

Flucht in Daten und Zahlen

Die Themen Flucht und Asyl sind ohnehin Themengebiete, um die sich zahlreiche Fragen drehen. Warum verlassen Menschen ihre Heimat? Wie lange sind sie unterwegs? Wer schafft es bis nach Deutschland? Was haben die hier Angekommenen für Aussichten? Wie geht es weiter für jene, die bleiben können, und wie für die, die abgeschoben werden? Wie viele werden überhaupt abgeschoben? Nur ein Bruchteil dieser Fragen lässt sich darstellen oder wird in Form von Daten erhoben. Das selbst gesetzte Ziel der Gruppe *Refugee Datathon Munich* ist es, die vorhandenen und zugänglichen Daten nutzbar zu machen und die, die nicht vorhanden sind, in Erfahrung zu bringen. Ein Bericht zur Entstehung und dem Vorgehen des *Refugee Datathon Munich*. Von Suny Kim



Der *Refugee Datathon Munich* entstand Ende 2015. Im Angesicht des immer wichtiger werdenden Flucht-Themas, wollten Menschen aus dem Fachbereich der IT etwas beitragen. Da lag die Überprüfung von zugänglichen Daten nahe. Zur Illustration unseres Vorgehens eignen sich die Anerkennungsquoten des Bundesamts für Migration und Flüchtlinge (BAMF). Eine realistische Situation sah im Jahr 2015 folgendermaßen aus: Eine Unterstützer*in braucht die Quoten für eine Präsentation. Sie kennt jemandem im BAMF und fragt per E-Mail nach den

Quoten. Im besten Fall erhält sie daraufhin eine Antwort, die ein paar PDFs enthält. Daraufhin überträgt die Unterstützer*in die Zahlen in Excel und fertigt ein Diagramm an, das dann in einer Präsentation gezeigt wird.

Freier Zugang zu Informationen
als zentraler Grundsatz

Das ist jedes Mal aufs Neue ein erheblicher Aufwand und das Ergebnis erreicht die Wenigsten. Dabei hätten diese Daten für so viele Menschen Relevanz, zum Beispiel die Unterstützer*in, die sich fragt, wie die Chancen für Geflüchtete aus dem Jemen stehen oder die Aktivist*in, die sich Sorgen macht, weil viele aus Syrien Kommende nur noch subsidiären Schutz bekommen und diese Beispiele kratzen nur an der Oberfläche dessen, wo solche Daten wichtig sind. Deshalb geht der *Refugee Datathon Munich* anders vor: Wir fragen mit dem Verweis auf das Informationsfreiheitsgesetz beim BAMF nach den Daten. Dafür benutzen wir die Open Data Plattform *fragenstaat.de*. Alle Daten, die wir erhalten, werden dann ins Netz gestellt. In der Regel geht das BAMF erst nach der dritten Anfrage dazu über, die monatlichen Asylgeschäftsstatistiken regelmäßig auf der Homepage zu veröffentlichen. Wir ziehen die Zahlen dann automatisiert aus dem PDF, so erzeugen wir einfache und maschinenlesbare Files (CSV, comma separated values), welche wir auf *GitHub* hochladen. *GitHub* ist

Suny Kim
 arbeitet im
 Fachbereich IT und
 ist Teil des Kollektives
 Refugee
 Datathon Munich

eine Plattform für Open Source Software, auf die jede*r Zugriff hat und die die freie Weiterverarbeitung der Inhalte ermöglicht. Wir verarbeiten diese Daten selbst weiter und machen daraus Histogramme, die wir auf refugee-datathon-muc.org zur Verfügung stellen. Auf diese Weise können alle Interessierten an die Zahlen kommen. Unser Interesse ist es, dass sich dieses Wissen verbreitet, was auch der Beweggrund zum Verfassen und die große Hoffnung hinter diesem Artikel ist.

