

Das Porträt: Peter Gerjets

Dr. Peter Gerjets ist Professor für Lehr- und Lernforschung am Leibniz-Institut für Wissensmedien (IWM). Dort leitet er die Arbeitsgruppe „Multimodale Interaktion“. Gerjets studierte Psychologie in Göttingen, wo er auch promoviert wurde. 1996 ging er an die Universität des Saarlandes, wo er sich habilitierte. Seit 2002 ist er leitender Wissenschaftler am IWM und seit 2003 Professor an der Eberhard Karls Universität Tübingen.

Es war der 9. Januar 2007, als auf der Macworld Conference & Expo in San Francisco Apple-Chef Steve Jobs das neueste Produkt seines Unternehmens vorstellte: ein Mobiltelefon, das zugleich Zugang zum Internet bot und darüber hinaus eine Reihe von Multimedia-Tools – das iPhone.

Jobs wusste offensichtlich sehr wohl, dass er an diesem Januartag dem Publikum eine technische Revolution präsentierte. Seine Ansprache begann er mit den Worten: „Every once in a while a revolutionary product comes along that changes everything.“ Jobs sollte recht behalten: Mobiltelefon, Touchscreentechnologie und Internetfähigkeit in einem Gerät zu vereinen, hat nicht nur den Markt der Mobiltelefone verändert. Kein Konsumprodukt seit der Erfindung des Automobils hat unsere Gesellschaft und unsere Kultur so revolutioniert wie das Smartphone. Wir kommunizieren nicht nur anders, wir verabreden uns anders, wir kaufen anders, wir arbeiten anders und informieren uns anders. Das Smartphone hat unsere Art zu denken verändert, unser Miteinander, unser Selbstbild.

Neben der Internetfähigkeit war es vor allem der berühungsempfindliche Bildschirm des iPhones, der nicht nur unseren Zugang zu Informationen revolutioniert hat, sondern auch unseren Umgang mit ihnen und die Art und Weise, wie wir sie verarbeiten.

Ziele, Motive und emergentes Bewusstsein

Peter Gerjets erforscht seit Jahren die Auswirkungen dieser technischen Revolution auf unsere kognitiven Möglichkeiten und Emotionen. Seine Arbeiten befassen sich mit dem Einfluss von Multi-Touch-Oberflächen auf das menschliche Lernverhalten, mit der Optimierung multimedialen Lernens, der multimodalen Interaktion in virtuellen Realitäten oder der kognitiven Verarbeitung visualisierter Information.

Begonnen hat Peter Gerjets seinen akademischen Lebensweg mit einem Studium der Philosophie und Sprache. „Allerdings fand ich den rein philosophischen Ansatz bald unbefriedigend, weil das letztlich immer ein Operieren in Prämissen und Konklusionen ist“, erinnert sich Gerjets. „So bin ich dann auf die psychologische Handlungstheorie gestoßen und habe begonnen, parallel Psychologie zu studieren. Dort habe ich mich stark mit Motivations- und Kognitionspsychologie befasst.“

Gerjets Doktorarbeit *Zur Verknüpfung psychologischer Handlungs- und Kognitionstheorien. Die strukturalistische Konstruktion intertheoretischer Bänder am Beispiel von Rubikontheorie der Handlungsphasen und ACT*-Theorie* war noch weitgehend wissenschaftstheoretisch ausgelegt. „Bei kognitiven Systemen“, erläutert der Wissenschaftler die Fragestellung seiner Promotionsschrift, „gibt es so etwas wie ein Motivationsloch.“ Kognitive Systeme würden deklaratives Wissen repräsentieren, auch prozedurales Wissen, also das Wissen um Handlungsabläufe, und seien aufgrund ihrer Architektur in der Lage, Probleme zu lösen. Allerdings hätten solch wissenschaftsbasierte Systeme keine wirklichen Ziele, sondern nur Repräsentationen von Zielen. „Das heißt“, so Gerjets, „dass ich einer

Maschine Ziele einprogrammieren kann, von diesen bildet sie dann Unterziele, greift auf Wissen zurück und kann Regeln anwenden, aber letztlich fehlt ihr die intrinsische Motivation.“ Maschinen würden letztlich nur Programme abarbeiten, Menschen aber hätten Gründe, etwas zu tun, und wollten etwas erreichen.

