OpenSanctions

Quelloffene Datenbank von internationalen Sanktionen

Schlussbericht

Zuwendungsempfänger: Friedrich Lindenberg

22.3.2022

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen **01IS21S48** gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurze Darstellung der Aufgabenstellung und Motivation

Was war Deine Motivation? Welches Problem wolltest Du mit Deinem Projekt lösen? Wie war die geplante Vorgehensweise zur Problemlösung (auch Angabe der wichtigsten Meilensteine)?

OpenSanctions ist eine Ressource für Journalist*innen, Aktivist*innen und Finanzdienstleister. Das Projekt sammelt Profile über Personen und Firmen, die Ziel internationaler Sanktionen sind. Hinzu kommen Einträge zu Personen, die aufgrund von kriminellem Verhalten oder als politische Mandatsträger*innen (sog. PEPs) ins öffentliche Interesse gerückt sind.

Solche Profile sind ein wichtiges Werkzeug, um in großen Datenmengen systematisch Hinweise auf Interessenkonflikte, Korruption und Geldwäsche zu finden. Wer zum Beispiel journalistisch relevante Hinweise in einem geleakten E-Mail-Postfach, oder in der Firmendatenbank einer Steueroase finden will, benötigt eine Liste von öffentlichen Figuren, deren Erwähnung den Anlass für weitere Nachforschungen schaffen würde.

Zudem stellen Sanktionen eine Art quantifizierte Geopolitik dar: Sie dokumentieren in ungewöhnlicher Klarheit, welche außenpolitischen Ziele unterschiedliche Staaten verfolgen.

Weil die Daten jedoch nicht leicht auffindbar sind, ist ihre Nutzung bisher aufwändig. Konsolidierte Datenbanken, die von Firmen wie Dow Jones oder Refinitiv angeboten werden, sind für Journalist*innen und Aktivist*innen unerschwinglich. OpenSanctions macht dieses Material deshalb für alle interessierten Nutzer*innen einfach, gratis und strukturiert nutzbar.

Das Projekt erzeugt die notwendige Datenbank durch den Betrieb einer Vielzahl von Skripten, die Informationen aus öffentlichen Quellen - Sanktionslisten von Staaten, internationalen Organisationen, aber auch Listen von zivilgesellschaftlichen Gruppen und Einträge aus dem Wikidata-Projekt - in einen gemeinsamen Datensatz vereinen.

Die Skripte lesen Einträge aus den Datenquellen und überführen sie in ein gemeinsames Datenformat, FollowTheMoney (FtM). FollowTheMoney ist ein Datenmodell aus dem Kontext des

Anti-Korruptions-Journalismus, das wichtige Konzepte - etwa Personen, Firmen, Eigentum, Familienbeziehungen oder Sanktionierungen - strukturiert beschreibt. Die Einträge aus verschiedenen Quellen werden dadurch vergleichbar und können als einheitlicher Datensatz verwaltet und abgefragt werden.

Als nächsten Schritt de-dupliziert OpenSanctions die Daten aus verschiedenen Quellen. So kommen einzelne Personen und Firmen - etwa ein bestimmter Oligarch - oft auf mehreren Listen vor. Um zu vermeiden, dass etwa eine Suche nach dem Namen dieser Person ein Dutzend unterschiedlicher Einträge produziert, nutzt OpenSanctions einen aufwändigen Werkzeugkasten, der die Deduplizierung der Listen ermöglicht und damit integrierte Profile der Sanktionsziele schafft. Diese Profile fassen die verfügbaren Daten und relevanten Entscheidungen zu einer physischen Person oder Firma in einem Eintrag zusammen.

Schließlich sollen die Einträge mit Informationen aus anderen Referenzdaten-Quellen wie Wikidata, Firmenregistern und geleakten Datenbanken angereichert werden. So können zusätzliche Verbindungen und Informationen zutage gefördert werden, die für eine Recherche von Belang sein könnten.

Die so erzeugte Datenbank sollte im Verlauf des Projektes in Form von Datenexporten bereitgestellt werden, die sich zum Import in relevante Datenanalyse-Tools eignen: die Rechercheplattformen OCCRP Aleph und ICIJ DataShare, aber auch Microsoft Excel. Es stellte sich jedoch heraus, dass die Veröffentlichung der Informationen als Webseite mit Suchfunktion und als API-Schnittstelle von zentraler Bedeutung für die Adoption der Datenbank durch Dritte war.

