

G-PRIVACY 2024

2024 정부·공공·기업 개인정보보호&정보보안 컨퍼런스

# AI 시대의 통합 개인정보보호 전략

주식회사 파수 | 최필준 팀장

2024.03

FAS00

# G-PRIVACY 2024

2024 정부·공공·기업 개인정보보호&정보보안 컨퍼런스

AI 시대의 통합 개인정보보호 전략

## contents

I 개인정보 데이터 현황

II 개인정보 보안 관리의 현주소

III AI시대에 맞는 개인정보 보호 전략

# I

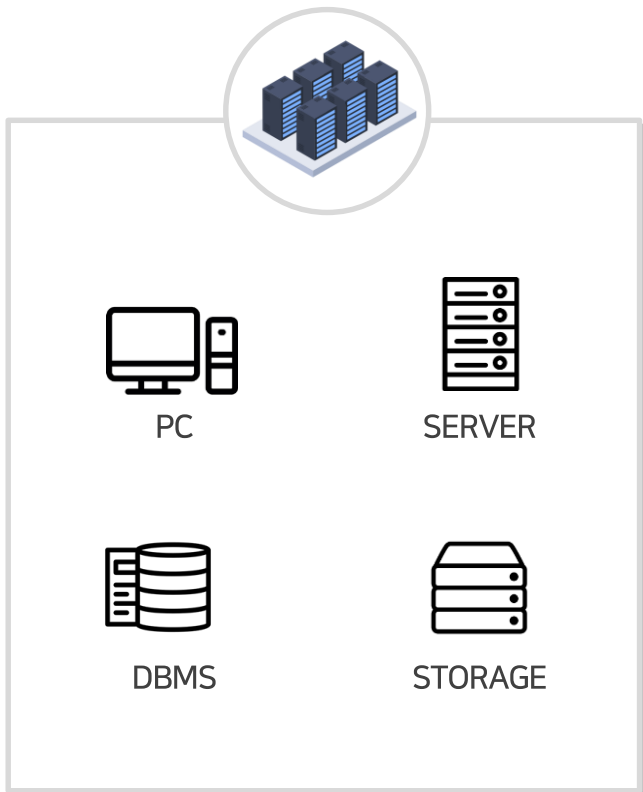
## 개인정보 데이터 현황

FAS00

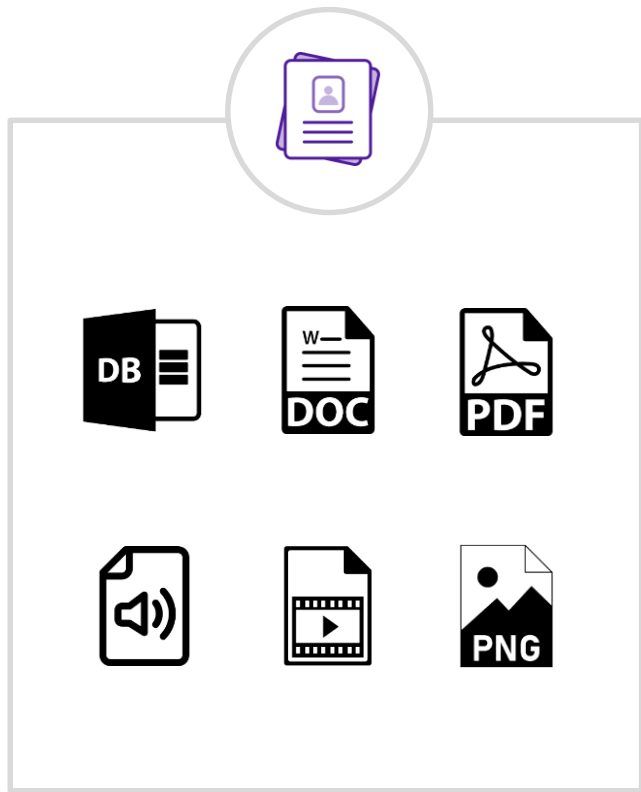
# I-1 개인정보 데이터 현황

## 보유 개인정보의 식별

### 개인정보 보관 장치



### 개인정보 저장 형태



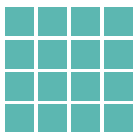
### 개인정보 식별 유형

구분	설명
인적 사항	• 성명, 주민번호, 주소, 연락처, 출생지 등
신체적 정보	• 얼굴, 홍채, 음성, 지문, 키, 몸무게 등
정신적 정보	• 사상, 신조, 종교, 정단, 노조 가입, 가치관 등
사회적 정보	• 학력, 병역여부, 직장, 전과, 범죄기록 등
재산적 정보	• 급여, 대출, 소유주택, 자동차, 계좌번호 등
기타 정보	• 이메일, GPS기록, 통화내역, 로그파일 등

# I-2 개인정보 데이터 현황

## 정형화 정도에 따른 개인정보 구분

### 정형 데이터 (Structure Data)



ID	name	dept_name	salary
22222	Einstein	Physics	95000
12121	Wu	Finance	90000
32343	El Said	History	60000
45565	Katz	Comp. Sci.	75000
98345	Kim	Elec. Eng.	80000
76766	Crick	Biology	72000
10101	Srinivasan	Comp. Sci.	65000
58583	Califieri	History	62000
83821	Brandt	Comp. Sci.	92000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
76543	Singh	Finance	80000

### 반정형 데이터 (Semi-Structured Data)



```

- <employees>
- <person id="1392">
  <name>John Smith</name>
  <dob>1974-07-25</dob>
  <start-date>2004-08-01</start-date>
  <salary currency="USD">35000</salary>
</person>
- <person id="1395">
  <name>Clara Tennison</name>
  <dob>1968-03-15</dob>
  <start-date>2003-05-16</start-date>
  <salary currency="USD">27000</salary>
</person>
</employees>
    
```

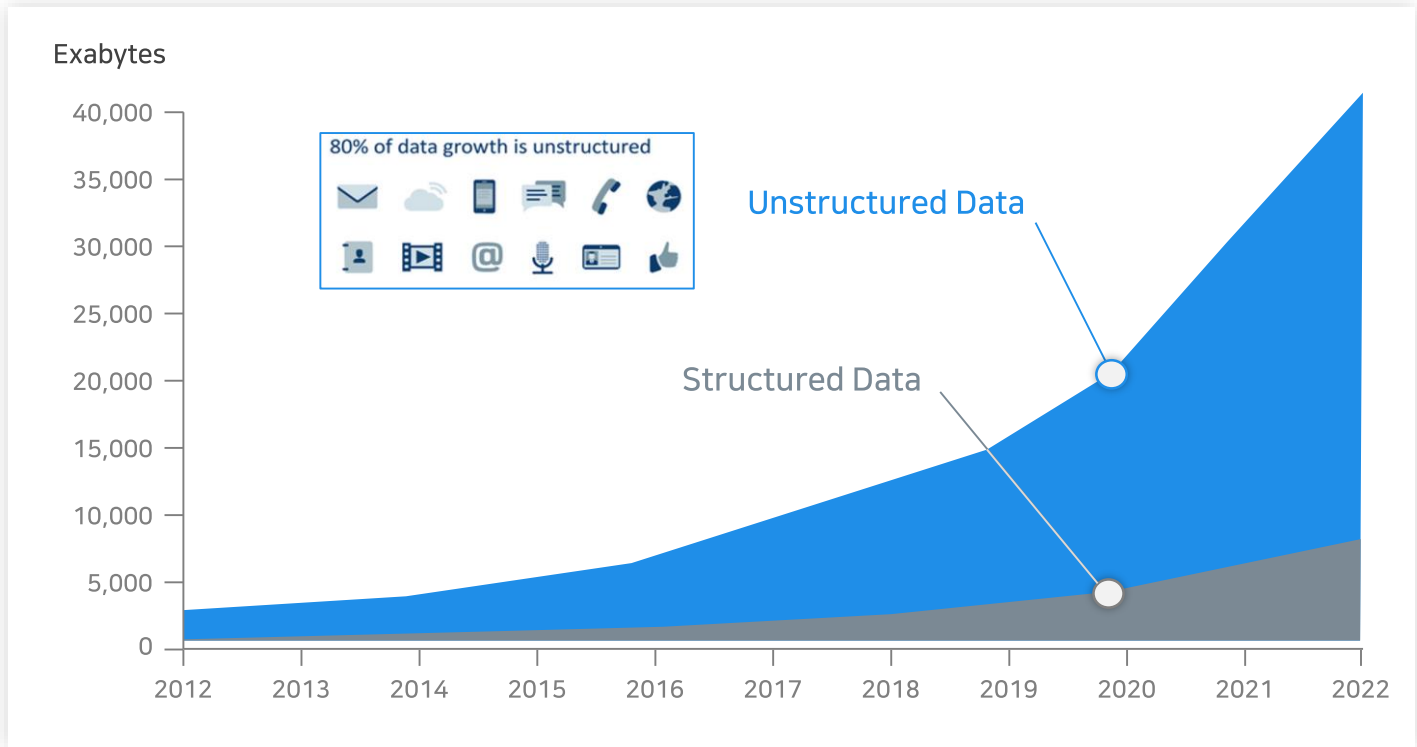
### 비정형 데이터 (Unstructured Data)