Deshalb gebe es auf der anderen Seite psychologische Handlungskontrolltheorien, die zu erklären versuchten, wie wir unsere Bedürfnisse, Motive und Intentionen entwickeln und sie zur Handlungsregulation nutzen. „Bei beiden Ansätzen geht es letztlich um Handlungstheorien – einmal kognitiv gedacht, einmal motivational gedacht“, erklärt Gerjets. „In meiner Doktorarbeit habe ich versucht, beide Bereiche zusammenzubringen, allerdings auf eine sehr formale Art und Weise.“

Mit Blick auf das menschliche Bewusstsein stelle sich bei komplexen kognitiven Architekturen die Frage, wann oder an welcher Stelle höherstufige Prozesse aus einem solchen System emergieren. „Unser Gehirn ist ja letztlich auch nur Physiologie, aber irgendwie schafft es das Gehirn, emergent Bewusstsein zu produzieren.“ Die Frage sei also, was geschehen müsse, damit man einem künstlichen kognitiven System Motivationen, Ziele und Handlungssteuerung zuschreiben könne.

Dichtende KI

„Im Moment interessiert mich dieses Thema wieder brennend“, betont Gerjets, „da wir uns ganz aktuell mit der Nutzung von künstlicher Intelligenz beim kreativen Schreiben beschäftigen.“ Inzwischen gebe es Systeme wie den Sprachgenerator GPT-3, der mittels Deep Learning in der Lage sei, künstliche Texte zu erzeugen, die sich nur schwer von menschlichen Texten unterscheiden ließen. GPT-3 ist der Nachfolger von GPT-2 und wurde von OpenAI entwickelt, einem Labor für künstliche Intelligenz mit Sitz in San Francisco. „Man braucht nur einen Textanfang von fünf Worten zu schreiben“, erzählt Gerjets sichtlich angetan, „und das Programm produziert einen geschliffen narrativen Text, der teilweise witzig ist und einen Plot hat.“ Trainiert werde das Programm auf Basis des Internets mit Milliarden von Verknüpfungen, die Voraussagen über die Wahrscheinlichkeit von Wortfolgen ermöglichen. „Der Witz ist: Die Texte sind teilweise wirklich gut, doch tatsächlich steckt nichts dahinter, die haben keine Intention, keine Absicht“. Besonders deutlich werde das bei Sachtexten. Die von GPT-3 geschriebenen Texte seien zwar sehr gut formuliert, teilweise besser als von menschlichen Autoren, letztlich aber vollkommener Unsinn: „Wir benutzen das Programm für ein Creative-Writing-Projekt, das wir zusammen mit dem Deutschen Literaturarchiv Marbach durchführen. Dabei geht es darum, wie gut GPT-3 Kurzgeschichten, Gedichte oder Songtexte verfassen kann, wenn ein entsprechender Anfang vorgegeben wird“, erläutert Gerjets. „Ziel ist es, herauszufinden, ob das im Kontext kreativen Schreibens nützlich wäre oder ob ein solches Programm Autoren eher blockiert, weil es in Millisekunden kreative Texte produziert.“

Auf die Idee zu diesem Projekt sei er gekommen, als er das erste Mal GPT-2 nutzen konnte und zusammen mit seinen Kindern versuchte, einen kleinen Krimi zu schreiben, in dem ein Lastwagen eine Rolle spielt, der auf einen Bauernhof fährt. „Ich begann damit, dass der Lastwagenfahrer eine Pause macht und das System schrieb dann weiter: ‚Und er grillte seine Würstchen auf den Eingeweiden seines Lastwagens.‘ Das hat mich schwer beeindruckt. Ich fand die Idee unheimlich kreativ. Der Plot hat zwar nie richtig funktioniert, aber für einzelne Formulierungen oder unerwartete Wendungen hat das durchaus Potenzial.“

Im Grunde benutze das System das semantische Verständnis von Personen, die schreiben. „Menschen, die Texte verfassen, haben einen wirklichen Körper und wirkliche Erfahrungen“, erläutert der Wissenschaftler weiter. „Auf Basis dieses geklauten Embodiments generiert das System dann auf einer rein sprachlichen Ebene Texte, ohne dass es selbst dieses Embodiment hat. GPT-3 benutzt uns Menschen quasi als Sensoren zur Welt, um zu lernen, wie Menschen denken, handeln und sprechen.“