Beitrag des Projektes zu den Zielen der Förderinitiative „Software-Sprint"

Wer ist die Zielgruppe für Deine Lösung? Wie profitiert sie von den Ergebnissen Deines Projekts? Welche Bezüge gibt es zu den Themenfeldern und Zielen des Software Sprints?

Die Zielgruppen für die Nutzung der Daten war breit definiert: zum einen zivilgesellschaftliche Akteure - Journalisten und Aktivisten - die mit dem Thema Sanktionen befasst sind. Zunächst war dabei an Sanktionsdaten als Recherche-Werkzeug gedacht, also etwa als Suchliste für die Analyse von Daten-Leaks. So wurden OpenSanctions-Daten etwa im Kontext der internationalen Recherche-Projekte "Pandora Papers" und "Swiss Secrets" genutzt, um Hinweise auf die Präsenz relevanter Personen in den Datensätzen zu finden. Das lateinamerikanische Journalistennetzwerk "El Clip" hat OpenSanctions zudem an ihr bislang unveröffentlichtes Recherchetool angebunden.

Im Rahmen der Invasion der Ukraine durch die russische Föderation wurden Sanktionen überraschend selbst zum Nachrichtenthema. So wurde OpenSanctions von verschiedenen Medien (Correctiv, RND, NZZ, Quartz) als Quelle für tagesaktuelle Recherchen genutzt.

Neben Journalisten wurde die Datenbank auch von einigen Aktivisten, etwa Transparency International in Russland genutzt. Dazu kommen einige öffentliche Einrichtungen wir Anti-Korruptions-Einheiten, zum Beispiel in einer internationale Entwicklungsbank.

Ein weiteres Ziel des Projektes war, die Sanktionsdaten auch für Firmen im Bereich Finanztechnologie (Fintech) und Regulatorische Technologie (Regtech) attraktiv zu machen. Dazu kamen im Projektverlauf auch einige Firmen, die Recherchesoftware (sog. OSINT-Lösungen) bereitstellen.

Insgesamt haben uns innerhalb des Förderzeitraums 35 kommerzielle Nutzer kontaktiert, die Interesse an der Weiternutzung des bereitgestellten Datensatzes zeigten und die Datenbank zum Teil jetzt nutzen.

Ausführliche Darstellung der Ergebnisse

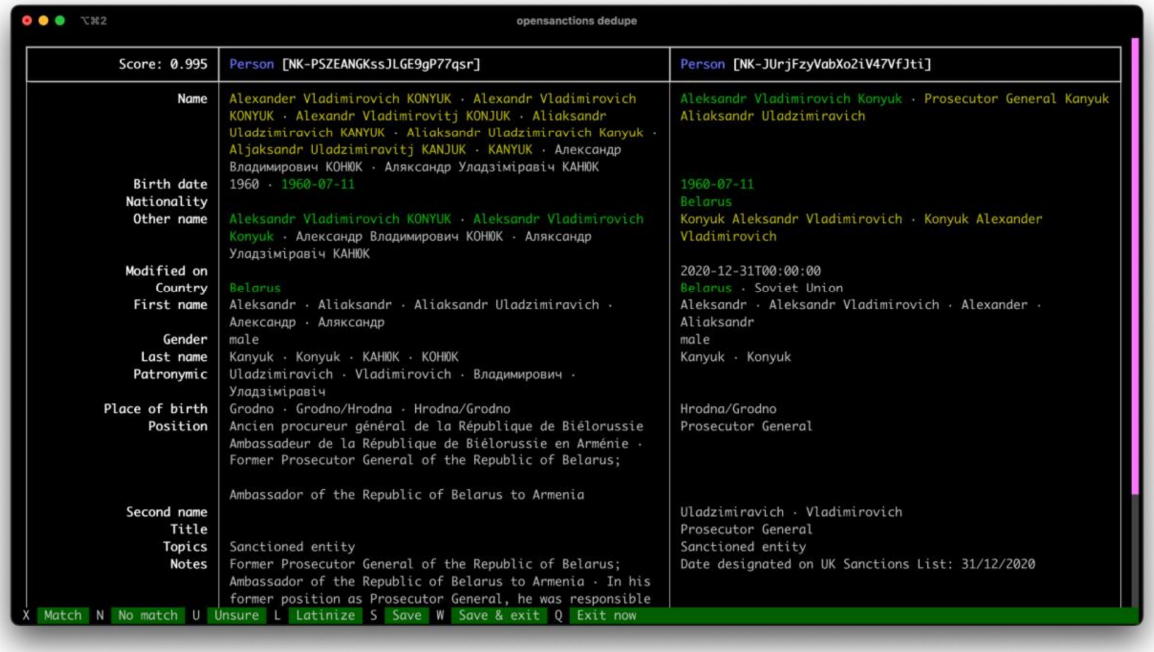
Welche konkreten Ergebnisse hast Du erzielt? Konnten alle Meilensteine erreicht werden? Welche zusätzlichen Erkenntnisse hast Du aus der Projektarbeit gewonnen, auch im Hinblick auf die Begleitung durch die Open Knowledge Foundation?

