AI 기반 무역거래 자금세탁방지 솔루션 Anti-TBML 시스템 개발

- 최종발표-

하나금융기

Mentor

김신웅 기술사님

CONTENTS

01

프로젝트 개요

- 1. Anti-TBML이란?
- 2. 주요 기능
- 3. 목표 시스템
- 4. R&R
- 5. 협업 방식
- 6. 개발 일정 & WBS

02

사용 기술

- 1. AWS RDS
- 2. MySQL
- 3. Google Vision API
- 4. Edit Distance

03

결과물

- 1. 시스템 구조
- 2. 상세 기능
- 3. 시연 영상

04

결론

- 1. 기대 효과
- 2. 개선 방안

01

프로젝트 개요

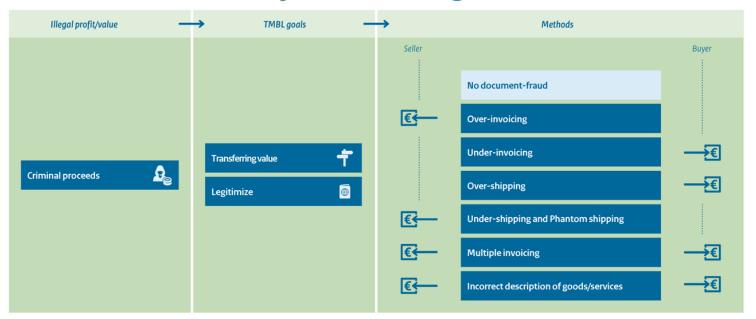
01. Anti-TBML이란?

프로젝트 개요

TBML (Trade-Based Money Laundering)

: 무역금융 과정에서 일어날 수 있는 범죄 자금 조달 행위

Trade Based Money Laundering



- 무역 거래를 활용하여 수익을 은 닉, 또는 이동시키는 방식으로 이뤄 해당 수익은 자금 세탁, 밀반입, 테러 와 같은 범죄 자금으로 활용 - Anti-TBML은 이러한 **범죄 자금 조** 달 행위를 사전에 탐지하고 차단하기 위한

시스템

02. 주요 기능

프로젝트 개요

데이터베이스

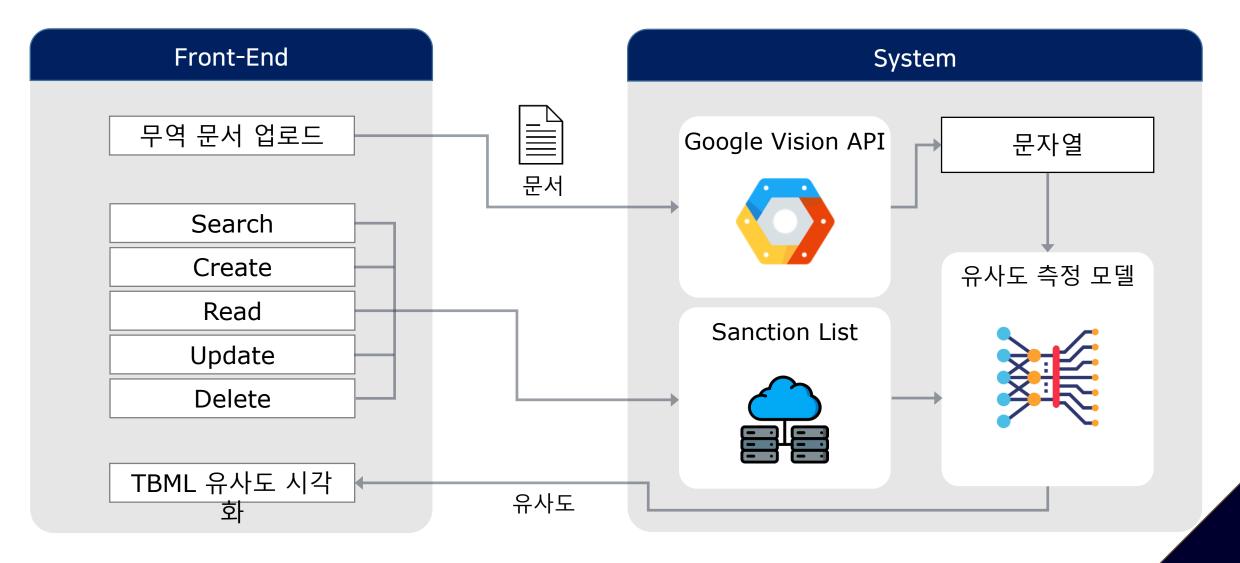
- 미국 재무부 홈페이지에 있는 sanction list (제재 리스트)를 크롤링하여 데이터베이스에 저장
- 개인·국가·기업·선박 등 다양한 데이터를 저장할 수 있는 테이블을 구성

