

Sicherheitsempfinden in sozialen Medien und Suchmaschinen – ein realistisches Abbild der Kriminalitätsbelastung?

Von Johannes Rieckmann und Jan-Lucas Schanze

Um die subjektive Kriminalitätsfurcht der Bevölkerung zu messen, werden bislang vor allem Befragungen durchgeführt. In Zeiten der zunehmenden Digitalisierung könnten aber auch Daten aus sozialen Medien und Suchmaschinen dazu beitragen, mehr über die subjektiven Empfindungen der Bevölkerung zu bestimmten Deliktfeldern zu erfahren. Dieser Beitrag zeigt, dass die Auswertung von Daten aus sozialen Medien und Suchmaschinen zwar nicht geeignet ist, die tatsächliche Kriminalitätsbelastung in Deutschland widerzuspiegeln. Sie kann aber durchaus als kostengünstige Ergänzung der bisherigen Erhebungsmethoden dienen, um die empfundene Belastung abzubilden.

Die Zahl der Internetnutzer in Deutschland – derzeit nutzen rund 80 Prozent der Gesamtbevölkerung zumindest gelegentlich das Internet¹ – nimmt stetig zu. Damit wird das Internet auch für Kriminalitätsforscher zu einer immer interessanteren Datenquelle. Bisher ist die Polizeiliche Kriminalstatistik (PKS) des Bundeskriminalamts, flankiert von gelegentlichen und in der Vergangenheit meist nur regionalen Dunkelfeld- und Kriminalitätsfurchtbefragungen, die Hauptquelle der Datenerhebung zur Kriminalitätsbelastung der Bevölkerung. Analysen von Online-Daten könnten vor allem die Perspektive auf die Wahrnehmung von Kriminalität bereichern oder auf lange Sicht sogar eine alternative und eigenständige Datenquelle für Kriminalitätsforscher werden. Der vorliegende Beitrag² evaluiert die Tauglichkeit solcher über das Internet erhobenen Daten als Ergänzung – oder gar als kostengünstigen Ersatz – klassischer Kriminalitätsfurchtstudien. Betrachtet werden hierzu zum einen das Suchmaschinenverhalten der Nutzer in Deutschland und zum anderen ihre Äußerungen in verschiedenen sozialen Medien. Dabei soll ausdrücklich nicht die Kriminalitätsentwicklung zum Zweck der vorausplanenden Kriminalitätsprävention und -bekämpfung (*Predictive Policing*) prognostiziert werden. Vielmehr soll die nutzerorientierte Medienanalyse einen ersten deskriptiven Befund zum Sicherheitsempfinden der Internetnutzer in Deutschland liefern.

¹ Van Eimeren, B., Frees, B. (2014): Ergebnisse der ARD/ZDF-Onlinestudie 2014. 79 Prozent der Deutschen online – Zuwachs bei mobiler Internetnutzung und Bewegtbild. Media Perspektiven 7-8/2014. In einer Befragung des WISIND-Projekts mit gut 12 000 Befragten liegt der entsprechende Wert bei rund 75 Prozent.

² Der Bericht entstand im Rahmen des Forschungsprojekts *Ein Wirtschaftswissenschaftlicher Sicherheitsindikator für Deutschland (WISIND)*. Es wird vom Bundesministerium für Bildung und Forschung in der Förderlinie *Gesellschaftliche Dimensionen der Sicherheitsforschung* ermöglicht. Die Konzeptionierung des WISIND-Projekts und die Erhebung und Analyse der WISIND-eigenen Daten sind in Zusammenarbeit mit Martin Kroh, Mathias Bug, Kristina Meier, Eric van Um und Nina Wald sowie den Mitarbeitern des Brandenburgischen Instituts für Gesellschaft und Sicherheit (BIGS) entstanden. Die Autoren danken darüber hinaus Enrique Fernandez, Martina Kraus, Magda Pchalek und Bartosz Walenda für ihre Unterstützung während dieser Prozesse.

Anschließend wird dieser Befund mit der Messung der Kriminalitätsfurcht im WISIND-Projekt abgeglichen.³

Soziale Medien und Suchmaschinen als Datenquellen

Dieser Bericht verwendet einerseits Daten zur Zahl von Internet-Einträgen zu bestimmten Kriminalitätsbereichen in sozialen Medien. Konkret handelt es sich dabei um Facebook, Twitter, Foren, Blogs sowie um Texte, die im Kommentarbereich von Youtube-Videos verfasst wurden. Diese dienen dem Ausdruck von Empfindungen und Meinungen sowie der Kommunikation und haben daher eine in erster Linie *expressive* Funktion.

Andererseits werden Daten aus der in Deutschland marktführenden Suchmaschine Google verwendet. Suchmaschinen haben eine in erster Linie *explorative* Funktion, da sie der Einholung von Informationen dienen. Google bietet mit Google Trends eine Plattform für den Export von Daten zum Aufkommen von Suchbegriffen über Raum und Zeit. Google Trends erfuh bereits mehrere wissenschaftliche Anwendungen zur zeitnahen rückblickenden Betrachtung verschiedener Phänomene⁴ und wird im Folgenden als Quelle für Suchmaschinendaten verwendet.

Motivation der Nutzer

Während Suchmaschinen primär der Informationsbeschaffung dienen, werden soziale Medien vor allem für den Informations- und Meinungsaustausch verwendet. Die Nutzung beider impliziert eine unmittelbare oder mittelbare persönliche Betroffenheit des Nutzers. Es ist davon auszugehen, dass sich ein großer Teil der Bevölkerung im Internet über als bedrohlich empfundene Ereignisse oder Umstände informiert. Ein Teil davon kommuniziert diese in sozialen Medien und interessiert sich dabei auch für Schutzmaßnahmen (zum Beispiel Alarmanlagen oder Pfefferspray), wodurch Daten für eine Abbildung des Bedrohungsempfindens der Bevölkerung abseits von Befragungen entsteht. Die Hypothesen zur konkreten Motivation der Nutzer sind jedoch spekulativ, denn Datenspuren im Netz zeigen zwar, *was* Nutzer tun, aber nicht, *warum* sie dies tun.⁵

³ Siehe dazu auch den Beitrag von Bug, M., Kraus, M., Walenda, B. in diesem DIW Wochenbericht.