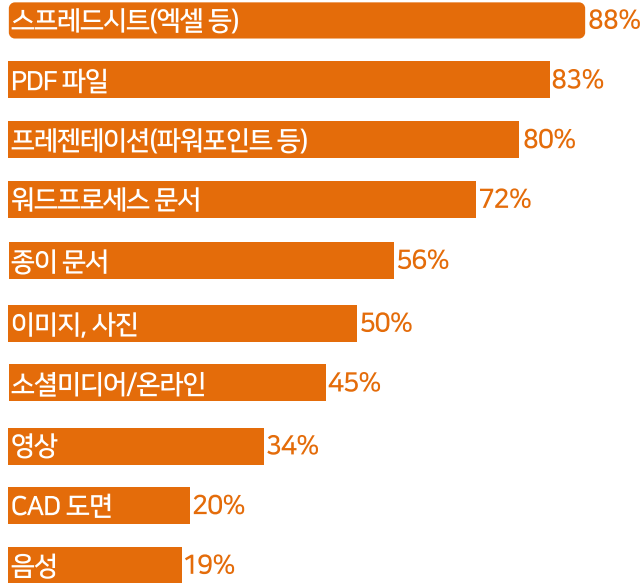
# I-3 개인정보 데이터 현황

## 비정형 데이터 전망

데이터는 매년 폭발적으로 증가하고 있으며,  
구조화되지 않은 **비정형 데이터는 매년 60~80% 증가 추세**



### 비정형 데이터 종류별 현황





## I-4

## 개인정보 데이터 현황

## AI 서비스 &amp; 비정형 데이터

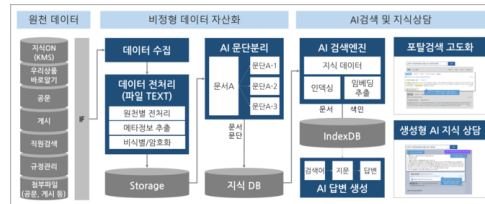
## AI 서비스에 비정형 데이터 활용

○○은행이 **비정형 데이터를 AI모델을 활용**해 자연어 처리 기반의 지식검색 및 상담 서비스를 구현하는 AI 지식 상담 시스템 도입 (2023.09)

2022년부터 상품설명서, 규정, 공문, 게시 등 **1000만건 이상의 비정형 데이터를 AI 학습이 가능한 형태로 자산화**하고, 자연어

처리 기반의 통합검색 및 상담이 가능한 시스템 구축을 추진해왔다.

AI 지식상담 서비스를 통해 직원들이 원하는 정보에 보다 쉽고 빠르게 접근할 수 있게 되고, 결과적으로 고객 서비스의 질을 크게 높일 수 있을 것으로 ○○은행은 기대하고 있다.



## 미래를 바꿀 AI 서비스



IDC

시장조사업체 IDC는 "초거대 AI를 포함한 전 세계 AI 시장 규모가 '24년 5,543억 달러 (약 700조 원)에 달할 것"

Gartner

글로벌 IT 리서치 기업 Gartner는 "2026년까지 약 1억 명의 사용자가 AI와 함께 일할 것이며, 기업의 혁신을 이끌 것"이라고 전망

McKinsey

컨설팅 업체 McKinsey는 "생성 AI를 통한 자동화로 업무 시간이 60~70% 절감하게 될 것"이라고 예측

# II

## 개인정보 보안 관리의 현주소

FAS00



## II-1

## 개인정보 보안 관리의 현주소

## 정해진 패턴 위주의 개인정보 식별

## 정규표현식 기반 개인정보 검출

주민등록번호 (Wd{6}[,-]?[1-4]Wd{6})(Wd{6}[,-]?[1-4])

운전면허번호 (Wd{2}-Wd{2}-Wd{6}-Wd{2})

전화번호 (Wd{2,3}[,-]?Wd{2,4}[,-]?Wd{4})

이메일 (([Ww!-\_W.])\*@([Ww!-\_W.])\*W.[Ww]{2,3})

나이/생년월일 (Wd{0,4})(년생|월생|세살)

여권번호 ([a-zA-Z]{1}[a-zA-Z]{2})Wd{8}

계좌번호 ([0-9,W-]{3,6}W-[0-9,W-]{2,6}W-[0-9,W-])

건강보험번호 [1257][~.[:space:]][0-9]{10}

## 정규표현식 검출 불가 예시

- 충북 그린리더로는 청주시의 '최상희' 씨, 제천시 '이은영' 씨가 선정됐다.
- 최상희 씨(개신주공 3단지 그린빌 아파트 관리소)는 산책하며 쓰레기 줍기(플로깅), 종이팩 수거, 자원수집 경진대회 참여 등 입주민과 다양한 행사를 진행했다. 또한, 환경 실천 포스터를 제작, 게시판에 부착하여 적극 홍보했다.

체중 609kg으로 '세계에서 가장 무거운 10대'라는 타이틀을 가졌던 사우디아라비아 남성이 10여년 만에 500kg 넘게 체중을 감량했다.

영국 대중지 더 선은 27일(현지시간) 17세 때 체중이 609kg이었던 사우디의 칼리드 모센 알샤에리 29가 비만 치료와 엄격한 식이요법, 운동 등을 통해 현재 몸무게를 63kg까지 감량했다고 전했다.

매체에 따르면 외출이 자유로운 여느 10대 청소년들과 달리 칼리드는 비대한 몸집 탓에 3년간 침대를 벗어나지 못했다.

지난 2013년 칼리드 사연을 알게된 사우디 압둘라 국왕이 치료 지원을 약속했고, 30여명의 의료진과 민방위 대원들로 구성된 구조팀이 꾸려졌다.

II-1

개인정보 보안 관리의 현주소

정해진 패턴 위주의 개인정보 식별

문제점 1

동음이의어, 동일한 패턴 탐지 불가

사전에 정의한 패턴만 인식하므로 동음이의어 및 같은 패턴에 대한 유연한 대응이 어렵다.

같은 패턴 다른 의미
비행기 탑승 제한 중량이 150Kg이다.
살이 5Kg 빠졌다.
나는 몸무게 60Kg이다.