Ein wesentlicher Teil der Arbeit an OpenSanctions bestand darin, die unterschiedlichen Quellformate in eine neue, kohärente Datenstruktur zu überführen. Der Ansatz des Projektes war, alle Bereinigungen vollständig reproduzierbar und transparent zu machen. Um die Verarbeitung nachvollziehbar zu machen, wird der ganze Prozess öffentlich einsehbar bei GitHub Actions ausgeführt, und alle entstehenden Fehlermeldungen oder Warnungen werden auf der OpenSanctions-Webseite in der Beschreibung des Datensatzes veröffentlicht.

Eine hilfreiche Komponente für die Reproduzierbarkeit waren Normalisierungs-Tabellen, die Ländercodes, Adressen oder Zeitangaben vereinheitlichen. Für deren systematische Anwendung habe ich die Python-Bibliothek datapatch veröffentlicht.

Neben der strukturellen Angleichung der Daten gab es noch ein zweites, komplexes Problem zu lösen: Die verschiedenen Datenquellen - vor allem die Sanktionslisten verschiedener Staaten - enthalten immer wieder die gleichen Personen und Organisationen. Zum ehemaligen irakischen Präsidenten Saddam Hussein, zum Beispiel, kennt das System Einträge aus 13 Quellen. Stattdessen sollte eine konsolidierte Ansicht angeboten werden.

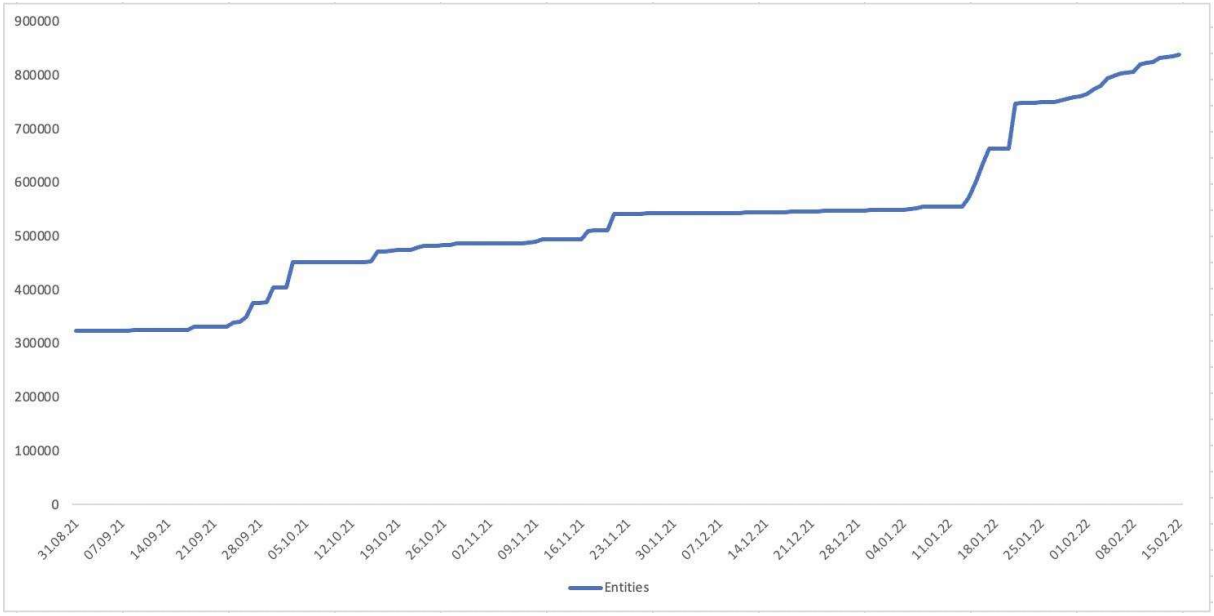
Dazu habe ich die Python-Bibliothek Nomenklatura weiterentwickelt, die mögliche Duplikate erkennt und auch verschmelzen kann. Nomenklatura enthält Funktionalität, die zwei Entitäten etwa auf der Grundlage von Namen, Geburtsort, Identifikationsnummern oder Adressen vergleicht. Zudem verwaltet Nomenklatura Entscheidungen darüber, welche Entitäten als identisch oder unterschiedlich klassifiziert wurden.



