

Die Macht der Autorenschaft

Auswirkungen der Kenntnis über die Autorenschaft (generative KI vs. Mensch) auf die Glaubwürdigkeit von Informationstexten.

Clara Schmidt
Lisa Opferkuch
Margit Rosenberg
Marina Neubert

Forschungskolloquium
Wintersemester 2023/24
Prof. Dr. phil. Michael Meng
Pia Bsdurrek



FORSCHUNGSSTAND

Generative Künstliche Intelligenzen (KI), die auf umfangreichen Sprachmodellen wie ChatGPT, Google Bard, Claude 2 oder Bing Chat basieren, haben in beeindruckendem Tempo die Sprachverarbeitung revolutioniert. Inzwischen sind diese Anwendungen weitverbreitet und werden in nahezu allen Bereichen genutzt. Mit zunehmender Verbreitung kommt auch die Debatte der Identifizierbarkeit von KI-generierten Inhalten gegenüber Mensch-verfassten Inhalten auf. Für einen Menschen kann es eine Herausforderung darstellen, den Ursprung der Inhalte zu identifizieren. Aber hat die Kenntnis über die Autorenschaft überhaupt einen Einfluss auf die Glaubwürdigkeit der Inhalte? Zahlreiche Studien haben sich bereits mit dieser aktuellen Debatte befasst, wobei zwei Untersuchungen für die vorliegende Arbeit von besonderer Relevanz sind.

Die im Jahr 2023 von Huschens et al. durchgeführte Studie „Do You Trust ChatGPT“¹ untersuchte die Wahrnehmung der Glaubwürdigkeit von ChatGPT mithilfe eines Fragebogens. Die Ergebnisse der Studie zeigten, dass die Teilnehmenden sowohl KI-generierte als auch Mensch-verfasste Inhalte als ähnlich glaubwürdig bewerteten. Die zweite für diese Studie relevante Forschung war die Validierungsstudie zur „General Attitudes towards Artificial Intelligence Scale“ (GA AIS)², bei der die Einstellung der Befragten gegenüber KI anhand von 20 Items auf einer fünfstufigen Likert-Skala abgefragt wird. Diese beiden Studien bilden die Grundlage unserer Forschung.

FORSCHUNGSFRAGE

Welchen Einfluss hat die Kenntnis über die Autorenschaft (KI vs. Mensch) auf die Glaubwürdigkeit von Informationstexten?

HYPOTHESEN

H1 Rezipienten schätzen KI-generierte Informationstexte als weniger glaubwürdig ein im Vergleich zu Mensch-verfassten Informationstexten.

H1A Rezipienten schätzen KI-generierte Informationstexte als weniger kompetent ein im Vergleich zu Mensch-verfassten Informationstexten.

H1B Rezipienten schätzen KI-generierte Informationstexte als weniger vertrauenswürdig ein im Vergleich zu Mensch-verfassten Informationstexten.

H1C Rezipienten schätzen KI-generierte Informationstexte als weniger klar ein im Vergleich zu Mensch-verfassten Informationstexten.

H1D Rezipienten schätzen KI-generierte Informationstexte als weniger einnehmend ein im Vergleich zu Mensch-verfassten Informationstexten.

H2 Jüngere Altersgruppen schätzen KI-generierte Informationstexte als glaubwürdiger ein als ältere Altersgruppen.

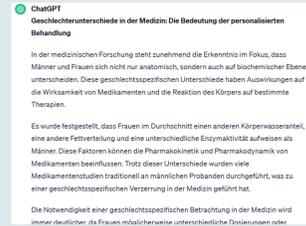
H3 Menschen, die eine positive Einstellung gegenüber KI haben, empfinden KI-generierte Informationstexte als glaubwürdiger, als jene, die eine negative Einstellung gegenüber KI haben.

H4 Menschen, die eine positive Einstellung gegenüber KI haben, nutzen generative KI häufiger, als jene, die eine negative Einstellung gegenüber KI haben.

H5 Menschen, die generative KI häufiger verwenden, finden KI-generierte Informationstexte glaubwürdiger, als jene, die generative KI seltener verwenden.

METHODE

Die Studie wurde mittels **Fragebogen** durchgeführt. Im Fragebogen erhielt jeder Proband als Stimuli jeweils einen Sachtext über Bienenkommunikation und einen Sachtext über Medizin. Das visuelle Erscheinungsbild jedes Textes machte jeweils einmal die „Autorenschaft ChatGPT“ und einmal die „Autorenschaft Mensch“ für die Probanden kenntlich. Die Anordnung der Themen und der Autorenschaft wurde randomisiert.



