



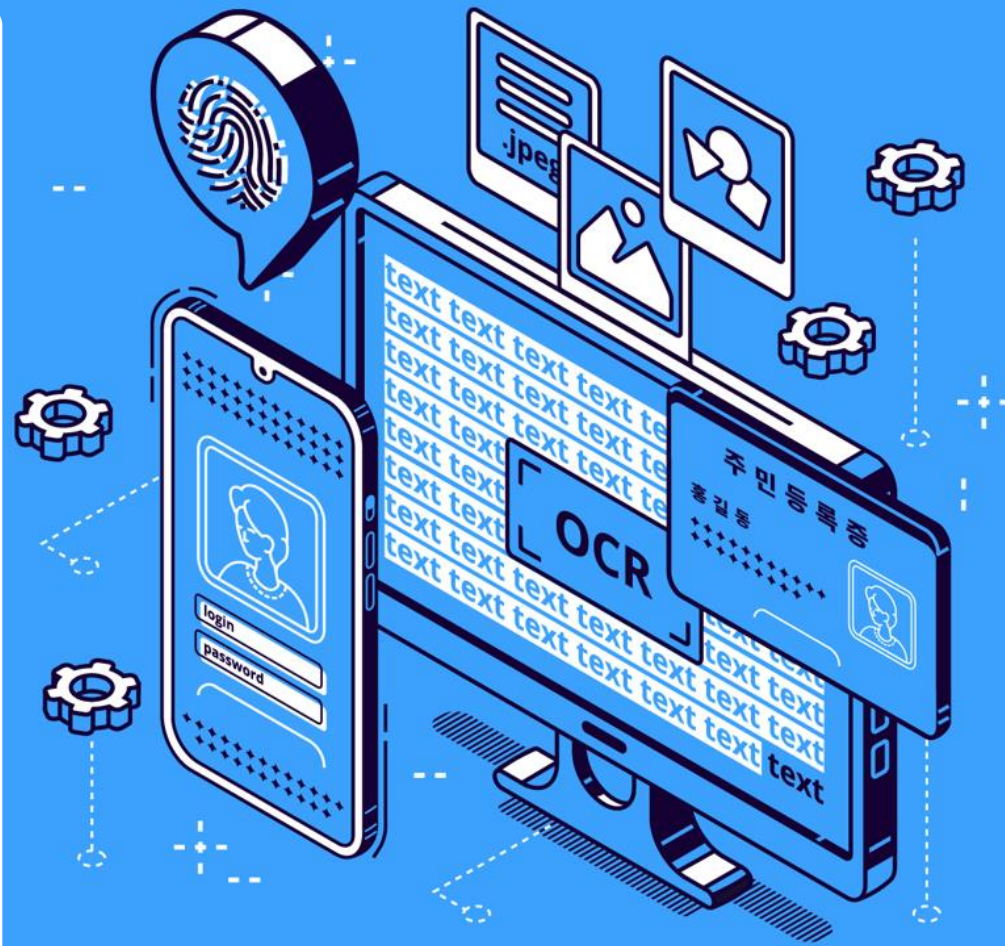
**PISFAIR2022**  
PERSONAL INFORMATION SECURITY FAIR