⁴ Insbesondere zu nennen ist Google Flu Trends (GFT), mit dem in den USA Grippewellen vorausgesagt werden sollen. Wenn potentielle Grippekranken zunächst per Suchmaschine im Internet nach Symptomen suchen („googlen“), wird über Google Trends die regionale wie temporale Häufung dieser Suchanfragen festgestellt und über diese Rückschlüsse auf anrollende Infektionswellen gezogen, noch bevor die Patienten zum Arzt gehen. Für eine Kritik an GFT siehe Lazer, D. et al. (2014): The Parable of Google Flu: Traps in Big Data Analysis. *Science* Vol. 343, 1203-1205.

⁵ Mahrt, M., Scharnow, M. (2013): The Value of Big Data in Digital Media Research. *Journal of Broadcasting & Electronic Media* Vol. 57 (1), 20-33.

Datenerhebung in sozialen Medien

Die Daten in den sozialen Medien wurden im Auftrag des DIW Berlin von der Firma Beck et al. Services GmbH, einem auf Datenerhebung aus entsprechenden Quellen spezialisierten Auftragnehmer, erhoben. Die Datenerhebung erstreckte sich über einen Zeitraum von viereinhalb Monaten, zwischen dem 12. Juni und dem 31. Oktober 2014. Datengrundlage für die Analyse sind alle vom Nutzer öffentlich mitgeteilten und damit für jedermann einsehbaren Einträge und Profilinformationen in den oben genannten Medien. Auf Basis einer Liste von Suchwörtern zu zehn definierten Deliktfeldern⁶ wurden täglich mittels eines sogenannten Crawlers automatisiert die deutschsprachigen Einträge in den fünf genannten Netzwerken ge- und durchsucht. Die Auswahl der Begriffe, ihre Berücksichtigung in unterschiedlichen Schreibweisen sowie die Verwendung von zum Ausschluss führenden Begriffen (sogenannte *killer terms*) stellte eine hohe Treffgenauigkeit sicher und befreite die erhobenen Daten von Beifängen. Zum Deliktfeld Körperverletzung wurden beispielsweise alle neu eingetragenen Text-Fragmente mit Schlüsselwörtern wie *Schlägerei*, *Körperverletzung* oder *verprügelt* gezählt; Beiträge mit dem Schlüsselwort *Gewalt*, die sich auf Orte außerhalb Deutschlands (Ukraine, Irak) oder auf abstrakte Begriffe (höhere Gewalt) bezogen, wurden dabei als Beifang identifiziert und nicht berücksichtigt.

Ausschließlich öffentlich einsehbare Beiträge wurden auf Vorhandensein der Suchbegriffe untersucht; die Erfassung nicht für jedermann einsehbarer Nachrichten und Einträge auf Profelseiten war von vornherein technisch ausgeschlossen. Für eine regionale Darstellung der Daten wurden Ortserwähnungen in den Textfragmenten aufgenommen sowie nach Ortsnennungen der Nutzer gesucht. So kann etwa eine Erwähnung von Berlin auf dem Profil („Lebe in: Berlin“) genauso zur Verortung eines Internet-Beitrages über Körperverletzung führen, wie die Nennung des Ortes im Beitrag selbst („Körperverletzung in Berlin“).⁷ Ergebnisse einer automatisierten Feststellung der Tonalität der Beiträge (positiv, negativ, neutral) erwiesen sich für die empirische Analyse nur als bedingt brauchbar und fanden daher keinen Eingang in diesen Bericht. Weitere Variablen wie Alter, Geschlecht oder andere Charakteristika der Internetnutzer wurden nicht erhoben. Der Datenschutz der Nutzer ist über die Anonymisierung der Beiträge voll-

⁶ Hierbei handelt es sich um die Deliktfelder Internetkriminalität, Diebstahl und Einbruch, Drogenkriminalität, Körperverletzung, Organisierte Kriminalität, politisch motivierte Straftaten, Raub, religiöser Fundamentalismus, Sexualdelikte sowie Tötungsdelikte.

⁷ Die Ortsnennung innerhalb des Beitrags ermöglicht dabei tendenziell die zuverlässige Lokalisierung des diskutierten Geschehens.

Kasten

Probleme der Datenerhebung im Internet

Im Zusammenhang mit der Analyse von Daten aus sozialen Medien, oftmals als *Big Data* bezeichnet, wird vermehrt der rein datengesteuerte, nicht theoriebasierte Ansatz mancher Forschungsarbeiten kritisiert.¹ Die bloße Verfügbarkeit und Größe neuer Datenquellen ist jedoch kein ausreichendes Argument für die Notwendigkeit von Forschung, sie muss immer mit der Feststellung von Problemen und Qualitätsfragen einhergehen; und wenn möglich Bezüge zu klassischer Datenerhebung herstellen.

Der erste und offensichtlichste Nachteil der Arbeit mit im Internet generierten Daten ist die mangelnde Repräsentativität der Daten. Die Gruppe der Internetnutzer in Deutschland ist zwar mittlerweile zahlenmäßig sehr groß, stellt aber dennoch ein systematisch verzerrtes Abbild zugunsten jüngerer Generationen dar. In der Generation 60plus nutzen weniger als die Hälfte der Deutschen das Internet. Diese Abweichung von der Grundgesamtheit verstärkt sich bei der Nutzung von sozialen Medien: Drei Viertel der Befragten unter 30 Jahren ist dort aktiv, aber nur fünf Prozent der Internetnutzer über 70 Jahren.²

Bei Google Trends ist weniger die Datenbasis als vielmehr die Datenaufbereitung problematisch. Google ist ein kommerzielles Unternehmen und kein wissenschaftlicher Datenanbieter, die Intransparenz der internen Verfahren zur Datenerstellung bei Google-Trends spiegelt dies wider. Nachteilig für wissenschaftliche Analysen ist die Ausweisung der Ergebnisse, die nicht in absoluten Suchvorgängen, sondern in relativen Suchvolumina

ausgegeben werden, immer in Bezug zu einem Maximalwert, der auf 100 festgelegt wird. Diese Normalisierung des Suchvolumens entsteht laut Google durch die Verrechnung eines Suchwortes mit allen Suchvorgängen in einer Region und macht die Ergebnisse vergleichbar mit anderen Regionen. Unterschreiten die gesamten Suchanfragen eines bestimmten Wortes einen von Google nicht öffentlich definierten Schwellenwert, werden die Ergebnisse entweder gar nicht (Suchvolumen = 0) oder nur für längere Zeitintervalle (Monats- statt Wochenbasis) angezeigt. Dieser komplexe Vorgang wird von Google nicht öffentlich gemacht. Ein weiteres großes Problem sind auch nicht nachprüfbar Veränderungen der Google-Algorithmen über die Zeit, die die Replizierbarkeit von Daten beeinträchtigen.³ Veränderungen von Suchvolumina, insbesondere wenn sie über lange Zeiträume festgestellt werden,⁴ sind nicht immer ein belastbarer Beleg für eine tatsächliche Veränderung des Suchverhaltens.

