

- Helene Hahn für ihre Arbeit *Digital identification systems and the right to privacy in the asylum context: An analysis of implementations in Germany*
- Silke Meiser für ihre Arbeit *Der Umgang mit der Vertrauensfrage*
- Martin Schmitt für seine Arbeit *Die Digitalisierung der Kreditwirtschaft. Computereinsatz in den Sparkassen der Bundesrepublik und der DDR, 1957-1991.*

erschieden in der *FifF-Kommunikation*,  
herausgegeben von *FifF e.V.* - ISSN 0938-3476  
[www.fiff.de](http://www.fiff.de)

Wir bedanken uns herzlich bei allen Teilnehmerinnen und Teilnehmern für die eingereichten Arbeiten und gratulieren den Gewinnerinnen und Gewinnern. Die Laudationes wurden von Rainer Rehak und Stefan Hüwe gelesen. Der Weizenbaum-Studienpreis ist ungenutzbar zu finden.



## Weizenbaum-Studienpreis – Laudatio für den ersten Preis

# Hendrik Heuer: Users & Machine-Learning-Based Curation Systems

## Dissertation an der Universität Bremen

Immer mehr verlagern sich unsere Informationsquellen von den „klassischen“ Medien – Presse, Rundfunk, (lineares) Fernsehen – zu Informationsangeboten im Internet, vor allem in den sozialen Medien. Betreiber sind große Konzerne – gerne als GAFAM abgekürzt –, die in der Regel nicht allein nach der bestmöglichen Informiertheit ihrer Nutzer:innen streben, sondern, wie jüngst die Facebook-Whistleblowerin Frances Haugen eindrucksvoll belegte, vor allem wirtschaftliche Interessen verfolgen.

Welche Informationen den Nutzerinnen und Nutzern auf welche Weise präsentiert werden, kann dabei einen erheblichen Einfluss auf die öffentliche Meinung haben. Politikerinnen und Politiker wie Donald Trump setzen dieses Mittel gezielt ein, um die Öffentlichkeit zu beeinflussen. Auch in Europa werden solche Entwicklungen immer deutlicher erkennbar, etwa durch die Aktivitäten von Mitgliedern vor allem rechtspopulistischer Parteien. Damit kommt den Betreibern der Plattformen eine erhebliche Verantwortung zu. Shoshana Zuboff schreibt in ihrem bekannten Buch vom „Überwachungskapitalismus“:

*„Durch die erfolgreiche Durchsetzung des Anspruchs auf Freiheit und Wissen sowie auf seine strukturelle Unabhängigkeit vom Menschen – weder als Angestellte noch als Kundschaft – manövriert uns der Überwachungskapitalismus heute mittels der radikalen Indifferenz, die diese Ansprüche sowohl bedingen als auch ermöglichen und aufrechterhalten, einer Gesellschaft entgegen, in der der Kapitalismus nicht länger als Mittel inklusiver politischer und ökonomischer Institutionen funktioniert. Stattdessen müssen wir den Überwachungskapitalismus als das anerkennen, was er ist: eine zutiefst antidemokratische soziale Kraft.“*

Gleichzeitig erleben wir gerade eine beschleunigte Entwicklung der Künstlichen Intelligenz und des Machine Learning. Folgerichtig wird beides von der Sozial-Media-Industrie verbunden und kombiniert. So wird entsprechend trainierte, vermeintlich unbestechliche und neutrale Künstliche Intelligenz für die Kuratierung von Informationsangeboten in sozialen Netzen eingesetzt. Das erscheint naheliegend – doch sollten Chancen und Risiken zunächst kritisch untersucht werden, und dabei auch: Chancen für wen, Risiken für wen?

Diese Untersuchung hat Hendrik Heuer unternommen. Der erste Preis des diesjährigen Weizenbaum-Studienpreises geht deswegen an seine Dissertation *Users & Machine Learning-Based Curation Systems*, in der er die besonderen Herausforderungen untersucht, die sich durch automatisierte Kurationsysteme ergeben, die mittels Machine Learning trainiert werden.

Indem er den Fokus dabei auf die Menschen legt, die mit diesen Systemen zu tun haben, gelingt ihm eine originelle Herangehensweise, weitab rein rechnerischer Überlegungen. Vielmehr führt er in 10 Kapiteln teils aufwändige Studien mit einer beeindruckenden Methodenvielfalt durch, von qualitativen Interviews über quantitativ ausgewerteten Fragebögen bis zu automatisierten Evaluationsverfahren.

Als Ergebnis dokumentiert er verschiedene Aspekte journalistischer Praxis, von Youtube-Empfehlungen aus Anwender-Sicht, der Verlässlichkeit von Nachrichtenempfehlungen bis zu Erklärungsformen und Prüfverfahren.

Ein wesentliches Ergebnis seiner Arbeit ist, dass diese Systeme einerseits von Praktikerinnen und Praktikern deutlich kritischer wahrgenommen werden als ihre Wirkung vermuten lässt, aber dass es andererseits bei Nutzerinnen und Nutzern eine Lücke gibt zwischen den möglichen und den gewünschten Erklärungen für die Präsentation konkreter Inhalte, eine Lücke, die er den *explanatory gap* nennt.

Die Darstellung erfolgt entlang von fünf Forschungsfragen, die in den Kernkapiteln der Arbeit diskutiert und exemplarisch beantwortet werden:

1. Wie wird Machine Learning von Praktikerinnen und Praktikern verstanden, die ein Kurationsystem trainieren?
2. Welche Annahmen treffen Nutzer eines ML-basierten Kurationsystems, in das sie keine Einblicke haben und das sie nur als Black Box betrachten können?
3. Wie zuverlässig sind die Eingabedaten, mit denen Menschen ein ML-basiertes Kurationsystem trainieren? Und wie hängt diese Zuverlässigkeit mit dem beschränkten Verständnis dieser Systeme zusammen?