Im Nomenklatura-Werkzeug werden Einträge dedupliziert und die Ergebnisse manuell verifiziert.

Um eine hohe Qualität des Datensatzes sicherzustellen, wurde ein manueller Prüfungsschritt eingeführt, mit dem die Richtigkeit eines Matches zwischen zwei Personen oder Firmen aus unterschiedlichen Listen durch einen Menschen bestätigt wird. Zur Duplikationsstrategie gibt es auch einen detaillierten Blogpost und technische Dokumentation.

Im Laufe des Projekts habe ich die Anzahl der verwendeten Datenquellen langsam erhöht. Neben den zunächst enthaltenen Sanktionslisten aus Amerika, Europa, Schweiz, Ukraine, Kanada, Großbritannien, Frankreich und den UN konnten wir zusätzliche Quellen aus Argentinien, Australien, Israel, Japan, Kyrgyzstan, Kazakhstan, Russland und Südafrika erschließen. Daneben haben wir einige thematisch verwandte Listen, etwa die Vergabevertrag-Verbotslisten aller internationalen Entwicklungsbanken und eine Liste von Blockchain-Konten, die im Zusammenhang mit Finanzkriminalität stehen, eingebunden.



Entwicklung der Anzahl der Einträge (Personen, Firmen) in OpenSanctions im Projektzeitraum

Den größten Datenzuwachs hat das Projekt allerdings nicht durch zusätzliche Sanktionslisten erfahren, sondern in einem verwandten Themenfeld: Informationen über Politiker*innen. Deren Profile sind aus ähnlichen Gründen relevant, die auch für Sanktionen gelten: Taucht ein*e Politiker*in (oder engsten Mitstreiter*innen) etwa in einem Leak auf, dann ist das häufig der Anlass für eine journalistische Recherche.

Doch ein systematisches Verzeichnis aller Politiker*innen gibt es nicht. Jede Nation der Welt hat andere Institutionen, die nicht immer Webseiten haben, und auf diesen Webseiten dann auch nicht immer preisgeben, wer denn aktuell Finanzminister*in sei.

Ich habe deshalb Kontakt mit dem Gründer des Projekts EveryPolitician.org aufgenommen, das ab 2015 Abgeordnete in aller Welt erfasste. Obwohl das Projekt mittlerweile inaktiv ist, hat er letztes Jahr angefangen, neue Scraper zu entwickeln, die zunächst alle Kabinette der Welt erschließen und in Wikidata eintragen sollen.

Weil unsere Projekte eine sehr ähnliche Zielsetzung haben, haben wir beschlossen, gemeinsam an der Integration aller in Wikidata erfassten Politiker*innen in OpenSanctions zu arbeiten. So konnten bislang beinahe 80.000 relevante Personen aus über 250 Ländern und Territorien identifiziert werden.

Zum Projektstart ging eine einfache Webseite online, auf der Anwender*innen die produzierten Datensätze in verschiedenen Formaten herunterladen konnten. Im Verlauf des Projekts kam dann noch eine umfangreiche Dokumentation der verwendeten Datenstruktur und der Methodik hinzu. Aktuelle Entwicklungen wurden in einem Blog dokumentiert.

Bedauernswerterweise ist das Feedback der Nutzer*innen zu den bereitgestellten Dateiformaten bisher spärlich gewesen: Zwar wurden einige spezifische Erweiterungen vorgeschlagen, aber ich hatte zu dem Thema auf mehr Dialog gehofft.