Eine zusätzliche Verzerrung entsteht durch eine temporär erhöhte Medienaufmerksamkeit bei bestimmten Themen, die die Suchanfragen künstlich in die Höhe treibt – ein Problem, mit dem auch *Google Flu Trends* zu kämpfen hat.⁵ Die Beeinflussung des Suchverhaltens von Internetnutzern, etwa aufgrund kommerzieller Interessen von Firmen, die sich mehr Klicks auf den eigenen Seiten und den Absatz von Produkten wünschen, ist im Bereich der inneren Sicherheit und Kriminalität – zumindest für die hier abgefragten Suchworte – dagegen eher nicht zu erwarten.

¹ Welker, M., Kloß, A. (2014): Soziale Medien als Gegenstand und Instrument sozialwissenschaftlicher Forschung. In: König, C., Stahl, M., Wiegand, E. (2014), a. a. O.

² Van Eimeren, B., Frees, B. (2014), a. a. O., 380, 387.

³ Lazer, D. et al. (2014), a. a. O., 1205.

⁴ Google Trends stellt Daten rückblickend bis ins Jahr 2004 zur Verfügung.

⁵ Lazer, D. et al. (2014), a. a. O., 1204.

ständig gewährleistet, die Daten sind nicht personenbezogen oder personenbeziehbar.⁸

Erhebung von Suchmaschinen-Daten (Google Trends)

Die Analyse der Suchmaschinennutzung beschränkt sich auf den Marktführer Google. Dieser verfügte im Dezember 2014 über einen Anteil von 95 Prozent aller

Suchanfragen in Deutschland,⁹ die Nutzung von Suchmaschinen in der Gruppe der 55,6 Millionen Internetnutzer (*Onliner*) ist laut der ARD/ZDF-Onlinestudie 2014 mit 82 Prozent sehr weit verbreitet.¹⁰ Eine Analyse von Daten via Google Trends stellt also nahezu eine Vollerhebung aller Suchanfragen in Deutschland dar. Die Suchanfragen werden dabei nicht von Forschern lokalisiert, sondern von Google auf verschiedenen räumlichen Aggregationsebenen zur Verfügung gestellt.¹¹

⁸ Zum Datenschutz bei der Datenanalyse in sozialen Medien vgl. Tschering, R. (2014): Soziale Medien – Gegenstand und Instrument der Forschung – Rechtliche Aspekte. In: König, C., Stahl, M., Wiegand, E. (2014): Soziale Medien – Gegenstand und Instrument der Forschung. Wiesbaden.

⁹ Vgl. de.statista.com/statistik/daten/studie/167841/umfrage/marktanteile-ausgewaehlter-suchmaschinen-in-deutschland/.

¹⁰ Van Eimeren, B., Frees, B. (2014), a. a. O., 387.

¹¹ Über die genauen Mechanismen der Lokalisierung durch Google gibt es keine detaillierten Informationen. Es ist zu vermuten, dass dies über die IP-Adressen der Suchanfrager geschieht.

Im Vergleich zur Datenerhebung in den sozialen Medien gestaltet sich die Datensammlung bei Google Trends recht unkompliziert. Die Bedienungsfläche verfügt über eine Reihe von Anwendungsfunktionen, die etwa den Vergleich mehrerer Suchwörter, Länder oder Zeiträume untereinander ermöglicht. Der Nutzer erhält für Deutschland räumliche Informationen auf Bundesländerebene und kann durch die Wahl von Kategoriefiltern die Ergebnisse bereinigen. Für den vorliegenden Bericht wurden mögliche Beifänge¹² bei nicht eindeutigen Suchwörtern durch die Wahl der Kategorie „Gesetz und Regierung“ ausgeschlossen (Kasten).

Ergebnisse der Analyse

Ein Vergleich sämtlicher im Erhebungszeitraum erfassten Beiträge in allen untersuchten sozialen Medien zeigt deutliche Spitzenreiter in der Aufmerksamkeit der Nutzer (Tabelle 1). Einträge zu Diebstählen und Einbrüchen wurden fast 300 000 Mal erfasst, dicht gefolgt von Beiträgen zu religiösem Fundamentalismus. Einträge zu Tötungsdelikten, Körperverletzung und Drogenkriminalität folgen mit weitem Abstand. Ein Vergleich der im Nutzerinteresse stehenden Deliktfelder mit der polizeilichen Kriminalstatistik sowie den Ergebnissen der im WISIND-Projekt durchgeführten Dunkelfeldstudie nährt die Vermutung, dass soziale Netzwerke allein keine verlässlichen Rückschlüsse auf die tatsächliche regionale Bedro-

hung zulassen. Die Anzahl der Nennungen im Verhältnis zu anderen Delikten spiegelt nicht die Bedrohungslage in Deutschland wider, sondern gibt eher Hinweise auf Stimmungsbilder und empfundene Bedrohung. So führte etwa die Berichterstattung zur mit gewaltsamen Ausschreitungen verbundenen Hogesa („Hooligans gegen Salafisten“)-Demonstration Ende Oktober 2014 in Köln zu einem bundesweit stark erhöhten Aufkommen von Begriffen, die als Schlüsselbegriffe für die Bedrohungsbereiche religiöser Fundamentalismus und politisch motivierte Kriminalität ausgezählt wurden.

Von den insgesamt 1,2 Millionen erhobenen Einträgen konnten bislang gut 18 Prozent in über 7 300 Orten bis auf Gemeindeebene lokalisiert und den Landkreisen in Deutschland zugeordnet werden. Die Ergebnisse wurden auf die Bevölkerungszahl in den jeweiligen Gebietskörperschaften angepasst und in Einträge pro 100 000 Einwohner umgerechnet.¹³

Die Einträge zu konkreten Deliktformen sind unterschiedlich gut lokalisierbar. Nutzer, die über Raub, politische Straftaten, Einbrüche und Diebstahl schrieben, verbanden ihre Einträge häufiger mit Ortsnennungen als Nutzer, die zu Internetkriminalität oder Sexualdelikten schrieben. Intuitiv ist dies durch eine weitgehende Ortsungebundenheit¹⁴ der Internetkriminalität erklärbar. Eine Ortsnennung bei Internetkriminalität ist insofern eine kaum wertvolle Information. Bei Sexualdelikten ist zu vermuten, dass aus Gründen wie Schamhaftigkeit, dem häufigen Fehlen von Zeugen, Pietät und Opferschutz seltener eine Nachricht mit Orten verbunden wird. Deutlich wird durch die Lokalisierbarkeit von Einträgen in sozialen Medien aber auch, dass ein Vergleich mit der tatsächlichen regionalen Verteilung von Kriminalität mit einer systematischen Schiefelage zu kämpfen hat: Manche Delikte lassen sich besser regional zuordnen als andere.