# 이미지 개인정보 검출부터 행정문서 분류까지

**COMTRUE**<sup>®</sup>  
Technologies

(주)컴트루테크놀로지

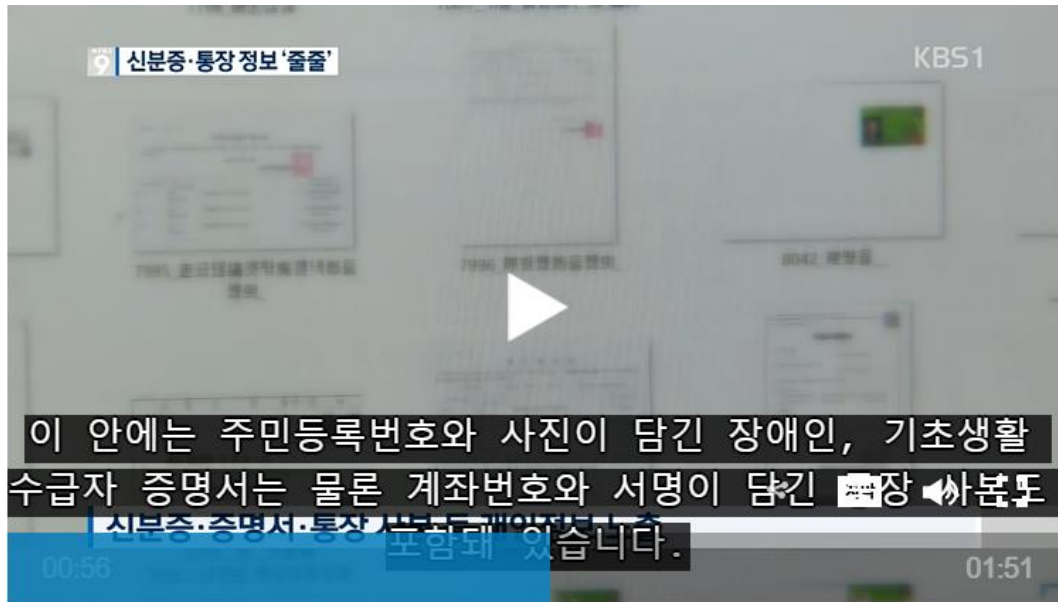
컴트루테크놀로지 [www.comtrue.com](http://www.comtrue.com)



# 이미지 형태의 개인정보 유출 위험성



구청 홈페이지 신분증·통장 정보 '줄줄'... 암호만 걸어놓았어도!



신분증 복사본 같은 **이미지 파일**을  
암호화할 수 있는 시스템을  
갖추고 있지 못하다고 털어놨...

KBS NEWS 보도자료 (<https://news.kbs.co.kr/news/view.do?ncd=4009247>)

# 개인정보 노출탐지 유형 확대, 이미지 추가



## | KISA 개인정보 노출 탐지 시 2019년부터 탐지유형 확대 및 이미지 추가 적용

- 이미지 파일 형태 개인정보를 신규로 탐지
- 기존 4종 (주민등록번호, 여권번호, 운전면허번호, 외국인등록번호)  
-> 8종으로 확대 (휴대전화번호, 계좌번호, 신용카드번호, 건강보험번호)

## | 공공기관 개인정보 관리수준 진단 대비 솔루션

- 공공기관의 개인정보 관리체계 및 유출 예방활동 진단 제도
- 적용대상 : 중앙행정기관, 산하 공공기관, 지방자치단체, 지방공기업 등
- 이미지 개인정보 진단 이슈

➡ 고유식별정보 및 카드번호, 계좌번호, 휴대폰번호, 건강보험번호 포함

➡ 이미지 포함 검출

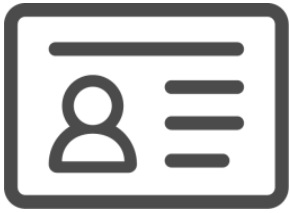
### 다. 숫자로 보는 개인정보 노출

공공부문에서의 개인정보 노출 탐지 추이를 살펴보면 개인정보 보호법 상 주민등록번호 등 주요 개인정보 처리에 대한 규제가 강화되고, 교육 등을 통해 개인정보 처리자의 인식이 제고됨에 따라 개인정보 노출 홈페이지 수는 지속적으로 감소세를 보이고 있다. 다만, 2019년에는 개인정보 탐지 유형을 고유식별정보 4종(주민등록번호, 여권번호, 운전면허번호, 외국인등록번호)에서 8종으로 확대(휴대전화번호, 계좌번호, 신용카드번호, 건강보험번호 추가)하고 이미지 파일 형태의 개인정보를 신규로 탐지함에 따라 노출 탐지 홈페이지 수가 875개로 다소 증가하였다.