웹 (Front-End & Back-End)

- sanction list에 대한 CRUD 및 검색 작업 수행
- 사용자가 무역 문서를 업로드
- OCR을 이용하여 해당 문서의 내용을 문자열 단위로 추출
- 추출된 문자열과 text matching 알고리즘 및 모델을 이용하여 TBML과의 유사도 측정

03. 목표 시스템

프로젝트 개요



04. R&R

프로젝트 개요



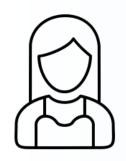
서준교(PM)

데이터베이스 구축 및 Front-End & Back-End 개발



최기훈

text matching 및 유사도 측정 알고리즘 설계 및 평가



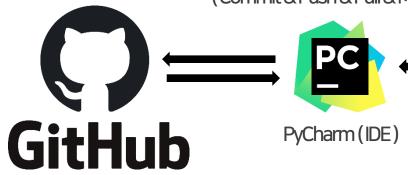
조현아

Front-End & Back-End 개발

05. 협업 방식

프로젝트개요

Contribution (Commit & Push & Pull & Merge)



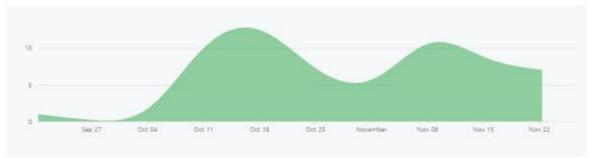
Source & Version Control

•	Joonkkyo threshold 이상의 유	사도를 가진 word의 boundary만 출 c50d1e9 3 days ago	© 62 commits
	.idea	코드 정리	10 days ago
	assets	navbar 문제 해결	6 days ago
	config	주말 전 수정	10 days ago
	document_inspection	threshold 이상의 유사도를 가진 word의 boundary만 출	3 days ago
	media/documents	navbar 문제 해결	6 days ago
	model	GOOGLE_VER.py 수정	6 days ago
	sanction	navbar 문제 해결	5 days ago
	templates	navbar 문제 해결	6 days ago
	manage.py	Initial commit	2 months ago
	requirements.txt	requirements.txt 생성	12 days ago

Sep 20, 2020 - Nov 26, 2020

Contributions: Commits ▼

Contributions to master, excluding merge commits









프로젝트 개요

Schedule

발표준비

9/4 9/11 9/18 9/25 10/2 10/9 10/16 10/23 10/30 11/6 11/13 11/20 11/27 8 / 28 자료및기술조사 데이터베이스구축및연동 Django기반Front-End&Back-End개발 Text matching 및 유사도 측정 알고리즘 개발 통합테스트&디버깅

프로젝트 개요

<요구사항 및 기능 정의 >

TASK	TITLE & TASK	산출물	담당	자	DUE	DAY	상태	2020년 8월		9윌		1	0윌		11월
No.	TITLL & INSK	022	정	부	START	END	75 VII	1 2 3 4	1	2 3 4	1	2 (3 4	5	1 2 3
1	요구사항 정의 & 명세화				08.24(월)	09.07(월)									
1.1	계획				08.24(월)	08.29(星)									
1.1.1	요구사항 파악에 필요한 정보 탐색		ALL		08.24(월)	08.29(星)	완료								
1.1.1.1	TBML 관련 배경 지식 확장 및 자료 조사	관련URL	ALL		08.24(월)	08.26(수)	완료								
1.1.1.2	기능 벤치마킹(구글링, OCR)	관련URL	최기훈		08.27(목)	08.30(일)	완료								
1.1.1.3	기능 벤치마킹(논문, 아티클) 1. NLP 2. OCR	관련자료	최기훈		08.27(목)	08.30(일)	완료								
1.2	정의 & 명세화				08.31(월)	09.07(월)									
1.2.1	요구사항정의서 작성	요구사항정의서			08.31(월)	09.03(목)	완료								
	RFP, 계획서, 기획서 검토 → 주요 요구사항 도출		최기훈		08.31(월)	09.03(목)	완료								
	요구사항정의		ALL		08.31(월)	09.03(목)	완료								
1.2.2	요구사항명세서 작성	요구사항명세서			09.04(금)	09.07(월)	완료								
	요구사항정의서 검토		서준교		09.04(금)	09.05(토)	완료								
	요구사항별 상세내용 작성		ALL		09.06(일)	09.07(월)	완료								
2	기능 정의 & 명세화				09.08(화)	09.13(일)									
2.1	계획				09.08(화)	09.11(금)									
2.1.1	요구사항명세서 검토				09.08(화)	09.09(수)	완료								
2.1.2	구현기능 범위 확정				09.08(화)	09.09(수)	완료								
2.2	기능 정의				09.10(목)	09.11(금)									
2.2.1	요구사항명세서 기반 기능 정의				09.10(목)	09.11(금)	완료								
2.2.2	구현기능 범위 조정					09.11(금)	완료								
2.3	기능 명세화				09.11(금)	09.11(금)	완료								
2.3.1	기능정의서 검토				09.11(금)	09.11(금)	완료								