Wichtig für Analysen der sozialen Medien ist die Auswahl mehrerer Netzwerke und Quellen, um nicht durch die Auswahl einer einzigen Quelle die Varianz der Daten weiter einzuschränken.¹⁵ Die meisten Beiträge des verwendeten Datensatzes stammen aus Facebook, dem am meisten genutzten sozialen Netzwerk in Deutschland.¹⁶ Twitter-Meldungen, Foreneinträge und Blogs stehen auf Platz zwei bis vier, Youtube-Kommentare machen nur einen sehr geringen Teil der Stichprobe aus. Interes-

¹² Beifänge im Kontext des Suchwortes Mord wären etwa: Bezug zur Literatur und Film (etwa „Tatort“ oder „Mord im Orientexpress“), für Suchwort (Computer-)Virus: Krankheiten (Ebola).

Tabelle 1

Erscheinungshäufigkeiten von Delikten in sozialen Medien¹

Delikt	Nennungen	Darunter: Lokalisiert	Anteil der lokalisierten Delikte in Prozent
Diebstahl und Einbruch	293 038	57 275	19,5
Religiös Fundamentalistisch	288 643	42 672	14,8
Körperverletzung	180 133	37 011	20,5
Tötungsdelikte	147 627	21 604	14,6
Drogenkriminalität	109 260	16 155	14,8
Raub	107 366	31 253	29,1
Internetkriminalität	53 280	2 639	5,0
Sexualdelikte	43 217	3 770	8,7
Politische Straftaten	40 384	8 947	22,2
Organisierte Kriminalität	20 478	3 362	16,4
Insgesamt	1 283 426	224 688	17,5

¹ In der Zeit vom 12. Juni bis 31. Oktober 2014.

Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trends; Berechnungen des DIW Berlin.

Spitzenreiter der Aufmerksamkeit in sozialen Medien sind Einträge zu Diebstählen und Einbrüchen sowie zu religiösem Fundamentalismus.

¹³ Anderenfalls hätte sich für Ballungsgebiete ein verzerrtes Bild ergeben. So wurden etwa allein Berlin mehr als 45 000 Einträge zugeordnet.

¹⁴ Der Tatbegehungsort ist in aller Regel fern des Erfolgsortes der Tat.

¹⁵ Mahrt, M., Scharnow, M. (2013), a. a. O., 25.

¹⁶ Vgl. Busemann, K., Gscheidle, C. (2012): Dabei sein ist alles – zur Nutzung privater Communities. Ergebnisse der ZDF-Studie Community 2011. In: Media Perspektiven 7–8, 380–390.

sant ist der Medienmix für verschiedene Deliktarten. Die Beiträge über Internetkriminalität stammen in der deutlichen Mehrzahl aus Foren, in denen sich Nutzer Hilfe für die Beseitigung von entstandenen Schäden oder Ratschläge zu Schutzmöglichkeiten vor Internetkriminalität einholen können (Abbildung 1). Einträge zu Raubdelikten werden überdurchschnittlich oft bei Facebook diskutiert und verbreitet, und Fundamentalismus, eines der im Beobachtungszeitraum wichtigsten betrachteten Diskussionsthemen, wird gleichermaßen auf Blogs und in Foren, häufiger aber bei Twitter und Facebook erwähnt.

Vergleich von sozialen Medien und Google-Trends

Wie stark unterscheidet sich das expressive vom explorativen Informationsverhalten im Internet? Eine Rangfolge der Häufigkeit von Suchwörtern bei Google Trends und ein Vergleich mit den Wortaufkommen aus der Analyse der sozialen Medien für den Zeitraum von Juni bis Oktober 2014 kann darüber Aufschluss geben (Tabelle 2). Die beiden häufigsten Begriffe aus beiden Datenquellen wurden jeweils auf den Wert 100 fixiert, die anderen Begriffe folgen im Verhältnis zu die-

Tabelle 2

Ordinale Bedeutungsabstufung in sozialen Medien und Suchmaschinen¹

Skala von 1 bis 100

Delikt	Google Trends	Soziale Medien
Tötungsdelikte (Mord)	100	50
Diebstahl und Einbruch	26	100
Körperverletzung	26	61
Drogenkriminalität (Drogen)	18	37
Raub	9	37
Internet-Kriminalität (Virus)	9	18
Sexualdelikte (Vergewaltigung)	9	15
Religiös Fundamentalistisch (Islamist/Salafist)	5	99
Organisierte Kriminalität (Menschenhandel)	2	7
Politische Straftaten (Rechts- & Linksextremismus)	2	14

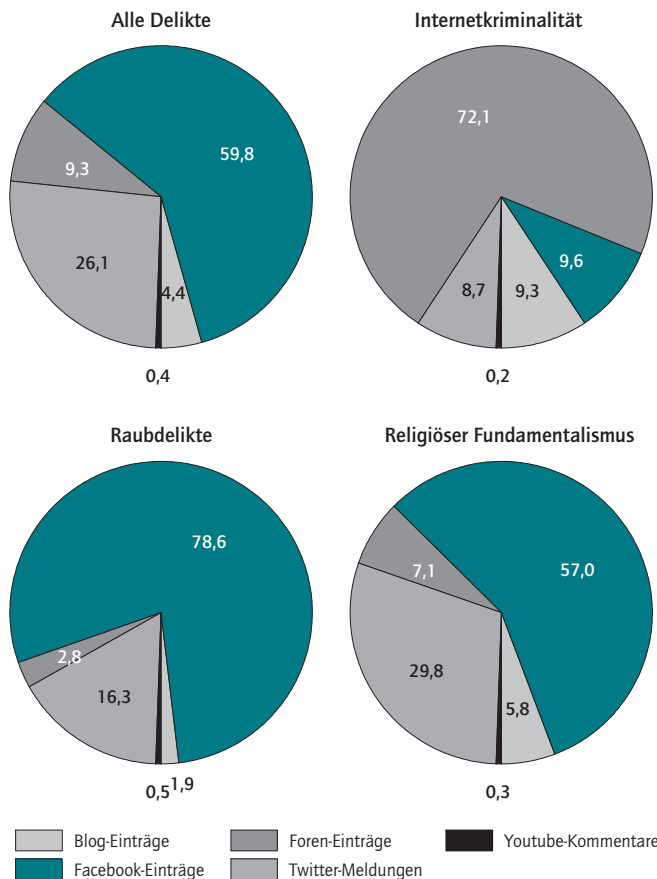
¹ In der Zeit vom 12. Juni bis 31. Oktober 2014. Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trends; Berechnungen des DIW Berlin.

