Richtlinie zum „Software-Sprint“

OpenRedact – Deutsche Texte von personenbezogenen Daten bereinigen

Schlussbericht

Zuwendungsempfänger:

Berrang, Langhabel, Ostendorff und Ostendorff GbR

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01IS19S25 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)?

Der Wunsch von Unternehmen, Kund*innendaten vollumfänglich für ihre geschäftlichen Interessen zu nutzen, widerspricht dem Bedürfnis der Nutzer*innen nach Datenschutz. Das Bewusstsein für Datenschutz ist seit der Einführung der europäischen Datenschutzgrundverordnung gestiegen und stellt eine Herausforderung für bestehende Datensätze sowie das Erschließen neuer Geschäftsfelder dar. Staatliche Stellen müssen bislang enormen Aufwand betreiben, um ihre Daten für die Veröffentlichung zu anonymisieren, denn die händische Bereinigung dieser Daten ist aufwändig und durch den erheblichen Umfang kaum realisierbar.

Das Ziel von OpenRedact (ursprünglicher Name: Schwärzer) ist die Stärkung der Transparenz der Verwaltung durch den Einsatz von technischen semi-automatisierten Anonymisierungs-Applikationen, um die Verwaltung bei Informationsfreiheitsgesetzanfragen und bei der Veröffentlichung von Gerichtsentscheidungen zu entlasten. Auch Firmen und Kund*innen können von einer Anonymisierungslösung profitieren, in dem Daten in neuen Geschäftsfeldern genutzt oder Kundendaten besser geschützt werden können.

Die geplante Vorgehensweise war es, sowohl eine prototypische Anonymisierungs-Applikation zu implementieren, als auch einen juristischen Leitfaden zur Anonymisierung zu schreiben. Um eine größtmögliche Nachnutzbarkeit zu ermöglichen haben wir geplant die Logik der Applikation in

mehrere funktionale Komponenten aufzuteilen, die als eigenständige Open Source Projekte bestehen können. Diese sind ein Modul zum Extrahieren und Modifizieren des Textinhalts von Dokumenten, ein Modul zum Erkennen von personenbezogenen Daten, und ein Modul zur Anonymisierung von Daten.

Die wichtigsten Meilensteine waren:

- MS0: Entwicklung einer juristischen Richtlinie für Anonymisierung.

  Auswertung von Anonymisierungsmöglichkeiten, Empfehlungsleitfaden für die Anonymisierung und Pseudonymisierung bei Urteilen, evt. Aufsatz zum Thema der Anonymisierung und Pseudonymisierung von Urteilen und der Frage, müssen diese überhaupt anonymsisiert werden.

- MS1: Erstellen eines Anonymisierungs-Framework das gegebene Textpassagen anonymisiert, pseudonymisiert beziehungsweise mit Methodiken der Differential Privacy verschleiert.

  Textpassagen können je nach Kategorie mit Hilfe von unterschiedlichen Methoden anonymisierten werden. Folgende Anonymisierungstechniken sollen unterstützt werden (leicht nach schwer): Schwärzen, Pseudonymisieren, Generalisieren, Differential Privacy.

- MS2: Regelbasierte und statistische Identifikation von personenbezogenen Daten mit verfügbaren Sprachmodellen.

  Mindestens 10 verschiedene Arten von personenbezogenen Daten können erkannt werden. Mindestens 5 davon sollen mit hohem F2-Maß erkannt werden.

- MS3: Implementation der Web App.

  Die App kann sowohl über eine grafische Nutzeroberfläche als auch eine API angesprochen werden. Der Fokus der grafischen Nutzeroberfläche liegt dabei auf einfacher Nutzbarkeit. Das Tool kann deutsche Texte in verschiedenen Formaten (mindestens jedoch einfacher Text) analysieren. Der Anonymisierungsvorschlag bzw. die identifizierten personenbezogenen Daten können manuell durchgegangen und angepasst werden. Der Benutzer erhält mit der Web App auch die Möglichkeit die Performance auf von ihm gegebenen Testdaten zu messen. Die Performance wird insbesondere über Indikatoren wie “false positive” und “false negative” Raten angegeben.

In der Planungsphase zu Beginn der Förderung haben wir folgenden ursprünglich angegebenen Meilensteinen, welche der Optimierung der Erkennung von personenbezogenen Daten dienen, niedrige Priorität gegeben.

- MS4: Erstellen eines Benchmark Corpus für die Identifikation der personenbezogenen Daten und messen der Performance unter Angabe der “false positive” und “false negative" Raten.

- MS5: Verbessern der Performance des statistischen Sprachmodells durch Transfer Learning und durch generieren von Trainingsdaten.

Am Ende steht ein Prototyp, der beispielhaft zeigt, wie Anonymisierung und Pseudonymisierung funktionieren könnte.