프로젝트 개요

<프론트엔드 및 백엔드개발>

TASK	TITLE & TASK	사ᄎᄆ	담당	당자	DUE	DUE DAY		2020년 8월			9월				10월				11	윌		
No.	IIILE & INSK	산출물	정	부	START	END	상태	1	2	3 4	1	2	3 4	1	2	3	4 5	1	2	3	4	
3	개발				09.14(월)	11.06(금)																
3.1	데이터베이스 개발				09.14(월)	09.24(목)																
3.1.1	데이터베이스 구축				09.14(월)	09.22(화)																
3.1.1.1	MySQL 개발 환경 구축		서준교		09.14(월)	09.15(화)	완료															
3.1.1.2	Sanction List 정보를 저장할 데이터베이스 설계	DB 파일	서준교		09.16(수)	09.17(목)	완료															
3.1.1.3	Sanction List 데이터 전처리 및 DB에 데이터 삽입	DB 파일	서준교		09.17(목)	09.18(금)	완료															
3.1.3	데미터베이스 연동				09.22(화)	09.24(목)	완료															
3.1.3.1	Django 프로젝트에 DB를 연동하여 데이터 접근		서준교		09.22(화)	09.24(목)	완료															
3.1.3.2	DB파일 원격으로 연동시키기		서준교		09.22(화)	09.24(목)	완료															
3.2.1	회원가입 및 로그인				09.14(월)	09.21(월)	완료															
3.2.1.1	사원번호, 비밀번호를 통한 회원가입 기능 구현		조현아		09.14(월)	09.15(화)	완료															
3.2.1.2	등록된 사원번호로 로그인 / 로그오프 기능 구현		조현아		09.16(수)	09.17(목)	완료															
3.2.1.3	ID / PW 찾기 기능 구현		조현아		09.18(금)	09.21(월)	완료															
3.2.2	데이터베이스 관리				09.25(금)	10.06(화)	완료															
3.2.2.1	Sanction List 카테고리별 조회 기능 구현		서준교		09.25(금)	09.29(화)	완료															
3.2.2.2	Sanction List 데이터에 대한 CRUD 구현		서준교		10.05(월)	10.07(수)	완료															
3.2.3	무역서류 업로드				10.05(월)	10.09(금)	완료															
3.2.3.1	TBML 유사도 검출을 위한 서류 업로드 기능 구현		조현아		10.05(월)	10.09(금)	완료															
3.2.4	무역서류 스캔				10.12(월)	11.06(금)	완료															
3.2.4.1	Google API (OCR)를 이용한 스트링 추출		조현아	서준교	10.12(월)	10.14(수)	완료															
3.2.4.2	Sanction List와 유사한 내용에 대한 boundary box 기능 구현		조현아	서준교	10.15(목)	10.23(금)	완료															
3.2.4.3	무역 서류 일괄 스캔 기능 구현 (optional)		조현아	서준교	10.26(월)	10.30(금)	완료															
3.2.5	부트스트랩 템플릿 적용				09.14(월)	11.06(금)	완료															
3.2.5.1	화면 UI 개선 (부트스트랩 템플릿)		조현아	서준교	09.14(월)	11.06(금)	완료															
3.2.2	프로젝트 중간 점검				10.26(월)	10.27(화)	완료															