© DIW Berlin 2015

Begriffe zu Diebstahl, Einbruch und Körperverletzung rangieren sowohl bei Google Trends als auch in den sozialen Medien im Kontext von Kriminalitätsfurcht weit vorne.

Abbildung 1

Delikteinträge nach Quellen in sozialen Medien
Anteile in Prozent



Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trends; Berechnungen des DIW Berlin.

© DIW Berlin 2015

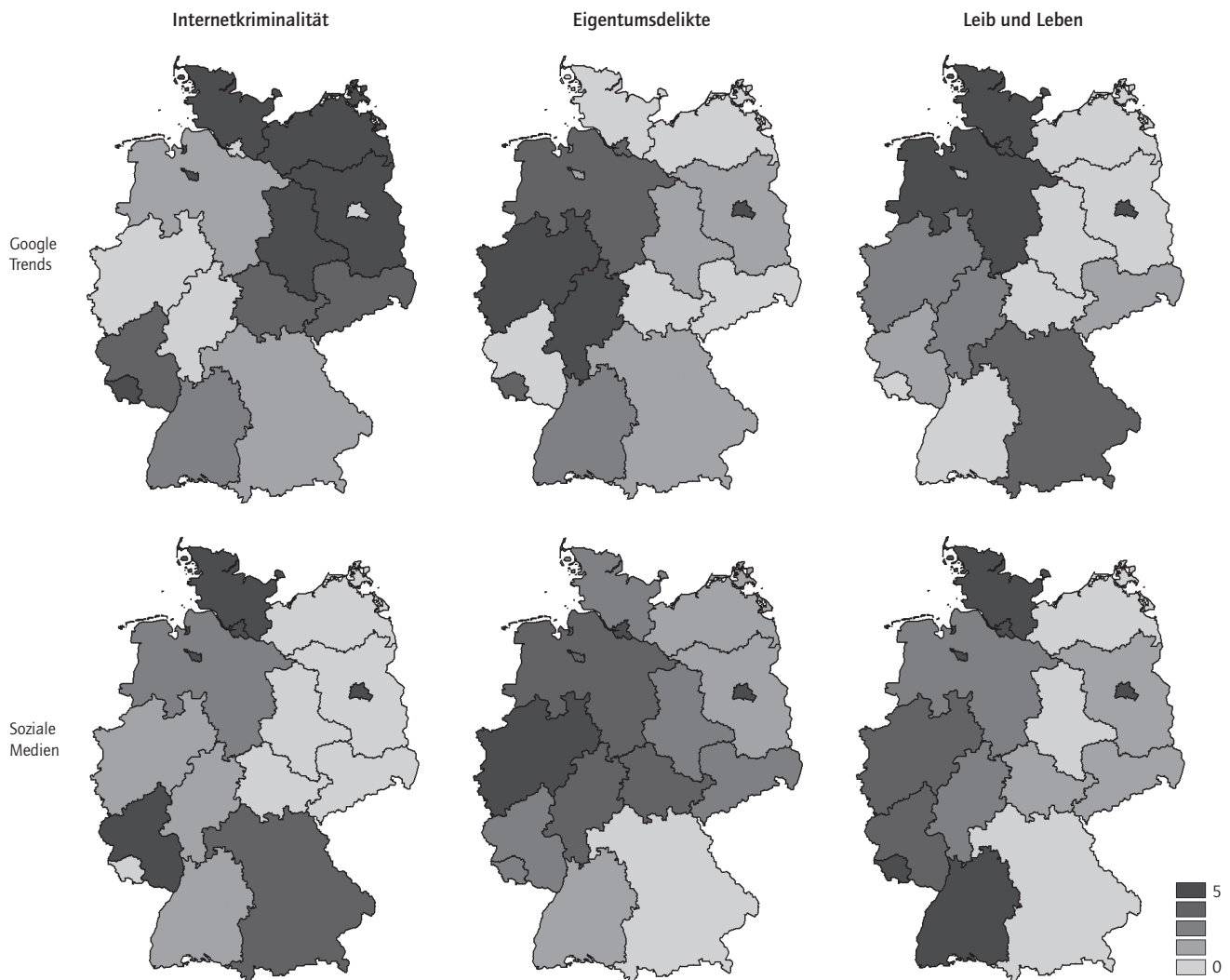
Die relative Häufigkeit der Einträge zu den verschiedenen Delikten variiert je nach Quelle: So überwiegen bei den Raubdelikten Facebook-Einträge deutlich.

sem Maximalwert. Bei Google Trends zeigt sich ein deutliches Überwiegen von Suchanfragen nach dem Schlagwort „Mord“, die nachfolgenden Worte sind abgeschlagen. Die sonstige Rangfolge ist ähnlich, „Diebstahl“, „Einbruch“ und „Körperverletzung“ rangieren bei Google Trends wie auch in den sozialen Medien weit vorne. Offenbar suchten die Nutzer in Deutschland in diesem Zeitraum bei Google kaum nach Suchworten, die im Zusammenhang mit religiösem Fundamentalismus stehen (Islamist, Salafist) – die eindeutigste Abweichung zu den sozialen Medien.

Zum weiteren systematischen Vergleich der beiden Datenquellen wurden drei Bedrohungs-Kategorien gebildet: Die erste Kategorie enthält Delikte aus dem Bereich Internetkriminalität, die zweite bündelt Begriffe aus dem Bereich Eigentumsdelikte und die dritte fasst

Abbildung 2

Räumliche Verteilungen der Suchvolumina in Google Trends (oben) sowie der Einträge in sozialen Medien (unten) in Quintilen



Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trends; Berechnungen des DIW Berlin.