Beitrag des Projektes zu den Zielen der Förderinitative „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 Haupt-Zielgruppe für die semi-automatisierte Lösung ist die öffentliche Verwaltung, da durch die semi-automatisierte Erkennung von personenbezogenen Daten eine zeitsparende Anonymisierung ermöglicht werden kann. Dadurch wird der öffentlichen Verwaltung die Realisierung von eGovernance und Open Data sowie der zeitgleiche Schutz von personenbezogenen Daten in Einklang gebracht.

Weitere Zielgruppen sind Unternehmen und Journalisten. Diese profitieren vom Projekt indem sie ihre Daten auf vielfältigere und umfassender Weise nutzen können unter gleichzeitiger Wahrung ihrer Verpflichtungen die Daten zu schützen. Auf der anderen Seite profitieren die Kunden beziehungsweise Bürger, indem ihre Daten besser anonymisiert und damit geschützt werden können.

Auf technischer Ebene haben wir OpenRedact in unabhängige Komponenten gegliedert. Jede einzelne Komponente erfüllt eine eigenständige Aufgabe und kann dadurch unabhängig von OpenRedact einen Mehrwert liefern. Zielgruppe für diese Komponenten sind Software Entwickler*innen die in den Bereichen Anonymisierung, Bearbeiten von Dokumenten und Erkennung von Daten in Dokumenten arbeiten. Auf diesem Wege fördern wir die langfristige Nachnutzung der im Projekt entstandenen Software.

In einem interdisziplinären Team haben wir mit dem Projekt an den Zielen Transparenz und Datenschutz gearbeitet, welche die ganze Gesellschaft betreffen. Transparenz und Datenschutz stehen mit den Zielen des Software Sprints im Einklang, und wir leisten mit OpenRedact einen Beitrag diese Ziele mit freier Open Source Software erreichbar zu machen.

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?

Die Teilnahme am Prototype Fund hat es OpenRedact ermöglicht, die Idee von einer offenen teil-automatisierten Anonymisierungs-Applikation umzusetzen. Im Laufe der letzten sechs Monate hat OpenRedact technische und juristische Lösungen entwickelt, die helfen können den Aufwand und die Kosten für die Veröffentlichung von Dokumenten als Open Data zu senken.

Mit Hilfe unserer Web-App können PDFs und andere Dokumente von personenbezogenen Daten bereinigt werden. Bei der Entwicklung des Prototyps haben wir besonderen Wert auf die Nachnutzbarkeit gelegt. Alle Komponente sind eigenständig nutzbar. Wir zeigen mit unserem Prototypen die technischen Möglichkeiten von teil-automatisierter Anonymisierung und bieten eine Grundlage für die zukünftige Zusammenarbeit und Nutzung des öffentlichen Sektors.

Der Anonymisierungsleitfaden soll Behörden und der Justiz das Vorhaben von Open Data unter Wahrung des Datenschutzes ermöglichen. Grundsätzlich sind personenbezogene Daten zu anonymisieren. Insbesondere der Anonymisierungsleitfaden soll auch über die Projektgrenzen hinweg eine juristische Diskussion zu Open Data und Datenschutz anstoßen.

Darüber hinaus haben unsere Arbeiten gezeigt, dass beim Einsatz von OpenRedact branchenspezifische Anforderungen beachtet werden müssen. Aus diesem Grund haben wir Anpassungsmöglichkeiten besonders berücksichtigt. Beispielsweise können die Erkennungs-Methoden der Komponente zur Erkennung von personenbezogenen Daten an den genauen Anwendungsfall angepasst werden.

Die Meilensteine 0 bis 3 konnten wie geplant erreicht werden. Die Meilensteine 4 und 5, denen in der Planungsphase zu Beginn der Förderung eine niedrige Priorität gegeben wurde, wurden nicht angegangen. Der Grund für deren niedrige Priorität und das Nicht-bearbeiten ist, dass es sich bei diesen Meilensteinen um eine für den Prototypen nicht notwendige Optimierung der Erkennung von personenbezogenen Daten handelt. Stattdessen haben wir eine Möglichkeit geschaffen, schnell und einfach regelbasierte Methoden zur Erkennung von personenbezogene Daten selbst zu schreiben, was wir vor dem Einsatz auf einer neuen Art von Dokumenten als notwendig erachten. Die von der Jury vorgegebene Angabe von “false positive” und “false negative” Raten ist dennoch möglich, und zwar auf Basis des jeweils bearbeiteten Dokuments. Die Performance der automatischen Erkennung von personenbezogenen Daten wird auf Basis der manuellen Korrekturen bestimmt und angezeigt.

