



UNIVERSITÄT ZU LÜBECK
INSTITUT FÜR
MEDIZINISCHE ELEKTROTECHNIK

PS4640-KP04 Technikethik

Algorithmic Bias, Datenherkunft, -qualität und -kontext

Christian Herzog, Sabrina Breyer, Arne Sonar



Einleitendes Video: Algorithmic Bias Explained

Einleitung

In einem Blog-Post berichtet der Statistiker und Informatiker Michael Jordan 2018 eine Anekdote aus seinem eigenen Leben. Eine Routine-Ultraschalluntersuchung wies am Herzen seiner damals noch ungeborenen Tochter mehrere weiße Punkte auf – Punkte, die im vorherigen Jahrzehnt als Prädiktoren für ein erhöhtes Risiko für das Down Syndrom identifiziert wurden. Die damals vom ärztlichen Personal vorgeschlagene potenzielle Diagnoseprozedur ist für den Fötus vergleichsweise gefährlich, so dass Michael Jordan ein paar Berechnungen und Nachforschungen anstellte. Das Ergebnis: Die mittlerweile höhere Auflösung modernerer Ultraschallgeräte führte dazu, dass verstärkt weißes Rauschen der Bildpunkte als Indikatoren für die Möglichkeit gehalten wurden, dass Föten am Down Syndrom leiden. Tatsächlich hatte das ärztliche Personal nach der Anschaffung der neuen Ultraschallgeräte eine Erhöhung entsprechender Diagnosen feststellen können.

Kurze Zeit später wurde Michael Jordan Vater einer gesunden Tochter. Diese Episode verdeutlicht eine besondere Gefahr statistischer Analysen: Die Übertragung von Analysen, die anhand von Daten einer bestimmten Herkunft und Qualität getätigt wurden, auf Situationen und Kontexte, die möglicherweise ganz andere Datenqualitäten aufweisen, ist nicht trivial. Im anekdotischen Beispiel wird vor allem deutlich, dass es sehr leicht passieren kann, dass systematische Fehler auf längere Zeit unentdeckt bleiben. Der Fall zeigt ganz besonders deutlich, dass hier möglicherweise zahlreiche Föten unnötigerweise einem Risiko ausgesetzt worden sind, wenn nicht sogar viel Schlimmeres noch passiert sein könnte.

Automatisierte Systeme könnten hier noch viel größeren Schaden anrichten. Zum einen neigen die Nutzer:innen automatisierter Systeme tendenziell dazu, diesen im übertriebenen Maße zu vertrauen (Automation Bias, vgl. Goddard, Roudsari, and Wyatt 2011). Zum anderen sind automatisierte Lösungen häufig im Wesentlichen Softwaresysteme, die sich besonders leicht replizieren lassen. Ein systematischer Fehler würde so leicht hundert-, oder tausendfach multipliziert, während sich Muster menschlicher Missinterpretationen auch leicht nur auf einzelne, verantwortliche Individuen beschränken.

In Bezug auf Künstliche Intelligenz, die anhand von Daten trainiert wird, bedeutet dies, dass die Datenqualität gut überwacht werden muss, um eine Übertragbarkeit auf aktuelle, neuartige und konkrete Anwendungsfälle zu gewährleisten. Ein an den alten Bilddaten trainierter Algorithmus hätte also auf den höher aufgelösten, aber verrauschten Bilddaten evtl. sehr systematisch eine stark erhöhte Tendenz dazu gehabt, das Down Syndrom zu diagnostizieren.