Entgegen der ursprünglichen Pläne habe ich mich schließlich dazu entschlossen, die Daten auch inhaltlich auf der Webseite zugänglich zu machen. Um eine Suchfunktion und die Darstellung einzelner Einträge zu ermöglichen, habe schließlich ich einen alleinstehenden API-Server entwickelt, der die OpenSanctions-Daten in ElasticSearch durchsuchbar macht. Die Darstellung der Einträge auf

der Website und die Suchfunktion spielten eine Schlüsselrolle bei der Adoption des Projektes. Entgegen meiner Erwartungen wird die Webseite mittlerweile auch alleinstehend als Recherche-Werkzeug von Journalist*innen genutzt - nicht nur als “Kontrastmittel” in Kombination mit anderen Datensätzen.

Mithilfe einer zudem bereitgestellten Implementation der OpenRefine Reconciliation API können Nutzer*innen zudem eine ganze Liste von Firmen oder Personen auf relevante Einträge zu prüfen.

Dazu kommt eine Matching-API, mit der etwa Firmen ihren eigenen Kundenstamm mit der Datenbank abgleichen können. Damit solche Kundendaten nicht in die Projekt-Infrastruktur wandern müssen, ist der API-Server als Docker-Image abgepackt. Das kann einfach auf dem Server der Anwender*innen installiert werden und holt sich jede Stunde aktuelle Daten von OpenSanctions ab.

Weil die Qualität des Matchers für die kommerzielle Nutzung von OpenSanctions von großer Bedeutung ist, will ich die erzeugten Scores in Zukunft durch ein Regressionsmodell generieren, das die manuell vorgenommenen Deduplikations-Entscheidungen als Trainingssatz nutzt. Das Modell produziert bisher gute Erkennungswerte, soll aber noch weiter geprüft werden, bevor es in den Produktiveinsatz übergehen kann.

Ein Expertenfeature der Webseite ist die Rohdatenansicht mit der Anwender*innen genau nachvollziehen können, aus welcher Quelle welches Attribut eines Eintrags stammt.

Über den Projektzeitraum haben etwa 40.000 Menschen die Webseite besucht - davon etwa 10.000 im letzten Monat. Es ist also ein wachsendes Interesse zu beobachten.

Eine wesentliche geplante Funktion ist im Projektzeitraum nicht fertig geworden: die Anreicherung der OpenSanctions-Daten mit zusätzlichen Verbindungs- und Kontextinformationen wie etwa den OffshoreLeaks oder verschiedenen Firmenregistern. Anders als bei der Einbindung zusätzlicher Datenquellen sollten hierbei Datenquellen nicht “en gros” übernommen werden, sondern es sollten Ausschnittsweise Verbindungen und Details importiert werden, die sich auf bestehende Einträge in der OpenSanctions-Datenbank beziehen (etwa die Firmengeflechte eines Sanktionsziels wie Arkady Rotenberg in verschiedenen Steueroasen, die als Teil der OffshoreLeaks/Panama Papers öffentlich wurden).

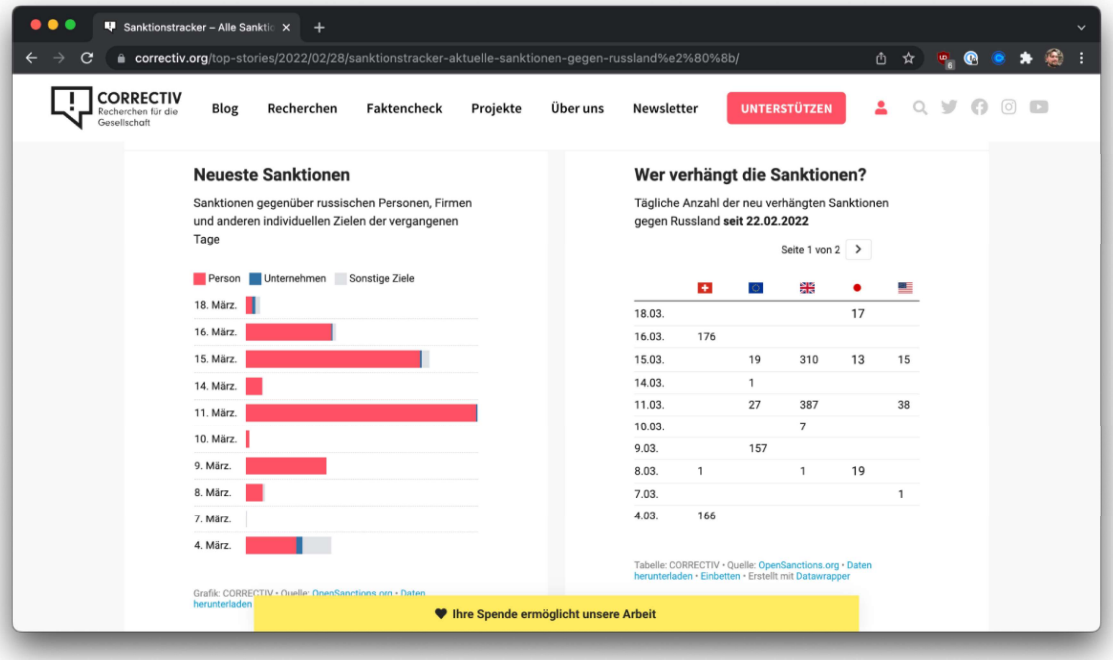
Während die Umsetzung dieser Funktion geplant und auch vom Aufwand her machbar war, habe ich mich entschlossen, in der zweiten Hälfte des Projektzeitraumes stattdessen meinen Fokus auf die Entwicklung von API und Webseite, Kommunikation mit Anwendern und die Dokumentation des bestehenden Projekts zu legen.