프로젝트 개요

<모델 및 알고리즘 개발 >

TASK	TITLE & TASK	사츠므	담당	자	DUE	DAY	상태	2020년 8월				9월			10월				게 e	
No.	TITLE & TASK	산출물	정	부	START	START END 25 1	1 2	2 3	4	1 :	2 :	3 4	1 1	2	3 .	4 5	1	2	3 4	
3.4	유사도 측정 모델 구축				09.07(월)	11.13(금)	완료													
3.4.1	OCR 후처리 모델				09.07(월)	10.07(수)	완료													
3.4.1.1	논문 리서치		최기훈		09.07(월)	09.18(금)	완료													
3.4.1.2	데이터 전처리 및 모델 구성	모델	최기훈		09.21(월)	09.25(금)	완료													
3.4.1.3	모델 유효성 검증 및 결과물 후처리	JSON or XML	최기훈		09.28(월)	10.07(수)	완료													
3.4.2	Sanction List와 유사도 측정 모델				10.12(월)	10.30(금)	완료													
3.4.2.1	논문 리서치		최기훈		10.12(월)	10.16(금)	완료													
3.4.2.2	모델 구성	모델	최기훈		10.19(월)	10.23(금)	완료													
3.4.2.3	모델 유효성 검증	JSON or XML	최기훈		10.26(월)	10.30(금)	완료													
3.4.3	모델 이식				11.02(월)	11.13(금)	완료													
3.4.3.1	방법 리서치		최기훈		11.02(월)	11.04(수)	완료													
3.4.3.2	IT 팀과 조율		최기훈		11.09(월)	11.13(금)	완료													
4	테스트				11.09(월)	11.20(금)														
5	Closing				11.23(월)	11.27(금)														

02

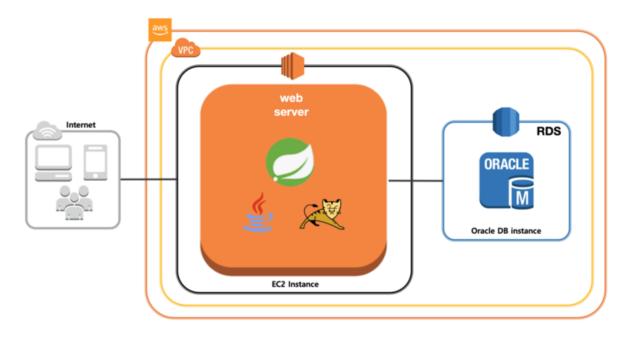
사용 기술

01. AWS RDS

사용기술

Amazon Relational Database Service (RDS)

- AWS에서 제공하는 **클라우드에서 관계형 데이터베이스를 쉽게 설치, 운영 및 확장**할 수 있는 웹 서비스
- 애플리케이션 내에서 관계형 데이터베이스의 설정, 운영, 스케일링을 단순케 하도록 설계됨



- AWS EC2가 서비스되고 있고, RDS 서비스 가 EC2와 연동되어 클라이언트가 EC2에 접 속하여 데이터베이스 정보를 조회하는 구조
- 사용 목적: <u>원격 개발 환경에서 데이터베이스</u> <u>에 상시 접근할 수 있어야 해야 하기 때문</u>에, 클라우드 환경에서 안정적으로 데이터베이스 서비스를 제공하는 AWS RDS 사용

02. MySQL

사용기술

데이터베이스 설계 (MySQL)

- Id, name, type, program으로 총 네 가지의 field로 구성
- type는 제재 대상의 종류(선박,기업,인물 등), program은 제재 프로그램의 명칭을 의미
- Id에 auto_increment 옵션을 적용하여 sanction list에 대한 index 관리에 용이하도록 함

```
mysql> desc sanction;
                                      | Default | Extra
 Field
           Туре
                           Null | Key
  id
            int
                           NO
                                  PRI
                                                   auto_increment
                                        NULL
            varchar(130)
                           YES
                                        NULL
 name
            varchar(20)
                           YES
                                         NULL
 type
                           YES
           varchar(60)
 program
 rows in set (0.01 sec)
```

<sanction 테이블구조>

02. MySQL

사용기술





<미국재무부 sanction list >





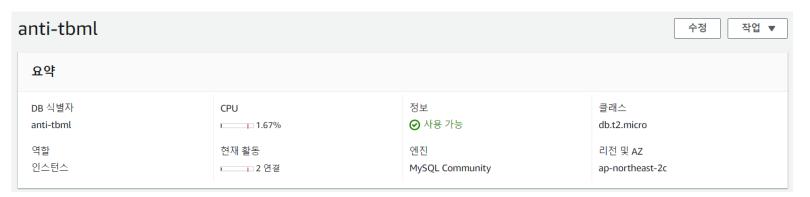
< Django back-end >

```
FIRST DIVISION OF THE SYRIAN ARAB ARMY
XINJIANG PRODUCTION AND CONSTRUCTION CORPS
SUN, Jinlong
PENG, Jiarui
HAMUD, Faruq
TAGWIREI, Kudakwashe Regimond
SAKUNDA HOLDINGS
PARTHIA CARGO LLC
MAHDAVI, Amin
DELTA PARTS SUPPLY FZC
AL SHIBL, Luna
SAATI BIN MOHAMAD NAWZAD, Mohamad Ammar
ISMAIL, Samer
DALAH, Ghaith
SAQR, Fadi
IBRAHIM, Yasser Hussein
TONEY OF THE SYRIAN ARAB ARMY
```