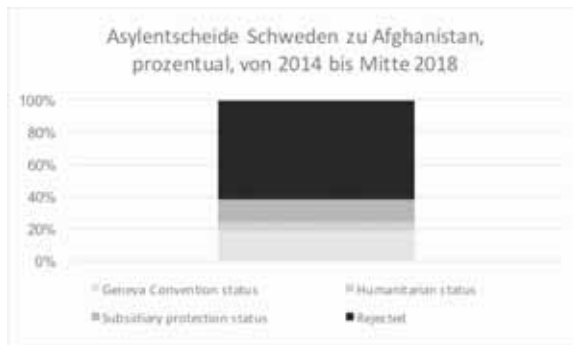
Das Problem mit Leerstellen

Grundlegende Fragen, die sich stellen, sind die danach, welche Daten es überhaupt gibt und wer diese hat. Zu manchen Themen gibt es naturgemäß so gut wie keine Informationen. Über Menschen ohne Aufenthaltsstatus oder die Ausbildung Geflüchteter. Zu diesen Themen gibt es keine systematischen Erhebungen, allenfalls gibt es Stichproben. Daten zu erheben ist eine Wissenschaft für sich und übersteigt unsere Möglichkeiten bei Weitem. Diesem Problem versuchen wir beizukommen, indem wir möglichst direkt an die Quellen herantreten. Vor allem im Umgang mit Daten muss der Schutz der Persönlichkeit beachtet werden. Es wäre durchaus interessant, wie viele Geflüchtete in der Psychiatrie behandelt werden, und vielleicht gibt es im Klinikum auch Daten, aber eine Anfrage hätte aus guten Gründen wenig Aussicht auf Erfolg.

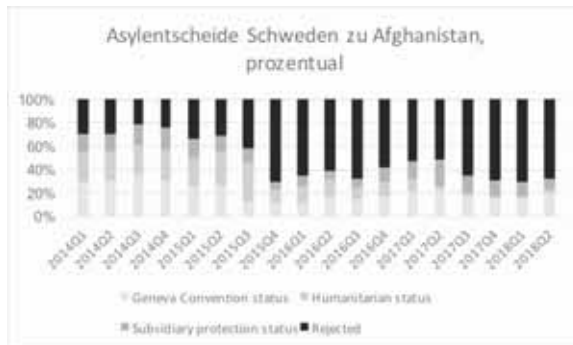
Das wachsende Interesse an Daten bringt Entscheidendes ans Tageslicht

In den letzten Jahren sind Datenportale ein großes Thema. Es wird Arbeit hineingesteckt und das Angebot wird immer besser. Eine gute Quelle ist *Eurostat*, das Statistische Amt der Europäischen Union. Hier gibt es auch viele Informationen zu Asyl und Migration. Die Mitgliedsländer liefern die Daten selbst. Die Eurostatdaten kann man in einem Datenbankbrowser ansehen oder über Schnittstellen herunterladen. Sie sind für jeden nutzbar. Am Beispiel der Anerkennungsquote von Schweden soll das illustriert werden. Die Datenquelle heißt 'migr_asycdfstq': 'Migration and Asylum', 'First instance decisions on applications by citizenship, age and sex Quarterly data (rounded)'. Wenn man im Netz danach sucht, findet man den Datenbankbrowser. Hinter der Variable „GEO“ versteckt sich das europäische Land, in dem der Asylantrag gestellt wurde. Wir setzen den Zeitraum auf von 2014 startend bis in die Gegenwart. Citizen beinhaltet das Herkunftsland. Im Bezug darauf filtern wir nach dem Kürzel „AF“ für Afghanistan. Außerdem wählen wir alle Entscheidungsarten aus. Die anderen Parameter lassen wir so,

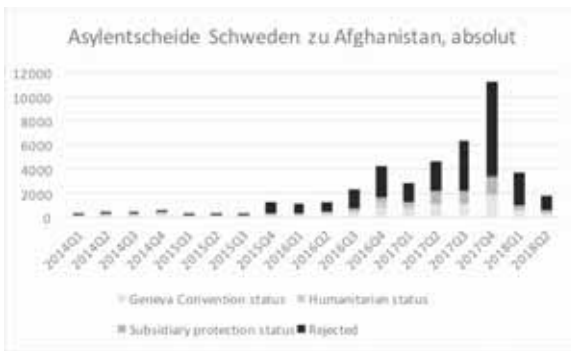
wie sie sind, klicken auf den Download Button und wählen Excel als Format. Excel ist ungeeignet für große Datenmengen oder Automatisierung, aber in diesem Beispiel tut es, was wir brauchen. Die Daten sind nicht so angeordnet, wie wir sie wollen, aber es sind nur 18 Quartale mit je sechs Werten, damit kann man arbeiten.



