

Im Folgenden eine kurze Beschreibung der Klassen und Schnittstellen:

DocumentCheckService

Hierbei handelt es sich um ein einfaches Interface, welches die CheckDocument- Methode definiert. Es ist darauf ausgerichtet, in Zukunft wahlweise um weitere Methoden ergänzt zu werden. Die CheckDocument- Methode erwartet ein CheckEvent und gibt ein CheckResultEvent zurück, sodass ein Framework, welches Kafka Topics empfängt und versendet diese Schnittstelle nutzen kann.

IBANChecker

Die Implementierung von DocumentCheckService, welche ein Gegebenes pdf- Dokument nach IBANs durchsucht und diese prüft.

Einige Annahmen wurden getroffen:

- Dokumente werden nur im pdf- Format angenommen. Sollen künftig noch weitere Dokumentarten unterstützt werden, muss dies zusätzlich implementiert werden. Durch die Verwendung der Schnittstelle FileTextParser und einer switch-case-Struktur ist dies jedoch recht bequem zu realisieren. Bisher ist nur ein PDFTextParser implementiert.

- Ich habe mich an der Struktur von Rechnungen nach der Beispieldatei Testdaten_Rechnungseinreichung.pdf orientiert, um die IBANs nach folgendem Prozess zu finden:

1. Download der pdf-Datei von der URL
2. Konvertierung der pdf- Datei in eine Textstring
3. Unterteilung des Textstrings nach dem Wort „IBAN“; die Idee ist, dass eine IBAN als solche bezeichnet sein muss
4. Die Zeichen unmittelbar nach „IBAN“ werden nach der IBAN durchsucht

Das Programm prüft IBANs zunächst auf Validität, danach auch Blacklist- Status. Werden IBANs auf der Blacklist gefunden, werden diese IBANs in der Nachricht der CheckResultEvents zurückgegeben. Werden keine Blacklist- IBANs gefunden, aber es existieren IBANs. Die nicht valide sind (z.B. falscher Ländercode oder falsche Prüfziffer), werden diese zurückgegeben und die Prüfung schlägt ebenfalls fehl.

Zum Sicherheitsaspekt ist zu sagen, dass ein Event, welches nur eine URL zu einem Dokument liefert nicht gerade ungefährlich ist. Ein weiteres Modul, welches die Sicherheit einer URL vor dem Zugriff sicherstellt, sollte implementiert werden. Ferner können durch diese Struktur auch nur ungesicherte Dateien verarbeitet werden; zur Verarbeitung von z.B. Passwortgeschützten Dateien fehlt die Information.

IBAN und IBANCountryCode

Eine einfache Klasse, um eine IBAN zu repräsentieren. Ich habe hier darauf verzichtet, auf vorhandene Implementierungen aus öffentlichen Bibliotheken zurückzugreifen.

Implementiert ist eine Prüfmethode, um die Validität einer IBAN zu verifizieren: Prüfung auf gültigen Ländercode, Länge der IBAN in Abhängigkeit davon und Prüfziffer bzw. -Summe.

Das Enum IBANCountryCode enthält alle gültigen Ländercodes; es ist denkbar in einer zukünftigen Implementierung diese Liste von einer zuverlässigen, aktuellen Quelle zu beziehen, anstatt sie auf diese Weise hardgecoded im Programm pflegen zu müssen, wenn sich Definitionen ändern.

IBANBlacklistLoader

Ein Interface für einen Loader einer (oder mehrerer) Blacklist(s) für IBANs.

Da in der gegebenen Aufgabe nicht spezifiziert wurde, wie eine geblacklistete IBAN definiert ist, bzw. woher die Blacklist kommt, habe ich hierzu keine Implementierung geschrieben. Lediglich für die Unit- Tests habe ich simple Testimplementierungen vorgenommen, um die Prüfmethode zu testen. Denkbar wäre eine Implementierung, die sich entweder eine vorhandene Blacklist von einer sicheren Quelle aus dem Netz runterlädt oder aber eine in der eigenen Datenbank gepflegte Blacklist abgreift.

FileTextParser und PDFTextParser

FileTextParser ist die generische Schnittstelle einer Klasse, die eine Datei von einer URL liest und in Textformat (ein String) umandelt. PDFTextParser ist eine Implementierung für PDF- Dateien.

Testfälle

Um die Module zu testen, habe ich die bereitgestellte Datei Testdaten_Rechnungseinreichung.pdf genutzt. Ich habe sie in meine Dropbox hochgeladen und von dort über eine URL in das Programm geladen. Ich werde zu Demonstrationszwecken die Datei auch in meiner Dropbox belassen, damit die Unit- Tests nicht fehlschlagen.

Getestet wurden sämtliche Fehlerfälle, wie auch ein fehlgeschlagener Test und ein Test, der zu positiver Rückmeldung führt.