문제점 2

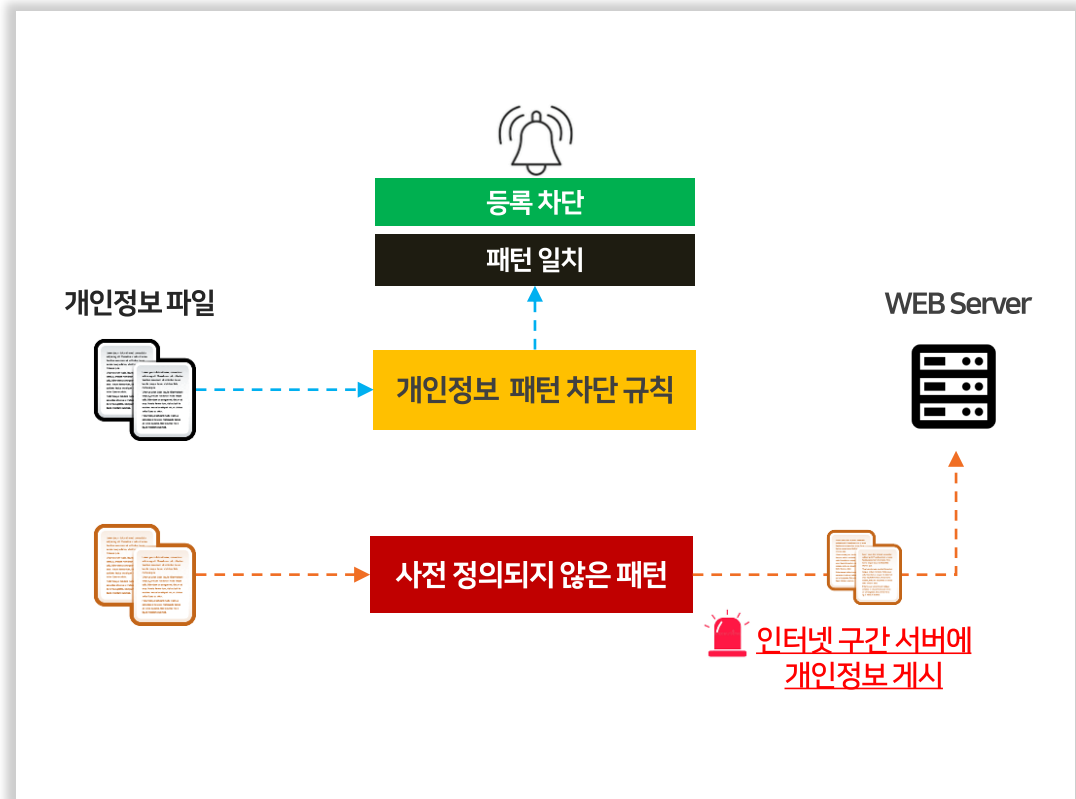
사전 정의하지 않은 패턴 검출 불가

사전에 정의하지 않은 패턴이 등장하면 탐지할 수 없다.

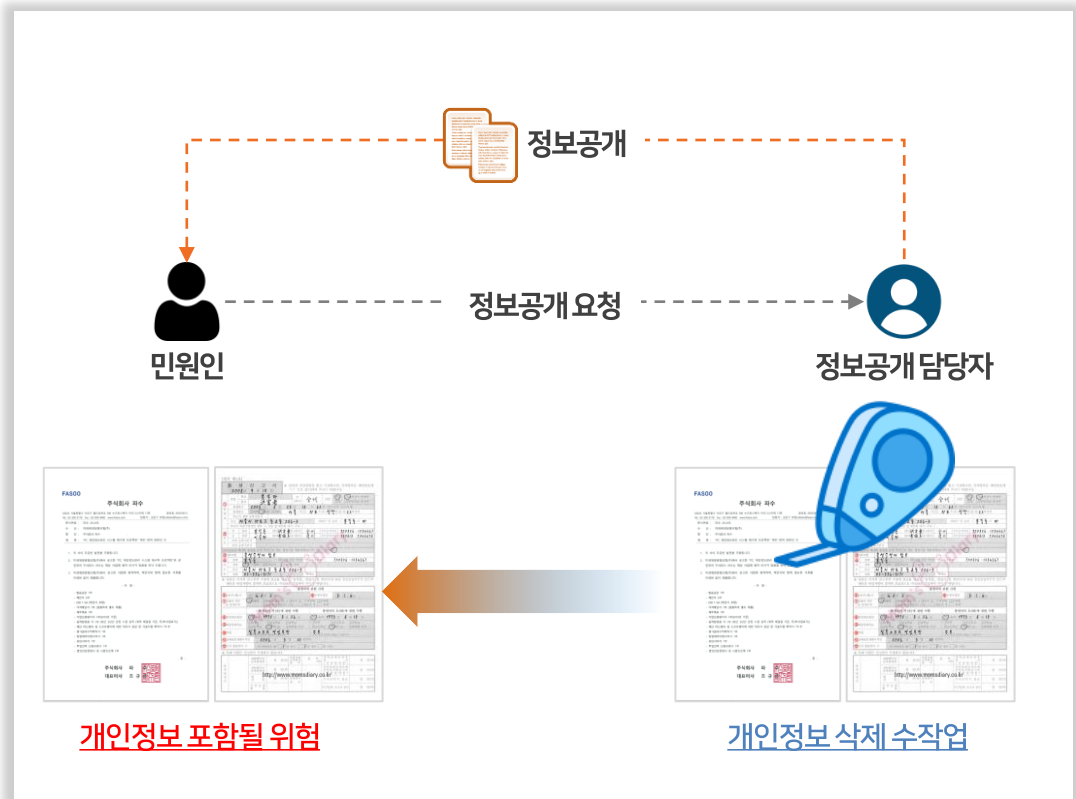
패턴을 특정하기 어려운 경우
010-1234-5678
010에 1234 뒷자리는 5678
칼리드 모센 알 샤에리

# II-2 개인정보 보안 관리의 현주소 개인정보 유출 위험 - 외부 제공 시

## 웹 서버 파일 등록 시



## 민원인 정보공개 요청 시



## II-2 개인정보 보안 관리의 현주소

# 개인정보 유출 위험 - AI 서비스 활용 시

### AI 서비스 - 업무 활용 사례

#### “5시간 걸리던 데이터 업무 1분만에”... 챗GPT 활용하는 직장인들

25년 차 개발자인 김용선 씨(49)는 최근 대화형 인공지능(AI) 서비스 ‘챗GPT’를 활용해 업무 시간을 단축했던 경험을 떠올리며 이같이 말했다. 김 씨는 “코딩을 위한 프로그래밍 연산 공식”을 구하려고 한 달 넘게 구글링(구글 검색)만 하기도 했는데 지금은 챗GPT에 요구하니 1분도 안 돼 답을 내놓는다”며 “20년 넘는 경력을 가진 나도 멘토처럼 모시며 일하고 있다”고 했다.

핀테크 업계에서 11년째 근무 중인 이모 씨(35·여)는 이달 초 동료가 퇴사하는 바람에 떠안게 된 추가 업무를 챗GPT로 해결했다고 털어놨다. 이 씨는 “어떤 프로그래밍 코드를 사용해야 할지” 챗GPT가 알려줘 그대로 따라 했다. 2, 3시간에 걸쳐야 만들 수 있는 주간보고서를 이제 는 손 안 대고 자동으로 만들 수 있게 됐다”고 했다.

3주 전부터 해외영업 부서에서 일하고 있는 직장인 이성혁 씨(28)는 “영어에 자신이 없는데 일손이 부족하다는 이유로 갑자기 발령이 나 막막하고 고민이 많았다”며 “다행히 한글로 쓴 사업 계획서나 이메일을 챗GPT가 영어로 자연스럽게 옮겨줘 해외 파트너들과 수월하게 일을 진행할 수 있었다”고 말했다.

출처 : 동아일보 2023.02.22

### AI 서비스 - 정보유출 사례

#### “챗GPT에 묻다가 기밀 샌다” 기업마다 정보보안 골머리

국내 기업들이 챗GPT 보안으로 골머리를 앓고 있다. 챗GPT가 각종 업무에 도움이 된다는 사실이 알려지면서, 일부 임직원이 핵심 기밀 같은 대외비 자료를 챗GPT에 입력하는 사례가 급증하고 있기 때문이다. 이 정보들은 모두 챗GPT 운영사인 오픈AI에 전송된다.



