Rainer Knyrim
Rechtsanwalt und Partner bei Preslmayr Rechtsanwälte

Wie Big Data Katastrophenhilfe unterstützt

Interview mit Mila Romanoff, Rechts- und Datenschutzspezialistin für „Global Pulse", die Big Data-Initiative des UN-Generalsekretärs. Liudmyla (Mila) Romanoff ist für die Erschaffung nachhaltiger Mechanismen für öffentlich-private Partnerschaften sowie für die verantwortungsbewusste Verarbeitung von Big Data im Bereich der globalen Entwicklung und humanitären Einsätze verantwortlich.

Datenschutz konkret: Können Sie erklären, was UN Global Pulse ist?
Mila Romanoff: UN Global Pulse ist eine Innovationsinitiative, die innerhalb des Vorstandsbüros des UN-Generalsekretärs entstanden ist. Wir werden als „Labor" („Lab") organisiert, das am verantwortungsbewussten Einsatz von Big Data zum Nutzen der Öffentlichkeit, der Entwicklungshilfe und humanitärer Fälle arbeitet. Aber von Anfang an: UN Global Pulse wurde 2009 als Reaktion auf die damalige Weltwirtschaftskrise entwickelt. Die Weltwirtschaftskrise ließ klar erkennen, dass der öffentliche Sektor mehr Echtzeitinformationen benötigt, um wirtschaftliche Herausforderungen oder Themen der Lebensmittelsicherheit, denen Bevölkerungsgruppen insbesondere in Entwicklungsländern entgegensehen, zu verstehen.

Datenschutz konkret: Wie funktioniert UN Global Pulse?
Romanoff: UN Global Pulse ist eine Innovationswerkstatt. Wir erforschen die Nützlichkeit von Big Data zum Zweck der Reaktion auf Entwicklung und humanitäre Themen. Wir erforschen Möglichkeiten, Echtzeit-Big Data aufzubauen, um so humanitären Organisationen zu helfen, zB das Planen nach einer Naturkatastrophe zu verbessern – was letztendlich Leben retten kann. Wir haben beispielsweise erforscht, wie anonymisierte mobile Daten oder öffentliche Social Media (wie zB öffentliche Tweets) verwendet werden können, um zu einem besseren Verständnis von Konversationsthemen in betroffenen Katastrophengebieten zu gelangen, also wie die Menschen dort über Preisschwankungen von Grundnahrungsmitteln oder Arbeitslosigkeit reden.

Datenschutz konkret: Wie ist UN Global Pulse strukturiert?
Romanofff: UN Global Pulse hat seinen Sitz in der UN-Zentrale in New York sowie zwei weitere Labs auf Landesebene in Kampala (Uganda) und Jakarta (Indonesien). Diese Labs fungieren grundsätzlich als regionale Drehscheiben für Asien und Afrika. Wir führen auch Projekte in anderen Ländern durch, bei denen wir uns auf Entwicklungsländer konzentrieren. Unsere Teams bestehen aus politischen Experten, Forschungskoordinatoren, Kommunikationsexperten sowie Rechts- und Datenschutzexperten. In den Labs arbeiten auch mehrere Datenwissenschaftler, Datenanalytiker und Dateningenieure; das sind technische Professionisten, die genau wissen, wie mit spezifischen Daten umzugehen ist. Sie erstellen Taxonomien und entnehmen Schlagwörter, welche in verschiedene Analysewerkzeuge eingegeben werden, um verschiedene Datenarten zu verarbeiten und auszuwerten.

Datenschutz konkret: Was ist Ihre Funktion bei UN Global Pulse?
Romanoff: Ich bin die Rechts- und Datenschutzverantwortliche im New Yorker Lab. Meine Aufgabengebiete umfassen einerseits die Kenntnis darüber, wie die Daten, zu denen wir Zugang erhalten, verarbeitet werden; andererseits sorge ich dafür, dass diese Daten verantwortungsbewusst und in Übereinstimmung mit den vertraglichen Verpflichtungen, unter denen wir den Datenzugang erhalten haben, verarbeitet werden. Das schließt auch Forschung und Entwicklung von Datenschutzempfehlungen, Richtlinien und „Best Practice" bei der Verwendung von Big Data in Entwicklungsprojekten ein.