Dabei ist das System ziemlich erfolgreich. In einer Studie haben Gerjets und sein Team literarische Texte von GPT-3 fortschreiben und von Menschen bewerten lassen, ob sie von einer KI stammen oder nicht. „Heraus kam: Die Leute konnten das kaum unterscheiden. Die Verwechslung lag bei 40 %, dafür waren sie sich ihrer Einschätzung sehr sicher. Hohe Verwechslungsraten bei hoher Overconfidence zeigen jedoch, dass die künstlichen Texte ziemlich glaubwürdig sind und sehr nah an dem dran, was man Menschen zutraut.“

An expositorischen Texten, so Gerjets, würde das System jedoch scheitern, da ihm das notwendige Weltwissen fehle. Literarische Texte, insbesondere etwa Gedichte mit freier Form, könnte es nach Belieben generieren. Das zeige exemplarisch auch das Projekt „Gedichtgrube“ des Bloggers Lukas Diestel, der GPT-2 aus seinem Blutzuckerspiegel alle 5 Minuten Gedichte generieren lässt, die dann von den Lesern entdeckt oder verworfen werden können (das Projekt ist einsehbar unter falschgefuehle.de).

Handlungssteuerung und Motivation bei Multimedia-Texten

Nach seiner Promotion begann Gerjets, Fragen der Handlungssteuerung, Handlungsregulation und der kognitiven Kontrolle auf Mediensettings anzuwenden. „Handlungssteuerung bedeutet ja auch, Strategien einzusetzen, um zu erkennen, was man eigentlich macht, ob das gut ist, was man macht, ob es funktioniert“, erklärt er. „Zugleich heißt es aber auch, zu erkennen, ob man sich ablenken lässt, Motivationsprobleme hat oder ob die Ausdauer ausreicht.“

Ein wichtiges Thema in diesem Zusammenhang waren die seinerzeit intensiv diskutierten Hyperlinks, die nicht nur Inhalte vertiefen und erläutern, sondern auch das Potenzial haben, von den eigentlichen Inhalten abzulenken. „Wir haben damals eine Lernumgebung mit zwei Themen gebaut, einem echten Thema und einem Ablenkungsthema. Das echte Thema war Stochastik – motivational für die meisten eher schwierig –, das wir eingebettet hatten in Fragen zu Attraktivität und Partnerwahl.“

Tatsächlich ließen sich die Probanden durch Links zu dem für sie überaus interessanten Thema „Partnerwahl“ erheblich von den Lernaufgaben ablenken. Sogar Versuchspersonen, die den Hyperlinks nicht folgten, waren in ihren Ergebnissen schlechter als die Vergleichsgruppe ohne Hyperlinks. „Das hinter dem Hypertext liegende Thema aktiviert offensichtlich Erwartungen, Fragen und Entscheidungsprobleme, die den Lernprozess und das sich dabei bildende Kohärenzmodell unterbrechen, auch wenn die Links überhaupt nicht benutzt werden.“

Diese zunächst mit rein psychologischen Methoden gewonnenen Ergebnisse konnten Gerjets und sein Team 20 Jahre später durch EEG-Untersuchungen neuroanatomisch bestätigen und erklären. „Letztlich können Hyperlinks oder auch Bilder in Multimedia-Texten zu einer kurzfristigen Überlastung des Arbeitsgedächtnisses führen“, erklärt der Wissenschaftler. „Und das ist natürlich ein Phänomen, das uns in vielen Bereichen interessiert, vom Internetsurfen bis zum Lesen von Wikipedia-Artikeln.“

»Die Leute denken, sie könnten mit der Vielfalt digitaler Angebote umgehen, sie können es aber letztlich nicht.«

Selbstverständlich habe das kuratierte Lesen traditioneller analoger Texte auch Nachteile, insbesondere, was die Informationstiefe und Informationsvielfalt angehe, es habe aber auch Vorteile hinsichtlich der Verarbeitung der präsentierten Information. „Die Leute denken, sie könnten mit der Vielfalt digitaler Angebote umgehen, sie können es aber letztlich nicht.“ Das zeigten auch Untersuchungen zu multiplen Dokumenten. Tatsächlich würden sich die Menschen einbilden, sie könnten die darin enthaltenen Informationen gut gewichten und einordnen sowie die Motivation und Expertisen der Autoren erkennen. Tatsächlich könnten sie das aber nur sehr schlecht.

