

# „ES FEHLT DIE KRITISCHE EVALUATION DER SYSTEME“

*Suchmaschinen wie Google, Bing oder DuckDuckGo sind heute die Gatekeeper zu Informationen, zum Beispiel über Politiker und deren Aktivitäten. Philipp Schaer von der TH Köln fordert von ihren Anbietern mehr Transparenz, Fairness und Heterogenität*

INTERVIEW: VERONIKA RENKES

**Herr Prof. Schaer: Sie erforschen die Interaktion von Algorithmen und Nutzern bei der Informationssuche in Suchmaschinen. Warum ist das wichtig?**

Bei meiner Forschung steht der Mensch im Zentrum beziehungsweise die Interaktion von Mensch und Maschine vor allem bei der Informationssuche. Schon seit ein paar Jahrzehnten suchen wir Informationen nicht mehr analog, sondern digital mit Suchmaschinen, und das in einer rasanten Geschwindigkeit. Hinter den Suchmaschinen stehen gigantische Softwaresysteme und Algorithmen des Information Retrieval. Die sind darauf angelegt, aus einer unübersichtlichen Menge an Objekten den Nutzern eine relevante Teilmenge in Form einer gerankten Liste zu empfehlen. Um die relevanten Informationen zu ermitteln, setzen heutige Suchmaschinen auf Methoden des Maschinellen Lernens und großer Sprachmodelle. Dadurch kann sich heute jeder selbst seine für ihn wichtigen Informationen beschaffen.

**Das ist doch eigentlich positiv?**

Grundsätzlich schon, aber was fehlt, ist die kritische Evaluation der Systeme. Falschinformationen und verzerrte Darstellungen werden nicht immer identifiziert. Wenn bei Google ein Suchauftrag eingegeben wird, erscheint eine Vielzahl unterschiedlichster Dokumente, in denen die gesuchte Wortkombination auftaucht. Die Frage ist, wie bringt man das Suchergebnis in eine sinnvolle Ordnung?

Der große Erfolgsfaktor von Google war dabei: Was am populärsten ist, wird ganz nach oben an die ersten Stellen der Suchergebnisse geschoben – also etwa Webseiten, die oft verlinkt, oder Suchanfragen, die angefragt oder geklickt werden. Dadurch entstehen aber sich selbst verstärkende Systeme und ein Matthäus-Effekt – was populär ist, wird noch populärer – wovon leider auch Falschinforma-

mationen profitieren. Das bedeutet: Die Relevanz, die das Ranking suggeriert, entsteht aus einer früheren Popularität und weniger aufgrund einer tatsächlich gegenwärtig gegebenen.

**Ein Ziel Ihrer Forschung ist, mehr Transparenz, Qualität und Fairness bei der Informationssuche mit Suchmaschinen zu erreichen. Warum?**

Open Access und frei verfügbare Informationen etwa in digitalen Bibliotheken sind für Wissenschaft und Gesellschaft wichtig. Wir erforschen und entwickeln Werkzeuge, mit denen wir diese selbst verstärkenden Systeme objektiv evaluieren können. Nimmt man zum Beispiel die bekannten Falschinformationen von Google über Bettina Wulff und ihre angebliche Rotlichtmilieuvorgangheit oder über Angela Merkel und den ihr angehafteten jüdischen Lebenslauf bei Bing: Uns interessiert, ob dies dramatische Einzelfälle oder Phänomene des Gesamtsystems sind.

Wir wollen aber nicht nur evaluieren, sondern auch dafür sorgen, dass eine gut nutzbare und zugleich verlässliche Suche ermöglicht wird – gerade auch bei der akademischen Suche. So entwickeln wir zum Beispiel in einem Projekt der Deutschen Forschungsgemeinschaft mit ZB MED, dem Informationszentrum Lebenswissenschaften Köln, einen Dienst, der explizit auf Preprints von wissenschaftlichen Erkenntnissen setzt, die es noch nicht durch das Peer-Review-Verfahren geschafft haben. Gerade in Covid-Zeiten war das wichtig, denn so konnten in der Wissenschaft Preprint-Informationen genutzt werden, um Phänomene zu erklären, die plötzlich auftauchten. Denn in der Pandemie-Situation konnte man ja nicht warten, bis alle Reviewer eines Journals ihre Gutachten eingereicht hatten. Ein Spagat zwischen Aktualität und Verlässlichkeit.