BIENEN: KLEINE HELDEN DER NATUR

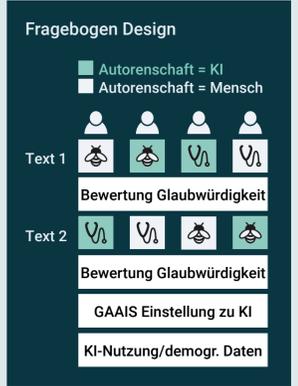
19. Oktober 2023 Maximilia Müller

Bienen, faszinierende Bestäuber und Honigproduzenten, spielen eine zentrale Rolle im Gleichgewicht der Natur. Ein Bienenvolk besteht aus Königin, Arbeiterinnen und Drohnen, die gemeinsam ein hochorganisiertes System bilden. Die Arbeiterinnen sind Altkolonen, sammeln Nektar, produzieren Honig und bestäuben Pflanzen.

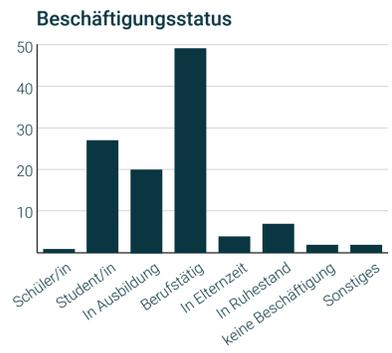
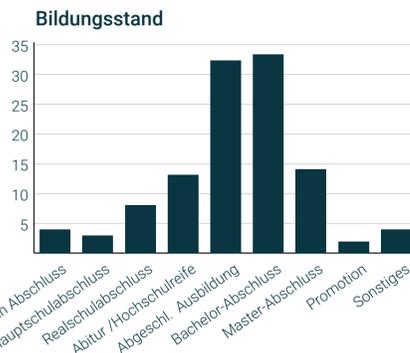
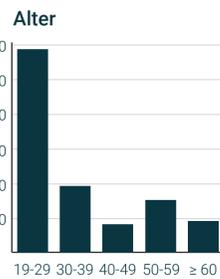
Besonders beeindruckend ist die Bienenkommunikation. Durch komplexe Tänze und Gerüche teilen sie Informationen über die Lage von Nahrungsquellen mit. Doch trotz ihrer Bedeutung stehen Bienen vor Herausforderungen wie Pestizidinsatz, Lebensraumverlust und dem Klimawandel.

Die Konsequenzen betreffen nicht nur die Bienen, sondern auch unsere

Nach dem Lesen der Informationstexte wurden vom Probanden Adjektive zur Einschätzung der Glaubwürdigkeit des Textes als mehr bzw. weniger zutreffend bewertet. Die dabei verwendeten Adjektive wurden der Studie „Do You Trust ChatGPT“ von Huschens et al. aus dem Jahr 2023 für diese Arbeit entnommen. Anschließend wurden die Probanden zu ihrer allgemeinen Einstellung bezüglich KI anhand des GA AIS befragt. Abschließend wurden die KI-Nutzung und einige demografische Daten erhoben.



TESTPERSONEN



AUSWERTUNG

Vor der Datenauswertung wurden die erhobenen Daten bereinigt und aufbereitet sowie die interne Konsistenz der verwendeten Skalen überprüft. Hierzu wurde Cronbachs Alpha berechnet. Je größer Cronbachs α ist, desto größer ist die Reliabilität der Gesamtskala. Die Ergebnisse zeigten eine gute interne Konsistenz über alle Skalen an (α -Werte zwischen 0,831 und 0,883).

Zur Überprüfung der Hypothesen wurden ungepaarte t-Tests sowie Pearson-Korrelationen durchgeführt. Hierbei wurden für H1A-D, H3 und H5 jeweils Tests zum Thema Biene sowie zum Thema Medizin durchgeführt.

Die **t-Tests** zeigten keine signifikanten Unterschiede zwischen den Gruppen KI und Mensch für H1A ($p = 0,686$), H1B ($p = 0,888$), H1C ($p = 0,995$) und H1D ($p = 0,602$) zum Thema Biene und für H1C ($p = 0,479$) und H1D ($p = 0,541$) zum Thema Medizin.

Bei H1A und H1B zum Thema Medizin zeigten sich hingegen **signifikante Unterschiede** zwischen den Gruppen bezüglich der Kompetenz ($t(95,33) = -3,570, p < 0,001$) sowie der Vertrauenswürdigkeit ($t(107) = -3,203, p = 0,002$). Rezipienten schätzen KI-generierte Informationstexte als signifikant weniger kompetent sowie vertrauenswürdig ein im Vergleich zu Mensch-verfassten Informationstexten.

Die **Korrelationen** zeigten keine signifikanten Ergebnisse für H2 ($p = 0,346$), H3 ($p = 0,751$) und H5 ($p = 0,316$) zum Thema Biene und für H2 ($p = 0,684$) und H5 ($p = 0,399$) zum Thema Medizin.