Zielgruppe, Nutzen und mögliche Weiterentwicklungen

Welcher Nutzen ergibt sich für die Zielgruppe aus den Ergebnissen Deines Projekts? Welche weitergehenden Effekte ergeben sich aus der Open-Source-Stellung der Ergebnisse? Gibt es Ideen für die Weiterentwicklung Deiner Lösung und Pläne zu deren Umsetzung?

OpenSanctions hat im Projektzeitraum eine nennenswerte Nutzung durch Dritte erfahren. Dazu gehören Journalist*innen, die die Daten zur Recherche in Registerdatenbanken eingesetzt haben, oder über Ereignisse in der Sanktionspolitik (insb. im Kontext des Ukraine-Kriegs) schreiben wollten. Zudem nutzen NGOs, zum Beispiel in Russland, der Ukraine, Kuba und in Kolumbien, die

OpenSanctions dazu, den Datenbestand von eigenen Datenprojekten durch Sanktions-Informationen anzureichern.



Die Effekte der Arbeit im Projektzeitraum genau zu beziffern ist schwer, weil das Ziel des Projekts die instrumentelle Befähigung von Journalist*innen und anderen Nutzer*innen ist. Das erschwert zum einen kausale Zuschreibungen, zum anderen entsteht hier oft ein enormer Zeitverzug. Das journalistische Verbundprojekt “Suisse Secrets” etwa wurde im Februar 2022 veröffentlicht, nutzte aber bereits seit April 2021 eine Vorversion der OpenSanctions-Datenbank, um sanktionierte Personen unter den Eigentümern von Konten bei der Credit Suisse zu identifizieren. Wie viele der identifizierten Bankkonten ohne die Datenbank unerkannt geblieben wären, ist jedoch unmöglich zu sagen.

Eine besonders lautstarke Nutzergruppe hingegen waren Firmen, die OpenSanctions genutzt haben: die Datenbank wurde vor allem von Startups im Bereich Open Source Intelligence und Know-Your-Customer/Compliance sondiert und zum Teil bereits in Produkte eingebunden.

Mir war von Anfang an wichtig, zumindest die Instandhaltung des Projektes über den Förderzeitraum hinaus finanziell zu untermauern. Deshalb rief die erste Version der Projektwebseite interessierte Firmen dazu auf, sich durch ein Sponsoring am Weiterbetrieb zu beteiligen. Nach halber Projektlaufzeit hatte das jedoch nur zu sehr begrenzter Resonanz geführt.

Deshalb habe ich im Dezember die Lizenzierung des Datensatzes von CC-BY auf CC-Attribution-NonCommercial geändert, mit besonderen Ausnahmeregeln für Medienunternehmen. Daraufhin meldeten sich bei mir zunehmend Firmen, die sich bereit erklärten, zur Nutzung der Daten eine monatliche Nutzungsgebühr zu entrichten. OpenSanctions befindet sich damit auf einem plausiblen Weg zur finanziellen Nachhaltigkeit.

