definition der EU, ist Künstliche Intelligenz ein maschinengestütztes System, das aus Eingaben ableitet, wie es Ereignisse und Inhalte erzeugt. Es kann mit unterschiedlichem Grad an Autonomie betrieben werden und zeigt Anpassungsfähigkeit. Etwas deutlicher wird diese allgemein gehaltene Erklärung anhand der Definition von Maschinellern Lernen: „Aufgrund möglichst vieler Textdaten erkennt und erlernt ein Computer die häufig sehr komplexen Muster und Gesetzmäßigkeiten bestimmter Phänomene. Daraufhin können die aus den Daten gewonnenen Erkenntnisse verallgemeinert werden und für



Webseite von Transkribus  
www.transkribus.org

neue Problemlösungen oder für die Analyse von bisher unbekanntem Daten verwendet werden.“ (<https://fortext.net/ueber-fortext/glossar/machine-learning>, abgerufen am 1.9.2024).

Voraussetzung für maschinelles Lernen ist die durch den Menschen angestoßene Fütterung der selbstlernenden Programme mit Informationen. Denn das macht KI im Kern: Muster in zur Verfügung gestellten Daten erkennen und wahrscheinlichkeitsbasierte Regeln daraus entwickeln, um diese dann auf neues Material oder Eingaben anzuwenden. Wie genau diese Regeln abgeleitet werden, also wie das Programm (auch Algorithmus genannt) lernt, erschließt sich dem menschlichen Bediener nicht. Er ist aber in der Lage, Fehlentwicklungen durch neue Eingaben und Anweisungen gegenzusteuern, also dem Programm vorzugeben, wie die gewünschte Lösung auszusehen hat, wodurch sich der Algorithmus weiterentwickelt und verbessert.

Um (hand-)schriftliche Dokumente in digitale Druckschrift umzuwandeln, führt Transkribus zusätzliche Schritte durch, die nötig sind, um die eigentliche Erkennungsleistung vorzubereiten. Zunächst wird das eingelebte und digital zur Verfügung gestellte Dokument binarisiert, d. h. in ein Schwarz-Weiß-Bild umgewandelt. Relevante Inhalte liegen sodann schwarz auf weißem Hintergrund vor. Inhaltlich Irrelevantes, wie Flecken oder von der Dokumentenrückseite durchscheinende Schrift, wird ausgeblendet. Es folgt das Zuschneiden des Dokuments auf den Druckspiegel und die Begradigung von schräg verlaufenden, gekrümmten oder welligen Schriftzeilen. Bei der Regionensegmentierung werden die schwarzen, gedruckten Bereiche klassifiziert, also nach Text, Abbildungen und Tabellen unterschieden. Die so entstandenen Textblöcke werden anschließend in die korrekte Anzahl an Zeilen segmentiert. Erst jetzt kann die eigentliche Texterkennung, die Transkription, beginnen. Dieser letzte und entscheidende Schritt funktioniert anhand von sogenannten „Modellen“, die zunächst angeleitet werden müssen. Dabei kann der Fokus frei gewählt werden, ein Modell ist also zunächst ein unbeschriebenes Blatt und kann für verschiedenste Schriftarten trainiert werden. Dazu bedarf es einer händischen, korrekten Transkription, mit der das Modell gefüttert werden kann, so daß der Algorithmus in der Lage ist, Muster zwischen Ausgangstext und Transkription zu erkennen. Im Idealfall sind dies mehrere hundert Seiten; bei Modellen, die auf eine individuelle Handschrift zugeschnitten sind, 5000 bis 15 000 Wörter. Die Algorithmen können also je nach Anwendungsfall spezialisiert werden. So ist es mit der

