

# Fake News und Desinformation als Herausforderung für die Gesellschaft

## Das CAIS: Forschung zur gemeinwohlorientierten digitalen Transformation

Das Center for Advanced Internet Studies (CAIS) hat das Ziel, an der verantwortlichen Gestaltung des durch die Digitalisierung ausgelösten sozialen, politischen und wirtschaftlichen Wandels aktiv mitzuwirken und den Dialog zwischen Wissenschaft und Gesellschaft zu fördern. Das CAIS versteht sich als Ort innovativer interdisziplinärer Forschung und als Impulsgeber für eine kritische Öffentlichkeit, die sich den Chancen und Risiken der Digitalisierung stellt und sich über ein normatives Leitbild für das Leben in einer digitalen Gesellschaft verständigen will.



© Michael Schwettmann

**Prof. Dr. Christiane Eilders**

Center for Advanced Internet Studies (CAIS)  
Wissenschaftliche Direktorin  
[christiane.eilders@cais-research.de](mailto:christiane.eilders@cais-research.de)

### Entwicklungen der Medienbranche in den nächsten Jahren

Das CAIS beobachtet einen Bedeutungszuwachs für generative KI auf der Basis von Large Language Model (LLMs) im Medienbereich. Das betrifft zum einen den generellen Bedeutungsverlust von Medien als Informationsquellen zugunsten von KI. So suchen Menschen zunehmend Nachrichten direkt über ChatGPT. Zum anderen betrifft das eine Verschiebung innerhalb der medial vermittelten Information – und zwar im Bereich der journalistischen Medien, der Social Media Plattformen und der Nutzerkommentare unter Nachrichtenwebsites. LLMs sorgen bei der direkten KI-basierten Nachrichtensuche, in den journalistischen Medien, aber auch auf Journalismus-nahen Kanälen von Plattformen dafür, dass nicht mehr alle Texte von Menschen erstellt werden und Texte schneller generiert werden, sei es in der Nachrichtenberichterstattung oder in anderen Textgattungen. Gleichzeitig sind diese Texte nicht mehr über die journalistische Arbeit qualitätsgesichert. Sie können unabsichtliche Fehler enthalten, die nicht auf den ersten Blick erkennbar sind. Sie können aber auch absichtlich als Fake News oder als Teile von Desinformationskampagnen erzeugt werden, um die öffentliche Meinung beeinflussen. Das kann sowohl von anderen Staaten gesteuert werden als auch von Akteuren im eigenen Staat, die Interesse an Irritation und Unruhe haben.

Das Gleiche gilt für Social Media Inhalte, die von offen nicht-journalistisch arbeitenden Akteuren generiert werden, etwa von Influencern oder Aktivistinnen und Aktivisten. Hier bestehen selbstverständlich noch weitreichendere Möglichkeiten für manipulative Einflussnahme durch generative KI als in den meist qualitätsgesicherten journalistischen Medien. Bei der Verbreitung von KI-basierten Fake News und Desinformationen spielen auch Nutzerkommentare unter Nachrichtenwebsites eine wichtige Rolle. Viele Kommentare sind durch LLMs oder mit LLM-Unterstützung erzeugt worden. Sie können ebenso wie die ursprünglichen Nachrichten ein falsches Bild der öffentlichen Meinung zu einem Thema zeichnen und damit Einfluss auf die Meinungsbildung gewinnen. Wegen des Konformitätsdrucks haben dann diese vielen künstlich generierten Kommentare die Chance, die öffentliche Artikulation abweichender Meinungen zu entmütigen, da die besonders häufig sichtbaren Meinungen für die öffentliche Meinung gehalten werden.

Es ist bekannt, dass Fake News und Desinformation auf Empörung und Hass setzen oder zumindest emotionale Reaktionen auslösen und daher besonders schnell verbreitet werden. Sie stehen in Verdacht, das Vertrauen in staatliche Institutionen und in als staatsnah wahrgenommene Akteure zu schwächen. Das kann die Ordnungskräfte oder aber die Parlamente betreffen, aber auch Krankenwagenbesatzungen und Feuerwehr. In jedem Fall gilt: Fake News und Desinformation werden in dem Maße problematisch, in dem diese Inhalte schnell an sehr große Kreise von Nutzerinnen

und Nutzer weiterverbreitet und als öffentliche Meinung wahrgenommen werden. Da in demokratischen Regimen die öffentliche Meinung relevant für politische Entscheidungen ist, können diese KI-gestützten Falschinformationen durchaus einen Einfluss auf Politikentscheidungen haben. Die Niederschwelligkeit der Erzeugung und Verbreitung von Fake News und Desinformation mithilfe generativer KI bedeutet damit in digital vernetzten Umgebungen ein erhebliches Einflusspotenzial, wenn es darum geht, demokratische Regimes zu irritieren und sogar zu destabilisieren.