의사

“환자 이름과 진료 기록을 입력했으니, 보험사에 보낼 양식을 만들어줘 ”



회사 임원

“회사 핵심 전략을 파워포인트로 만들어줘 ”



삼성전자 반도체 부문 직원

“반도체 프로그램 설계가 잘못된 것 같은데 고쳐줘 ”

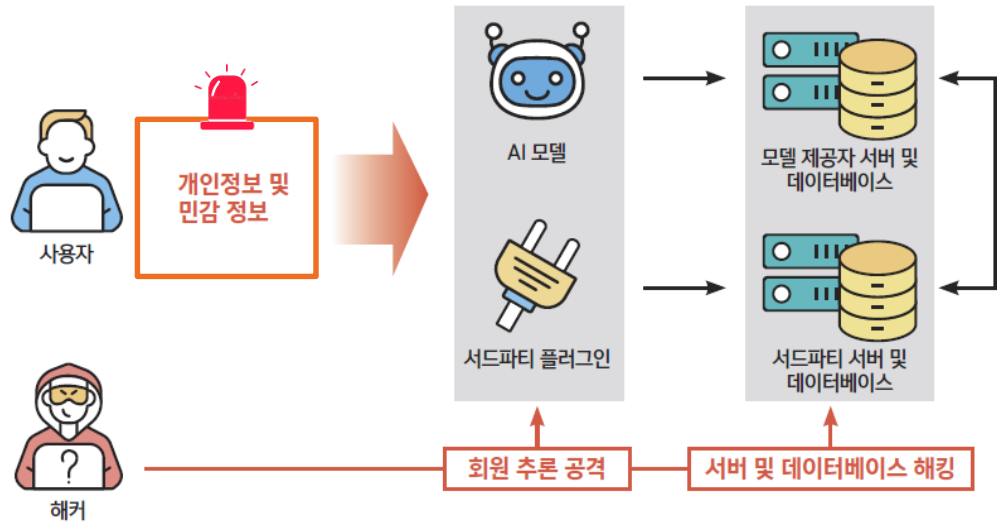
출처 : 조선일보 2023.04.03

## II-2 개인정보 보안 관리의 현주소

# 개인정보 유출 위험 - AI 서비스 활용 시

### AI 서비스 - 데이터 유출 시나리오

사용자의 민감정보 입력 및 해커에 의한 데이터 유출



출처: 챗GPT 등 생성형 AI 활용 보안 가이드라인, 국가정보원

### AI 서비스 - 대표 보안위험

5가지 데이터 유출 위험

훈련 데이터 유출

URL, 개인정보, 연락처, API 키와 같은 민감 정보들이 모델 내에서 데이터를 처리하고 응답 생성 과정에서 외부에 노출

데이터 불법 처리

최근 국가들은 대규모 언어모델 기반 생성형 AI 기술과 관련, 데이터의 불법 처리에 대해 우려와 서비스 중단으로 대응

기밀 유출

개인 또는 기관의 기밀 정보를 AI 모델에 제공 하면, 이러한 정보가 다른 사용자에게 유출될 위험이 있음

대화 기록 유출

일부 AI 서비스에서 타인의 대화 기록(이름, 결제정보, 카드 정보 등)이 다른 사용자에게 보이는 버그 현상이 발생한 사례

해킹 공격

해커가 해당 모델 혹은 서비스 제공자의 데이터베이스를 해킹 하거나 모델에 대한 회원 추론 공격을 수행하여 기밀성 침해



# III

## AI시대에 맞는 개인정보 보호 전략

FAS00

# III-1 AI시대에 맞는 개인정보 보호 전략

## Compliance - 가명정보 처리 가이드라인

### 개인정보보호위원회 가명정보 처리 가이드라인 (2024.02)

부록1

#### 3. 텍스트정보

① 규칙기반 개인정보 단순 삭제 혹은 마스킹

- ▶ 사전에 텍스트 내 개인식별(가능)정보들을 정의하고 정의된 형태(포맷)에 기반하여 해당 정보를 삭제하거나 마스킹, 대체 처리 등의 방법으로 제거

② 스크리빙(Scrubbing)

- ▶ 원 텍스트의 내용과 구조를 보존하면서 축적해서 파싱을 통하여 혹은 파싱 이후 개인식별(가능)정보만을 제거(마스킹 혹은 대체)하는 것으로 이 경우 다수의 정보 주제와 해당 속성들 사이의 명확한 연관성이 없어질 수 있음
- 즉, 특정 정보 주제가 어느 속성을 지칭하는지 알 수가 없게 될 수도 있음
- 단순 삭제 혹은 마스킹 방법과 유사하며 수작업이 아닌 자동화 SW를 이용한다는 점이 다름

③ 정규표현식(Regular Expression)

- ▶ 문자나 혹은 문자열의 일정한 패턴을 표현하는 일종의 형식 언어

	영역	텍스트 예시	정규표현식
1	전화번호	010-1111-2222	^([0-9]{3,4})-[0-9]{3,4}\$
2	이메일주소	aa@bb.net	^[0-9a-zA-Z]{1,}@[0-9a-zA-Z]{1,}([0-9a-zA-Z]{1,})?\$
3	IP주소	111.111.111.111	^([0-9]{1,3})\.([0-9]{1,3})\.([0-9]{1,3})\.([0-9]{1,3})\$
4	주민등록번호	220101-111111	^[0-9]{6}([0-9]{1,4})\$

※ 상시 오픈나 혹은 문자 소자 혼합 등에 따라 누락되는 경우가 존재하므로 추가 검증이 필요할 수 있음

④ 주석 달기(Annotation)

- ▶ 주어진 텍스트를 논리적으로 분할한 후 분할된 단어(들)에 주석을 추가하는 기법

종류	설명
규칙(Rule) 기반	• 구문 문법의 규칙에 따라 텍스트를 토대로 :사전 정의된 수의 단어(들)으로 나누는 것으로 고급 규칙의 경우 정규표현식을 사용하여 정의
사전(Dictionary) 기반	• 개인식별(가능)정보들을 미리 사전으로 정의한 후 개체명 인식(NER, Named Entity Recognition) 기술을 이용하여 주어진 텍스트(이름, 주소, 전화번호 등)와 일치 시킴 • 사전기반의 경우 AI 기술을 활용하여 각 단어들을 자동으로 인식하지만 모든 단어(들)을 완벽하게 인식하는 것은 어려우며, 따라서 맥락(context)에 의한 개인식별 가능성에 대한 조치는 어려울 수 있음

113

가명정보 처리 가이드라인

⑤ AI 기반 텍스트정보 가명처리

- ▶ 규칙, 정규표현식 등을 통한 개인정보 검출 및 마스킹은 정확도 측면에서 한계가 있을 수 있으며, 이를 보완하기 위해 딥러닝 기술 등을 적용한 자연어 처리 언어 모델을 통해 사전에 정의되지 않은 패턴의 개인정보를 검출하고 마스킹할 수 있음
- ▶ 학습방법에 따라 다양한 형태의 인공지능 기반 개인정보 검출 기법 존재 (HMM, MEM, CRFs, structural SVM, Deep-Learning)
- ▶ 규칙기반의 유연성 부족을 해결하기 위해 패턴이나 규칙을 수동이나 반자동으로 작성하고 인공지능을 통해 사전(dictionary)을 확장하여 규칙에 적용되는 개인식별가능정보를 새롭게 검출할 수 있음(단계별 규칙을 순차적으로 적용하여 가중치에 따라 인명, 지명, 조직명 등으로 범주를 결정하는 등)

⑥ 텍스트를 테이블 형식으로 변환

- ▶ 주어진 텍스트를 구문 문법의 규칙에 따라 파싱(parsing, 정해진 규칙에 따라 문장의 구문을 분할)한 다음 분할된 각 세그먼트들을 열과 행이 있는 테이블 형태로 정렬한 후 나머지 데이터들은 삭제
- 변환된 테이블은 기존 정형 데이터에 대한 가명처리를 적용
- 실제로는 텍스트 자체의 복잡성으로 인하여 구조화된 테이블로의 변환이 불가능할 수도 있음

114

#### ⑤ AI 기반 텍스트정보 가명처리

- ▶ 규칙, 정규표현식 등을 통한 개인정보 검출 및 마스킹은 정확도 측면에서 한계가 있을 수 있으며, 이를 보완하기 위해 딥러닝 기술 등을 적용한 자연어 처리 언어 모델을 통해 사전에 정의되지 않은 패턴의 개인정보를 검출하고 마스킹할 수 있음
- ▶ 학습방법에 따라 다양한 형태의 인공지능 기반 개인정보 검출 기법 존재 (HMM, MEM, CRFs, structural SVM, Deep-Learning)
- ▶ 규칙기반의 유연성 부족을 해결하기 위해 패턴이나 규칙을 수동이나 반자동으로 작성하고 인공지능을 통해 사전(dictionary)을 확장하여 규칙에 적용되는 개인식별가능정보를 새롭게 검출할 수 있음(단계별 규칙을 순차적으로 적용하여 가중치에 따라 인명, 지명, 조직명 등으로 범주를 결정하는 등)

