

Wie man Diskriminierung auf Online-Plattformen entdeckt und bekämpft: Was wir von Airbnb, Uber und anderen lernen können

Michael Luca und Dan Svirsky

Die zunehmende Verlagerung der Märkte auf Online-Plattformen hat das Potenzial, Transaktionen effizienter und fairer zu machen. Untersuchungen aus der Frühphase haben auch darauf hingedeutet, dass die relative Anonymität der Online-Transaktionspartner zu weniger Diskriminierung führte. In welchem Maße dieses Versprechen eingehalten wird, hängt jedoch vom Plattformdesign ab. Mit der Verlagerung von immer mehr Märkten und Transaktionen ins Internet entwickeln sich auch die Marktplätze weiter, und die Plattformdesigner suchen beständig nach neuen Möglichkeiten, Vertrauen zwischen fremden Transaktionspartnern zu fördern. Dabei sind die Plattformen im Laufe der Zeit unterschiedliche Wege gegangen, selbst innerhalb einzelner Branchen. Die konzeptionellen Entscheidungen im Plattformdesign bestimmen, wie effizient und integrativ Märkte tatsächlich sind.

Wenn vertrauensbildende Maßnahmen Diskriminierung fördern ✕ Anders als frühere Online-Marktplätze haben sich Plattformen wie Airbnb entschieden, Namen und Fotos von potenziellen Transaktionspartnern online zu stellen. Dies mag in der Absicht geschehen sein, Vertrauen zu stiften und Geschäftsbeziehungen zwischen Fremden zu erleichtern, aber es öffnet auch der Diskriminierung Tür und Tor. Untersuchungen zeigen inzwischen in einer Reihe von digitalen Plattformen, von Stellenmärkten über Kreditbörsen bis hin zum Immobilienmarkt, rassistische oder ethnische Diskriminierung. Möglich wird dies vor allem durch zwei grundsätzliche Möglichkeiten, die Plattformen bieten. Erstens geht es um die Präsentation von Merkmalen, die Rückschlüsse auf die Zugehörigkeit zu bestimmten Gruppen ermöglichen und bewusste oder unbewusste Diskriminierung auslösen können. Dazu gehören zuallererst Fotos, aber auch subtilere Indikatoren wie Namen. Zweitens können Verkäufer immer freier entscheiden, mit welchen

KEYWORDS

Marktdesign, Plattformdesign, Diskriminierung, Feldstudien

AUTOREN

Michael Luca

Lee J. Styslinger III Associate Professor of Business Administration, Harvard Business School, Boston, MA, USA, mluca@hbs.edu

Dan Svirsky

Data Scientist, Uber Technologies, Inc, Boston, MA, USA, dsvirsky@uber.com





Plattform-Manager sollten mehr Bewusstsein für Diskriminierung entwickeln und das Problem proaktiv angehen.