Die Aufzeichnungen über die Handynutzung konnten nachweisen, dass die Bevölkerung nicht auf die Hochwasserwarnungen reagiert hatte.

Datenschutz konkret: Es gab ein Projekt über Hochwasser in Mexico. Was war das Ziel dieses Projekts?
Romanoff: Das Ziel dieses Projekts war es, zu erforschen, ob Aufzeichnungen anonymisierter Handydatennutzung oder Einzelverbindungsnachweise zeigen könnten, wie Bevölkerungsgruppen ihre Verhaltensmuster in einer Katastrophensituation verändern. Ziel war es, zu beweisen, dass Detailaufzeichnungen anonymisierter Einzelverbindungsnachweise eine wertvolle Informationsquelle für Katastropheneinsätze sein könnten. Die mexikanische Regierung hat uns Hochwasserinformationen, Unterlagen über Niederschlag und Aufzeichnungen über Bevölkerungsschutz bereitgestellt, die mit Aufzeichnungen anonymisierter Handynutzung verglichen und analysiert wurden.

Datenschutz konkret: Und was war das Ergebnis?
Romanoff: Tatsächlich ist es so, dass Handydaten – wenn anonymisiert und bis zu einem gewissen Umfang aggregiert – zeigen können, dass eine Abnormalität im Gange ist. Insbesondere kann man mit Hilfe der Daten sehen, wohin sich Menschen nach den Überschwemmungen bewegt haben und wie lange es dauert, bis diese wieder in die „Normalität" zurückkehren. So eine Kenntnis könnte humanitären Organisationen helfen, rechtzeitig zu agieren und die betroffenen Gemeinden zu unterstützen.

Eine weitere Schlussfolgerung aus diesem Projekt war, dass die von der Regierung einige Tage zuvor ausgesprochenen Hochwasserwarnungen die Bevölkerung nicht in Bewegung gebracht hatten. Das bedeutet, dass die Menschen nicht auf die Warnhinweise reagiert haben. Als dann aber das Hochwasser eintrat, begannen die Menschen plötzlich, ihre Verhaltensmuster zu verändern. Ein solcher Einblick kann der Regierung helfen, humanitären Helfern und Noteinsatzteams zu zeigen, wie effektiv ihre Katastrophenwarnungen sind.

Datenschutz konkret: Können Sie erklären, was ein Einzelverbindungsnachweis ist?
Romanoff: Natürlich, ein Einzelverbindungsnachweis ist ein digitaler Datensatz eines Signals, das von einem Sendemasten zu einem anderen übermittelt wird und anzeigt, dass ein Anruf oder eine SMS gesendet wurde; dieser beinhaltet aber nicht den Inhalt eines Anrufs oder einer SMS. Der Einzelgesprächsnachweis wird in einem Projekt lediglich in anonymisierter Form verwendet, die Person, die den Anruf tätigt, wird nicht identifiziert. Mitunter beinhaltet dieser Einzelgesprächsnachweis die Zeit, die Dauer und den Standort des Sendemasts des Anrufs. Man kann es als anonymisierte Meta-Daten sehen.

Datenschutz konkret: Beinhaltet der Einzelverbindungsnachweis auch die Handynummer?
Romanoff: Das Telekommunikationsunternehmen kennt natürlich die Telefonnummer. Bei Forschungsprojekten und Initiativen wie unserer erhalten wir aber nie Zugang zu Telefonnummern, und in diesem speziellen Projekt hatten wir nicht einmal Zugang zu dem Datenmaterial. In diesem Projekt wurden die Forschungsarbeit und Analyse der Daten nur von Forschern innerhalb der Firewall des Unternehmens durchgeführt.

Datenschutz konkret: War es also das Telekommunikationsunternehmen selbst, das die Forschung durchführte?
Romanoff: Es handelte sich dabei um Wissenschaftler, die individuelle Geheimhaltungserklärungen mit dem Telekommunikationsunternehmen unterzeichnet und die Datenanalyse hinter der Firewall des Unternehmens durchgeführt hatten.