Da diese negativen Konsequenzen von generativer KI im Medienbereich als Gefährdung der Demokratie wahrgenommen werden, wurden und werden auch Gegenstrategien entwickelt. Diese Gegenstrategien gehören ebenso wie die Verbreitung von Fake News und Desinformationen zu den Entwicklungen im Medienbereich, auf die wir uns einzustellen haben. Das umfasst eine Zunahme der Nachfrage nach und des Angebots an Faktenchecks in den journalistischen Medien. Das umfasst aber auch die Verbreitung und niederschwellige Nutzung von – ebenfalls KI-unterstützten – Tools zur Detektion von Fake News und bot-generierten Kommentaren. Im Idealfall könnten Nutzerinnen und Nutzer so selbst dazu beitragen, dass Falschmeldungen entlarvt und dass die Verzerrungen der öffentlichen Meinung abgemildert werden.

### Schwerpunkte der CAIS-Aktivitäten

Das CAIS untersucht mit Blick auf das Ideal der informierten Meinungsbildung und der demokratischen Willensbildung und Entscheidungsfindung, wie die digitale Transformation gemeinwohlorientiert und menschenfreundlich gestaltet werden kann. Wir erforschen und erproben Möglichkeiten, wie Fake News und Desinformation entdeckt werden können und ihre Verbreitung verhindert werden kann. Es ist zu vermuten, dass auf beiden Seiten immer bessere Tools für Manipulation und Detektion entwickelt werden und die ressourcen-stärkere Seiten diesen Wettlauf gewinnt. Indem das CAIS Tools entwickelt, die Nutzerinnen und Nutzer befähigen, kritisch mit Informationen umzugehen und ihren Wahrheitsgehalt zu prüfen, trägt es zu einer Chancenverbesserung der Nutzerseite bei.

Dieses Ziel steht auch hinter unserer Forschung zur Verbesserung von Medien- oder Digitalkompetenz. Wir untersuchen, wie Kinder und Jugendliche, aber auch Erwachsene, in die Lage versetzt werden können, KI zu verstehen und selbst anzuwenden. Diese Kompetenzen sollen die Nutzerinnen und Nutzer sensibilisieren und einen reflektierten Umgang mit medialen Inhalten und ihrer Weiterverbreitung ermöglichen.

Weiter untersuchen wir, welche Rolle Falschinformationen für die Polarisierung der Gesellschaft spielen. An den Rän-



© Michael Schwellmann

dern des politischen Spektrums können sich – nicht zuletzt durch die Verbreitung von Falschinformationen – Gruppen von Gleichgesinnten bilden, die eine Abspaltung von der moderateren Mehrheitsgesellschaft vorantreiben. Wir wollen aufklären, nach welchem Muster sich die Meinungen in einer Gesellschaft verschieben, wenn Falschinformationen verbreitet werden. Unter welchen Bedingungen etwa entwickeln sich zwei Gruppen in Bezug auf ein Streithema auseinander, wann dagegen kommt es zu einem Rechts- oder Links-Ruck beider Lager? Ziel ist es auch, festzustellen, welche Art der Botschaft zu einer Mäßigung der Meinungen auf beiden Seiten, also einer De-Polarisierung, führt und inwiefern hierbei generative KI nützlich sein könnte.

### Taskforce „Künstliche Intelligenz im Bildungswesen“

Die Verbreitung von KI in allen Bereichen des gesellschaftlichen Lebens ruft eine gewisse Verunsicherung hervor. Die Politik hat daher erheblichen Beratungsbedarf. Das CAIS hat – angeregt durch die Politik – eine Taskforce „KI im Bildungswesen“ ins Leben gerufen, in der die Einsatzmöglichkeiten von KI für Lehren und Lernen diskutiert und Handlungsempfehlungen an die Landesregierung formuliert wurden. Die Medienbranche ist dabei indirekt über die Notwendigkeit der Vermittlung von Medienkompetenz angesprochen. So kann der Journalismus zur Aufklärung des Publikums beitragen, indem die Manipulationsmöglichkeiten erläutert und Fake News und Desinformationen entlarvt werden.

In Bezug auf Methodeninnovationen spielt die generative KI nicht nur dann eine Rolle, wenn es um die Entdeckung von Manipulationsversuchen geht. Sie kann auch genutzt werden, um allgemeinere inhaltliche Muster in journalistischen Medien und auf Social Media zu rekonstruieren. Mithilfe von LLMs können Konfliktlinien und Interpretationsweisen von aktuellen Streit-Themen, aber auch Vielfalt und Lösungs-ideen in gesellschaftlichen Diskursen frühzeitig entdeckt werden.