Fertig ist natürlich nichts. Neben einer schier unendlichen Liste an Datenquellen, die ich gerne einbinden würde, gibt es auch noch eine Reihe anderer Projekte, die vermutlich erst nach der Förderphase zum Zug kommen werden:

Zum einen ist da die Anreicherung von Sanktionsdaten mit Fakten aus anderen, inhaltlich verwandten, Datenbanken. Dazu gehören neben Wikidata zum Beispiel die OffshoreLeaks-Datenbank von ICIJ, oder sogenannte Beneficial Ownership-Datenbanken in denen die Eigentümer*innen von Firmen dokumentiert sind. Dies ist besonders relevant, weil Firmen, die anderen sanktionierten Firmen oder Personen gehören in der Regel selbst als sanktioniert gelten. Die bestehende Infrastruktur zur De-Duplikation wird hier eine Rolle spielen. Die Software muss jedoch zuvor Matching-Kandidaten durch API-Aufrufe beschaffen.

Umgekehrt wollte ich auch gerne die Eintragung aller Sanktionsziele als Items in Wikidata in Angriff nehmen. Für die zentralen Akteur*innen gibt es natürlich bereits Einträge, aber hier eine Vollständigkeit zu erzielen, ist ein Ziel für die Zukunft.

Schließlich bleibt auch die politische Analyse der Sanktionsdaten unerledigt. Was können uns Sanktionziele über die Außenpolitik verschiedener Staaten verraten? Etwa über die amerikanische Fokussierung auf iranische Nonproliferation, oder die europäischen Experimente in der Bestrafung des Putin-Regimes? Diese Fragestellungen sind zum Ende des Förderzeitraums durch den Einmarsch der russischen Armee in die Ukraine in den Vordergrund getreten, und die Bereitstellung von Kerndaten zur Bewertung der weltweiten Sanktionspolitik wird vermutlich in Zukunft einen zentralen Arbeitsbereich von OpenSanctions darstellen.

Hat die Arbeit in dem Projekt Dich in Deiner persönlichen, fachlichen Weiterentwicklung unterstützt?

Das Projekt hat mir persönlich erlaubt, sowohl in technischer wie auch in fachlicher Hinsicht viel dazu zu lernen. Zum einen habe ich den Projektzeitraum genutzt, um mich mit einigen neuen Technologien vertraut zu machen (z.b. asynchrone Python-Entwicklung, Implementation von Duplikationsverfahren, Wikidata-Semantik).

Zum anderen konnte ich viel über inhaltliche Aspekte der Sanktionspolitik verschiedener Länder, und der Anforderungen an Know-Your-Customer und Due Diligence-Verfahren in regulierten Industrien lernen. Dazu kamen zahlreiche lehrreiche Gespräche zu Themen wie Geldwäscheprävention, Kryptowährungen, Recherchestrategien.

Kurze Darstellung der Arbeiten, die zu keiner Lösung geführt haben

Gab es Arbeiten bzw. Lösungsansätze, die nicht weiter verfolgt wurden? Was waren die Hintergründe, und wie bist Du alternativ vorgegangen?

Ein gescheitertes Experiment war die Auswertung von sog. Debarment-Listen verschiedener Entwicklungsbanken. Betrügt etwa ein Anbieter die Weltbank im Rahmen eines Vergabeverfahrens, so wird er (unter Umständen) von dieser auf einer öffentlichen Sanktionsliste genannt. Diese Liste wird dann von anderen Entwicklungsbanken - etwa der Asiatischen Entwicklungsbank, der Inter-Amerikanischen Entwicklungsbank, der EIB oder des EBRD - bei eigenen Vergabeverfahren berücksichtigt.

Ich habe die Hypothese verfolgt, dass diese sog. Cross Debarment-Policy sehr lückenhaft umgesetzt würde. Zusammen mit einer Journalistin haben ich deshalb Vergabeverfahren mehrer Entwicklungsbanken mit relevanten Ausschlusslisten verglichen. Das Ergebnis war jedoch der Albtraum jedes Journalisten: das Cross Debarment scheint zu funktionieren - wer von einer der großen Entwicklungsbanken als unzuverlässig eingestuft wird, taucht danach auch in den Auftragsbüchern der anderen nicht mehr auf.