Datenschutz konkret: Können Sie auch etwas über ein Social-Media-Projekt erzählen?
Romanoff: In Indonesien wurde beispielsweise ein Projekt mit Social Media durchgeführt. Wir haben hier mit dem UN-Welternährungsprogramm und dem indonesischen Ministerium für Entwicklung zusammengearbeitet. Dieses Projekt bezog sich auf Nahrungsmittel und Landwirtschaft. Wir haben Schlagwörter aus öffentlichen Tweets analysiert, um Veränderungen bei Lebensmittelpreisen zu erkennen. Die Idee hinter diesem Projekt war es, zu sehen, ob öffentlich zugängliche Social Media als Echtzeitindikatoren für offizielle Lebensmittelpreise gelten können. Kurz, das Projekt beinhaltete die Erstellung eines statistischen Modells für einen täglichen Preisindikator für vier Lebensmittel: Rind, Huhn, Zwiebeln und Chili. Die Preise wurden öffentlichen Tweets entnommen und mit den offiziellen Lebensmittelpreisen, die das Ministerium für Entwicklung zur Verfügung stellte, verglichen. Das Projekt bewies, dass die Daten aus öffentlich zugänglichen Social Media bezüglich der Lebensmittelpreise tatsächlich eng mit jenen der offiziellen Preisstatistik korrelierten. Eine meiner Meinung nach wichtige Erkenntnis hierbei war, dass das Zurverfügungstellen derartiger Echtzeitinformationen bei politischen Entscheidungsfindungen, in Bezug auf Lebensmittelsicherheit und sogar andere wirtschaftliche Belange (besonders in Entwicklungsländern) zu unterstützen vermag.

Datenschutz konkret: Nachdem wir bereits zwei Projekte besprochen haben, würde ich nun gerne die datenschutzrechtlichen Herausforderungen in Bezug auf Big Data ansprechen. Es gibt Leute, die etwa fürchten, dass ihre Daten trotz Anonymisierung rückverfolgt werden können, wenn ihre Daten in Big Data-Studien verwendet werden.
Romanoff: Zunächst möchte ich die Grundstruktur der Arbeit von Global Pulse und wie Datenschutz sich in diese Arbeit einfügt, erklären. Bei Global Pulse haben wir Methoden und Richtlinien entwickelt, um die besten Techniken zu finden, die in unseren Labs implementiert werden können, wenn wir Big Data-Projekte durchführen.

Am Ende werden wir eine Reihe von Datenschutzempfehlungen oder Richtlinien für Interessensgruppen des öffentlichen und privaten Sektors und Wissenschaftler erstellen, welche aufgenommen oder in zukünftigen Studien verwendet werden können, bzw tatsächlich als Ausgangspunkt, auf dem auf- und ausgebaut werden kann. Auch für die Arbeit innerhalb der Labs führen wir ein Privacy Impact Assessment vor jedem neuen Projekt durch und haben hierfür bestimmte Geheimhaltungs- und Datenschutzprinzipien.

Datenschutz konkret: Können Sie einige davon nennen?
Romnaoff: Ja, sie sind in einer Kurzfassung auf unserer Webseite zu finden. Wir haben auch interne Richtlinien, die entsprechend detaillierter sind. Ein wichtiger Grundsatz unserer Arbeit ist, dass wir auf keine persönlichen Daten zugreifen. Falls wir doch auf derartige Daten zugreifen, muss dies mittels informierter Zustimmung erfolgen. Wenn wir beispielsweise über Twitter-Daten reden, gibt es dort natürlich eine Fülle an personenbezogenen Informationen und Daten. Es sind jedoch alles öffentliche Daten und durch unsere Kooperation bekommen wir infolge einer Serie von Due-Diligences Zugang zu den Daten. Aufgrund vertraglicher Verpflichtung mit unseren Datenüberlassern stellen wir sicher, dass die Daten rechtmäßig und angemessen gesammelt und verarbeitet werden. Wir haben uns dazu verpflichtet – und ich halte dies für sehr wichtig –, nie zu versuchen, bereits anonymisierte Daten zu reidentifizieren.

Tweets, die „privat“ gesetzt oder gelöscht wurden, werden in unseren Projekten nicht verwendet.