그림 5 공공부문 개인정보 노출 탐지·삭제(홈페이지 수 기준)



# 다양한 형태의 이미지 내 개인정보 보호



고유식별정보 外  
961230-2111111



바이오정보  
지문 등



행정문서  
가족관계증명서 등

# 이미지 개인정보보호를 위한 OCR 기술



신분증, 카드, 개인/기업 행정문서, 병원진료비영수증, 계약서, 동의서, 체크박스, 필기체 등 다양한 영역 검출

# 인공지능 OCR 레퍼런스



질병관리청

이미지 개인정보(OCR) 검출

삼성증권



고객제출서류 문서자동분류

kt M mobile

신분증 OCR



인공지능 사물인식(PET병분류)

kakao

신분증 OCR



캡차(Captcha)인식

# 자체 개발 인공지능 OCR을 활용하는 이유



ex. 주민등록번호

111111-2111111

기  
존  
O  
C  
R

오픈소스 OCR의 한계 (한글인식)

이미지 왜곡, 저화질 문제

스캔된 이미지에 최적화



인공지능 OCR 자체개발, 인식률 개선

문서 영역, 글 영역 검출 신경망 개발

모바일 촬영본, 영상에서도 검출

자  
체  
개  
발  
O  
C  
R

# PC, 웹 서버, 파일서버 및 다양한 시스템 연동



PC 이미지 개인정보 검출



홈페이지 이미지 개인정보 검출



파일서버 이미지 개인정보 검출



# Case. PC 내 이미지 개인정보보호



질병관리청

## PC 이미지 개인정보(OCR) 검출

- 2020년 도입
- 약 3,000 Copy

### 에이전트



개인 PC 별 검사

검출/암호화/삭제/격리

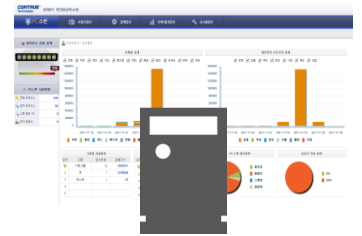
출력제어/매체 제어



개인정보 검색 결과 전송  
출력/매체 이용 결과 전송

정보보안 정책 전송

### 관리서버

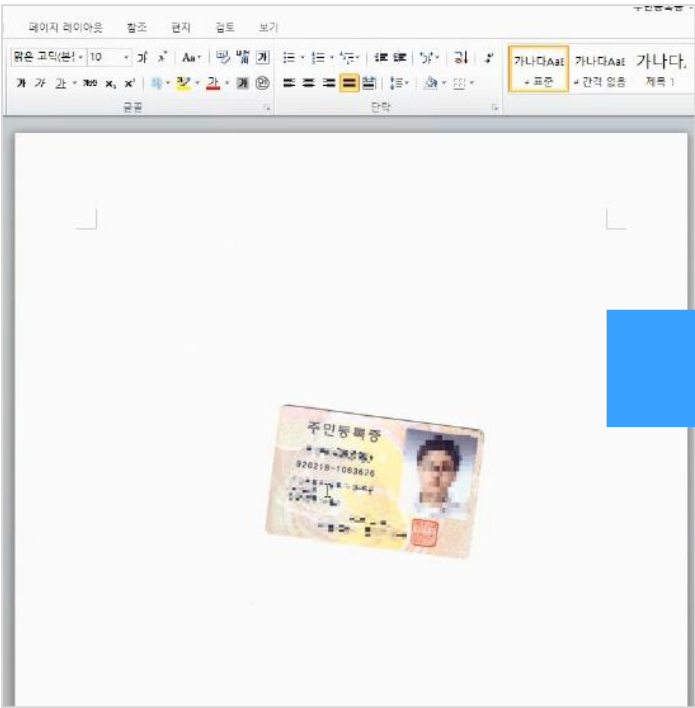


각 PC 별 검출 내역 확인

각 PC별 출력/매체 사용 현황

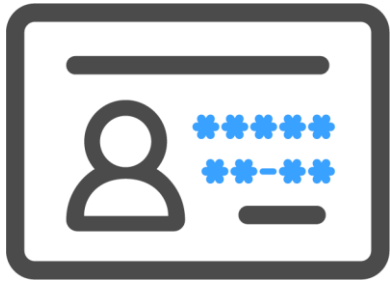
통계 기능을 통한 보고서 출력

# 이미지, 문서파일 내 개인정보 검출



파일명		주민등록증.docx
개인정보 검출	1 건	
키워드 검출	0 건	
의심문서 검출	0 개	
번호	종류	검출항목
1	주민번호	<a href="#">920218-1</a>

# 이미지 개인정보 후처리 방식



비식별화 (마스킹)



암호화



완전삭제

# Case. 홈페이지 내 이미지 개인정보보호



웹서버 이미지 개인정보(OCR)  
검출 및 필터링

제주ON

Jeju

제주특별자치도

Jeju Special Self-Governing Province