© DIW Berlin 2015

In den neuen Bundesländern wird bei Google häufiger nach Begriffen zu Internetkriminalität gesucht, in den alten Bundesländern werden diese vor allem in den sozialen Medien erwähnt.

Delikte zusammen, die sich gegen Leib und Leben richten. Zwar stellt Google Trends auch Daten unterhalb der Bundeslandebene zur Verfügung, diese sind jedoch lückenhaft und nicht für jeden Landkreis erhältlich. Daher wurden für den Vergleich die lokalisierten Daten aus den sozialen Netzwerken auf Bundeslandebene aggregiert, Google-Trends-Daten wurden für den Zeitraum Juni bis Oktober 2014 abgerufen (Abbildung 2 zeigt einen Vergleich der räumlichen Verteilung).¹⁷

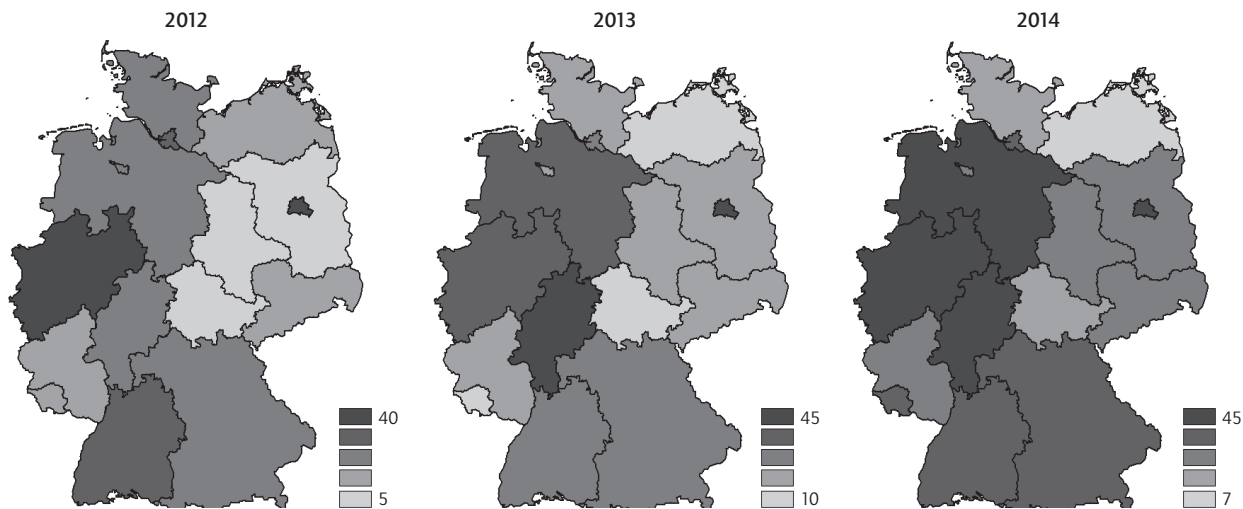
¹⁷ Ein direkter Vergleich der Farbschattierungen zwischen den Karten (Google Trends und soziale Medien) ist nicht möglich, da sich die Farbtöne nach

Frappierend ist der Gegensatz, der für das Deliktfeld Internetkriminalität zwischen den sozialen Medien und Google Trends entsteht. Bei Google sind Suchanfragen aus nord- und ostdeutschen Bundesländern häufiger vertreten, im Verhältnis dazu weisen die süd- und westdeutschen Bundesländer ein geringeres Suchvolumen auf – was auch auf eine Parallele zur Messung von Kriminalitätsfurcht im WISIND-Projekt hindeutet, wo

unterschiedlichen Maximalwerten richten. Somit sind nur Rückschlüsse auf die relativen Häufigkeiten der Begriffsennungen innerhalb der einzelnen Karten aussagekräftig.

Abbildung 3

Abfragen bei Google Trends zu Eigentumsdelikten¹



¹ 2012 bis 2014 in gleich großen Intervallen.

Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trend;. Berechnungen des DIW Berlin.

© DIW Berlin 2015

Die Suchanfragen zum Deliktfeld Eigentumsdelikte stiegen zwischen 2012 und 2014 (von links nach rechts) in den meisten Bundesländern tendenziell an.

im Norden Deutschlands in der Tendenz höhere Furchtwerte festgestellt werden.¹⁸ Bei der expressiven Internetnutzung mit sozialen Medien ist es annähernd umgekehrt. Auch wenn die Werte aus den sozialen Medien mit Vorsicht behandelt werden müssen – nur fünf Prozent der Gesamteinträge konnten für dieses Deliktfeld lokalisiert werden – ist dieser Gegensatz doch für weitere Forschungsarbeiten interessant.

Im Bereich Eigentumsdelikte liegen die Länder Hessen, Nordrhein-Westfalen und Berlin bei Google Trends mit einem hohen Suchvolumen zu Wörtern wie „Einbruch“, „Diebstahl“ und „Alarmanlage“ mit Niedersachsen an der Spitze, in sozialen Medien zeichnet sich ein etwas deutlicheres Nord-Süd-Gefälle bei gleichzeitig geringeren Unterschieden zwischen West- und Ostdeutschland ab.

Bei Delikten, die sich gegen Leib und Leben richten, lassen sich sowohl Gemeinsamkeiten als auch Unterschiede der Datenquellen und Regionen ausmachen. So kommen die Begriffe dieses Deliktfelds in den nordwestlichen Bundesländern (mit Ausnahme Bremens) und in Berlin besonders häufig vor. In Baden-Württemberg,

Bremen und im Saarland wird bei Google deutlich seltener nach entsprechenden Begriffen gesucht als in Bayern; in den sozialen Medien ist es genau umgekehrt.

Vergleich der Daten mit tatsächlicher Kriminalitätsbelastung und Kriminalitätsfurcht

Google-Trends erlaubt über die Auswahl eines Zeitraums den Vergleich mehrerer Jahre. Für alle Bundesländer wurden auf Grundlage der wöchentlichen oder monatlichen Daten Jahresmittelwerte gebildet.¹⁹ Die Suchanfragen zum Deliktfeld Eigentumsdelikte stiegen zwischen 2012 und 2014 in den meisten Bundesländern tendenziell flächendeckend an (Abbildung 3). Mit Ausnahme von Mecklenburg-Vorpommern und Schleswig-Holstein haben die Google-Nutzer in jedem Jahr häufiger nach Suchworten rund um das Thema Eigentumsdelikte gesucht als im jeweiligen Vorjahr,

¹⁹ Gerade für das Suchwort „Einbruch“ wäre die Feststellung der Saisonalität des Suchvolumens ein weiterer interessanter Ansatzpunkt, auf den hier nicht näher eingegangen werden kann. Die Beschränkung der Analyse sozialer Medien auf den Zeitraum Juni bis Oktober 2014 erlaubt keinen verlässlichen Rückschluss auf saisonale Änderungen von Stimmung und Interesse der Internetnutzer.