„Wir forschen letztlich über Aufmerksamkeit, vor allem über Aufmerksamkeitsprobleme und Aufmerksamkeitsbelastungen“, fasst Gerjets zusammen. Die Angebote im Netz, sofern sie kostenlos seien, würden letztlich alle über Aufmerksamkeitsökonomie finanziert. „Die reichsten Techkonzerne sind deshalb reich, weil sie es geschafft haben, unsere Aufmerksamkeit zu binden“, betont der Psychologe. Die Algorithmen der großen Netzanbieter wollten nicht unser Wissen vertiefen, sondern unsere Aufmerksamkeit auf kommerzielle Inhalte lenken. Um unsere Aufmerksamkeit abzubauen, reiche allein schon das Bewusstsein, einen Text digital zu lesen. „Der Leseduktus ist einfach ein anderer. Wir haben ein Behavior Setting ‚Buchlesen‘, das uns gar nicht dazu bringt, anzunehmen, dass am Ende des Absatzes auf einen anderen spannenden Text verwiesen wird, sondern wir lesen im Deep-Reading-Modus weiter. Das ist im Netz oder mit einem Lesegerät grundlegend anders.“ Es sei daher wichtig, diesen Deep-Reading-Modus zu schulen.

Handnähe und die lernpsychologischen Möglichkeiten der VR

Informationsaufnahme, das machen die Arbeiten von Peter Gerjets und seinem Team deutlich, ist mehr als nur das Decodieren von Buchstaben. Menschen nehmen Informationen mit ihrem ganzen Körper auf, auch wenn sie lesen. Es ist ein Unterschied, ob man ein Buch oder mit einem Tablet liest, ob man einen Hypertext oder an einem Touchscreen studiert. Sogar die Position des Bildschirms ist von erheblicher Bedeutung. „Wenn ich einen Touchscreen in einem Museum an eine Wand hänge“, erläutert der Forscher, „versetzen sich die Besucher in eine Art Berieselungsmodus. Ganz anders, wenn man den Bildschirm um 90 Grad dreht und als Tisch präsentiert. Dann entsteht eine Arbeitsatmosphäre und die Leute setzen sich viel aktiver mit den Inhalten auseinander.“ Um solche Phänomene zu erforschen und die Ergebnisse unmittelbar anwenden zu können, arbeitet Gerjets eng mit Museen zusammen, etwa dem Herzog Anton Ulrich-Museum in Braunschweig oder der Dokumentation Obersalzberg.

Eine entscheidende Rolle beim Lesen spielen die Hände. Der Grund dafür sind offensichtlich unterschiedliche neuronale Verarbeitungspfade, der wahrnehmungsorientierte ventrale und der handlungsorientierte dorsale Verarbeitungsweg. „Die

Idee ist nun, dass man mit Touch-Interfaces den dorsalen Verarbeitungsweg stärker anspricht, auch weil dieser zentral für das Arbeitsgedächtnis ist.“ Offensichtlich, so Gerjets, würden visuelle Informationen besser verarbeitet, wenn man das entsprechende Objekt anfassen könne.

Das gelte nicht nur für Objekte und Bilder, sondern auch für Texte. Studien hätten gezeigt, dass Menschen selbst sinnlose Sätze eher als sinnvoll erachteten, wenn sie sie in der Hand hielten. „Die Autoren dieser Studie sind übrigens der Ansicht“, ergänzt Gerjets, „dass diese Handnähe sogar eine Verschlechterung der semantischen Verarbeitung mit sich bringt, was mit Blick auf das kritische Textlesen auf Smartphones ein interessanter Aspekt sein könnte.“ Der Glaube an Verschwörungstheorien und Fake News habe zwar vor allem sozialpsychologische Gründe. „Doch Menschen verfahren nach einfachen Heuristiken – und der Handnähe-Effekt könnte das noch unterstützen“, so Gerjets.

In seinen aktuellen Studien beschäftigen sich Gerjets und sein Team mit den lernpsychologischen und kognitionswissenschaftlichen Aspekten virtueller Realitäten (VR), die die Möglichkeiten von Touch-Interfaces noch einmal erweitern. „Adaptive Lese- und Lernumgebungen, in denen sich die Benutzer tatsächlich in Wissensräumen bewegen können, bieten faszinierende Möglichkeiten“, betont der Wissenschaftler, etwa auch für eine Erschließung von Archiven, die man mithilfe von VR dreidimensional und körperlich erlebbar machen könnte. Gerjets fasst zusammen: „Letztlich könnte virtuelle Realität es uns ermöglichen, die Vorteile traditioneller haptischer Wissensaneignung mit den Möglichkeiten digitaler Informationsvermittlung zu kombinieren.“



Dr. Alexander Grau arbeitet als freier Kultur- und Wissenschaftsjournalist u. a. für „Cicero“, „NZZ“ und den Deutschlandfunk.