비회원 로그인

로그인

Language

도정뉴스

신문고

소통/참여/인권

정보공개

제주소개

일자리/중소기업

분야별정보

관광정보

제주자치도에 바란다

도지사 핫라인

관광불편민원접수

공직자비리신고(공직자  
부동산 투기 신고 포함)

공익신고

甲의 부당행위신고센터

부정청탁 및 금품등  
수수행위 신고

궁금합니다

신고센터

선물신고센터

예산낭비신고센터

부정불량식품신고

공무원부조리신고(감사위원  
회)

지방기업 규제 애로 신고센터

국민신문고

국민신문고 제안신청

국민신문고 공모제안

국민신문고 나의제안

관광불편민원접수

HOME > 신문고 > 도지사 핫라인 > 관광불편민원접수

제주특별자치도에서는 공정하고 안심할 수 있는 제주관광을 실현하고자 관광불편 사항에 대한「신고센터」를 설치·운영하고 있습니다.

본 게시판은 실명제로 운영되는 「게시판」으로 제주 여행 중 느낀 불편사항에 대해 자유롭게 의견을 제시하여 주시면 적극 개선하겠습니다.

전체(1718건)

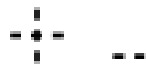
이동

전체 1477 건 :: 현재 1 페이지 / 전체 148 페이지

No.	제목	담당	작성자	작성일	답변	조회
1,698	제주상성렌트카 횡포 및 비방 신고, 보상요. N	자지도 교통정책과	손ㅇㅇ	2021.08.30	답변중	34
1,697	택시 불친절 신고합니다. N		허광호	2021.08.30	대기중	20
1,696	제주스타렌트카 환불정상적으로 되게 해주. N		장ㅇㅇ	2021.08.30	완료	5
1,695	성산항 여객 터미널의 불친절 N		오ㅇㅇ	2021.08.30	대기중	50



# 이미지 파일 다운로드 차단



개인정보가 포함된 이미지



게시물의  
첨부파일 다운로드 차단



원문

830422-1185600



83\*\*\*\*-\*\*\*\*\*

## 개인정보 영역 비식별화(\*치환)

# 공공기관 클라우드 전환 정책 배경



2025년까지 모든 행정·공공기관 정보시스템 인터넷  
기반 자원 공유(클라우드)로 전환

**- 행정공공기관정보자원인터넷기반자원공유(클라우드) 전환통합추진계획발표 -**

□ 오는 2025년까지 모든 행정·공공기관의 정보시스템이 클라우드 기반 통합관리 운영 환경으로 전환되어 다양한 국민수요에 대응하고 안정적인 대국민 서비스를 제공하게 된다.

□ 행정안전부(장관 전해철)는 ‘행정·공공기관 정보자원 클라우드 전환·통합 추진계획’을 7월 27일 발표하고, 올해부터 5년 동안 행정기관과 공공기관이 운영 중인 모든 정보시스템(10,009개)을 클라우드로 전면 전환·통합한다고 밝혔다.



