Soziale Netzwerke und Algorithmen: Merkmale und Risiken

5.1 Was Algorithmen sind – und warum sie für die Soziale Arbeit relevant sind

Algorithmen zerlegen Aufgaben in Teilschritte und geben Anweisungen, wie diese durchzuführen sind: „wenn a, dann b“. Sie liegen in Software(-systemen) vor, in der Regel viele gemeinsam mit Datenstrukturen. **Künstliche Intelligenz** ist dabei als Oberbegriff zu verstehen, unter den wissensbasierte Systeme/Expertensysteme und **maschinelles Lernen** (als Teilbereich davon Predictive Analytics) fallen; alle diese Systeme benötigen große Datenmengen – **Big Data**, charakterisiert über die Kern-V-Begriffe *Volume* (Datenmenge), *Velocity* (Geschwindigkeit/Echtzeit) und *Variety* (Vielfalt der Datenformen) (Beranek 2021, Kap. 2).

Entscheidend ist: Algorithmen entfalten ihre gesellschaftliche Wirkkraft erst durch ihren **Einsatz in sozialen, ökonomischen, administrativen und rechtlichen Bereichen**. Algorithmische Entscheidungsfindung (ADM) ist dann menschenrechtsrelevant, wenn durch sie Grund- und Menschenrechte verletzt, Ungleichheiten erzeugt oder verfestigt werden (Beranek 2021).

5.2 Diskriminierung durch algorithmische Systeme

Beranek (2021, Kap. 5.2) unterscheidet zwei Wege, auf denen diskriminierende Elemente in den Code gelangen:

- **Präferenzbedingte Diskriminierung:** Ungleichbehandlung beruht auf persönlichen Vorurteilen der (am Systembau und an der Datenlage) Beteiligten.
- **Statistische Diskriminierung:** Ungleichbehandlung entsteht über Ersatzinformationen, wenn unvollständige Daten über Personen vorliegen – beim maschinellen Lernen treten anstelle einzelner Ersatzvariablen ganze Modelle mit gewichteten Variablenbeziehungen.

Katharina Zweig (2019, zit. n. Beranek 2021) benennt fünf Ursachen für Diskriminierung in ADM-Prozessen – u.a. explizit/implizit in Daten enthaltene Diskriminierungen, fehlende Daten ganzer Personengruppen, vorenthaltene sensible Informationen und falsches Weitertrainieren von Systemen „in freier Wildbahn“.

Zentral ist der **Akkumulations- und Verstärkungseffekt**: Systeme, die mit historischen Daten trainiert werden, übernehmen die in den Daten enthaltenen Vorurteile und verstärken sie. „Wo Personengruppen in Daten nicht vorkommen, können deren Eigenschaften nicht mitgelernt werden“ (Zweig 2019, S. 215, zit. n. Beranek 2021).

Beispiele aus Beranek (2021):

- **Versteckte Schönheitsnormen:** Smartphones bearbeiten Selfies bereits bei der Aufnahme („Schönes Porträt“, Real-Time-Enhancement): Hautglättung, Augenvergrößerung, Hautaufhellung – westliche Schönheitsideale werden technisch ausgeliefert und sind für Nutzende kaum erkennbar. Mädchen nutzen zusätzlich Filter, um „wahrgenommene Defizite des eigenen Körpers auszugleichen“ (Götz 2019). 55 % der befragten Patient*innen einer Studie wollten 2018 ihr Erscheinungsbild in Selfies verbessern – nicht mehr wie Prominente aussehen, sondern „wie eine retuschierte Version ihrer selbst“ (Rajanala/Maymone/Vashi 2018).
- **Rassistische ADM:** In den USA diskriminierten Strafverfolgungs- und Gesundheitsalgorithmen systematisch Menschen mit dunkler Hautfarbe – der zitierte Gesundheitsalgorithmus wählte weiße Patient*innen häufiger für „High Risk Care Management“ aus, weil er auf früheren Behandlungskosten aufbaute; sozial Benachteiligte nutzen aber seltener medizinische Leistungen (Obermeyer et al. 2019).
- **Sexistische ADM:** Googles Übersetzungstool übersetzte geschlechtsneutrale Begriffe zunächst männlich; Amazons Bewerbungs-AI bewertete Bewerbungen von Frauen auf technische Stellen systematisch schlechter (Beranek 2021).

Für Kinder und Jugendliche heißt das: Die Auseinandersetzung mit Schönheitsidealen und dem eigenen Körper, die Verunsicherung gegenüber dem eigenen Aussehen und die Streberei nach unrealistischen Erscheinungen „können gerade im Jugendalter eine gesunde Entwicklung beeinträchtigen“ (Beranek 2021, S. 82).