< MySQL Database >

02. MySQL 사용기술

<AWS RDS >







1 2 3 4 ... 82 83 8291 sanction lists

8199	MORON HERNANDEZ, Santiago Jose	individual	VENEZUELA
8198	KALAI, Nader	individual	SYRIA] [SYRIA-CAESAR
8197	AL-ZUBAIDI, Khaled	individual	SYRIA] [SYRIA-CAESAR
8196	EBLA HOTEL	-0-	SYRIA
8195	CASTLE INVESTMENT HOLDING	-0-	SYRIA
8194	ZUBAIDI AND QALEI LLC	-0-	SYRIA
8193	TELEFOCUS CONSULTANTS INC	-0-	SYRIA

03. Google Vision API

사용기술

Google Cloud[©] Vision API

- REST 및 RPC API를 통해 선행 학습된 강력한 **머신러닝 모델 기반** API
- 이미지에서 **텍스트를 감지하고 추출**하기 위한 목적으로 사용
- 해당 문서에 대한 페이지, 블록, 단락, 단어, 줄바꿈 정보를 JSON
 형태로 반환





03. Google Vision API

사용기술

< JSON data from Google Vision API >

```
'ORIGINAL TEL : SWIFT : KOEX': {'place': [(1269, 66), (1506, 66), (1506, 171), (1269, 171)], 'danger': [], 'similar_word': []}, '*** This is a constituent and integral part of advice No
      A12WH191': {'place': [(133, 274), (1234, 268), (1234, 299), (133, 305)], 'danger': [], 'similar_word': []}, '( OP )': {'place': [(1347, 270), (1399, 270), (1399, 297), (1347, 297)], 'danger': [], 'similar_word': []}, '( OP )': {'place': [(1347, 270), (1399, 270), (1399, 297), (1347, 297)], 'danger': [], 'similar_word': []}, '( OP )': {'place': [(1347, 270), (1399, 270), (1399, 297), (1347, 297)], 'danger': [], 'similar_word': []}, '( OP )': {'place': [(1347, 270), (1399, 270), (1399, 297), (1347, 297)], 'danger': [], 'similar_word': []}, 'similar_wo
 nger': [], 'similar_word': []}, '* KRYEE ===': {'place': [(133, 331), (273, 331), (273, 341), (133, 341)], 'danger': [], 'similar_word': []}, 'PREPAID INDICATING LC NUMBER , NUMBER NIE OF
    APPLICANT': {'place': [(233, 362), (1107, 357), (1107, 380), (233, 385)], 'danger': [], 'similar_word': []}, '09981': {'place': [(234, 402), (310, 401), (310, 422), (234, 423)], 'danger'
      [], 'similar_word': []}, 'ETABLISHED IN SOUTH KOREA .': {'place': [(496, 402), (919, 400), (919, 422), (496, 424)], 'danger': [], 'similar_word': []}, 'COMMERCIAL': {'place': [(234, 482
   ), (396, 482), (396, 511), (234, 511)], 'danger': [], 'similar_word': []}, 'INVOICE': {'place': [(416, 482), (528, 482), (528, 511), (416, 511)], 'danger': [], 'similar_word': []}, 'IN':
   ['place': [(238, 1479), (267, 1479), (267, 1503), (238, 1503)], 'danger': [], 'similar_word': []}, '03': {'place': [(598, 482), (626, 482), (626, 511), (598, 511)], 'danger': [], 'similar
  _word': []}, 'ORIGINALS': {'place': [(646, 482), (792, 482), (792, 511), (646, 511)], 'danger': [], 'similar_word': []}, 'INDICATED': {'place': [(814, 482), (960, 482), (960, 511), (814,
511)], 'danger': [], 'similar_word': []}, 'AMOUNT': {'place': [(764, 1561), (861, 1561), (861, 1589), (764, 1589)], 'danger': [], 'similar_word': []}, ',': {'place': [(784, 1521), (789,
521), (789, 1549), (784, 1549)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'OF': {'place': [(696, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 649), (793, 675), (696, 675)], 'danger': [], 'similar_word': []}, 'ORIGIN': {'place': [(696, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649), (793, 649),
    [(881, 1561), (912, 1561), (912, 1589), (881, 1589)], 'danger': [], 'similar_word': []}, 'GOODS': {'place': [(283, 526), (361, 526), (361, 552), (283, 552)], 'danger': [], 'similar_word': []}, 'GOODS': {'place': [(283, 526), (361, 526), (361, 552), (283, 552)], 'danger': [], 'similar_word': []}, 'goods': [], 'similar_word': 
   ': []}, 'AND': {'place': [(1162, 1479), (1211, 1479), (1211, 1503), (1162, 1503)], 'danger': [], 'similar_word': []}, 'PAYMENT': {'place': [(448, 525), (560, 525), (560, 552), (448, 552)]
       'danger': [], 'similar_word': []}, 'TERMS': {'place': [(580, 525), (658, 525), (658, 551), (580, 551)], 'danger': [], 'similar_word': []}, 'BY': {'place': [(516, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546, 1521), (546,
       , 1549), (516, 1549)], 'danger': [], 'similar_word': []}, 'LETTER': {'place': [(729, 524), (826, 524), (826, 550), (729, 550)], 'danger': [], 'similar_word': []}, 'CREDIT': {'place': [(1
163, 1521), (1260, 1521), (1260, 1549), (1163, 1549)], 'danger': [], 'similar_word': []}, 'SIGNED': {'place': [(235, 694), (330, 694), (330, 715), (235, 715)], 'danger': [], 'similar_word
    : []}, 'STAMPED': {'place': [(236, 567), (347, 567), (347, 590), (236, 590)], 'danger': [], 'similar_word': []}, 'THE': {'place': [(1229, 1687), (1275, 1687), (1275, 1707), (1229, 1707)]
        'danger': [], 'similar_word': []}, 'BENEFICIARY': {'place': [(779, 692), (959, 691), (959, 713), (779, 714)], 'danger': [], 'similar_word': []}, 'COPY': {'place': [(234, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651), (297, 651)
```