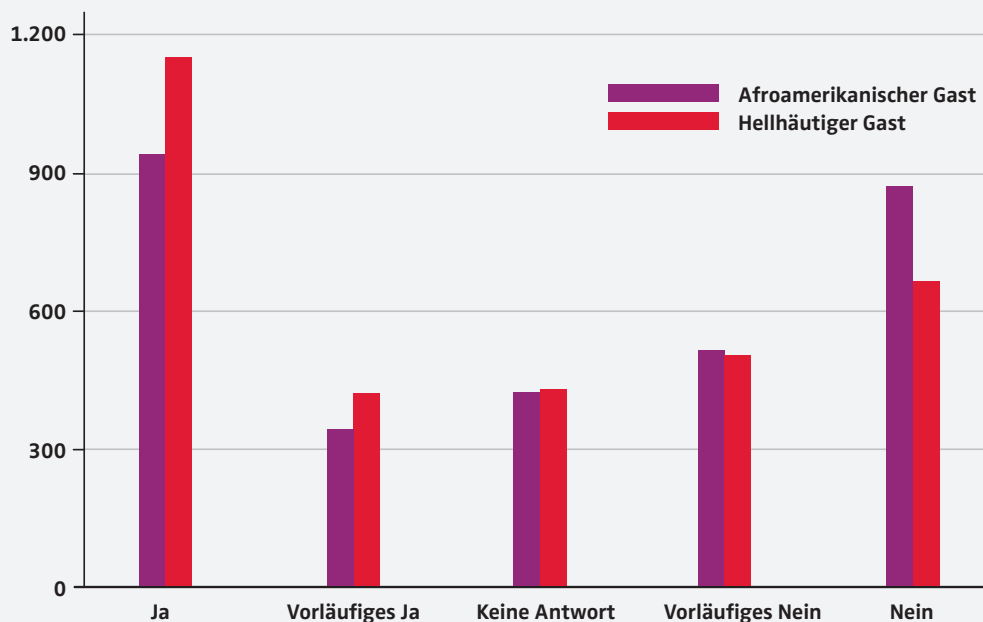
BOX 1

Untersuchungen zu Diskriminierung auf Airbnb

2015 haben wir zusammen mit einem Kollegen rassistische Diskriminierung auf Airbnb untersucht. Im Rahmen einer Studie, die sich auf die US-Städte Baltimore, Dallas, Los Angeles, St. Louis und Washington konzentrierte, erstellten wir 20 Nutzerprofile und versandten Mietanfragen an rund 6.400 Gastgeber. Die Profile und Anfragen waren identisch und unterschieden sich ausschließlich im Nutzernamen. Die Hälfte der Profile hatte Namen, die laut Geburtenregister unter Menschen mit weißer Hautfarbe üblich sind, die andere Hälfte hatte Namen, die unter Afroamerikanern weit verbreitet sind.

Entgegen den Aussagen von Airbnb fanden wir auf der Plattform Belege für beträchtliche Diskriminierung. Anfragen mit afroamerikanisch klingenden Namen waren um 16 % seltener erfolgreich als Namen, die Menschen weißer Hautfarbe vermuten lassen. Die meisten Gastgeber, die Anfragen von Inhabern afroamerikanisch klingender Profilen ablehnten, hatten noch nie einen Menschen mit dunkler Hautfarbe zu Gast; dies deutet darauf hin, dass manche Gastgeber dazu neigen, Menschen auf Basis ihrer Hautfarbe zu diskriminieren. Die Studienergebnisse wurden auf weitere Merkmale von Airbnb-Gästen erweitert. Dabei fanden die Wissenschaftler Diskriminierung unter anderem gegen Menschen mit Behinderungen, gleichgeschlechtliche Paare und Menschen mit Migrationshintergrund.

ABBILDUNG 1 > Antworten der Gastgeber nach Hautfarbe (auf Basis des Namens)



aus Edelman, B., Luca, M. & Svirsky, D. (2017)

Airbnb beteiligte sich zwar nicht an der Studie, hat aber als Reaktion auf die Untersuchungsergebnisse und auf wachsende Kritik von Nutzern und Regulierungsbehörden begonnen, das Problem anzugehen. Wir haben Airbnb Vorschläge zur Diskriminierungsbekämpfung vorgelegt, und das Unternehmen beauftragte eine Task Force damit, die von uns vorgeschlagenen und andere Maßnahmen zu bewerten. Dies alles hat letztlich zu einer Reihe von Änderungen im Plattformdesign geführt. Ein Beispiel für eine vielversprechende Änderung ist die seit Kurzem erweiterte „Instant Book“-Funktion, mit der qualifizierte Gäste eine Buchung abschließen können, bevor die Gastgeber ihr Profil zu sehen bekommen. Diese Funktion beugt nicht nur Diskriminierung vor, sie macht den Buchungsvorgang auch einfacher und bequemer. Als Reaktion auf unsere Arbeit sagte Airbnb außerdem zu, die Anwendung der Sofortbuchung auszuweiten; inzwischen gibt es mehrere Millionen von Objekten rund um den Globus, für die diese Funktion zur Verfügung steht. Allerdings hat Airbnb bis heute weder öffentlich gemacht, in welchem Ausmaß auf der Plattform nach wie vor Diskriminierung stattfindet, noch, welche Wirkung die Änderungen erzielt haben.