¹⁸ Siehe dazu auch den Beitrag von Bug, M., Kroh, M., Meier, K. in diesem DIW Wochenbericht.

wobei die Intensität des Anstiegs in ganz Deutschland seit 2014 abgenommen hat.

Die stark begrenzte Eignung der Suchmaschinendaten zur Abbildung der objektiven Kriminalitätsbelastung – nicht der Kriminalitätsfurcht – wird bei einem Vergleich mit den Daten²⁰ aus der PKS deutlich. Die PKS vermeldet für die Berichtsjahre 2012 und 2013 nur einen sehr moderaten Anstieg der Einbrüche und Diebstähle um 0,1 Prozent in Deutschland.²¹ Google Trends scheint auf Basis dessen keine direkten Rückschlüsse auf die tatsächliche Kriminalitätsbelastung in Deutschland zuzulassen. Auf der Bundesländerebene spiegelt Google Trends die Entwicklung im Zeitraum von 2012 bis 2013 aber immerhin elf Mal richtig. Ob sich aus diesen Zahlen jedoch eine robuste Reflektion oder gar Prognose ableiten lässt, ist mehr als fraglich.

Ein entsprechender Entwicklungsvergleich zwischen den Daten aus sozialen Medien und der PKS kann aufgrund des Erhebungszeitraums nicht vorgenommen werden, jedoch zeigen sich auch hier Unterschiede, nämlich räumliche zwischen Eintragungshäufigkeiten und PKS-Fallhäufigkeiten.

Für die Beurteilung der Eignung der hier gezeigten Ergebnisse aus sozialen Medien und Suchmaschinen für die Abbildung der Kriminalitätsfurcht (im Gegensatz zum objektiven Kriminalitätsaufkommen) sollen repräsentativ erhobene Umfragedaten zum Vergleich hinzugezogen werden. Im Rahmen des WISIND-Projekts des DIW Berlin wurden 12 000 Menschen in Deutschland zu ihrer Sorge, Opfer verschiedener Verbrechen zu werden, befragt.

Ein Vergleich mit den in einem weiteren Bericht der vorliegenden Wochenberichts-Ausgabe präsentierten Daten der Messung von Kriminalitätsfurcht²² für das Jahr 2014 zeigt, dass sich einige geographische Muster der zehn Deliktgruppen auch in der regionalen Darstellung des Kriminalitätsindikators subjektiver Furcht wiederfinden (Abbildung 4).²³ Dies gilt insbesondere für Schleswig-Holstein, den nördlichen Teil Niedersachsens und weite Teile Nordrhein-Westfalens sowie Berlin. Auch die Landkreise um Stuttgart²⁴ sowie die Regierungsbezirke Karlsruhe und Freiburg sind als Region in beiden Kar-

²⁰ Diese Aussage gilt unter Annahme eines sich nicht stark verändernden Dunkelfeldes.

²¹ Bundeskriminalamt (2012–2014): Jahrbuch Polizeiliche Kriminalstatistik.

²² Für eine grafische Darstellung des Kriminalitätsindikators subjektiver Furcht siehe Bug, M., Kroh, M., Meier, K. (2015), a. a. O.

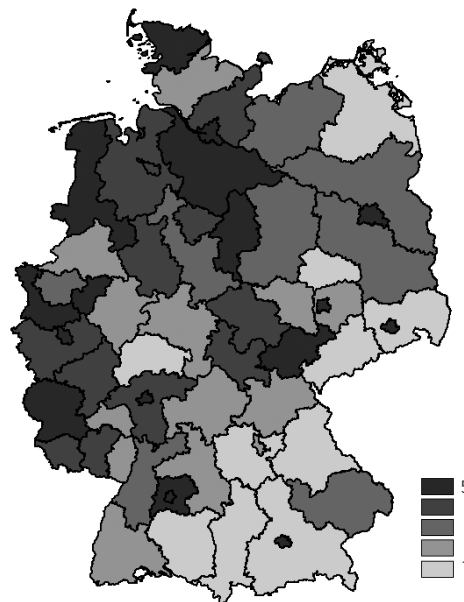
²³ Die Abbildung beruht auf einer Darstellung der regionalen Aktivität in sozialen Medien in Quintilen.

²⁴ Dieser besteht aus den Landkreisen Rems-Murr, Böblingen, Esslingen und Ludwigsburg.

Abbildung 4

Eintragungshäufigkeit¹ in sozialen Medien für alle zehn Delikte auf Regionsebene

Pro 100 000 Einwohner



¹ Juni bis Oktober 2014.

Quellen: Bug, M., Kroh, M., Meier, K., Rieckmann, J., van Um, E., Wald, N. (2015): WISIND-Datensätze: Soziale Netzwerke/Google Trend; Berechnungen des DIW Berlin.

© DIW Berlin 2015

Die räumlichen Verteilungen der Einträge zu den zehn Deliktgruppen zeigen Ähnlichkeiten mit der regionalen Darstellung des Kriminalitätsfurchtindikators.

ten sichtbar und in der Abstufung der Kriminalitätsfurcht und Internetaktivität identisch in beiden Karten.

Andererseits offenbart der Vergleich auch einige Unterschiede auf Regionalebene. Diese könnten zunächst auf die unterschiedlichen Beobachtungszeiträume zurückzuführen sein, zudem wird die schlichte Erwähnung von Themen aus Deliktfeldern der konkreten Frage nach Sorge vor verschiedenen Formen der Kriminalität gegenübergestellt.²⁵ In einigen Fällen sind gegensätzliche Einfärbungen der Regionen ersichtlich. Die Region Trier weist einen vergleichsweise geringen Wert der subjektiven Kriminalitätsfurcht auf, die Menschen dort machen sich also eher weniger Sorgen, Opfer eines Verbrechens zu werden. In den sozialen Medien liegt diese Region jedoch in der Spitzengruppe der Einträge, in

²⁵ Nicht zuletzt kann auch das Mediennutzungsverhalten, beeinflusst durch demographische Faktoren, eine große Rolle bei regionalen Unterschieden spielen.

denen über Kriminalität geschrieben wird. Der gleiche konträre Zusammenhang lässt sich auch in Mittelfranken, der Oberpfalz oder Niederbayern beobachten. Insgesamt reflektieren die aus dem Internet erhobenen Daten offenbar weit besser die Kriminalitätsfurcht als die tatsächliche Kriminalitätsbelastung und können ein guter Anhaltspunkt für regionale Unterschiede der Bevölkerungseinstellung sein.