'ACCEPTANCE': {'place': [(334, 1521), (496, 1521), (496, 1549), (334, 1549)], 'danger': [1.0], 'similar_word': ['ACCEPTANCE']},

ORIGINAL

TEL: SWIFT:KOEX

```
*** This is a constituent and integral part of advice No : A12WH191000040(OP)
:44E Port of Loading/Airport of Departure :PORT SOUTH KOREA
:44F Port of Discharge/Airport of Destinati:PORT ORAN
                                       :191210
:44C Latest Date of Shipment
:45A Description of Goods and/or Services :
        PORT ORAN (INCOTERMS 2010)
     FULL AUTOMATIC DISPOSALE SYRINGE PRODUCTION IN LINE FOR
     FOLLOWING TO THE PROFORMA INVOICE NO: DMT0916
     DATED:16/09/2019
:46A Documents Required
                                       :
     95 PCT OF THE AMOUNT OF THE L/C OF USD 190.000,00 PAYABLE
     AGAINST PRESENTATION OF THE FOLLOWING DOCUMENTS:
     COMMERCIAL INVOICE IN 03 ORIGINALS INDICATED AMOUNT, ORIGIN
     OF GOODS AND PAYMENT TERMS BY LETTER OF CREDIT, SIGNED AND
     STAMPED BY THE BENEFICIARY
     2/3 BILL OF LADING ORIGINALS ON BOARD TO THE ORDER OF BEA
     BRANCH TLEMCEN EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT
     PREPAID INDICATING LC NUMBER, NUMBER NIF OF APPLICANT
                     ESTABLISHED IN SOUTH KOREA.
     ORIGINAL CERTIFICATE OF ORIGIN (ORIGIN SOUTH KOREA) DELIVERED
     AND SIGNED BY CHAMBER OF COMMERCE OF BENEFICIARY COUNTRY
     COPY CERTIFICATE OF CONFORMITY SIGNED AND STAMPED BY
     BENEFICIARY
     COPY CERTIFICATE OF WARRANTY ETABLISHED BY THE MANUFACTURER
               TECHNO) .
     PACKING LIST SIGNED AND STAMPED BY THE SELLER (
      TECHNO) IN 2 COPIES
     BENEFICIARY'S ATTESTATION CERTIFYING THAT THE FOLLOWING
     DOCUMENTS HAVE BEEN SENT BY SHIP BAG (EXPRESS MAIL) UPS/DHL
     DIRECTLY TO THE APPLICANT:
     1/3 BILL OF LADING ORIGINAL ON BOARD TO THE ORDER OF
     BRANCH TLEMCEN EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT
*** This advice to be continued next page
                                                  (A12WH1910
```

.