Käufern sie ins Geschäft kommen wollen. Beide Optionen basieren auf Entscheidungsmöglichkeiten, die Plattformdesigner aktiv anbieten.

Wie Manager die Diskriminierung auf Online-Plattformen reduzieren können ✕ Selbst innerhalb einer Branche unterscheiden sich Plattformen oft in ihren konzeptionellen Merkmalen, was ein unterschiedliches Maß an Diskriminierung zur Folge haben kann. Das Ferienwohnungsportal HomeAway zeigt auf der Hauptseite der Suchergebnisse nur Fotos der Mietobjekte und veröffentlicht Fotos der Gastgeber erst in einer späteren Phase oder überhaupt nicht, während Airbnb traditionell Fotos der Gastgeber direkt auf der Hauptergebnisseite gezeigt hat. Als Reaktion auf unsere Untersuchung hat Airbnb diese Voreinstellung umprogrammiert und zeigt Fotos von Gastgebern und Gästen jetzt ebenfalls erst in einer späteren Phase. Dies ist ein gutes Beispiel dafür, mit welchen Schritten Plattformen Diskriminierung reduzieren können. Auf der Basis unserer früheren Untersuchungen und unserer Erfahrung mit Unternehmen prüfen wir Schritte zur Bekämpfung von Diskriminierung (siehe Abbildung 2).

> **Bewusstsein für Diskriminierungspotenzial auf Plattformen schärfen** ✕ Plattformen sollten ein Verständnis dafür entwickeln, wie Designentscheidungen und Algorithmen das Ausmaß der Diskriminierung auf einem Marktplatz beeinflussen können. Mit einem besseren Bewusstsein für diese Problematik können Manager das Problem proaktiv untersuchen und angehen. So hat zum Beispiel das zentrale Steuerungsteam des Fahrdiensteanbieters Uber eine Fairness Working Group eingerichtet, die Diskriminierungsprobleme untersuchen soll. Diese Gruppe ist unter anderem deshalb so wertvoll, weil sie interdisziplinär angelegt ist: Sie umfasst Ökonomen, Data Scientists, Rechtsanwälte und Produktmanager aus allen Bereichen des Unternehmens, die gemeinsam für mehr Fairness sorgen sollen. Insbesondere große Organisationen können von Arbeitsgruppen profitieren, die sich ausschließlich mit Diskriminierungsrisiken bei neuen Projekten befassen.

> **Diskriminierung auf Plattformen erfassen** ✕ Aktuell wissen viele Plattformen wenig über die Zusammensetzung ihrer Nutzerschaft in Bezug auf Hautfarbe, ethnische Gruppierung oder Geschlecht. Und es ist schwer, ein Problem zu bekämpfen, das man nicht erfasst hat. Ein regelmäßiger Untersuchungsbericht zu diskriminierungsgefährdeten Gruppen und deren Erfolg auf der Plattform ist eine entscheidende Maßnahme, um Probleme aufzudecken und zu lösen. Airbnb hat nach unserer Untersuchung damit begonnen, das Ausmaß der Diskriminierung auf der Plattform zu erfassen und Ökonomen gemeinsam mit Data Scientists mit diesem Thema betraut. Und auch Uber arbeitet mit seiner Arbeitsgruppe an der kontinuierlichen Überprüfung von Diskriminierung.



> **Sensible Daten zurückhalten** ✕ Eine einfache, aber wirksame Maßnahme besteht in vielen Fällen darin, dass Plattformen potenziell sensible Nutzerdaten wie Hautfarbe oder Geschlecht bis nach einem Geschäftsabschluss zurückhalten. Manche Plattformen, darunter Amazon und eBay, tun dies bereits. Was passiert, wenn Rückschlüsse auf die Hautfarbe im Vorfeld einer Transaktion möglich sind, zeigte sich auch bei eBay, obwohl dieser Online-Marktplatz die Hautfarbe nicht offen darstellt. Bei einer aktuellen Studie haben eBay-Nutzer Baseballkarten zum Verkauf angeboten und die Karten nicht auf einer neutralen Unterlage fotografiert, sondern mit der Hand vor die Kamera gehalten und dieses Foto gepostet; selbst dieses subtile Offenlegen der Hautfarbe hat beim Verkauf der Karten auf eBay zu Diskriminierung geführt.