Wenngleich dieser Fall besonders tückisch ist, so gibt es doch auch viele, nicht notwendigerweise lebensbedrohliche, aber dennoch sehr relevante Beispiele dafür, dass die Unklarheit darüber, was genau an Daten trainierte Algorithmen dazu veranlasst, eine bestimmte Ausgabe (Diagnose, Identifikation, etc.) zu geben, zu stark unerwünschten Folgen führen kann. Eine vereinfachte Sichtweise dieser Dinge besteht darin, KI-Algorithmen vorrangig als Algorithmen für die Mustererkennung zu verstehen. In diesem Sinne wird versucht, Muster aus den Trainingsdaten auch in den aktuellen Daten einer Anwendung wiederzufinden. Wäre das obige Beispiel eines aus der Künstlichen Intelligenz gewesen, wären die Trainingsdaten Bilddaten mit geringer Auflösung gewesen, die vergleichsweise wenig weißes Rauschen aufgewiesen hätten. Die vorhandenen weißen Artefakte wären eventuell als ein Prädiktor für das Down Syndrom erkannt worden. Höher aufgelöste, aber auch eventuell stärker verrauschte Bilddaten hätten ggfs. viele falsch-positive Erkennungen hervorgerufen. Nun handelt es sich hierbei um Spekulation – es bleibt sogar zu hoffen, dass Künstliche Intelligenz ein Mittel sein kann, mit dem die Fehleranfälligkeit und Missinterpretationen durch Menschen reduziert werden können. Allein, die Abhängigkeit statistischer Analysen – und damit auch trainierter Algorithmen – von den Datenbasen wird vom Beispiel eindrucksvoll illustriert.

Diese Abhängigkeit wird auch häufig im Kontext des Stichworts des „Algorithmic Bias“ (Koene 2017) angeführt – dies wird häufig mit der Tendenz eines Algorithmus gleichgesetzt, in Daten befindliche Artefakte zu reproduzieren.

Sind Daten objektiv?

Mit dem Stichwort „Algorithmic Bias“ wird häufig die Tatsache verbunden, dass bestimmte in Daten implizit oder explizit vorhandene kulturelle, technische oder anderweitige Aspekte von anhand dieser Daten trainierter Algorithmen identifiziert und eventuell reproduziert werden. Das Prinzip des „Garbage in – Garbage out“ beschreibt dies umgangssprachlich. Zum Teil wird der Begriff auch im engeren Sinn verwendet, der voraussetzt, dass die Verwendung von Algorithmen mit Biases tatsächlich zu Ungerechtigkeiten führt¹. Sind beispielsweise bestimmte Rassismen, (Gender-)Stereotypen, o.ä. in Textcorpora vorhanden, so wird ein trainierter Algorithmus im Sinne einer raffinierten Mustererkennung diese erkennen und – abhängig vom Einsatzzweck – ebenso wiedergeben.

Eine besonders viel-zitierte Studie der Wissenschaftler:innen Caliskan, Bryson, und Narayanan 2017, beispielsweise, konnte zeigen, dass ein sehr gebräuchlicher Machine Learning-Ansatz in der Lage war, bestimmte Tendenzen der menschlichen Sprache aus im Internet verfügbaren Textdaten akkurat wiederzugeben. Nicht alle davon sind problematisch. Da aber Geschlechterdiskriminierung, Rassismen und dergleichen leider noch weit verbreitete Phänomene sind, konnte der Algorithmus auch diese abbilden. Die Autoren schlagen in ihrem Artikel vor, diesen Effekt zu nutzen, um die Ursprünge dieser als Vorurteile identifizierten Tendenzen zu erforschen. Andererseits wird aber deutlich, dass Technologie, die auf derartigen Daten beruht nicht ohne Weiteres die gesamtgesellschaftlichen Herausforderungen adäquat adressiert. Im Gegenteil: Im großen Stil und ohne weitere Vorkehrungen ausgerollte derartige KI-basierte Technologie würde die Ungerechtigkeiten und Diskriminierungen in der Gesellschaft wohlmöglich verstärken, mindestens aber diesen nicht entgegenwirken.

Tatsächlich werden daten-basierte Algorithmen aber nicht selten als objektiver als Menschen beschrieben. Das mag potenziell möglich sein, jedoch ist dies nicht zu erreichen, wenn davon ausgegangen wird, dass die Datenbasen selbst ein Medium für die objektive Wiedergabe der Realität sind (Crawford 2013). Die Behauptungen gehen zum Teil sogar so weit, dass die Verwendung daten-basierter Algorithmen das Ende aller Theorie einläuten könnte (Anderson 2008).

Algorithmic Biases in Gesundheitsdaten – Ein Beispiel

Die sogenannte elektronische Patientenakte (engl. Electronic Health Record) ist ein Weg, mit dessen Hilfe automatisierte und objektive Diagnosen vereinfacht und begünstigt werden sollen. Auch hier zeigen sich jedoch besonders relevante Herausforderungen im Kontext potenzieller Biases (Gianfrancesco et al. 2018). Genau wie Algorithmen Rassismen und soziale Diskriminierungen verstärken können, können automatisierte Diagnosesysteme, sollten diese unter unerwünschten Tendenzen leiden, Ungerechtigkeiten im Gesundheitskontext verstärken.