 설록홈즈®  
**PrivacyCenter** on CLOUD V5.0



# 다양한 형태의 이미지 내 개인정보 보호



고유식별정보 外  
961230-2111111



바이오정보  
지문 등

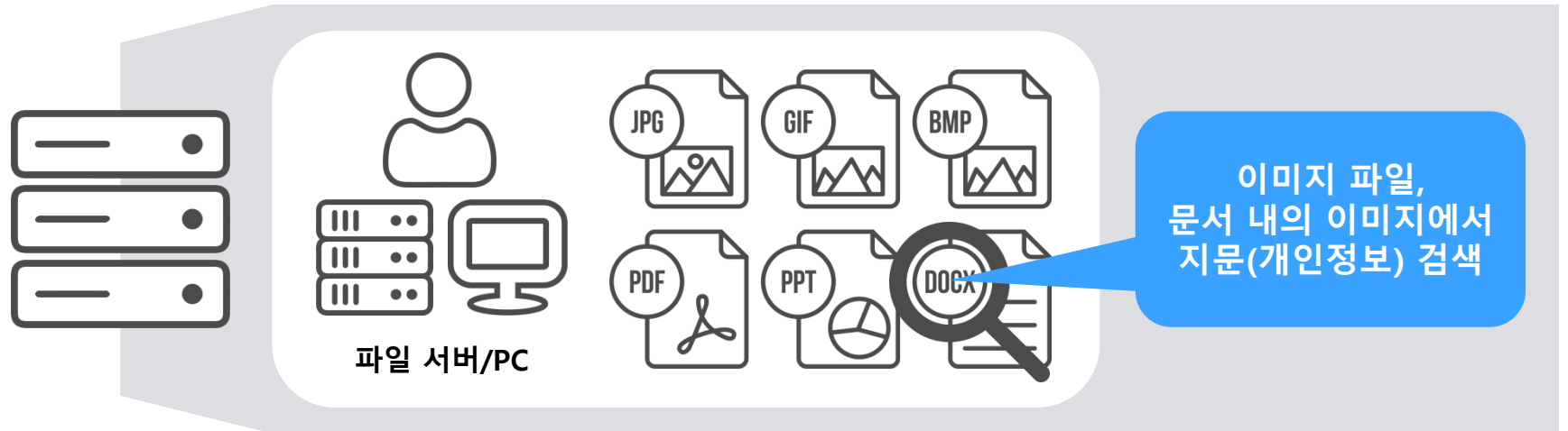


행정문서  
가족관계증명서 등

# Case. 서버 내 바이오정보 검출, 비식별화



서버 내 지문검출 및 파기  
(완전삭제, 검게칠하기)



# Case. 바이오 정보 검출 / 비식별화\_통신사 콜센터

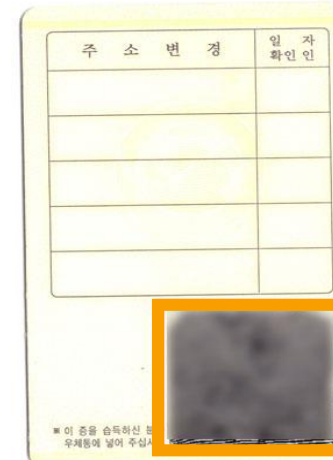
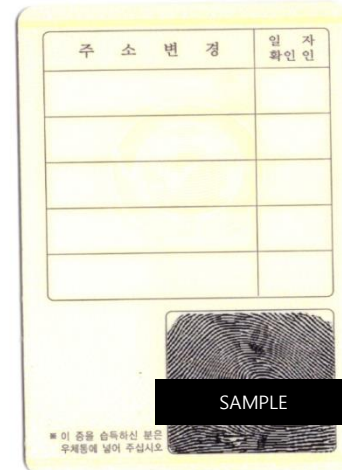