> **Transaktionen automatisieren, aber Algorithmen auf Diskriminierung prüfen** ✕ Automatisierung und Algorithmen können nützliche Werkzeuge gegen Vorurteile sein. Die in Box 1 beschriebene Sofortbuchung bei Airbnb ist ein gutes Beispiel dafür. Bei dieser Funktion entfällt der Schritt, in dem Gastgeber sich Namen und Fotos von Gästen ansehen und sich dann für oder gegen bestimmte Gäste entscheiden können. Airbnb hat die Zahl der Nutzer, die die Sofortbuchungsfunktion nutzen, inzwischen deutlich erhöht.

ABBILDUNG 2 > Antidiskriminierungsstrategien für das Plattformdesign



Darüber hinaus untersuchen immer mehr Experten, wie sich auch bei Algorithmen Vorurteile ausschließen lassen. Präferenzen oder Verhaltensweisen der Menschen zu ändern, ist oft schwierig; die Eingabeparameter oder die Ziele eines Algorithmus zu verändern, kann da deutlich einfacher sein. So hat zum Beispiel LinkedIn sein Recruiter-Tool, mit dem Arbeitgeber nach geeigneten Kandidaten suchen können, so verändert, dass das Geschlechterverhältnis der Suchergebnisse dem Geschlechterverhältnis des jeweiligen Berufs entspricht. Wenn also 30 % aller Data Scientists weiblich sind, dann sind auch 30 % der Kandidaten, die einem Unternehmen angezeigt werden, weiblich. Dieses Beispiel zeigt, wie Algorithmen den Gleichheitsgrundsatz und die Fairness einer Plattform beeinflussen können – und welche Spielräume und Möglichkeiten der Feinsteuerung Manager beim Entwickeln von Algorithmen

haben, denn die Zielquote ist nur eine von vielen Möglichkeiten, Fairness zu fördern.

> **Wie ein Entscheidungsarchitekt denken** ✕ Plattformen können auch über die Entscheidungsarchitektur Entscheidungen gezielt lenken, um Diskriminierung vorzubeugen. So wurde zum Beispiel durch Untersuchungen in den unterschiedlichsten Kontexten nachgewiesen, dass Menschen dazu neigen, sich für die Alternative zu entscheiden, die als Standard festgelegt ist. Dass es hier viel Spielraum gibt, zeigt das Beispiel mit der Sofortbuchungsoption bei Airbnb, die Gastgeber über die Änderung von Einstellungen selbst aktivieren müssen. Würde Airbnb diese Funktion zum Standard machen und die Gastgeber müssten sie bewusst deaktivieren, könnte das Unternehmen den Spielraum für

Diskriminierung verkleinern. Airbnb hat solche Ansätze immer wieder getestet.

Antidiskriminierungsrichtlinien sind ein weiteres Beispiel. Die meisten Plattformen verfügen über Regeln, die Diskriminierung verbieten, aber diese Vorschriften sind im Kleingedruckten versteckt, dem die meisten zustimmen, ohne es richtig zu lesen. Stattdessen könnten Marktplätze Antidiskriminierungsregeln zu einem relevanteren Zeitpunkt vorlegen, nämlich während der tatsächlichen Transaktion. Manche Menschen würden dann immer noch dagegen verstoßen, aber andere wären vielleicht dankbar für den Hinweis, weil sie sich des Problems überhaupt nicht bewusst sind.