In dieser Hinsicht gibt es mindestens drei Quellen möglicher Biases, die in Daten von elektronischen Gesundheitsakten auftreten können.

- **Fehlende Daten:** Patient:innen, die tendenziell eher selten das Gesundheitssystem in Anspruch nehmen und sich dabei auf heterogene Stellen verlassen, die nicht notwendigerweise ihre Daten in das übergreifende System einspeisen, könnten nur eine schlechtere datengetriebene Versorgung genießen.
- **Unrepräsentative Daten:** Minderheiten werden in klinischen Daten unterrepräsentiert sein, was zu falschen oder nicht aussagekräftigen Inferenzen für diese Gruppen führen kann.
- **Falschklassifizierungen:** Gruppen niedriger sozialer Schichten könnten tendenziell in Institutionen mit mehr Lehrpersonal oder schlechteren Versorgungsleistungen behandelt werden. Fehlerhafte oder vorurteilsbehaftete Dateneingaben und mangelndes Qualitätsmanagement könnten dazu führen, dass die Unterversorgung niedriger sozialer Schichten systematisch verstärkt wird.

¹ Vgl. Wikipedia, https://en.wikipedia.org/wiki/Algorithmic_bias, aufgerufen am 05.01.2021

Ansätze, die diesen Effekten entgegenwirken können, beschreiben die Autoren Gianfrancesco et al. 2018 ebenfalls. Beispielsweise wird empfohlen, medizinische Algorithmen in sozio-ökonomisch diversen Kontexten zu konzipieren, zu implementieren und zu erproben. Gleichzeitig sollten echte Datenanalysen durchgeführt werden und auf die Variablen hinsichtlich von Geschlecht und Ethnie kontrolliert werden.

Weitere Beispiele für Algorithmic Biases sind bereits im Themenmodul „Beispiele ethischer und gesellschaftlicher Implikationen der Künstlichen Intelligenz“ angeführt worden. Im Folgenden soll daher verstärkt auf die sogenannten ontologischen und epistemologischen Herausforderungen datenbasierter Algorithmen eingegangen werden.

Biases jenseits des Rassismus – Weitere Typen von Algorithmic Biases

Das Konzept des „Garbage in – Garbage out“ legt nahe, dass Tendenzen in Algorithmen im Wesentlichen auf bereits existierende Biases der Daten zurückzuführen sind. Diese Charakterisierung ist jedoch nicht erschöpfend. Folgende Quellen/Typen von Algorithmic Biases können identifiziert werden².