#### I (참고) 가명정보 제공시 법적책임 범위

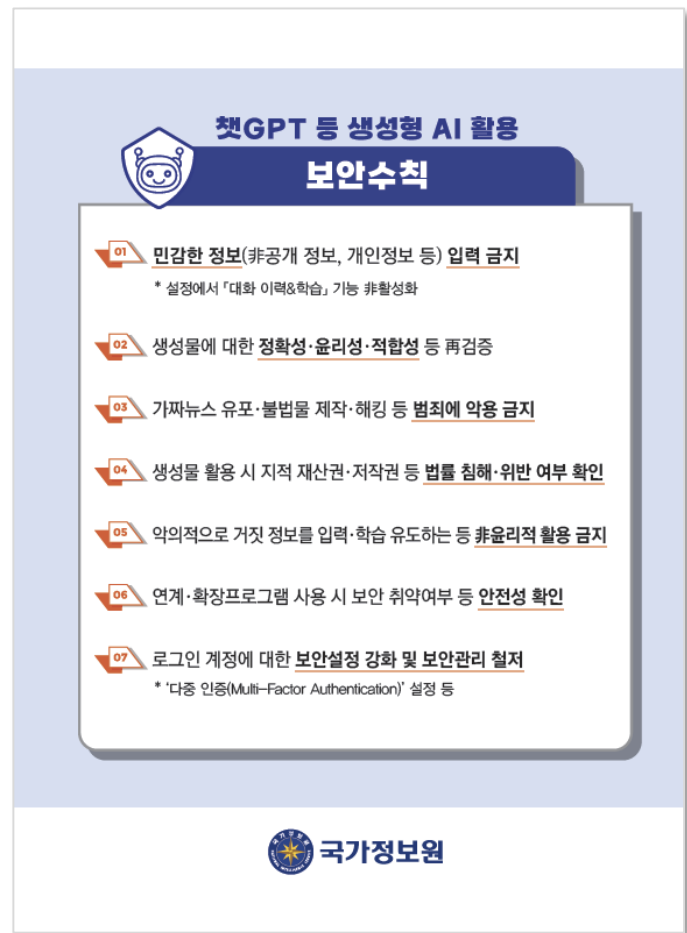
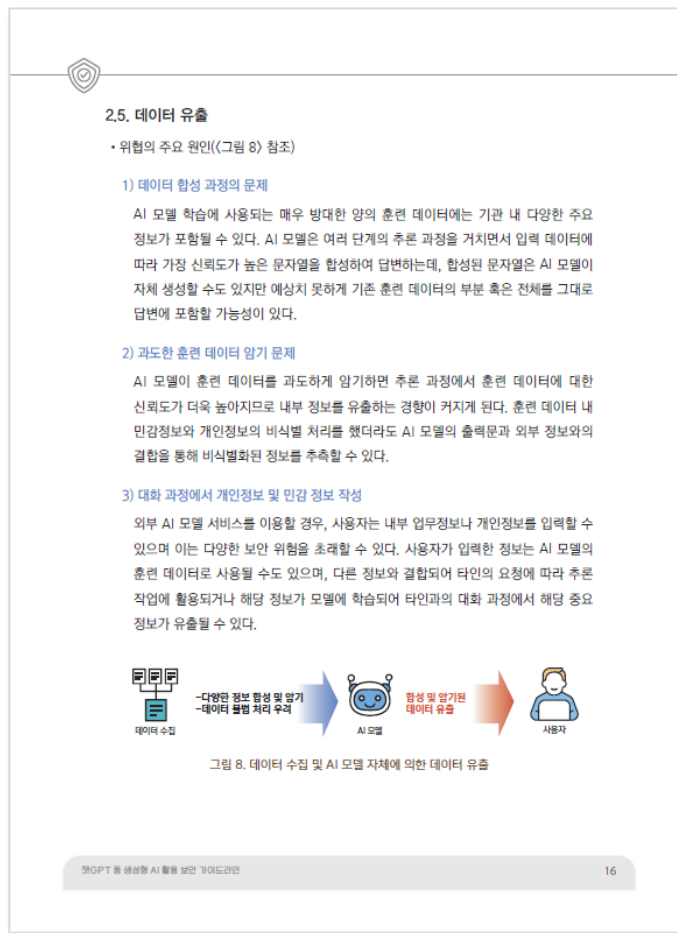
- ▶ 개인정보를 보호법에서 정한 처리 목적에 따라 가명처리하고 관련 안전조치 등 법률에서 정한 사항을 모두 준수하여 가명정보를 제공한 경우,
  - 가명정보를 제공받은 자가 가명정보 이용 과정에서 의도치 않게 특정 개인을 알아볼 수 있는 정보가 생성되었다는 사실만으로는 가명정보를 제공한 자에 대해 개인정보 보호법상 행정처분을 하지 아니함
  - ※ 단, 제공받은 자는 위 생성된 정보의 처리를 즉시 중지하고, 지체없이 회수·파기하여야 함
- ▶ 가명정보를 제공받은 자가 안전조치 미이행 등으로 가명정보를 유출하였거나 고의로 재식별 행위를 하는 경우, 해당 행위자만 제재함

# III-1 AI시대에 맞는 개인정보 보호 전략

## Compliance – 생성형 AI 활용 보안 가이드라인



### 국가정보원 챗GPT 등 생성형 AI 활용 보안 가이드라인 (2023.06)



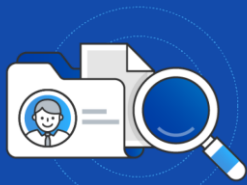
사전 정의된 패턴 기반의 검출 방식 문제점 해결을 위해

## 자연어 처리 언어 모델 적용 및 안전한 사용환경 조성



AI 기술 탑재  
(비정형 데이터 이해)

- 인공지능 기반으로 비정형 텍스트 혹은 문서에 포함된 개인정보의 문맥 이해
- 한글의 특성과 대한민국 개인정보에 최적화된 솔루션 구현



자연어 처리  
(Transformer)

- 트랜스포머 기술 기반의 자연어 처리 언어 모델 적용
- 다양한 형태의 API 적용으로 고객 시스템 마스킹 처리 연계 지원



안전한 접속  
(생성형AI 활용)

- 특정 IP, ID, 데이터 크기, 개인정보, URL 등 세부적인 차단 정책 설정
- 관리자는 중요 검사 기록에 대한 알림 메일 수신 후 정밀한 모니터링 진행



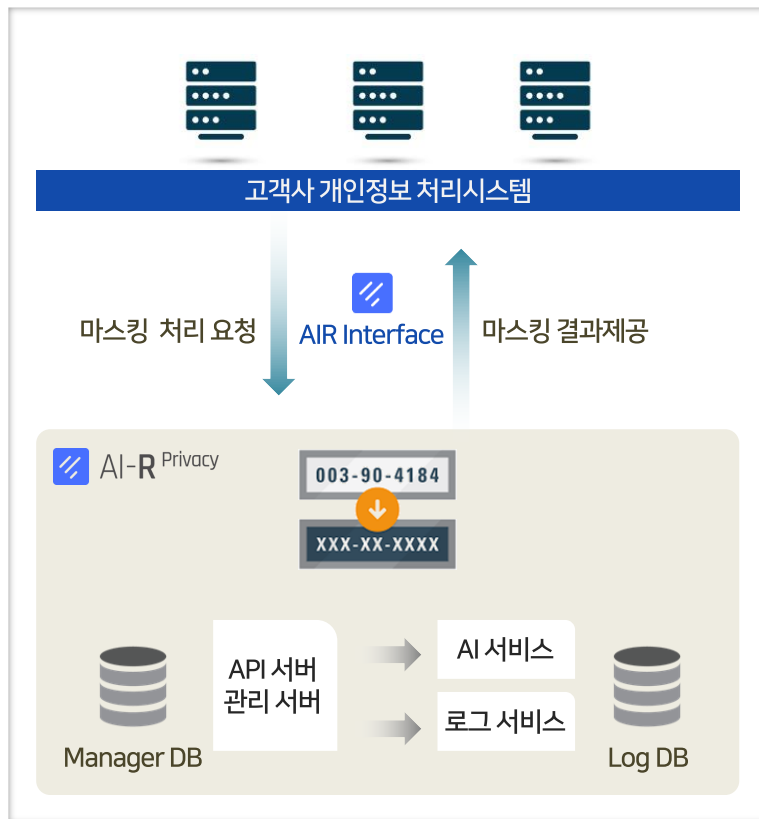
높은 정확도  
(Fine-Tuning)