Datenschutz konkret: Aber wer führt die Anonymisierung durch? Jene Stelle, die die Daten an UN Global Pulse übermittelt, oder gibt es noch jemanden dazwischen, der dies macht?
Romanoff: Bei Twitter-Daten erhalten wir beispielsweise Zugang zu öffentlichen Tweets. Normalerweise versucht niemand, den Personenzug dieser Daten herzustellen. Bei öffentlichen Tweets kann es jedoch passieren, dass wir darüber benachrichtigt werden, dass ein bestimmter Tweet „privat“ gesetzt oder gelöscht wurde. Diese Veränderung muss selbstverständlich respektiert werden, was auch durch die Kernpolitik des Datenüberlassers festgelegt ist. Wir müssen dann die Daten vernichten und dürfen diese in keinem Projekt verwenden.

Bei Einzelgesprächsnachweisen erhalten wir für gewöhnlich keinen Zugang zu den Daten. Aktuell gibt es keine allgemein gültige Strategie, die behauptet: „Das ist die einzig richtige Art zu anonymisieren.“ Es gibt allerdings einige Strategien, wie ein Forschungsprojekt mit anonymisierten Daten funktionieren kann. Die Daten können durch unsere Partner anonymisiert werden, etwa wenn es sich um ein Telekommunikationsunternehmen handelt – diese haben mitunter ihre eigenen Methoden, Daten zu anonymisieren. Diese Daten

können dann in anonymisierter Form hinter der Firewall des Unternehmens bleiben und würden in diesem Fall die Firmenräumlichkeiten nicht verlassen. Das ist dann der Fall, wenn das Forschungsprojekt durch Mitarbeiter dieses Unternehmens ausgeführt wird oder etwa durch Forscher einer Universität, die einer Geheimhaltungserklärung unterliegen. Falls die Daten das Firmengelände des Telekommunikationsunternehmens verlassen, sollten diese in aggregierter Form übermittelt werden, um sicherzustellen, dass das Risiko einer Reidentifikation auf ein Minimum reduziert wird.

Datenschutz konkret: Wie ist die datenschutzrechtliche Arbeitsweise strukturiert?
Romanoff: Wir haben kürzlich eine Datenschutz-Beratergruppe gegründet, die aus 25 Experten aus aller Welt besteht.

Rainer Knyrim im Interview mit Mila Romanoff in Washington

Datenschutz konkret: Sind darunter auch Juristen?
Romanoff: Der eigentliche Grund, warum wir diese Gruppe gegründet haben, war der, dass wir nicht nur Rechts- und Datenschutzexperten, sondern auch Ingenieure, politische Entscheidungsträger, Experten des Entwicklungssektors, Behörden und Forscher miteinbeziehen wollten. Wir haben Anwälte, reine Datenschutzexperten, Wissenschaftler, Experten bezüglich Datenidentifikation und Datentechniker in dieser Datenschutz-Beratungsgruppe versammelt. Wir versuchen, diese Gruppe nicht nur in geografischer Hinsicht, sondern auch in Bezug auf beruflichen Hintergrund und Erfahrung ausgewogen zu gestalten.

Datenschutz konkret: Gibt es bereits Ergebnisse dieser Gruppe, etwa in Bezug auf Ethik und Big Data, die sich auch abseits der rechtlichen Aspekte zu einem großen Thema entwickeln?
Romanoff: Die Gruppe hat zwei Funktionen: Einerseits berät sie uns in Bezug auf unsere internen Prozesse, Projekte und Abläufe, die wir entwickelt haben, andererseits fungiert sie auch als Anwalt für einen verantwortungsbewussten Umgang mit Big Data für Entwicklungs- und humanitäre Zwecke. Die Gruppe hat uns dabei geholfen, ein Big Data „Privacy Impact Assessment" zu entwickeln, das zeigen soll, ob die Daten in einer legitimen und angemessenen Art und Weise verwendet werden.

Datenschutz konkret: Sie haben also einen Rahmen für ein derartiges „Privacy Impact Assessment" erstellt?
Romanoff: Ja. Unser „Privacy Impact Assessment" (kurz „PIA") beschäftigt sich nicht nur mit personenbezogenen Daten, sondern berücksichtigt auch jene Daten, die in pseudoanomysierter oder aggregierter Form vorliegen.