- **Bereits existierende Tendenzen:** Vorurteile können nicht nur in Daten manifestieren, sondern können auch durch die sozialen und institutionellen Umstände des Entwicklungssystems, bestehend aus Management, Entwicklungsteam, regulatorischen und anderen Randbedingungen, entstehen. Solche Tendenzen können explizit und bewusst, oder implizit und unbewusst ihren Weg in die Algorithmen finden. In diesem Sinne ist diese Art des Algorithmic Bias auch nicht nur auf durch Daten trainierte Algorithmen beschränkt.
- **Technischer Bias:** Technische Biases entstehen durch Limitationen und Randbedingungen des Systems, welches entwickelt und implementiert wird. Solche Beschränkungen können durch simple Darstellungformate (alphabetische Sortierung) dazu führen, dass bestimmte Ausgaben bevorzugt wahrgenommen werden.
- **Emergender Bias:** Emergente Biases sind Tendenzen, die dadurch entstehen, dass sich Randbedingungen, bspw. sozio-kultureller Art, ändern, die technischen Artefakte aber nicht angepasst werden. Beispielhaft können hier Algorithmen verstanden werden, die eine veraltete Gesetzesgrundlage umsetzen.
- **Durch Korrelationen induzierter Bias:** Korrelationen in großen Datensätzen können auf z.T. auf unvorhersehbare Art von trainierten Algorithmen identifiziert werden, ohne dass diese Korrelationen eine sinnvolle Kausalität darstellen. Manchmal können diese Korrelationen auch mit Surrogatparametern entstehen, die selbst eine direkte Verbindung zu sensiblen Variablen wie Geschlecht oder Ethnie haben. Häufig, kann auch eine Korrelation ohne die Berücksichtigung des Kontexts dazu führen, dass in der Verallgemeinerung falsche Entscheidungen getroffen werden. Ein Beispiel dafür ist bereits im Themenmodul „Beispiele ethischer und gesellschaftlicher Implikationen der Künstlichen Intelligenz“ aufgeführte Algorithmus zur Prädiktion der Mortalität von Patient:innen mit Lungenerkrankungen. Hier wurde eine informelle Vorgehensweise im Krankenhaus nicht in den Daten erfasst, so dass Asthmatiker:innen systematisch aber unlogischerweise eine höhere Überlebenswahrscheinlichkeit attestiert bekamen – nur, weil diese stets direkt auf der Intensivstation behandelt wurden, vgl. (Caruana et al. 2015).
- **Durch nicht antizipierte Verwendungen entstehende Biases:** Technische Systeme – auch nicht daten-basiert trainierte Algorithmen – können so entworfen sein, dass sie auf eine gewisse Zielgruppe zugeschnitten sind. Wird ein solches System in einem anderen Kontext mit anderen Nutzer:innen eingesetzt, kann bereits die im System verwendete Sprache oder die Voraussetzung von tieferem technischen Sachverständnis bestimmte Personengruppen benachteiligen, oder gar ausschließen.

² Vgl. Wikipedia, https://en.wikipedia.org/wiki/Algorithmic_bias, aufgerufen am 05.01.2021

- **Durch Rückkopplungseffekte induzierte Biases:** In gewisser Weise als Meta-Bias induzierend, können Rückkopplungseffekte genannt werden. Wenn durch Daten trainierte Algorithmen Realitäten schaffen, die in neuen Datensätzen abgebildet zum erneuten Trainieren des Algorithmus eingesetzt werden, so kann ein selbstverstärkender Effekt entstehen. Systeme des Predictive Policing etwa werden eingesetzt, um Polizeistreifen in Gegenden mit erhöhter Kriminalität zu planen. Werden mehr Streifen in bestimmten Gegenden geplant, kann es zu einer verstärkten Identifikation von Verbrechen führen. In der Folge werden weitere Ressourcen auf diese Gegend konzentriert. In Verbindung mit Stadtteilen, die vorwiegend von bestimmten Ethnien und Personengruppen bevölkert werden, kann es auch hier schnell zu rassistischen und anderen diskriminierenden Tendenzen kommen.

Die obige Klassifizierung kann helfen, schon während des Designs von Algorithmen potenzielle Biases zu antizipieren und Strategien zu entwickeln, darauf zu reagieren. Möglichkeiten dafür können sowohl technischer als auch nicht-technischer Art sein.

Möglichkeiten der Adressierung

Die Probleme, welche durch Algorithmic Biases hervorgerufen werden können, sind schon seit einiger Zeit bekannt. In diesem Zuge sind bereits Bestrebungen im Gange, Standards und Normen auszuarbeiten, die Entwickler:innen einen Leitfaden an die Hand geben, diese Herausforderung entsprechend zu adressieren.

Die Arbeit der IEEE im Rahmen des sogenannten P7003 Standards zu Algorithmic Biases etwa, versuchen Entwicklungsteams mit Hilfe eines Standards die Möglichkeit zu geben, ihre Bestrebungen zur Eliminierung und Minimierung von Biases zu dokumentieren (Bird et al. 2020). Hierbei werden auch Tests, bzw. sogenannte Benchmarks für Datensätze eine Rolle spielen, die versuchen Quellen von subjektiven Informationen in den Daten zu identifizieren. Im Sinne des oben genannten Typs eines Bias durch nicht-antizipierte Nutzung soll der Standard auch eine klare Dokumentation des Anwendungsszenarios enthalten, für das ein Algorithmus entworfen und validiert worden ist. Des Weiteren müssen Strategien nachgewiesen werden, die einer Missinterpretation der Eingangs- und Ausgangsdaten durch die Nutzer:innen eines Algorithmus vorbeugt.