entsprechenden Vorarbeit beispielsweise möglich, Modelle für die Kurrent aus dem deutschen Sprachraum des 17. Jhds.; für die Handschrift von Goethe; für hebräische oder auch arabische Schrift aufzusetzen. Ausschlaggebend ist allein das zur Genüge dargereichte Lernmaterial in Form von transkribierten Vergleichstexten der entsprechenden Art. Dieses entscheidet auch darüber, wie hoch die Fehlerrate im Ergebnis ist, wobei Fehlerkorrekturen, wie schon erwähnt, zur Verbesserung des Modells beitragen – der Algorithmus „lernt“. Wie auch für den menschlichen Leser, bestehen Herausforderungen z. B. im fluktuierenden Schriftbild und in Schriftkorrekturen. Aber auch komplexe Seitengestaltungen und eine darauf zurückzuführende falsche Zeilen- und Abschnittserkennung stellen Fehlerquellen dar. Weiters werden Zierinitialien zu meist als Abbildung klassifiziert, wobei es eine Mustererkennung ohnehin schwer hätte, das inhaltlich Maßgebliche aus einem individuell entworfenen Wust an Zierlinien und Ornamentik zu extrahieren.

Wie aus der Anforderung der Modellerstellung klar wird, richtete sich Transkribus zunächst fast ausschließlich an Experten, die bereits mit dem Lesen alter Handschriften vertraut sind – namentlich Archivare, Ahnenforscher, Germanisten usw. Dementsprechend komplex, umständlich und wenig intuitiv kam die Ursprungsversion daher. Statt einer perfekten Transkription auf Knopfdruck, lieferte das Programm nach noch einigen händischen Vorbereitungen ein mitunter fehlerbehaftetes Produkt, das der Nachbearbeitung bedurfte, aber die Transkription von umfangreichen Schriften für Fachleute deutlich erleichterte. Modelle konnten und können auch per Datenwolke mit anderen Nutzern geteilt werden, was einiges an Vorarbeit spart, sofern ein passendes Modell für den zu transkribierenden Text zur Verfügung steht. Ein weiterer Vorteil von nicht nur photographisch eingelebten, sondern transkribierten, in digitaler Druckschrift vorliegenden Texten, ist die Suche nach Schlagworten. So wird die Recherche in historischen Schriftdokumenten drastisch vereinfacht und Geschichtsforschung ortsungebunden möglich gemacht.

Mit August 2023 startete Transkribus als Netz-Anwendung und richtet sich seitdem dezidiert auch an Personen ohne besondere Vorkenntnisse im Lesen historischer Schriften. Im Zentrum stehen nun die Benutzerfreundlichkeit, Einfachheit und Zugänglichkeit für Laien. Die Plattform kann jetzt wahlweise in einer kostenlosen Basisversion oder in umfangreicheren kostenpflichtigen Abonnements genutzt werden. Erstere beinhaltet 100 Seiten Transkription pro Monat und den Zugriff auf

einige der geläufigsten Modelle, d. h. bereits trainierte Algorithmen. Außerdem kann durch Aufheben einer persönlichen Netzseite die eigene Sammlung von in Maschinenschrift übersetzten Dokumenten anderen Transkribus-Nutzern zugänglich gemacht werden. Besonders begrüßenswert ist die geplante Lernplattform zum spielerischen Erlernen des Lesens historischer Handschriften. Ähnlich einem Rätsel, sollen durch Gegenüberstellung von Originaltext mit Transkription die menschliche Fähigkeit der Mustererkennung geschult werden. Der Vergleich mit anderen Transkribus-Nutzern fördert dabei den Ehrgeiz. Allerdings war das sogenannte „Transkribus Learn“, Stand September 2024, nicht (mehr) abrufbar. Ob sich die Umsetzung verzögert hat oder die Funktion in der Zwischenzeit wieder eingestellt wurde, ist unklar.

Die besagte Nebenanwendung von Transkribus konnte jedoch getestet werden und offenbarte überraschende Ergebnisse. Vier ausgesuchte Textbeispiele augenscheinlich verschiedener Schriftdeutlichkeitsgrade wurden der maschinellen Schrifterkennung unterzogen und sollen nun vorgestellt werden. Zunächst ein stilisierter Schriftzug von Rudolf Koch (Abb. 1.1). Mit Grundkenntnissen in Kurrent läßt sich dieser einfach entziffern; die KI erkannte jedoch erst gar keinen Text! Wie ich richtig vermutete, war der Ausschnitt zu klein gewählt. Mit einem verkleinerten Schriftbild (Abb. 1.2) war Transkribus in der Lage,



Transkribus-Test, Abb. 1.1



Transkribus-Test, Abb. 1.2

Schrift auszumachen, wobei das Ergebnis von drei Fehlern bei einem solch kurzen Absatz eher ernüchternd ausfiel.