Datenschutz konkret: Ich könnte mir vorstellen, dass all diese Daten und die Art, wie sie verwendet werden, auch missbraucht werden könnten. Ich denke beispielsweise daran, was passiert, wenn in einem Land Bürgerkrieg herrscht und eine Konfliktpartei diese Techniken anwendet, um herauszufinden, wo sich der Gegner gerade aufhält oder wohin die Menschen im Fall eines Angriffs flüchten, um dort erneut anzugreifen. Die Konfliktparteien könnten auch die Tweets analysieren, um an geheime Informationen zu gelangen oder einfach die Kommunikation „abzuhören".
Romanoff: Das ist genau der Grund, warum wir dieses Privacy Impact Assessment entwickelt haben, denn es zeigt nicht nur Risiken und Gefahren auf, sondern auch mögliche Menschenrechtsverletzungen, es berücksichtigt eben den Kontext. Daher kann ein Projekt in einem Land durchgeführt werden – sogar unter Miteinbeziehung öffentlicher Tweets –, um die Gespräche bezüglich HIV zu erforschen, während dieselbe Projektform in einem anderen Land mit einem anderen politischen Klima – selbst mit öffentlichen Tweets – nicht durchgeführt werden sollte. Wir müssen daher den kulturellen Hintergrund und religiöse Überzeugungen berücksichtigen, aber auch die allgemeine politische Stabilität bzw die politische Lage in dem jeweiligen Land, um festzustellen, ob das Projekt mit öffentlichen Informationen fortgeführt werden kann. Das ist also eine sehr gute Frage und tatsächlich ein sehr großer Teil des PIA.

Das Privacy Impact Assessment von Global Pulse wird demnächst veröffentlicht.

Datenschutz konkret: Ist dieses PIA bereits veröffentlicht oder wird es demnächst veröffentlicht?
Romanoff: Ja, wir haben es zuerst intern bei UN Global Pulse entwickelt, jetzt warten wir auf das Feedback unserer Datenschutz-Beratergruppe. Sobald dieses Feedback eingearbeitet ist, wird es in unseren Labs überprüft und der Öffentlichkeit zur Kommentierung zur Verfügung gestellt, bevor es tatsächlich als fertiges Werk veröffentlicht wird. Also ja, das fertige Werk wird veröffentlicht, so ist es vorgesehen.

Datenschutz konkret: Wie sieht also der Zeitplan aus? Wird es noch in diesem Jahr veröffentlicht?
Romanoff: Wir haben gerade die erste Phase des Feedbacks im März abgeschlossen. Wir haben ein fantastisches Feedback von unseren Experten bekommen, im nächsten Schritt werden wir die Empfehlungen einarbeiten. Es ist dann geplant, das PIA heuer zu veröffentlichen.

Dako 2015/32

Zum Thema

Über die Interviewpartnerin

Mila Romanoff war vor ihrer Tätigkeit bei „Global Pulse" als unabhängige Beraterin zweier ständiger Missionen für die Vereinten Nationen und für das Justizministerium der Ukraine tätig. Bevor sie sich bei den Vereinten Nationen beteiligte, arbeitete sie für Anwaltskanzleien in New York und der Ukraine, wobei sie sich auf Wirtschaftsrecht und Prozessführung spezialisierte. Als Datenschutzspezialistin und Menschenrechtsanwältin war sie in Militärkommissionsverfahren der US-Marinebasis in Guantanamo Bay, Kuba, involviert, wo sie sich auf die Themen Anwaltsgeheimnis, Verschwiegenheitspflicht, Datensicherheit und Privatsphäre/Datenschutz im Bereich der Kommunikation konzentrierte. Sie ist zugelassene Rechtsanwältin in New York. E-Mail: Romanoff@unglobalpulse.org

Global Pulse

- Harnessing big data for development and humanitarian action: www.unglobalpulse.org
- Jahresbericht 2014: www.unglobalpulse.org/2014-Annual-Report

Links

- Big data innovation as part of a Data Revolution in Africa: www.unglobalpulse.org/blog/african-data-consensus-and-big-data
- WFP And UN Global Pulse Show How Big Data Can Save Lives And Fight Hunger: www.unglobalpulse.org/WFP-GlobalPulse-Mobile-Data-Food

Klaus Steinmaurer
General Counsel T-Mobile Austria GmbH

Big Data @ Telekommunikationsnetzbetreiber

Ist Big Data für TK-Anbieter tabu? Der Beitrag definiert den unspezifischen Begriff von Big Data und leitet davon eine Hypothese ab, welche Bedeutung sich daraus für das Geschäftsmodell eines Telekommunikationsnetzbetreibers ergibt. Ein praktischer Anwendungsfall zeigt das rechtlich zulässige Potential von Big Data.