Hieraus lässt sich erkennen, dass das Standardisierungskonzept der IEEE vorrangig auf der Prozessebene angesiedelt ist. Dies erscheint aus der oben diskutierten Einsicht heraus, dass Datenbasen in der Regel keine vollständig objektive Quelle für das Algorithmen-Design sind, plausibel. Auch trägt dieser Ansatz der Einsicht Rechnung, dass implizite und unbewusste Biases durch die Entscheidungen und Zusammensetzung des Entwicklungsteams eingeführt werden können. Ziel ist es also, die Reflexion bzgl. der Herangehensweise während des Entwicklungszeitraumes und auch während der Nutzungsphase von Algorithmen fortwährend zu unterstützen und zu dokumentieren.

Eine Studie von Cowgill et al. 2020 hat darüber hinaus ergeben, dass nicht nur bessere und damit auch repräsentativere Datensätze dazu beitragen, dass daten-basierte Algorithmen weniger Biases propagieren, sondern auch, dass diversere Entwicklungsteams eine vorbeugende Wirkung haben. Die Studie stellt dabei klar, dass Randgruppen und Minderheiten nicht notwendigerweise per se weniger implizite Vorurteile in die Bewertung der Daten und die Konstruktion des Algorithmus einbringen. Stattdessen ergibt die Studie, dass unterschiedliche Personengruppen unterschiedliche Biases einbringen und ein diverseres Entwicklungsteam damit durchschnittlich geringer ausgeprägte unbewusste Biases in die Algorithmen integriert. Tatsächlich hat eine weitere Studie von Krishnan, Almadan, and Rattani 2020, sogar ergeben, dass bei der Verwendung von Standardalgorithmen zur Gesichtserkennung auch möglichst diverse Datensätze nicht vor Biases schützen (Wiggers 2020). Besondere Maßnahmen, angepasste Algorithmen und Validierungsprozesse sind also auch bei repräsentativen Datensätzen notwendig.

Natürlich gibt es auch technische Maßnahmen, die ein Entwicklungsteam dabei unterstützen können, Biases in Daten oder KI-Methoden zu identifizieren. Eine Studie von KI-Wissenschaftler:innen des

Massachusetts Institute of Technology, Intel und der Kanadischen KI-Initiative CIFAR hat beispielsweise eine Reihe vortrainierter KI-Modelle wie Googles BERT, XLNet, OpenAIs GPT-2 und Facebooks RoBERTa analysiert und dabei stark ausgeprägte Stereotypen festgestellt (Johnson 2020). Damit einher geht die Erforschung und der Einsatz von Metriken, die Biases in algorithmischen Ausgaben und Daten automatisch identifizieren können. Die Plattform StereoSet³ ist eine Plattform zur Messung von Biases in Sprachmodellen.

Ausblick

Technologie – auch auf Daten basierende Algorithmen – ist menschengemacht. So wie auch technologische Entwicklungen im Sinne ethischer Innovation positive moralische und gesellschaftliche Wirkung sucht bewusst zu realisieren, kann eine unreflektierte und unbewusste Herangehensweise dazu führen, dass nicht akzeptierbare Wertvorstellungen mit Technologie transportiert wird. Besonders kritisch wird dieser Aspekt, wenn die Technologie große Skaleneffekte erzielen kann, also eine geringe Anzahl an Entwickler:innen dafür verantwortlich ist, wie technische Systeme millionenfach verwendet werden. Algorithmic Biases sind eine der ganz konkreten Formen, wie sich unbewusste Wertvorstellungen in zum Teil diskriminierender Weise in Technik einbetten lassen. Technische Systeme sind daher in aller Regel keine objektiven und verwendungsneutralen Instrumente und ein erster Schritt zur Prävention negativer ethischer Implikationen besteht darin, dies zu erkennen und die eigene Sichtweise kritisch zu reflektieren.

Weiterführende Literatur

Crawford, K. (2013). The Hidden Biases in Big Data. *HBR Blog Network*, 9–10.

Koene, A. (2017). Algorithmic Bias. *IEEE Technology and Society Magazine*, June, 31–32.