ORIGINAL
TEL:
SWIFT: KOEX

*** This is a constituent and integral part of advice No.

*** This is a constituent and integral part of advice No.

*** This is a constituent and integral part of advice No.

*** This is a constituent and integral part of advice No. 44 € 44 E 44 E 44 E 44 C 45 X Port of Loading / Airport of Departure PORT SOUTH KOREA
PORT of Loading / Airport of Departure PORT SOUTH KOREA
PORT of Discharge / Airport of Destinate PORT ORAN
Legal to Discharge / Airport of Destinate PORT ORAN
Legal to Description of Statement of Destinate PORT ORAN
Latest Date of Shipment :191210 port oran (INCOTERMS 2010) : 45A FULL AUTOMATIC DISPOSALE SYRINGE PRODUCTION IN LINE FOR FULL AUTOMATIC DISPOSALE SYRINGE PRODUCTION IN LINE FOR FOLLOWING TO THE PROFORMA INVOICE NO DMT0916
DATED TO THE PROFORMA INVOICE NO DMT0916 DATED:16/09/2019 BSCEPANOF PROLETATION OF THE L/C OF USD 190,000,00 PAYABLE AGAINST PRESENTATION OF THE FOLLOWING DOCUMENTS: COMMERCIAL INVOICE IN 03 ORIGINALS INDICATED AMOUNT, ORIGIN COMMERCIAL INVOICE IN 03 ORIGINALS INDICATED AMOUNT, ORIGIN OF GOODS AND PAYMENT LEAMS BY LETTER OF CREDIT SIGNED AND SECOND AND PAYMENT LEAMS BY LETTER OF CREDIT SIGNED AND STAMPED BY THE BENEFICIARY 2/3 BILL OF LADING ORIGINALS ON BOARD TO THE ORDER OF BEA
2/3 BILL OF LADING ORIGINALS ON BOARD TO THE ORDER OF BEA
BRANCH TLEMCEN EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT
BRANCH THEOREM EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT
PREPAID INDICATING COMMERC NITH OF APPLICANT
PREPAID INDICATING COMMERC NITH OF APPLICANT
ESTABLISHED IN SOUTH KOREA. ORIGINAL CERTIFICATE OF ORIGIN (ORIGIN SOUTH KOREA) DELIVERED ORIGINAL CERTIFICATE OF ORIGIN (ORIGIN SOUTH KOREA) DELIVERED AND SIGNED BY CHAMBER OF COMMERCE OF BENEFICIARY COUNTRY AND SIGNED BY CHAMBER OF COMMERCE OF BENEFICIARY COUNTRY COPY CERTIFICATE OF CONFORMITY SIGNED AND STAMPED BY BENEFICIARY COPY CERTIFICATE OF CONFORMITY SIGNED AND STAMPED BY BENEFICIARY COPY CERTIFICATE OF WARRANTY ETABLISHED BY THE MANUFACTURER (
COPY CERTIFICATE OF WARRANTY ETABLISHED BY THE MANUFACTURER
(TECHNO) PACKING LIST SIGNED AND STAMPED BY THE SELLER (
PACKING LIST SIGNED AND STAMPED BY THE SELLER (
TECHNO) IN 2 COPIES BENEFICIARY'S ATTESTATION CERTIFYING THAT THE FOLLOWING BENEFICIARY ATTESTATION CERTIFYING THAT THE FOLLOWING DOCUMENTS HAVE BEEN SENT BY SHIP BAG (EXPRESS MAIL) DESIDED TO THE APPLICANT BY SHIP BAG (EXPRESS MAIL) DESIDE 1/3 BILL OF LADING ORIGINAL ON BOARD TO THE ORDER OF
1/3 BILL OF LADING ORIGINAL ON BOARD TO THE ORDER OF
BRANCH TLEMCEN EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT
BRANCH TLEMCEN EL KIFFANE NOTIFY APPLICANT MARKED FREIGHT *** This advice to be continued next page (A12WH1910 (A12WH1910 *** This advice to be continued next page

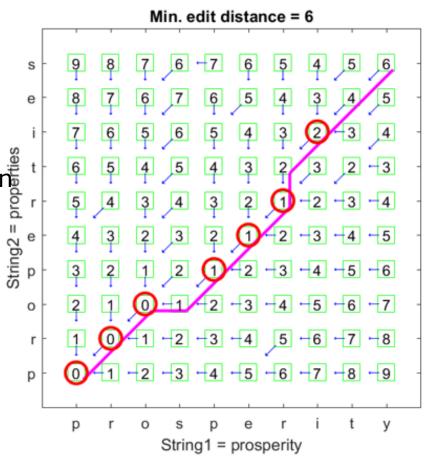
.