- Pre-Trained, Fine-Tuning 과정을 통해 자연어 처리의 높은 정확도 기록
- 문장 속 대한민국 행정구역 체계 분류 및 키 - 길이, 몸무게 - 중량의 정확한 구분

## III-3

AI시대에 맞는 개인정보 보호 전략

## 비정형 데이터 개인정보 보호 조치



## 검출 및 마스킹



오피스 파일 검출 및 마스킹  
비정형 텍스트(자연어) 검출 및 마스킹  
이미지 파일 OCR 기반 검출 및 마스킹

## AI Radar API 제공



고객시스템에서 유통되는 개인정보 데이터의 검출  
및 마스킹을 위한 AI Radar API 연동 지원  
배치 형태로 검출 및 마스킹 지원

## 다양한 산업군 적용



(기업) 구매내역, 상담 내역 등 마스킹  
(의료) 환자 진료정보(이미지)의 마스킹  
(공공) 외부 공개 문서의 개인정보 마스킹

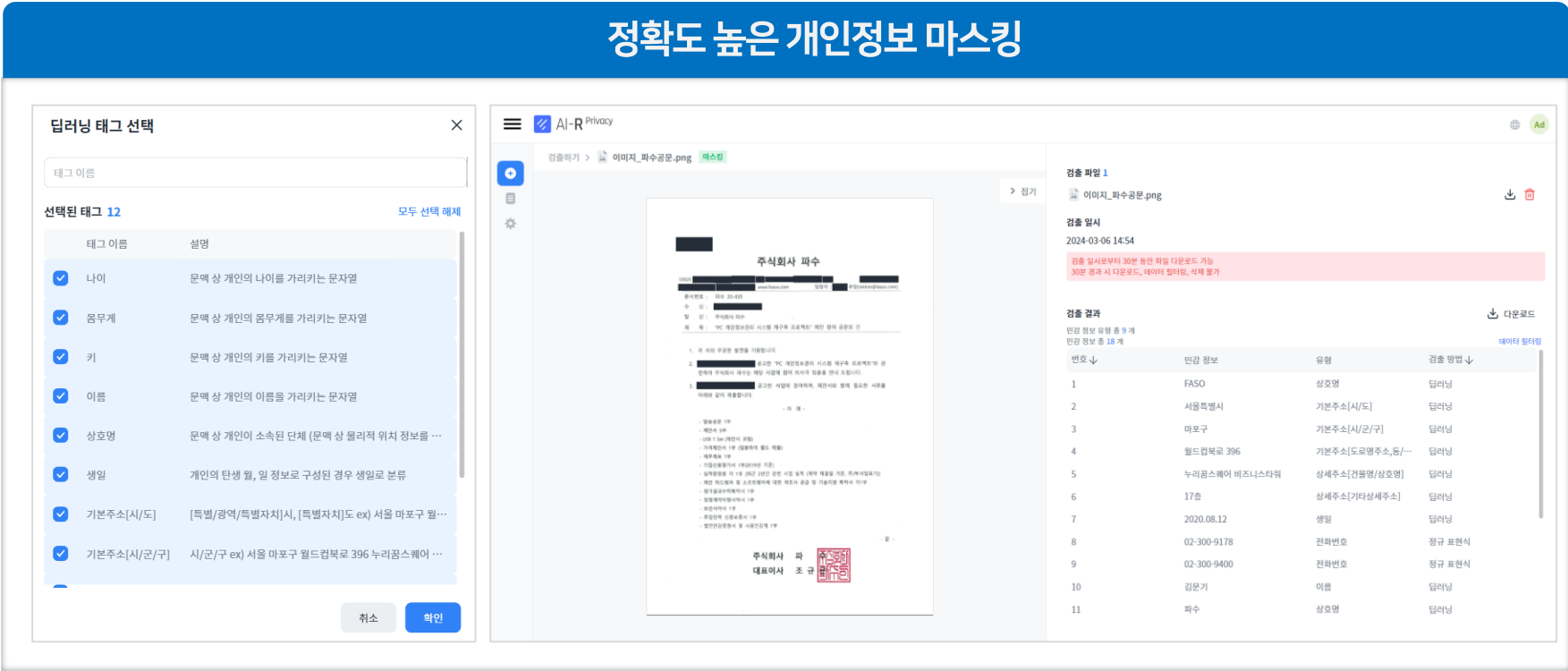


# III-3 AI시대에 맞는 개인정보 보호 전략

## 비정형 데이터 개인정보 보호 조치

개인정보 마스킹 정확도 향상을 위해

### 자연어 처리 언어 모델과 정규표현식을 조합하여 개인정보 마스킹 처리



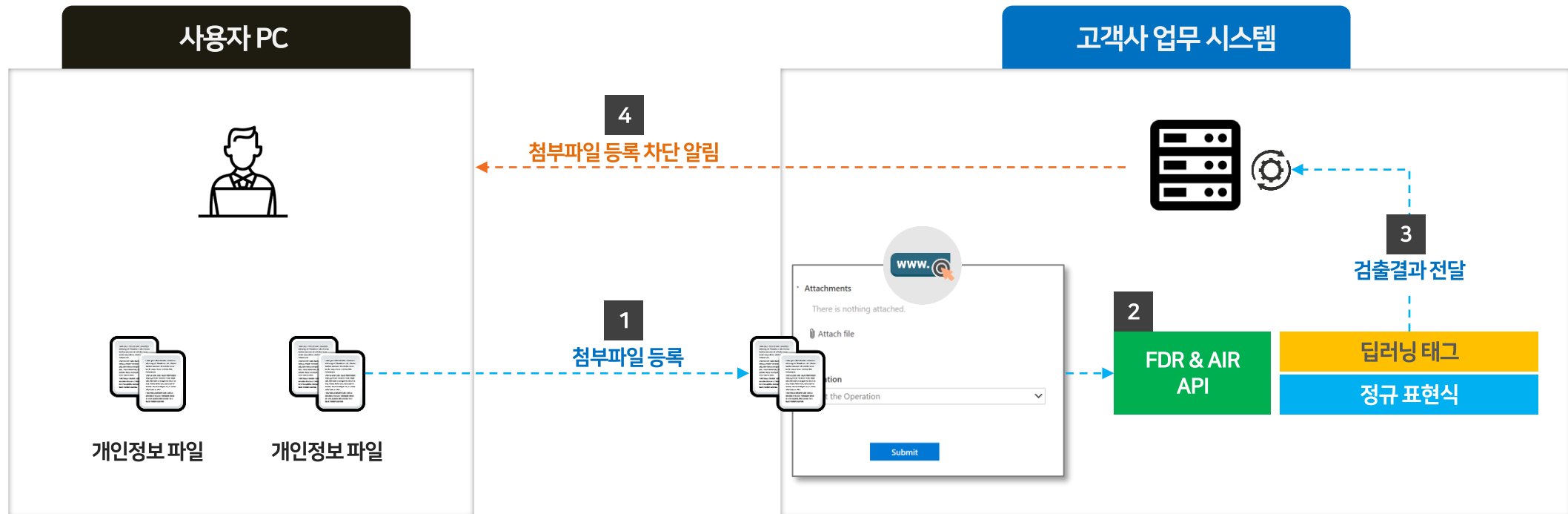
## III-3

AI시대에 맞는 개인정보 보호 전략

## 비정형 데이터 개인정보 보호 조치

PC의 문서, 이미지 등 비정형 데이터가

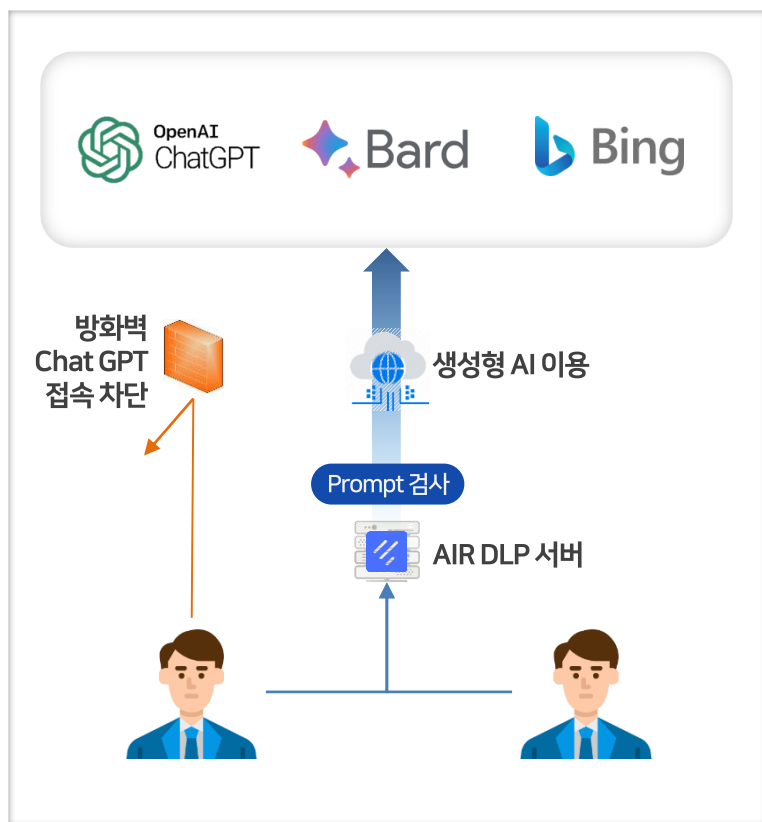
## 웹 서버 등 업무시스템에 등록 시 개인정보 검출 정확도 향상



## III-4

AI시대에 맞는 개인정보 보호 전략