Cramer, H., Garcia-Gathright, J., Springer, A., & Reddy, S. (2018). Assessing and addressing algorithmic bias in practice. *Interactions*, 25(6), 58–63. <https://doi.org/10.1145/3278156>

Panch, T., Mattie, H., & Atun, R. (2019). Artificial intelligence and algorithmic bias: Implications for health systems. *Journal of Global Health*, 9(2), 1–5. <https://doi.org/10.7189/jogh.09.020318>

Kirkpatrick, K. (2016). Battling algorithmic bias. *Communications of the ACM*, 59(10), 16–17. <https://doi.org/10.1145/2983270>

³ <https://stereoset.mit.edu>

Aufgaben

Vorbereitung

1. Schauen Sie das einleitende Video an.
2. Lesen Sie den Thementext.
3. Lesen Sie einen der weiterführenden Artikel. Beginnen Sie mit dem Text von Kate Crawford und lesen Sie nach Interesse weitere Artikel.
4. Erkennen oder antizipieren Sie mögliche Biases in Ihrem Stakeholder-Projekt? Welche Lösungsansätze schlagen Sie vor?
5. Erstellen Sie in Ihrer Gruppe eine Posterfolie, welche das Ergebnis der Vorbereitungsaufgaben darstellt.

Themensitzung

1. Erläutern Sie das Poster anderen Gruppen, die dasselbe Stakeholder-Projekt wie Sie bearbeiten.
2. Diskutieren Sie Ihre Resultate und aufkommenden Fragen im Plenum.

Literatur

- Anderson, Chris. 2008. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *Wired*, 1–2. <https://doi.org/10.1016/j.ecolmodel.2009.09.008>.
- Bird, Eleanor, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp, and Alan Winfield. 2020. "The Ethics of Artificial Intelligence: Issues and Initiatives." <https://doi.org/10.2861/644>.
- Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. "Semantics Derived Automatically from Language Corpora Contain Human-like Biases." *Science* 356 (6334): 183–86. <https://doi.org/10.1126/science.aal4230>.
- Caruana, Rich, Yin Lou, Johannes Gehrke, Paul Koch, Marc Sturm, and Noemie Elhadad. 2015. "Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-Day Readmission." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, 1721–30. <https://doi.org/10.1145/2783258.2788613>.
- Cowgill, Bo, Fabrizio Dell'acqua, Samuel Deng, Daniel Hsu, Nakul Verma, and Augustin Chaintreau. 2020. "Biased Programmers? Or Biased Data? A Field Experiment in Operationalizing AI Ethics." *EC 2020 - Proceedings of the 21st ACM Conference on Economics and Computation*, no. NeurIPS: 679–81. <https://doi.org/10.1145/3391403.3399545>.
- Crawford, Kate. 2013. "The Hidden Biases in Big Data [Blog Post]." *HBR Blog Network*, 1–4. http://blogs.hbr.org/cs/2013/04/the_hidden_biases_in_big_data.html.
- Gianfrancesco, Milena A., Suzanne Tamang, Jinoos Yazdany, and Gabriela Schmajuk. 2018. "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data." *JAMA Internal Medicine*. American Medical Association. <https://doi.org/10.1001/jamainternmed.2018.3763>.
- Goddard, Kate, Abdul Roudsari, and Jeremy C Wyatt. 2011. "Automation Bias – A Hidden Issue for Clinical Decision Support System Use." *International Perspectives in Health Informatics. Studies in Health Technology and Informatics* 164: 17–22.
- Johnson, Khari. 2020. "StereoSet Measures Racism, Sexism, and Other Forms of Bias in AI Language Models." *Venture Beat*, April 2020.
- Jordan, Michael I. 2018. "Artificial Intelligence — The Revolution Hasn't Happened Yet." *Medium*.
- Koene, Ansgar. 2017. "Algorithmic Bias." *IEEE Technology and Society Magazine*, no. June: 31–32.

- Krishnan, Anoop, Ali Almadan, and Ajita Rattani. 2020. "Understanding Fairness of Gender Classification Algorithms Across Gender-Race Groups." <http://arxiv.org/abs/2009.11491>.
- Wiggers, Kyle. 2020. "Study Indicates Neither Algorithmic Differences nor Diverse Data Sets Solve Facial Recognition Bias." *Venture Beat*, September 2020. <https://venturebeat.com/2020/09/28/study-indicates-neither-algorithmic-differences-nor-diverse-data-sets-solve-facial-recognition-bias/>.