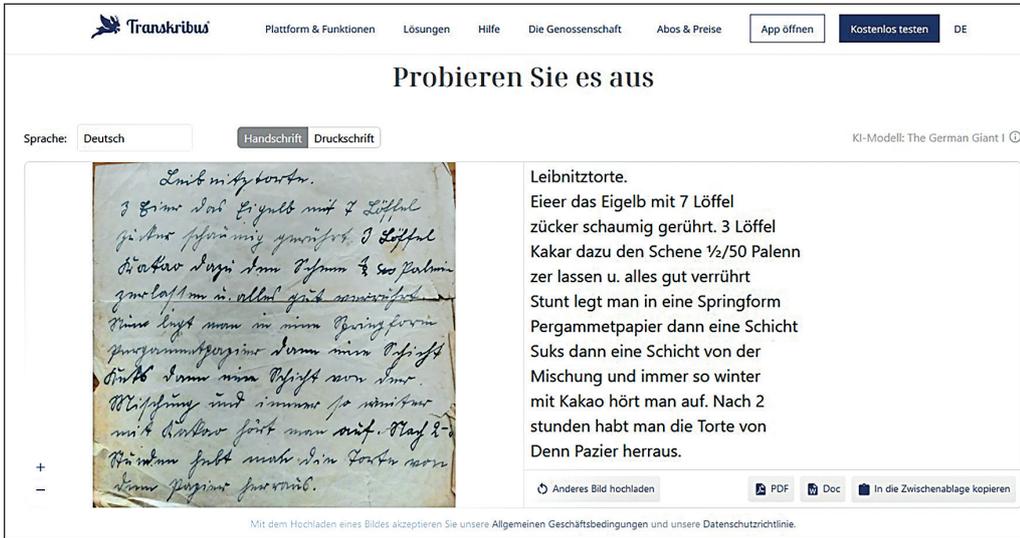
Ein zweiter Versuch, eine kurze Dyrif von Joseph von Eichendorff (Abb. 2), brachte keine wirklich bessere Transkription hervor. Mit viel Phantasie läßt sich der Ursprungs- vom Ausgabebetext her zusammenreimen; Verfasser und Jahresangabe werden größtenteils verschluckt. An dieser Stelle wird deutlich, wie wichtig ein passendes Textmodell ist: Das hier verwendete ist auf Kurrent spezialisiert, kann also den in lateinisch-unverbundener Handschrift verfaßten Part nicht erkennen.



Transkribus-Test, Abb. 2

Weiter geht es mit einem freier geschriebenen, aber für den Geübten immer noch gut zu lesenden Rezept für eine Leibnistorte (Abb. 3). Die Überschrift transkribiert der Algorithmus noch mühelos, dem Fließtext gemäß würde ich Transkribus eine schwere Degasthenie diagnostizieren. Das übersezte Rezept könnte jedoch mit dem nötigen Hintergrundwissen zu wahlweise der deutschen Schrift oder aber der Backkunst dennoch gelingen. Im genauen Vergleich zwischen Handschrift und Ergebnis ist es nachvollziehbar, wie Transkribus die entsprechenden Buchstaben und Wörter falsch erkannt hat. Es findet aber – zumindest in der Basisversion – kein Abgleich mit Wörterbucheinträgen statt, so daß Nonsens-Wortgebilde gleichberechtigt neben richtig transkribierten, sinnhaften Wörtern stehen.