Die Überprüfung der H3 zum Thema Medizin zeigte hingegen eine **signifikante, schwach positive Korrelation** zwischen der Einstellung und der Glaubwürdigkeit des KI-Textes ($r = 0,285, p = 0,031$). Je positiver die Einstellung des Probanden gegenüber KI, umso glaubwürdiger wurde der Text eingeschätzt. Die Überprüfung der H4 zeigte, dass die Einstellung und die Nutzung **signifikant, moderat negativ miteinander korrelieren** ($r = -0,414, p < 0,001$). Je positiver die Einstellung gegenüber KI, desto weniger häufig wird generative KI verwendet.

ERGEBNISSE

Die Ergebnisse zur **Glaubwürdigkeit von Informationstexten zum Thema Biene** zeigten, dass es keine signifikanten Unterschiede zwischen den Gruppen KI und Mensch gab, in Bezug auf die Einschätzung der Kompetenz, Vertrauenswürdigkeit, Klarheit und wie einnehmend der Text war. Die Ergebnisse zur **Glaubwürdigkeit von Informationstexten zum Thema Medizin** zeigten hingegen, dass es signifikante Unterschiede zwischen den Gruppen KI und Mensch in Bezug auf die Einschätzung der Kompetenz und der Vertrauenswürdigkeit gab. Rezipienten schätzen KI-generierte Informationstexte als signifikant weniger kompetent sowie vertrauenswürdig ein im Vergleich zu Mensch-verfassten Informationstexten beim Thema Medizin. Es wurden allerdings keine signifikanten Unterschiede bei der Einschätzung der Klarheit und wie einnehmend der Text war, ermittelt.

Bezüglich des **Alters** und der **Glaubwürdigkeit** zeigten sich keine signifikanten Ergebnisse, weder beim Thema Biene noch beim Thema Medizin. Es gab somit keine signifikante Korrelation zwischen dem Alter und der Glaubwürdigkeit von KI-generierten Informationstexten.

Beim KI Text mit dem Thema Biene gab es keine signifikante Korrelation zwischen der **Einstellung gegenüber KI** und der **Glaubwürdigkeit**. Beim KI Text mit dem Thema Medizin hingegen gab es eine schwache positive Korrelation zwischen der Einstellung und der Glaubwürdigkeit. Je positiver die Einstellung des Probanden gegenüber KI, umso glaubwürdiger wurde der Text eingeschätzt.

In Bezug auf die **Einstellung** und **Nutzungshäufigkeit von KI** zeigten sich signifikante Ergebnisse. Die Einstellung und die Nutzung korrelieren negativ moderat miteinander. Je positiver die Einstellung gegenüber KI, desto weniger häufig wird generative KI verwendet.

Bezüglich der **Glaubwürdigkeit** und der **Nutzungshäufigkeit** wurden keine signifikanten Ergebnisse ermittelt. Beim KI Text zum Thema Biene sowie zum Thema Medizin gab es keine signifikante Korrelation zwischen der Nutzung von KI und der Glaubwürdigkeit.

LIMITATIONEN

Die Studie weist einige Einschränkungen auf, die bei der Interpretation der Ergebnisse berücksichtigt werden müssen:

- Zunächst war der Screenshot für die mobile Anwendung nicht optimal gestaltet. Beim Heranzoomen ging die Kennzeichnung von ChatGPT verloren, was möglicherweise Auswirkungen auf die Beurteilung haben könnte.
- Ein weiterer Aspekt, der berücksichtigt werden muss, ist das Fehlen einer Kontrollfrage zur Wahrnehmung der Probanden bezüglich der verschiedenen Autoren. Dies könnte potenziell Einfluss auf die Validität der Ergebnisse haben.
- Die Altersanalyse erfolgte nicht differenziert in Gruppen wie „jung“ und „alt“.
- Ein weiterer Punkt betrifft die überwiegend weibliche Zusammensetzung der Stichprobe. Eine ausgeglichene Geschlechterverteilung ist repräsentativer.
- Zuletzt bleibt die Unsicherheit darüber, ob die Ergebnisse durch die Themen (Medizin und Biene) beeinflusst wurden.

FAZIT

Unter Verwendung der "General Attitudes towards Artificial Intelligence Scale" schätzten Probanden KI-generierte Texte umso glaubwürdiger ein, je positiver ihre Einstellung gegenüber KI ist. Gleichzeitig sank die positive Einstellung der Probanden gegenüber KI mit steigender Nutzungshäufigkeit von KI-Systemen.

Im Unterschied zur Studie von Huschens et al. zeigt diese Studie, dass KI-generierte Texte weniger glaubwürdig (kompetent, vertrauenswürdig) sind als Mensch-verfasste, jedoch nur bei dem Thema Medizin. Hier könnten weitere Forschungen ansetzen und untersuchen, ob das Thema einen Einfluss auf die Glaubwürdigkeit hat.

1 Huschens, Martin, Martin Briesch, Dominik Sobania and Franz Rothlauf (2023): Do You Trust ChatGPT? - Perceived Credibility of Human and AI-Generated Content.

2 Astrid Schepman & Paul Rodway (2022): The General Attitudes towards Artificial Intelligence Scale (GA AIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust, International Journal of Human-Computer Interaction.