5.3 Die Logik sozialer Netzwerke: Öffentlichkeit unter Algorithmik

Die Öffentlichkeit hat durch das Internet einen Strukturwandel erfahren: Digitale Öffentlichkeiten können auch von Privatpersonen geschaffen werden; eine „fünfte Gewalt“ aus den vernetzten Vielen betreibt Agenda-Setting „von unten“ (Hauser/Opilowski/Wyss 2019, zit. n. Beranek 2021). Gleichzeitig strukturieren ADM-Systeme Öffentlichkeit in wachsendem Maße:

- **Persönliche Öffentlichkeiten** (Schmidt 2009): Kommunikation im Netz ist persistent, duplizierbar, skalierbar und durchsuchbar – was heute gepostet wird, bleibt auffindbar.
- **Personalisierung und Filterblasen:** Aus Suchanfragen und Nutzungsdaten werden Trefferlisten individuell gefiltert (Pariser 2011). Studien deuten zwar darauf hin, dass die Wirkung oft überschätzt wird – für Gruppen, die sich in wenigen, gleichartigen Quellen informieren (etwa AfD-Wähler:innen, Sindermann et al. 2020), entstehen aber reale Parallelwelten und wahrnehmungsverweigernde Gemeinschaften (Antos 2019) (Beranek 2021).
- **Suchmaschinen als Gatekeeper:** Google hielt 2019 bei der Desktopsuche in Deutschland ca. 90 %, bei der mobilen Suche über 98 % Marktanteil. Suchmaschinen sind **Vertrauens- und Erfahrungsgüter**: Nutzer*innen können weder fehlende Treffer noch die fremden Relevanzkriterien beurteilen – sie müssen einen Vertrauensvorschuss leisten. Damit sind Betreiber „mitverantwortlich für die Rezeption und in der Folge Reflexion gesellschaftlicher Ansichten“ (Lewandowski 2015, zit. n. Beranek 2021).

5.4 Fake News – Merkmale und Verbreitungsmechanik

Fake News sind Falschmeldungen, die im *Look and Feel* journalistischer Beiträge daherkommen, inhaltlich aber nicht der Faktenlage entsprechen – teils mit bewusster Erfindung oder Verfälschung zur Erreichung politischer oder kommerzieller Ziele (Appel/Doser 2020, zit. n. Beranek 2021). Ihre Viralität folgt der Architektur der Plattformen (Schmid et al. 2017, zit. n. Beranek 2021):

1. **Narrative des Skandals und der Empörung:** Empörung ist attraktiver als nüchterne Berichterstattung.
2. **Instrumentalisierung von Empfehlungsalgorithmen:** Häufig geklickte, gelikte und kommentierte Beiträge werden prominenter platziert.
3. **Eigendynamiken im Rezeptionsverhalten:** Artikelempfehlungen werden ungelesen geteilt, weil die Überschrift anspricht.
4. **Fehlende Filterfunktion:** Der Weg von der Quelle zur Öffentlichkeit ist weggefallen.

Verschärft wird das durch **Emotionalisierung**: Je stärker emotional aufgeladen ein Tweet ist, desto schneller wird er weiterverbreitet; negativ emotionalisierte Beiträge provozieren mehr Reaktionen als neutrale (Stieglitz/Dang-Xuan 2013) – und genau diese Reaktionsintensität ist auf werbefinanzierten Plattformen *ökonomisch gewollt*, weil sie Nutzer*innen länger auf der Seite hält. **Social Bots** unterlaufen zudem die scheinbare Transparenz sozialer Medien: Sie geben sich als

Menschen aus und nutzen das Prinzip sozialer Bewährtheit (Anschluss an Mehrheitsmeinungen bei Unsicherheit) sowie „Medienwahrheiten“, die ohne faktischen Bezug durch gesellschaftlichen Konsens validiert werden (Graber/Lindemann 2018, zit. n. Beranek 2021).

5.5 Hate Speech – gruppenbezogene Menschenfeindlichkeit im Netz

Definition: Hate Speech umfasst abwertende, vorurteilgeleitete Aussagen, die sich gegen *Gruppen* richten – im Kern „eine Form der kommunikativen Herstellung menschlicher Minderwertigkeit“ (Sponholz 2018, S. 48, zit. n. Beranek 2021). Nicht die Emotion des Hasses, sondern negative Vorurteile gegenüber spezifischen Gruppen sind konstitutiv (Geschke et al. 2019). Abzugrenzen ist Hate Speech von individuellen Beleidigungen und Cybermobbing – wobei die Grenzen fließend sind und Cybermobbing ein eigenes Risiko für Jugendliche darstellt.