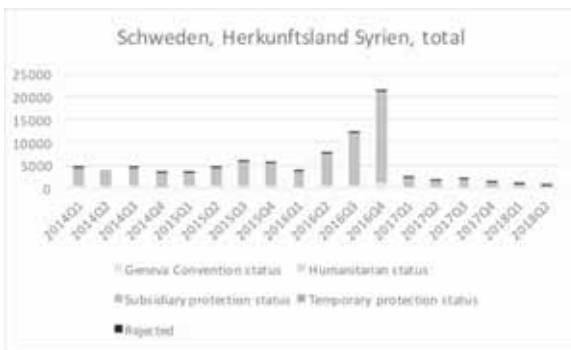
Die Anerkennungsquote des gesamten Zeitraums von Anfang 2014 bis Mitte 2018 liegt bei 40%. Das ist niedriger als in Deutschland (knapp 50%). Schweden galt doch als eher aufnahmefreundlich. Die Frage, was dahinter steckt, drängt sich uns auf. Wir finden durch die Zahlen heraus, dass Schweden seine Flüchtlingspolitik im Laufe des Jahres 2015 dramatisch verändert hat. Im November gab die Regierung eine Verschärfung der Asylpolitik bekannt. Die stellvertretende Regierungschefin Åsa Romson von den Grünen brach auf der Pressekonferenz in Tränen aus. Ministerpräsident Stefan Löfven sagte: „Es schmerzt mich, dass Schweden nicht in der Lage ist, auf diesem Niveau weitere Asylsuchende aufzunehmen.“ Also schauen wir die Anerkennungsquote noch einmal an, diesmal ihre prozentuale Entwicklung im Verlauf der Zeit.



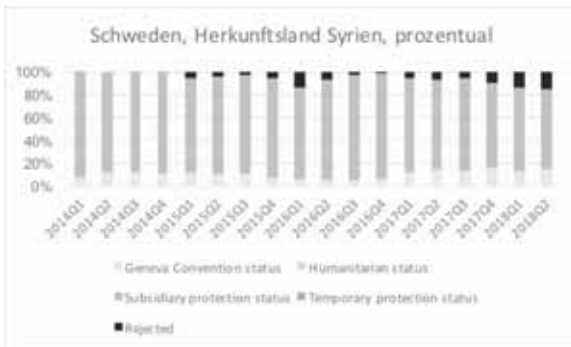
Sie zeigt einen deutlichen Einbruch Ende 2015. Zeigt sich darin der Politikwechsel? Wenn man die absoluten Zahlen heranzieht, stellt man fest, dass aus Afghanistan vorher sehr wenig Asylsuchende gekommen sind.



Vielleicht hilft es, ein Herkunftsland zu untersuchen, aus dem über viele Jahre mehr Asylsuchende gekommen sind, beispielsweise Syrien.



Hier ist die prozentuale Entwicklung viel stabiler. Es gibt keinen starken Abfall Ende 2015.



Interpretation von Daten

War der Eindruck also falsch, den wir bei der Analyse der Asylanträge aus Afghanistan bekommen hatten? Hatte die politische Umkehr einen deutlichen Einfluss auf die Anerkennungsquote? Nicht unbedingt. Es liegt der Verdacht nahe, dass besonders Entscheidungen über Asylanträge aus Afghanistan politischem Einfluss unterliegen. Um diese Vermutung zu erhärten, kann man wie im Fall von Schweden, Politikwechsel und die zeitliche Änderung analysieren oder die Quote in verschiedenen europäischen Ländern vergleichen. Es würde den Rahmen dieses Artikels aber sprengen und mit Excel würde ich das übrigens auch nicht mehr machen. Bei aller Begeisterung im Bezug auf Daten ist es wichtig, kritisch zu bleiben und sorgfältig zu arbeiten. Ohne Hintergrundwissen zum Thema entsteht schnell ein falsches Bild. Der Status 'subsidiärer Schutz' beispielsweise hat in verschiedenen Ländern verschiedene Bedeutung. Die Art der Zählung unterscheidet sich von Land zu Land und hat sich auch im Lauf der Zeit verändert. Statistik ist mühsam, man darf Zahlen nicht unbesorgt addieren und subtrahieren. Wenn ein Asylsuchender Anträge in mehreren Ländern stellt und wir über die Länder summieren, zählen wir diese Anträge doppelt. Wenn Daten aggregiert und korrigiert geliefert werden, ergibt die Differenz der Zahlen nicht den richtigen Wert für die Differenz der Zeiträume: Wenn es beispielsweise erst heißt, dass es im Januar 100 Anträge gab, und später, dass es in den Monaten Januar und Februar insgesamt 250 Anträge waren, inklusive Korrekturen für den Januar, dann ist unklar, wie viele Anträge es im Februar gab. Bei aller gebotenen Vorsicht kann man viel aus den Daten lernen. Sie können Vermutungen erhärten, Klischees widerlegen und insgesamt ein richtigeres Bild der Welt erzeugen. Dafür arbeitet der *Refugee Datathon Munich*.<