Besonders viel Zeit habe ich mit der Entwicklung von Werkzeugen zur De-Duplikation bzw. Verknüpfung der Einträge aus verschiedenen Listen verbracht. Während der direkte Vergleich zweier Sanktionsziele recht einfach zu implementieren ist, stellt sich der Quervergleich von zehntausenden Einträgen als Herausforderung heraus. Um ein sog. “blocking” (also eine generierung von Kandidatenpaaren für die Deduplikation) zu ermöglichen, habe ich einen kleinen Suchindex entwickelt. Das war kompliziert und nicht besonders erfolgreich, im Rückblick hätte ich direkt auf einen Suchindex wie ElasticSearch zurückgreifen sollen.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

Wo können sich Interessenten detailliert über Deine Projektergebnisse informieren (z.B. Webseite, GHitHub, Veröffentlichungen)?

Alle Ergebnisse des Projektes werden auf der Projektwebseite www.opensanctions.org bereitgestellt. Dort werden auch die produzierten Daten zum Download angeboten, und ist eine Suche und Darstellung einzelner Sanktionsziele möglich. Der Quellcode für alle Projektkomponenten ist auf der GitHub-Organisationsseite https://github.com/opensanctions gesammelt.

Weiter Veröffentlichungen umfassen z.B.:

- Mehrere Blogposts zu technischen und inhaltlichen Aspekten des Projektes.
- Ein Webinar zum Thema Netzwerkdatenanalyse, in Zusammenarbeit mit der Firma Linkurious.
- Ein Artikel des Global Investigative Journalism Network (GIJN) zum Thema Recherchen zu Finanzthemen, OpenSanctions.

Kurze Erläuterung zur Einhaltung der Arbeits- und Kostenplanung

Gab es im Projektverlauf Ereignisse, die eine Anpassung der Planung erforderlich machten – z.B. Mehr- oder Minderaufwand bei der Bearbeitung von Teilaufgaben?

Die Arbeits- und Kostenplanung für das Projekt hat sich als realistisch herausgestellt. Die ersten drei Monate habe ich mich sehr stark auf die im Antrag festgeschriebenen Meilensteine fokussiert, danach haben mich Gespräche mit Nutzern dazu bewegt, eine Re-Priorisierung der Aufgaben vorzunehmen und den Fokus der weiteren Arbeiten anzupassen. Anstatt auf die Anreicherung aus externen Datenbanken habe ich mich deshalb auf eine bessere Präsentation der Daten auf der Projektwebseite und auf die Entwicklung einer Daten-API konzentriert.

Besonders hilfreich hätte ich es gefunden, wenn es eine formelle Möglichkeit gäbe, notwendige Zuarbeiten von Dritten im Rahmen des Projekts innerhalb des Budgets abzurechnen (etwa eine Rechtsberatung, gestalterische Dienstleistungen).

Kurze Darstellung von etwaigen Ergebnissen bei anderen Stellen

Gab es Entwicklungen anderer Personen oder Institutionen, die Einfluss auf Deine Arbeiten und die Zielsetzung hatten? Wenn ja, worin bestand dieser und wie bist Du damit umgegangen?

Anders als im Antrag geplant, haben wir uns entwieden auf der OpenSanctions-Seite nicht nur Informationen zu Sanktionszielen anzubieten, sondern auch Daten über politische Mandatsträger*innen (sog. “Politically Exposed Persons”, PEPs) bereitzustellen. Diese Informationen dienen ähnlichen Zwecken wie die Sanktionsziele: die Präsenz einer Politikerin in einem journalistischen Leak oder dem Firmenregister eines anderen Landes ist oft Anlass für eine vertiefte Recherche.

Daten über politische Mandatsträger*innen werden, anders als Sanktionsdaten, selten von Regierungen als strukturierter Datensatz zum Download angeboten. Stattdessen müssen sie von Kabinetts- und Parlaments-Webseiten gescrapt werden. Dazu bin ich eine Kooperation mit einem Entwickler eingegangen, der seit 2015 im Rahmen des Projektes “EveryPolitician” eine große Zahl solcher Scraper entwickelt hat. Er hat mir Datensätze mit Parlaments- und Kabinettsmitgliedern aus der ganzen Welt zur Einbindung in OpenSanctions freundlicherweise zur Verfügung gestellt.

Während des Projektzeitraums wurde ich von einer Vielzahl von Personen und Organisationen freundlich in der Umsetzung des Projekts beraten, u.a. den Firmen OpenCorporates, DeepDiscovery, dmetrics, Ontotext, Senzing, C4ADS und der OpenOwnership-Kampagne.