4. Wie kann die Funktionsweise dieser Systeme erklärt werden? Hier führt die Erklärungslücke dazu, dass wesentliche Aspekte nicht berücksichtigt, weil nicht verstanden oder von Nutzer:innen und Nutzern in ihrer Bedeutung falsch eingeschätzt werden.

5. Wie kann die Qualität von Kurationsystemen geprüft werden? Vorgestellt wird ein halbautomatisiertes Prüfverfahren, das die Zuverlässigkeit von Vorschlägen zu bewerten hilft.

Zentral für die Entscheidung der Jury war die konsequente Einordnung von Machine-Learning-basierten Kurationsystemen in gesellschaftliche Kontexte, die eben nicht nur aus technischer Sicht, sondern vor allem in ihrem Zusammenhang mit den Menschen, die sie nutzen, und dem gesellschaftlichen Kontext zu betrachten sind. Eine Kernaussage ist, dass Machine-Learning-Systeme auch hier sozio-technische Systeme sind, die falsch oder

unzureichend trainiert werden können, wenn es hier überhaupt ein „richtig“ gibt. Dadurch ergibt sich eine Erweiterung der Perspektive von den Algorithmen hin zu den Daten, die nicht nur von der Maschine, sondern auch von den Menschen ausgewählt und interpretiert werden.

Wir gratulieren ihm zu einer hervorragenden, originellen und gründlichen Arbeit über ein gesellschaftlich hochrelevantes Thema. Es hat eine kritische Begleitung dringend nötig, wofür hier ein wichtiger Baustein gelegt wurde. Wir verstehen den Preis auch als Empfehlung, welche Arbeit wir in diesem Jahr als besonders lesenswert erachten. Dies ist bei dieser Arbeit ohne Zweifel der Fall.

**Herzlichen Glückwunsch, Hendrik Heuer, zum Weizenbaum-Studienpreis 2021.**



Hendrik Heuer

## Warum YouTubes KI gefährlich ist und was wir dagegen tun können

### Zusammenfassung der Dissertation „Users & Machine Learning-based Curation Systems“



1. Preis

*Fake News, Verschwörungstheorien und Filterblasen sind gravierende gesellschaftliche Probleme. YouTube wird immer wieder eine Mitschuld an diesen Problemen gegeben. Mehr als zwei Milliarden Menschen nutzen YouTube jeden Monat und 70 % der dort konsumierten Videos werden von einem KI-System vorgeschlagen. Meine Dissertation untersucht, warum die Künstliche Intelligenz (KI) auf YouTube gefährlich ist und was man dagegen tun kann. Die Arbeit liefert die Grundlage für ein ganzheitliches Verständnis der Systeme, die Videos und Nachrichten empfehlen. Dazu kombiniere ich ein technisches Verständnis mit Einblicken in die Perspektiven der Nutzer:innen, der Entwickler:innen sowie die Perspektive jener, die Daten für solche Systeme liefern. Die Untersuchung zeigt, dass Erklärungen von KI-Systemen für Menschen nicht verständlich sind. Als Alternative dazu beschreibt die Arbeit Untersuchungsverfahren, die sicherstellen können, dass KI-Systeme im Interesse der Öffentlichkeit funktionieren.*

KI ist ein wichtiges Teilgebiet der Informatik. Sie ermöglicht es Computern, Aufgaben zu erledigen, die normalerweise menschliche Intelligenz erfordern. Ein wichtiger Ansatz, der für viele der jüngsten Durchbrüche in der künstlichen Intelligenz verantwortlich ist, ist das so genannte Maschinelle Lernen. Anders als Logik-basierte KI-Ansätze versucht man beim Maschinellen Lernen, mittels statistischer Verfahren aus Daten zu lernen. Dank der Techniken des Maschinellen Lernens ist es mittlerweile möglich, Sätze von einer Sprache in eine andere zu übersetzen, Personen und Objekte in Bildern zu erkennen oder Autos autonom fahren zu lassen. Plattformen wie YouTube nutzen die Techniken des Maschinellen Lernens, um automatisch Videos auszuwählen, die für Nutzer:innen besonders interessant sind. Im Jahr 2019 wurden durchschnittlich 500 Stunden an Inhalten pro Minute bei YouTube hochgeladen. Auf den Tag gerechnet sind das 82,2 Jahre Videomaterial. Für ein Team menschlicher Expert:innen wäre es unmöglich, diese von Nutzer:innen hochgeladenen Inhalte zu überprüfen und zu kategorisieren. Social Media-Plattformen setzen daher KI-Techniken wie Maschinelles Lernen ein, um Videos zu kategorisieren und zu empfehlen. Solche Empfehlungssysteme sind sehr erfolgreich. 70 % der Videos, die die zwei Milliarden Menschen gucken, die YouTube jeden Monat nutzen, werden von diesem KI-System ausgewählt.



Hendrik Heuer beim Vorstellen seiner Dissertation

YouTube wurde von Journalist:innen immer wieder beschuldigt, Fake News und Verschwörungstheorien zu verbreiten. Ein wichtiges Beispiel für diesen Vorwurf sind die rechtsradikalen Ausschreitungen in Chemnitz im Jahr 2018. Bei einer Auseinandersetzung im Rahmen des Chemnitzer Stadtfestes wurde ein Mann durch Messerstiche tödlich und zwei weitere Personen wurden schwer verletzt. Rechte und rechtsextreme Gruppen instrumentalisierten den Vorfall mit Blick auf die vermeintliche Herkunft des Täters und riefen zu Aufmärschen auf. Einem Bericht der