- > **Wirkung von Plattformänderungen auf die Diskriminierung in Experimenten testen** ✕ Wenn Plattformen Möglichkeiten zur Messung von Diskriminierung gefunden haben, sollten sie diese mit entsprechenden Kennzahlen bei

»

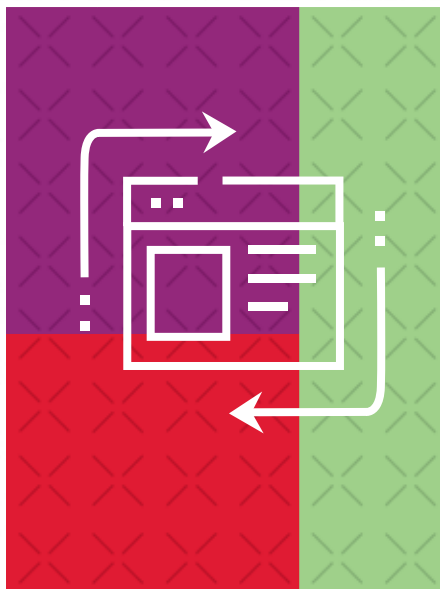
Marktplätze könnten Antidiskriminierungsregeln zu einem relevanteren Zeitpunkt vorlegen, nämlich während der tatsächlichen Transaktion.

«

Konzepttests mitberücksichtigen. So können sie die unterschiedlichen Wirkungen von Änderungen in Design und Funktion besser nachvollziehen. Airbnb hat zum Beispiel in einem Experiment in der Liste der Suchergebnisse keine Fotos der Gastgeber mehr gezeigt und Auswirkungen auf die Buchungszahlen untersucht. Seit unserer Studie hat Airbnb ein eigenes Team, das sich mit solchen Überlegungen beschäftigt. LinkedIn hat untersucht, wie sich die Änderungen an der Recruiter-Funktion ausgewirkt haben, und konnte keinen Effekt auf die Erfolgsquote der Stellenanzeigen von Arbeitgebern feststellen.

- > **Auswirkungen offen legen und transparent sein** ✕ Plattformen sollten Transparenz anstreben und gemeinsam mit einem breiten Spektrum an Stakeholdern Diskriminierungsprobleme ermitteln und lösen. So können sich Plattformmanager und Designer austauschen, bevor es zu einer Krise kommt; außerdem lässt sich auf diese Weise messen und nachverfolgen, wie erfolgreich bestimmte Maßnahmen sind. Dies ist besonders wichtig, weil ja nicht im Voraus klar ist, welche Wirkung bestimmte Änderungen an einer Plattform haben werden.

Die zunehmende Verbreitung von Online-Marktplätzen hat das Wesen vieler wirtschaftlicher Transaktionen verändert. Unsere Untersuchungen zeigen, welche tiefgreifende Auswirkungen Änderungen im Design einer Plattform haben können. Wer sich mit diesen Wirkungen auseinandersetzt und aktuelle Forschungsergebnisse nutzt, kann Märkte schaffen, die sowohl effizient als auch integrativ sind. ✕



LITERATURHINWEISE

Ayres, I.; Banaji, M.; & Jolls, C. (2015): „Race effects on eBay“, The RAND Journal of Economics, Vol. 46, 891-917. doi: 10.1111/1756-2171.12115

Edelman, B.; Luca, M.; & Svirsky, D. (2017): „Racial discrimination in the sharing economy: Evidence from a field experiment“, American Economic Journal: Applied Economics, Vol. 9(2), 1-22. <https://doi.org/10.1257/app.20160213>

Fisman, R.; & Luca, M. (2016): „Fixing Discrimination in Online Marketplaces“, <https://hbr.org/2016/12/fixing-discrimination-in-online-marketplaces>

Luca, M. & Bazerman, M. (2020): „The Power of Experiments: Decision-Making in a Data-Driven World“, <https://mitpress.mit.edu/books/power-experiments>

Morton, F. S.; Zettelmeyer, F.; & Silva-Risso, J. (2003): „Consumer Information and Discrimination: Does the Internet Affect the Pricing of New Cars to Women and Minorities?“, Quantitative Marketing and Economics, Vol. 1, 65-92. <https://doi.org/10.1023/A:1023529910567>