---

**„FÜR EINE DEMOKRATIE IST ES WICHTIG,  
VERSCHIEDENE INFORMATIONQUELLEN MIT  
UNTERSCHIEDLICHEN AUSRICHTUNGEN ZU HABEN.  
DIE GLEICHE HETEROGENITÄT SOLLTE AUCH FÜR  
SUCHANBIETER IM INTERNET EXISTIEREN“**

---

**Was hat Sie bei Ihrer Forschung besonders überrascht?**

Dass der Matthäus-Effekt ein durchgehendes Grundprinzip ist, auch bei der akademischen Suche. So landen sehr bekannte Forscher und Forschergruppen immer ganz oben in den Suchergebnissen. Andererseits gibt es jede Menge wissenschaftliche Fachliteratur, die inhaltlich sehr relevant ist, es aber in diesen Suchsystemen nie nach oben schafft – etwa, weil die Autoren nicht einer renommierten Forschungsgruppe oder einer bekannten Hochschule angehören. Doch nur, weil etwas populär ist, ist es nicht auch schon zwangsläufig relevanter für die Wissenschaft als etwas weniger populäres. Und auch bei politischen Informationen ist eine Heterogenität der Ergebnisse wichtig – und nicht immer ist nur das, was populär ist, ist auch inhaltlich relevant.

**Sie haben die Suchanfragenvorschläge für Politiker und politische Themen in Suchmaschinen untersucht. Ein Ergebnis war, dass es Verzerrungen gibt. Warum ist das so?**

Wir haben uns die Suchanfragenvorschläge von Google, Bing und DuckDuckGo angeschaut, also geschaut, was als Suchervervollständigung vorgeschlagen wird, wenn man die ersten Teile der Suchanfrage getippt hat. Die Grundannahme dieser Langzeitstudie war: Wenn sich Leute informieren wollen über Bundestagsabgeordnete oder politische Themen, dann geben sie zwei, drei Begriffe in die Suchmaschinen ein – also zum Beispiel Angela Merkel. Bevor die Suche richtig losgeht, klappt eine Vorschlagsliste auf. Diese schlägt weitere Suchbegriffe vor – wie etwa bei Bing damals einen vermeintlichen jüdischen Lebenslauf Angela Merkels. Das war schon verwunderlich. Wo kommt das her, fragten wir uns. Menschen, die etwas über Angela Merkels Politik erfahren wollten, wurden umgelenkt auf ganz andere Themen, die aus dem harten rechts-verschwörungstheoretischen Spektrum kamen.

Genau dasselbe passierte bei Bettina Wulff und ihrer angeblichen Rotlichtmilieuvergangenheit. Beides verbunden mit einem Matthäus-Effekt, der durch Aufmerksamkeit gespeist wird und dafür sorgte, dass der Vorschlag immer weiter in der Vorschlagsliste nach oben rückt. Diese Verzerrungen in der Darstellung und Diffamierungen von Politikern sind kei-

ne Einzelphänomene. Gibt man etwa die Namen weiblicher Politiker ein, dann erscheinen in unseren Untersuchungen mehr Suchvorschläge zu Themen aus dem persönlichen oder emotionalen Spektrum als bei männlichen Politikern. Man wird also als Nutzer von dem System gelenkt. Deshalb muss man sich ernsthaft die Frage stellen, inwieweit man den Suchsystemen trauen kann. Für eine Demokratie ist es wichtig, verschiedene Informationsquellen mit unterschiedlichen Ausrichtungen zu haben. Die gleiche Heterogenität sollte auch bei Suchanbietern im Internet herrschen.

**Wie kommt es zu diesen Verzerrungen der Suchvorschläge?**

Dazu gibt es viele Theorien. Geht man davon aus, dass die Suchvorschläge darauf basieren, was Menschen zuvor gesucht haben, dann sind die Verzerrungen letztlich ein Spiegel der Gesellschaft. Ein anderer Grund könnte sein, dass die

---

**PROF. DR. PHILIPP SCHAEER**



Foto: Thilo Schmüllgen/TH Köln

hat eine Schwerpunktprofessur für Data Science mit dem Fokus auf Information Retrieval, Search Engine Technology und Natural Language Processing an der Fakultät für Informations- und Kommunikationswissenschaften der Technischen Hochschule Köln. Zu seinen Forschungsgebieten zählen: Digital Libraries, Evaluation

of IR systems und Living Labs. Im Rahmen des Forschungsprojekts ESUPOL hat er sich mit dem Einfluss von Suchmaschinen auf die politische Meinungsbildung beschäftigt.

<https://ir.web.th-koeln.de/people/philipp-schaer/>  
[philipp.schaer@th-koeln.de](mailto:philipp.schaer@th-koeln.de)

Systeme zwar durchaus ein Spiegel der Gesellschaft sind, aber vielleicht auch überproportional widerspiegeln oder verzerren. Einige Theorien gehen auch davon aus, dass die Informationsanbieter gezielt manipulieren, um Einfluss zu nehmen – auch auf Wahlentscheidungen. Man denke an Cambridge Analytics. Wir haben versucht, dies mithilfe von Large Language Models, also statistischen Sprachmodellen, zu simulieren. Trainiert haben wir hierzu auf US-amerikanischen News-Webseiten. Wir haben Themen aus dem linken und rechten Spektrum gesammelt, unsere Large Language Models damit trainiert und dann den Grad der Verzerrung, wir sprechen hier von Bias, gemessen.