## 생성형 AI 서비스 이용 시 보호 조치



### 안전한 생성형 AI 활용



ChatGPT등의 생성형 AI 서비스에서 전송하는 데이터의 모니터링 및 통제, 민감정보 식별 및 후처리 정책 적용을 통해 안전한 생성형 AI 활용 환경 구축

### 세부적인 차단 정책 설정



특정 IP, ID, 데이터 크기, 개인정보, URL 등 세부적인 차단 정책 설정을 할 수 있으며, 중요 검사 기록에 대한 알림 메일 전송

### 직관적인 사용자 UX 제공



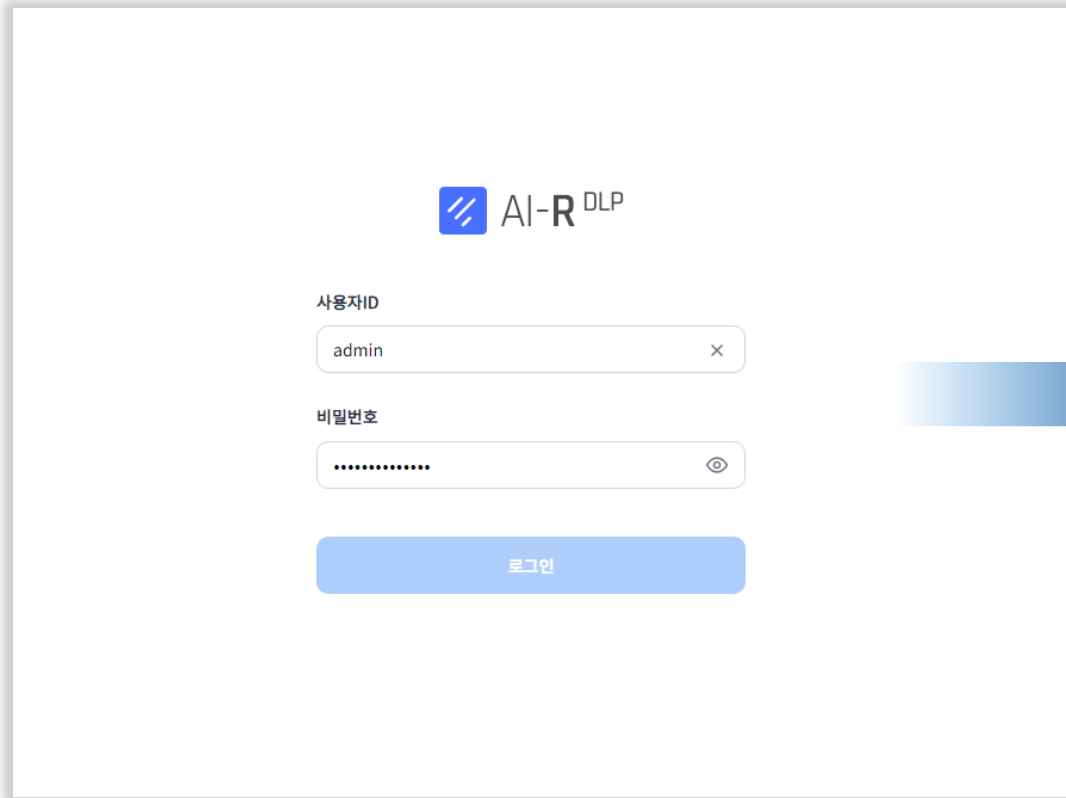
별도의 정책 관리 서버가 존재하며, 정책 설정 및 전송 로그 화면 등 다양한 관리자 기능을 직관적인 형태로 제공해 운영의 편의성 제공

III-4

AI시대에 맞는 개인정보 보호 전략

# 생성형 AI 서비스 이용 시 보호 조치

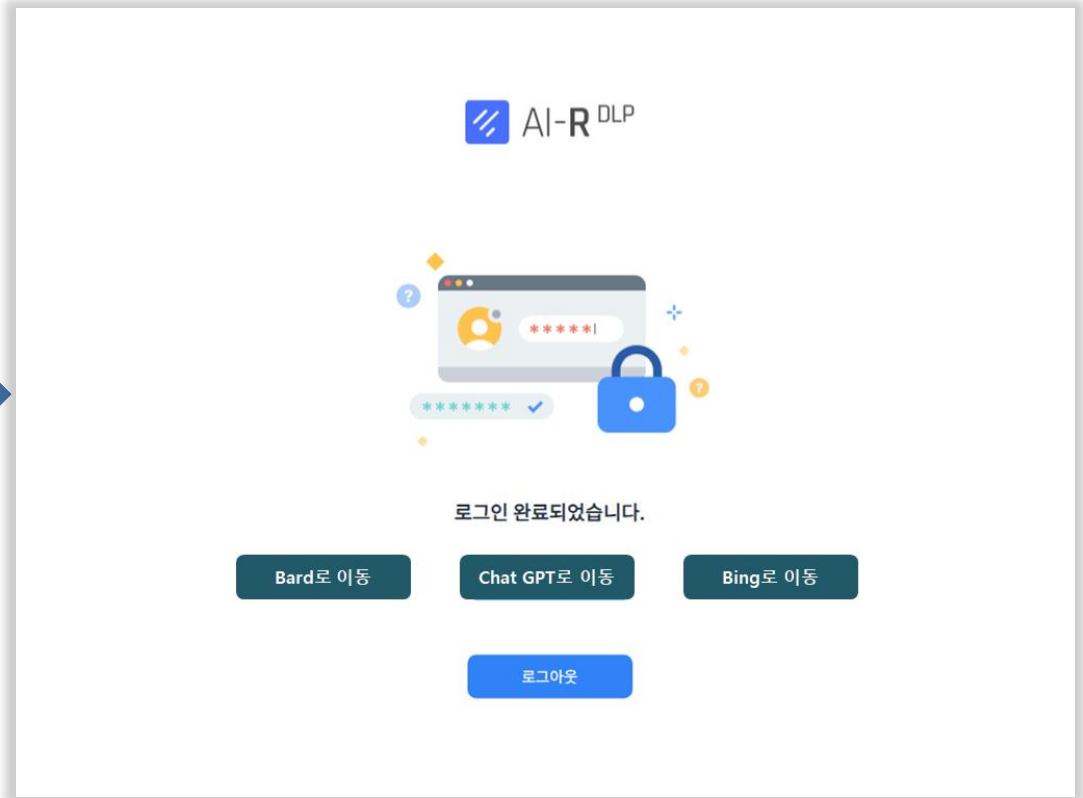
## AI-R DLP 시스템 로그인



The login screen for the AI-R DLP system features the logo at the top. Below it, there are two input fields: '사용자ID' (User ID) with the text 'admin' and a clear button (X), and '비밀번호' (Password) with masked characters and a toggle icon. A blue '로그인' (Login) button is positioned at the bottom.