Wesensmerkmale (Sponholz 2018): Öffentlichkeit, Diskriminierung (Gruppenbezug), Kommunikation – ergänzt um Intentionalität und Performativität. Als Aktypen unterscheidet Sponholz das lokutionäre (Hateful Speech), ilokutionäre (Intention) und perlokutionäre Modell (Dangerous Speech, Wirkungen).

Typische Elemente (Amadeu Antonio Stiftung): Gleichsetzungen, Verschwörungstheorien, De-Realisierung, Wir-/Ihr-Gegenüberstellungen, konstruierte Handlungszwänge und die Normalisierung bestehender Diskriminierungen (Baldauf et al. 2015, zit. n. Beranek 2021).

Wer betroffen ist (Geschke et al. 2019, zit. n. Beranek 2021) – Prozentanteile „sehr oft/oft“ von Hassrede betroffener Gruppen:

Gruppe	sehr oft / oft
Amtierende Politiker:innen	19 % / 33 %
Geflüchtete Menschen	19 % / 32 %
Politisch Andersdenkende	19 % / 31 %
Menschen mit Migrationshintergrund	18 % / 32 %
Muslimas und Muslime	18 % / 28 %
Menschen, die nicht dem aktuellen Schönheitsideal entsprechen	10 % / 22 %
Arbeitslose Menschen	9 % / 19 %
Homosexuelle Menschen	8 % / 19 %

Gruppe	sehr oft / oft
Frauen	7 % / 18 %

Motivlagen der Verbreiter:innen (Schmitt 2017): (1) Ausgrenzung und Abwertung einer Fremdgruppe zur Stärkung der eigenen Gruppenidentität, (2) Einschüchterung, (3) Machtdemonstration und Erlangen von Deutungshoheit, (4) Freude am Beleidigen und Erniedrigen.

Warum das Netz begünstigt: Der **Online-Enthemmungseffekt** entsteht durch fehlende soziale Hinweisreize (Gestik, Mimik) und durch **soziale Anonymität** – Nutzer:innen nehmen sich gegenseitig ohne individuelle Identität wahr, was Deindividuation und ein „Wir gegen Die“-Gefühl fördern kann (Kaspar 2017, zit. n. Beranek 2021). Hinzu kommt die **pluralistische Ignoranz**: In Gruppen hilft niemand, weil alle fälschlich annehmen, die anderen fänden das Verhalten „okay“ – die Schwelle der ersten Gegenwehr wird nicht überschritten.

5.6 Die Systematik im Überblick

```
flowchart TD
    A["Kommerzielle Plattformlogik<br/>Werbefinanzierung · Aufmerksamkeit als Ressource<br/>Nutzer:innen möglichst lange binden"] --> B["Empfehlungsalgorithmen<br/>Personalisierung · Ranking nach Interaktion<br/>Emotionalisierung erhöht Reichweite"]
    B --> R1["Filterblase & Fragmentierung<br/>Personalisierte Trefferlisten · Parallelwelten"]
    B --> R2["Fake News<br/>Empörungsnarrative · Social Bots<br/>De-/Rekontextualisierung"]
    B --> R3["Hate Speech<br/>Gruppenbezogene Abwertung · Anonymität<br/>Enthemmung · pluralistische Ignoranz"]
    B --> R4["Diskriminierung durch ADM<br/>Statistische & präferenzbedingte Verzerrung<br/>Historische Daten · kumulative Benachteiligung"]
    R1 & R2 & R3 & R4 --> F["Eingeschränkte Teilhabe, Meinungsbildung<br/>& Lebenschancen junger Menschen"]
```

5.7 Was das für die Soziale Arbeit bedeutet

Beranek (2021) zieht als Menschenrechtsprofession den Schluss, dass die Soziale Arbeit die Auslegung ihrer Professionstheorie auf **Menschenrechte im digitalen Raum** erweitern muss: Diskriminierungen durch ADM-Systeme können Teilhabe verhindern (Rechte auf Bildung, Art. 26, und auf kulturelle Teilhabe, Art. 27 AEMR), versteckte Werte und Normen transportieren und die Selbstwahrnehmung – v.a. Heranwachsender – beeinflussen. Schutzkonzepte reichen dafür nicht aus; nötig sind Ansätze, die junge Menschen **befähigen**, Plattformen, Algorithmen und ihre eigene

Rolle darin zu verstehen und zu gestalten. Genau hier setzt – als analytischer und zugleich handlungsorientierter Rahmen – die Medienpädagogik an (→ Kapitel 6).

Revision #1
Created 30 October 2026 11:41:31 by Shelly Pröhl
Updated 30 October 2026 11:42:18 by Shelly Pröhl