Die Verzerrungen sind sowohl in der Simulation als auch in unseren Web-Langzeitdaten signifikant, aber nicht in der erwarteten Effektstärke. Wie schon in der Theorie der Filter-Bubbles in sozialen Netzwerken zeigt sich in großen Langzeitstudien, dass die Effekte oft nur im kleinen Maßstab und für besondere Subgruppen zu beobachten sind. Trotzdem müssen wir diese Phänomene im Auge behalten, denn die Effektstärke könnte sich jederzeit ändern. Und wenn alles nur noch Google ist, dann gibt es keine alternativen Algorithmen mehr. Dann gibt es nur noch den Google-Algorithmus. Alternativen werden gerade dazu entwickelt, wie zum Beispiel OpenWeb-Search.EU., an dem 14 Forschungsinstitute aus sieben europäischen Ländern beteiligt sind.

#### Was machen Sie anders und wie decken Sie auf, ob Falschinformationen produziert und lanciert werden?

Wir fokussieren uns auf die Evaluation von Suchsystemen. Wir können uns live in Online-Systeme einklinken und führen dort unsere Versuchssettings durch – natürlich mit Genehmigung der Suchmaschinenbetreiber. Hierzu untersuchen wir zum Beispiel Standard-Ranking-Systeme und bauen ein anderes Ranking-System ein, das bewusst gegen den Matthäus-Effekt arbeitet. Wir messen, wie sich das auf

das Suchverhalten der Nutzer auswirkt und wie diese damit interagieren. Bleiben wir in der Welt der akademischen Suche: Hier können wir etwa evaluieren, ob sich das Suchverhalten signifikant ändert, wenn wir nicht mehr nur popularitätsbasiert „alte weiße Männer“ hervorheben, sondern zum Beispiel auf eine faire Repräsentation der Geschlechter im Suchergebnis achten.

Wir beschäftigen uns zudem mit Zugang zu Forschungsdaten und der nationalen Forschungsdateninfrastruktur, die aktuell im Aufbau ist. Hinzu kommen große Umfragestudien, etwa aus der Medizin oder den Sozialwissenschaften. Diese werden von Systemen wie Google Scholar nur unzureichend abgedeckt.

#### Warum ist Ihre Forschung wichtig für die Wissenschaft selber?

Google Scholar ist ein schöner Service mit einem Riesenindex, aber denkbar schlecht geeignet für wissenschaftliche Spezialrecherchen. Denn das System berücksichtigt etwa nicht die Feinheiten und Besonderheiten der Fachkulturen. Wir brauchen aber für die Wissenschaft spezialisiertere Lösungen, die sauber evaluiert und weiterentwickelt werden.

#### Was möchten Sie mit Ihrer Forschung für die Gesellschaft erreichen?

Wie eingangs erwähnt, steht der Mensch und seine Interaktion mit Information im Zentrum meiner Forschung. Die Befriedigung unserer Informationsbedürfnisse ist ein menschliches Grundbedürfnis. Die damit zusammenhängenden Suchprozesse besser zu verstehen und somit besser unterstützen zu können, ist ein allgemeines Ziel meiner Forschung. Denn ohne Informationen können wir als Individuen oder auch als Gesellschaft keine fundierten Entscheidungen treffen, sei es bei Wahlen, im Job oder anderen Lebensbereichen. //

Anzeige

**Die Messe für Bildung, Job und Gründung in Sachsen**

**Karriere Start**

**Auszug Rahmenprogramm**

- ▶ Frei zugängliches Fachprogramm mit Vorträgen, Diskussionsrunden und Workshops im Rahmen der Messe
- ▶ Bewerbungsmappencheck für Studenten, Absolventen, Ein- und Umsteiger ... für eine erfolgreiche Jobsuche
- ▶ Bewerbungstraining
- ▶ Vorstellung weiterführender Studiengänge
- ▶ Einstieg in die Berufswelt

Änderungen vorbehalten!

**19.–21. Jan. 2024 · MESSE DRESDEN**

Fr. 9–17 Uhr | Sa./So. 10–17 Uhr [www.messe-karrierestart.de](http://www.messe-karrierestart.de)

Die Messe als App!

Auf Facebook folgen!

Auf Instagram folgen!

#messekarrierestart2024  
ORTEC Messe und Kongress GmbH

Save the date!