Die Möglichkeit, die Menge der Daten aus sozialen Medien zu regionalisieren, macht die Daten auch über den bloßen Abgleich mit anderen Zahlenwerken hinaus interessant. Frappierend ist in den Daten die Auffälligkeit der Region Braunschweig-Salzgitter-Wolfsburg, die den eindeutigen Maximalwert der Eintragshäufigkeit zu Begriffen aus dem Bereich des religiösen Fundamentalismus aufweist – und das bereits Monate vor der Medienberichterstattung über Festnahmen junger Islamisten in Wolfsburg und über den in Braunschweig wegen eines unmittelbar drohenden islamistisch motivierten Anschlags kürzlich abgesagten Karnevalsanzug.²⁶

Fazit

Die Ergebnisse der Auswertung der Daten aus sozialen Medien und Suchmaschinen zeigen, dass diese aufgrund ihrer Natur und auch aus systematischen Gründen nicht geeignet sind, ein verlässliches Bild der

tatsächlichen regionalen Kriminalitätsbelastung zu reflektieren. Sie eignen sich aber durchaus, um ein Bild des subjektiven Empfindens der Bevölkerung von Kriminalität zu zeichnen und damit als kostengünstige Datenquelle klassische Befragungsformen zu Kriminalitätsfurcht zu flankieren. Die Analyse der Beiträge in sozialen Medien zeigt drei Beschränkungen für eine Interpretation der Daten als Hinweis auf die tatsächliche Bedrohungslage: Erstens werden die sozialen Medien stark durch Medieneffekte und Aufmerksamkeitswellen geprägt, typischerweise manifestiert sich dies im Weiterleiten von Medienauszügen oder Beiträgen anderer Nutzer (auch: *retweets*). Zweitens ist die Lokalisierung der Daten nicht einfach. Lokalisierbar sind die Daten nur, wenn Nutzer diese Informationen preisgeben.²⁷ Diese Bereitschaft ist jedoch auch abhängig vom Kontext der Beiträge, sodass manche Begriffe zu bestimmten Delikten sich besser lokalisieren lassen als andere. Drittens ist auch die Wahl der Internetplattform entscheidend, da sich der Medienmix sehr deutlich zwischen den verschiedenen Delikten unterscheidet. Vergleiche mit der tatsächlichen Kriminalitätsbelastung werden zusätzlich durch die Stichprobenverzerrung zugunsten junger Nutzer beschränkt, da diese in sozialen Medien deutlich aktiver sind.

²⁶ Die Karte zur Nennungshäufigkeit von Begriffen im Deliktfeld religiöser Fundamentalismus ist in diesem Wochenbericht nicht abgebildet.

²⁷ Auch von der offiziellen Schreibweise abweichende Ortsbezeichnungen stellen eine Herausforderung dar, die Zuordnung ist zeitintensiv und nur begrenzt automatisierbar.

Johannes Rieckmann ist Wissenschaftlicher Mitarbeiter in der Abteilung Entwicklung und Sicherheit am DIW Berlin | jriemann@diw.de

Jan-Lucas Schanze ist studentische Hilfskraft in der Abteilung Entwicklung und Sicherheit am DIW Berlin | jschanze@diw.de

PERCEPTIONS OF PERSONAL SECURITY IN SOCIAL MEDIA AND SEARCH ENGINES – A REALISTIC REFLECTION OF ACTUAL CRIME RATES?

Abstract: The most common method of measuring subjective fear of crime in the general population has traditionally been through surveys. With the spread of digital technologies, however, data from social media and search engines could now help researchers learn more about people's subjective perceptions of certain types

of crime. The present article will show that although the analysis of data from social media and search engines is not suitable as an indicator of actual crime levels in Germany, it can certainly be a cost-effective supplement to traditional methods of collecting data on perceived crime levels.

JEL: H79, K42, R12, Z18

Keywords: crime, crime statistics, indicator, inner security, perception, police statistics, security, social media, social network, search engine



DIW Berlin – Deutsches Institut
für Wirtschaftsforschung e.V.
Mohrenstraße 58, 10117 Berlin
T +49 30 897 89 -0
F +49 30 897 89 -200
82. Jahrgang

Herausgeber

Prof. Dr. Pio Baake
Prof. Dr. Tomaso Duso
Dr. Ferdinand Fichtner
Prof. Marcel Fratzscher, Ph.D.
Prof. Dr. Peter Haan
Prof. Dr. Claudia Kemfert
Dr. Kati Krähnert
Prof. Karsten Neuhoff, Ph.D.
Prof. Dr. Jürgen Schupp
Prof. Dr. C. Katharina Spieß
Prof. Dr. Gert G. Wagner

Chefredaktion

Sabine Fiedler
Dr. Kurt Geppert

Redaktion

Renate Bogdanovic
Andreas Harasser
Sebastian Kollmann
Dr. Claudia Lambert
Marie Kristin Marten
Dr. Wolf-Peter Schill

Lektorat

Alexander Eickelpasch
Dr. Anika Rasner
Dr. David Richter
Dr. Dirk Ulbricht
Dr. Lilo Wagne

Pressestelle

Renate Bogdanovic
Tel. +49-30-89789-249
presse@diw.de

Vertrieb

DIW Berlin Leserservice
Postfach 74
77649 Offenburg
leserservice@diw.de
Tel. (01806) 14 00 50 25
20 Cent pro Anruf
ISSN 0012-1304

Gestaltung

Edenspiekermann

Satz

eScriptum GmbH & Co KG, Berlin

Druck

USE gGmbH, Berlin

Nachdruck und sonstige Verbreitung –
auch auszugsweise – nur mit Quellen-
angabe und unter Zusendung eines
Belegexemplars an die Serviceabteilung
Kommunikation des DIW Berlin
(kundenservice@diw.de) zulässig.

Gedruckt auf 100 % Recyclingpapier.