## 지문 검출 및 마스킹 지원

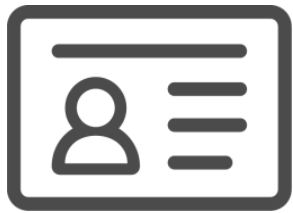


국내 유일 PDF 파일 내  
지문 비식별화 지원

지문	합계	상태	파일 후처리		지문 후처리
0	1	검출	암호화	완전삭제	미검출
0	297	검출	암호화	완전삭제	미검출
1	1	검출	암호화	완전삭제	처리
0	1	검출	암호화	완전삭제	미검출
0	297	검출	암호화	완전삭제	미검출



# 다양한 형태의 이미지 내 개인정보 보호



고유식별정보 外  
961230-2111111

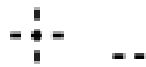


바이오정보  
지문 등



행정문서  
가족관계증명서 등

# 이미지 형태의 개인정보 유출 위험성



주 권 등 록 표  
( 등 본 )

2014년 11월 11일

대전광역시 서구청장

주민등록번호	880101-11111111
성명	홍길동
주소	서울시 마포구 ...
생년월일	1980-01-01
성별	남
직업	회사원
전화번호	02-1234-5678
이메일	honggil@gmail.com
가족관계	배우자: 홍길남
학력	고등학교 졸업
종교	불교
건강상태	건강
주요취미	독서, 여행
주요경력	2010.01 ~ 2012.12: 삼성전자 근무
주요경력	2013.01 ~ 2015.12: 삼성전자 근무
주요경력	2016.01 ~ 2018.12: 삼성전자 근무
주요경력	2019.01 ~ 2021.12: 삼성전자 근무
주요경력	2022.01 ~ 2024.12: 삼성전자 근무

대전광역시 서구청장

본 주민등록증은 주민등록번호와 성명, 주소, 생년월일, 성별, 직업, 전화, 이메일, 가족관계, 학력, 종교, 건강상태, 주요취미, 주요경력 등을 포함하고 있습니다.

880101-11111111

개인정보 검출

880101-1\*\*\*\*\*

개인정보 미검출

이름  
홍길동

주소  
서울시 마포구 ...

개인정보  
홍길동 집주소

위험!

# 개인정보 보호·활용 기술개발 챌린지 최우수상 수상!



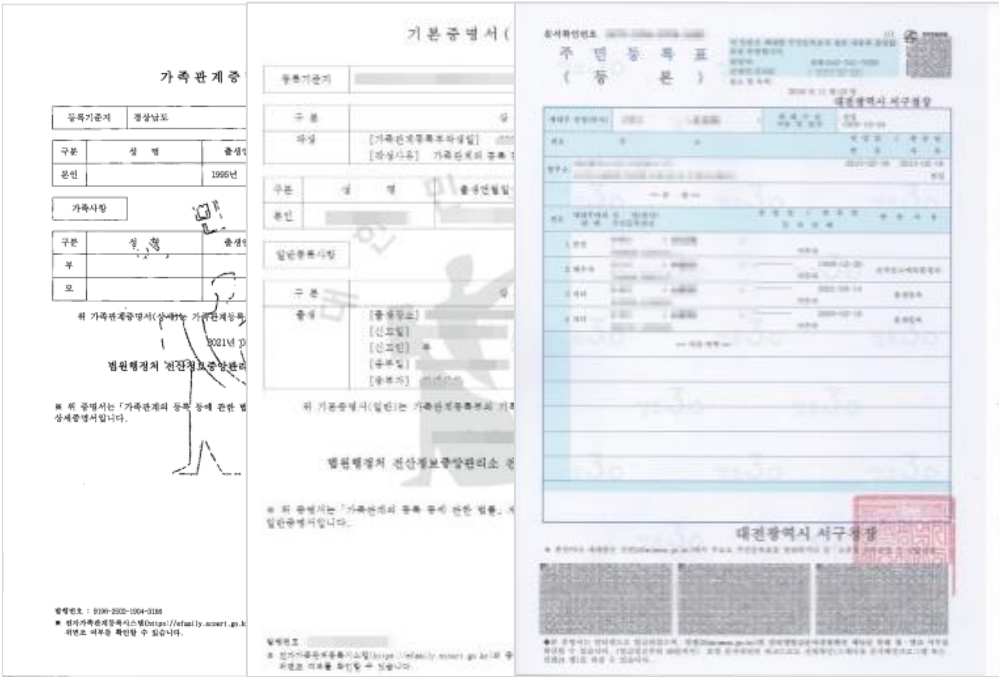
“인공지능 기반 이미지 문서 분류 및 개인정보 비식별화”