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

OpenRedact zeigt beispielhaft wie eine teil-automatisierte Anonymisierung in der öffentlichen Verwaltung eingesetzt werden kann, beispielsweise bei Informationsfreiheitsgesetzanfragen oder bei der Veröffentlichung von Urteilen, indem personenbezogene Daten geschwärzt werden. Mit dem Einsatz einer Anonymisierungs-Applikation kann die öffentliche Verwaltung entlastet werden, weil keine händische Anonymisierung erfolgen muss. Sicherlich bleibt eine abschließende Überprüfung notwendig, jedoch kann die teil-automatisierte Anwendung Informationsfreiheitsgesetzanfragen erleichtern.

Für die teil-automatisierte Anonymisierung von Dokumenten in der öffentlichen Verwaltung stellt OpenRedact derzeit einen Prototypen dar. Für den finalen Einsatz einer semi-automatisierten Anwendung bedarf es einer Untersuchung zum Einsatz von OpenRedact in der öffentlichen Verwaltung. Die Arbeit zeigt in technischer Hinsicht was mit Automatisierung und Methoden der künstlichen Intelligenz ermöglicht werden kann.

Durch die Open Source Veröffentlichung der eigenständigen Komponenten haben wir Lücken gefüllt, für die es noch keine Software-Bibliotheken gab. Wir hoffen dadurch langfristigen Nutzen für mögliche Anwender*innen geschaffen zu haben. Ohne die Open Source Veröffentlichung wäre dies nicht möglich. Mit dem Anonymisierungsleitfaden haben wir eine Grundlage für die weitere Diskussion von Anonymisierung amtlicher Werke geschaffen.

Fachlich war das Projekt sehr lehrreich. Im interdisziplinären Team haben wir unsere Kenntnisse für den gesamten Software-Entwicklungs-Zyklus zusammen getragen und uns dadurch nicht nur neues Wissen erarbeitet, sondern auch Bekanntes weitergegeben. Insbesondere auch die UX-Coachings haben unsere Weiterentwicklung unterstützt.

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?

Wir haben von der Bereitstellung einer Live-Demo des Prototypen aufgrund der Haftung bei Fehlern des Programmes abgesehen.

Des weiteren haben wir unsere Arbeiten an einer Kooperation mit einer Plattform für Informationsfreiheitsgesetzanfragen eingestellt. Ziel einer solchen Kooperation war es Trainingsdaten und echten Anwendungsdokumente zu erhalten. Der Grund dafür war, dass die verfügbaren Daten für uns nicht als Trainingsdaten geeignet waren, und wir dem Trainieren von Modellen des Maschinellen Lernens ohnehin eine niedrige Priorität gegeben hatten.

Anstatt einer Komponente die nur personenbezogene Daten in deutschen Dokumenten identifiziert, haben wir eine Komponente für allgemeine Named-entity recognition (NER) geschrieben. Wir haben gelernt, dass Erkennungsmethoden – um beste Ergebnisse erzielen zu können – spezifisch für einen Sprachraum oder eine Art von Dokumenten sein müssen. Statt den Fokus auf eine Art von Dokumenten zu legen, haben wir eine universell nutzbare Suite für NER gebaut, die das Ziel hat, dass Nutzer*innen schnell eigene Erkennungsmethoden für die spezifischen vorliegenden Daten schreiben können.

Die Suche nach Kooperationspartnern von denen wir Testdokumente erhalten und die Tester darstellen können, hat zu keinen Ergebnissen geführt. Kontaktierte offizielle Stellen, welche unsere Hauptzielgruppe darstellen, waren uns gegenüber nicht offen für dir Durchführung einer Testphase mit einen Prototypen – im Gegensatz zu einem fertigen Produkt. Auch die Corona-Pandemie hat die Kontaktaufnahme erschwert. Alternativ haben wir die Anforderungen der Software auf den Erkenntnissen basiert, die wir im Team erlangt und im UX-Coaching optimiert haben.

Kurze Angabe von Präsentationsmöglichkeiten für mögliche Nutzer

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

Auf unserer Homepage (https://openredact.org/) kann man sich über unser Projekt informieren. Medienwirksam sind unser Blogpost der Demo-Woche des Prototype Funds, unsere detaillierten technischen Blogposts auf Medium, und unsere Aktivität auf Twitter. Der Quellcode ist auf GitHub (https://github.com/openredact) veröffentlicht.

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?

Nach der initialen Planphase nach Förderbeginn wurden keine große Änderung des Plans mehr vorgenommen. Die Bearbeitungszeit der Teilaufgaben war stets im Rahmen des Plans.

Durch die frühen Abgabefristen für die Demo-Woche mussten einige Arbeiten vom August in den Juli vorgezogen werden.

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?

Wir haben uns mit Behörden über sog. Anonymisierungsleitfaden ausgetauscht. Die dadurch gewonnen Erkenntnisse wurden in der Erstellung unseres Anonymisierungsleitfadens berücksichtigt.