## 생성형 AI 서비스 선택

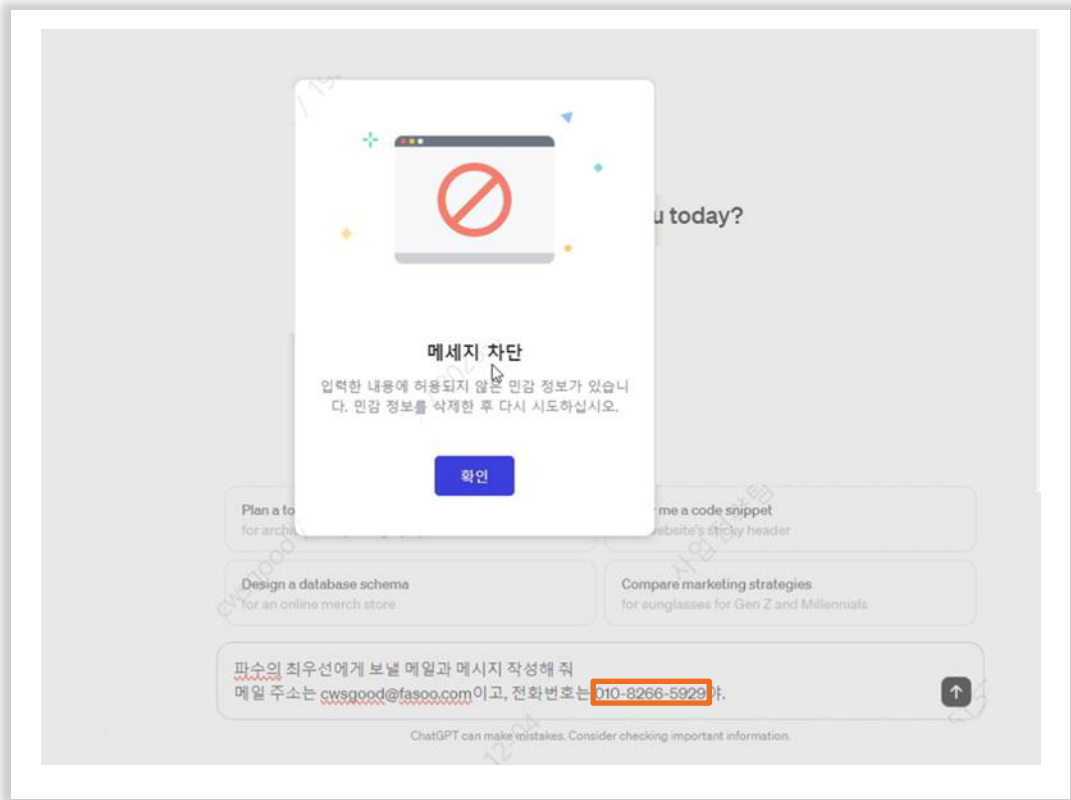


The service selection screen displays the AI-R DLP logo and a central illustration of a computer monitor with a lock icon. Below the illustration, the text '로그인 완료되었습니다.' (Login completed) is shown. At the bottom, there are three buttons for service selection: 'Bard로 이동' (Move to Bard), 'Chat GPT로 이동' (Move to Chat GPT), and 'Bing로 이동' (Move to Bing), along with a '로그아웃' (Logout) button.

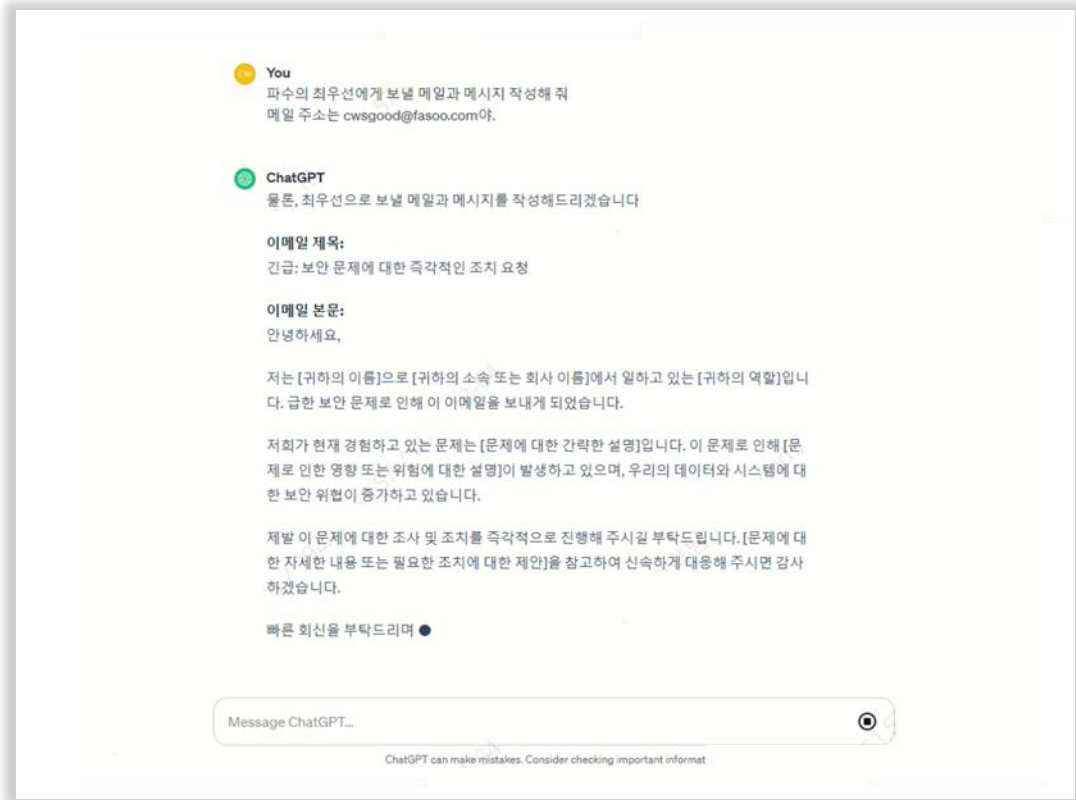
# III-4 AI 시대에 맞는 개인정보 보호 전략

## 생성형 AI 서비스 이용 시 보호 조치

### 민감정보 포함 시 전송 차단



### 민감정보 삭제 후 질의





## III-5 AI 서비스 보안통제 기록 모니터링

## 생성형 AI 사용 로그

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

로그인

로그아웃

회원가입

마이페이지

## 개인정보 검출/마스킹 로그

[illegible]

**G-PRIVACY 2024**

2024 정부·공공·기업 개인정보보호&정보보안 컨퍼런스

## **AI 시대의 통합 개인정보보호 전략**

---

서울특별시 마포구 월드컵북로 396 (상암동, 누리꿈스퀘어) 비즈니스타워 6·17층

02-300-9000 | [www.fasoo.com](http://www.fasoo.com)

2024 Fasoo. All rights reserved.