# 개인정보 행정문서 검출(분류) 소개 \*유일기능



신분증



행정문서



# 제2회 삼성금융 Open Collaboration 최우수상 수상



단계적 PoC 진행 후, 데모데이를 통해 삼성 금융 OC 전 과정 '최우수상' 수상

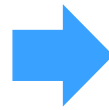


# Case. 행정문서 분류 디지털 창구 연동



개인정보 서류 스캔

<오프라인 창구>



삼성증권

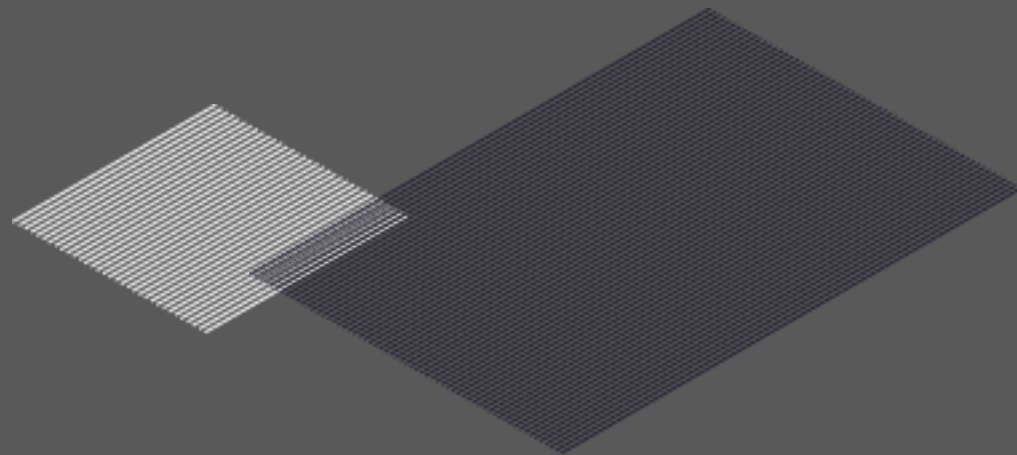


디지털 창구 프로그램 저장 / 열람

문서분류 신경망 연동 API에서 분류 수행

- 디지털 창구 프로그램에 자동 입력 -

# 다양한 시스템과의 연동을 통한 이미지 개인정보보호 조치



# 업무 시스템 내 개인정보 유출위협



웹 서버  
(홈페이지)

전자결재/그룹  
웨어

망연계  
솔루션

이메일  
시스템

문서 관리  
시스템

개인정보의 유출 위험이 있는 다양한 공개 업무 시스템

# 이미지스캔 구축 사례



질병관리청



산업통상자원부



곡성군  
GOKSEONG COUNTY



동작구



사천시



강동구



강서구



SAN CHEONG



안녕  
자연의 창녕



FULL LIFE

GIMPO



yeong yang 영양



경주  
김해



Jeju

제주특별자치도



송파구

SONGPA-GU



전라남도



대덕구



남해군



옥천군

OKCHON GUN



한국임업진흥원  
Korea Forestry Promotion Institute



증권



키움저축은행

감사합니다.

컴트루테크놀로지 부스(중앙 B03)

**COMTRUE**<sup>®</sup>  
Technologies

(주)컴트루테크놀로지