Einleitung

Big Data ist ein Begriff, der zuerst einmal negative Assoziationen weckt. Bei „Big" in Zusammenhang mit „Data", da ist die Brücke zu „Big Brother" nicht mehr weit. Zusammen mit Medienberichten zu NSA-Skandalen, Hackerangriffen auf große Unternehmen und öffentliche Institutionen werden solche unangenehmen Gefühle natürlich verstärkt.

Und dann ist da noch das „Big Business". Keine Fachtagung zum Thema, kein Artikel oder Buch dazu, wo nicht landauf, landab gepredigt wird, dass **Daten das neue Öl oder Gold der Zukunft** sind. Daten sind also dazu da, um Geschäfte zu machen.

Aber so einfach ist die Sache nicht. Zuerst einmal ist es notwendig, sich auf eine **klare Definition** von „Big Data" zu verständigen, insb dann, wenn es darum geht, dieses Thema in Relation zur Tätigkeit eines Netzbetreibers im Telekommunikationsbereich zu stellen. Der Beitrag definiert Big Data und stellt davon abgeleitet eine Hypothese auf, welche Bedeutung sich daraus für das Geschäftsmodell eines Telekommunikationsnetzbetreibers ergibt.

Dabei sind die Daten, die sich aus dem Betrieb eines Telekommunikationsnetzes ergeben, und der Zweck, zu dem sie dem Netzbetreiber vom Konsumenten zur Verfügung gestellt werden, näher zu betrachten. Hier könnte sich aus rechtlicher Sicht eine **differenzierte Betrachtung dieser Daten** und damit deren kommerzieller Verwertbarkeit ableiten lassen.

Big Data ist der Stein der Weisen, um aus etwas, das für sich allein nutzlos ist, Gold (bzw Geld) zu machen.

Der aus dem Kernzweck ableitbare Zusatznutzen, der sich aus der Verwendung von Telekommunikationsdaten für neue Dienste und Services ergeben kann, das **mögliche neue Geschäftsmodell** also, ist dann aus rechtlicher Sicht anhand der spezifischen telekommunikationsrechtlichen Bestimmungen zu überprüfen. Zusätzlich sind bei der Betrachtung neben branchenspezifischen Besonderheiten im Datenschutz auch Unterschiede in der vertraglichen Ausgestaltung des Kundenverhältnisses bei Telekommunikationsdienstleistungsverträgen gegenüber anderen Dienstleistungsverträgen im Internet zu beachten. Va aber stellt sich die Frage, ob Big Data-Anwendungen im Telekommunikationsbereich überhaupt mit den einschlägigen rechtlichen Vorgaben, insb § 96 iVm § 99 TKG vereinbar sind. Bestehen **Besonderheiten dieser Telekommunikationsdaten** gegenüber sonstigen Daten, die aus anderen Quellen stammen? Welche Bedeutung kommt der Anonymisierung, sowohl bei der datenschutzrechtlichen, aber auch bei der kommerziellen Einordnung, dabei zu? Wenn sich nämlich die erste Frage aus Unternehmenssicht zufriedenstellend beantworten lässt, ist immer noch zu überlegen, ob es nicht notwendig ist, die kommerzielle Bewertung differenziert zu betrachten.

Was ist Big Data?

Wie schon vorab zum Ausdruck gebracht, ist der **Begriff sehr unspezifisch**: Ein Alles-oder-Nichts, das heute auf fast alles angewendet wird, das mit größeren Datenmengen zu tun hat. Da ist schon die erste Unschärfe; wann spricht man von einer größeren Datenmenge? Meinen wir Giga-, Tera-, Peta- oder Zetabyte? Was tun wir eigentlich, wenn uns die griechischen Buchstaben ausgehen?

„Big" im Sinne von Big Data ist relativ. Was heute darunterfällt, ist morgen ganz normaler Standard für eine Heimanwendung.

Die reine Datenmenge hat also nur eine beschränkte Relevanz für die Beurteilung, ob es sich um eine „Big Data"-Anwendung handelt oder nicht. Da die Datenmenge aber Auswirkungen auf die Hardwaregröße hat, kommt es oft zu der fälschlichen Antwort, dass mit Big Data die Analyse großer strukturierter Datenmengen gemeint sein kann.