Für das letzte Beispiel wurde die KI mit einem, sagen wir nachlässig geschriebenen, kniffligeren Dokument auf die Probe gestellt (Abb. 4). Erstaunlicherweise fällt das Ergebnis nicht schlechter aus als bei den drei vorigen Beispielen: Der Sinngehalt des Textausschnitts erschließt sich zumindest grundlegend und auch die offensichtlichen Fehler könnte wohl nur ein sehr lesegeübter

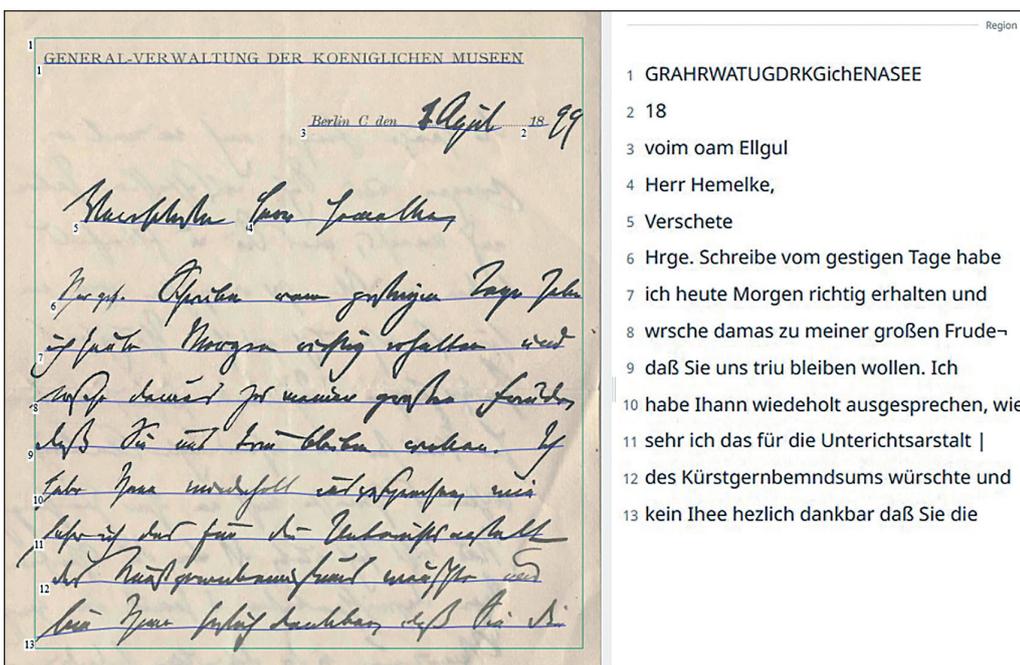


Transkribus-Test, Abb. 3

Kenner der Kurrent korrigieren. Interessant ist, wie die klar zu lesende lateinische Drucktype nicht korrekt erkannt wird, eben weil das Modell lediglich mit Handschriften trainiert wurde. Stattdessen „halluziniert“ Transkribus einen Buchstabenalat, der an WLAN-Passwörter erinnert.

Nach der überschwenglichen, wahrscheinlich allzu hohe Erwartungen schürenden Einleitung, bleibt als Fazit die Ernüchterung. Zumindest – und das muß deutlich betont werden – für die kostenlose Basisversion der Schrifterkennungs-KI Transkribus. Jedenfalls macht der kleine Anwendungstest deutlich, wie sich künstliche von menschlicher Intelligenz im Lesen bzw. Transkribieren von Handschrift unterscheidet: KI erkennt nicht aus dem

Kontext, sondern wendet erlernte Regeln auf neue (visuelle) Informationen an. Sie besitzt oder entwickelt kein reales Verständnis, sie kann nicht holistisch denken, sie besitzt keine Emotion oder Meinung. Durch diese „kalte“, schrittweise Anwendung von Blaupausen, ist ihr Ergebnis zwar nicht durch Bewertungen oder inhaltliche Annahmen verzerrt; der KI gelingt es somit aber auch nicht, ad hoc aus dem Zusammenhang eines sich wiederholenden Schriftbildes Verbindungen oder aus der Sinnhaftigkeit die nötigen Schlüsse zu ziehen. Die kreativen Lösungsansätze des menschlichen Geistes sind dem Algorithmus bei der Entschlüsselung schwierig zu lesender Handschriften fremd. Damit bleibt Transkribus aus meiner Sicht vorerst eine reine Mustererkennungsmaschine.



Transkribus-Test, Abb. 4