04. Edit Distance

사용기술

Edit Distance 알고리즘

- 두 문자열의 유사도를 판단하는 알고리즘
- 두 문자열이 같아지기 위해서 필요한 연산의 최소 횟수를 계산
- 시간 복잡도: O (MN), 유사도: (len (str1) table[len2][len1]) / length tance(s₁,s₂)



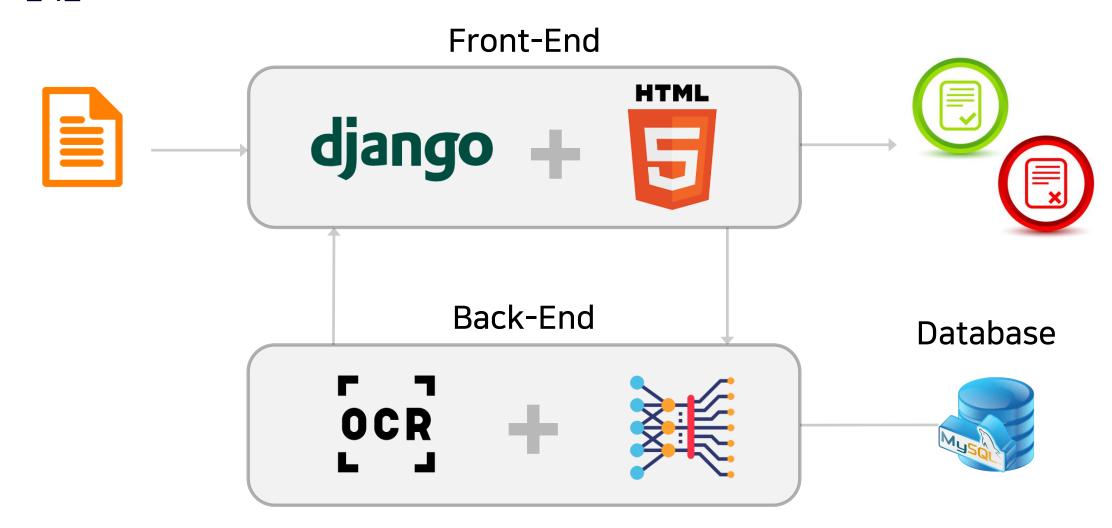
< Pseudo Code of Edit Distance Algorithm >

03

결과물

01. 시스템 구조

결과물



02. 상세 기능

결과물

프로젝트 핵심 기능

- 1. 로그인 및 회원가입을 통한 인증 기능
- 2. Sanction List에 대한 CRUD 및 검색 기능
- 3. 문서 스캔을 통한 유사도 측정 기능



< Django 프로젝트 구조>



04

결론

01. 기대 효과

결론

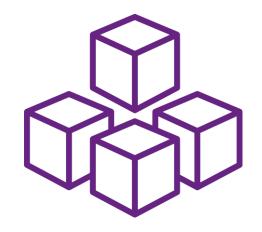
1. 금융 범죄 예방 효과

■ 자금 세탁 및 관련 금융 범죄에 연루되어 있을 가능성이 있는 문서를 사전에 탐지함으로써 **경제적 손실 방지 가능**



2. 데이터 사용 편의성

■ 웹에서 Sanction List에 대한 CRUD 작업이 클라우드 기반 DB와 실시 간 연동되어 데이터 확장이 매우 용이하여 **데이터 관리 편의성 제고**



02. 개선 방안

결론

1. 알고리즘 시간 복잡도 개선

- 시간 복잡도가 O(MN)인 Edit Distance 알고리즘에 의존하기 때문에 Sanction List의 수가 늘어날 수록 유사도 계산 시간이 급격히 증가
 - ⇒ 관련 알고리즘 개선 필요

```
def find_similar_word(word, sanc_list, thresh):
    similar_word = {}
    for sanc in sanc_list:
        similarity = dist_sen(word, sanc, thresh)

    if similarity > thresh:
        similar_word[sanc] = similarity

return list(similar_word.keys()), list(similar_word.values())
```

2. 자연어 처리 개선

범용적인 단어를 Sanction List에 존재하는 단어로 인식하는 경우가 존재
 ⇒ 품사를 기준으로 자연어 처리 개선 필요

Thank you ANTI-TBML