



PII Masking Dashboard

TEAM 9
개인정보가려조





목차

- 01 팀원 소개
- 02 개발 일정
- 03 시스템 아키텍처
- 04 주요 기능
- 05 주요 문제점 및 해결 방안
- 06 시연 영상
- 07 프로젝트 회고
- 08 Q & A

01 팀원 소개



PM

유주현

- 프로젝트 총괄, 의사결정

Front

김태희

- UI 레이아웃 구성 및 발표 자료 제작

이석진

- 데이터 정제 및 발표 자료 제작

Back

김상빈

- PII 탐지 및 마스킹 모듈 구현

윤유준

- Gemini API 연동, 시각화, 발표

02 프로젝트 개요



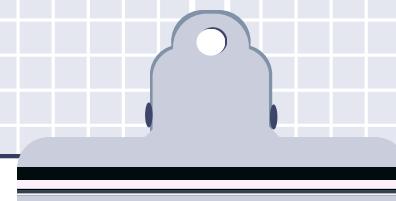
원인

- 개인정보 유출 사고 빈번해짐 → 법적 규제 강화
- 문서 내 개인정보를 가명처리하는 기술의 필요성 증가

목표

- 개인정보 마스킹 후 안전하게 관리
- XLSX 문서 내 개인정보 자동 탐지 후 마스킹
- 누구나 손쉽게 안전한 문서를 생성할 수 있도록 지원

03 개발 일정



DAY 1

▶ 프로젝트 기반 다지기

- 전체 목표 구체화 및 최종 확정 → R&R 분배
- PII 탐지 모듈 · 마스킹 모듈 개발
- 개인정보 별 민감도 매핑 · UI 기본 레이아웃

DAY 2

▶ 기능 확장 및 고도화

- LLM 기반 챗봇 서비스 추가
- UI 레이아웃 및 파일 업 · 다운로드 구현
- 코드 디버깅 및 리팩토링

DAY 3

▶ 통합 준비 및 안정화

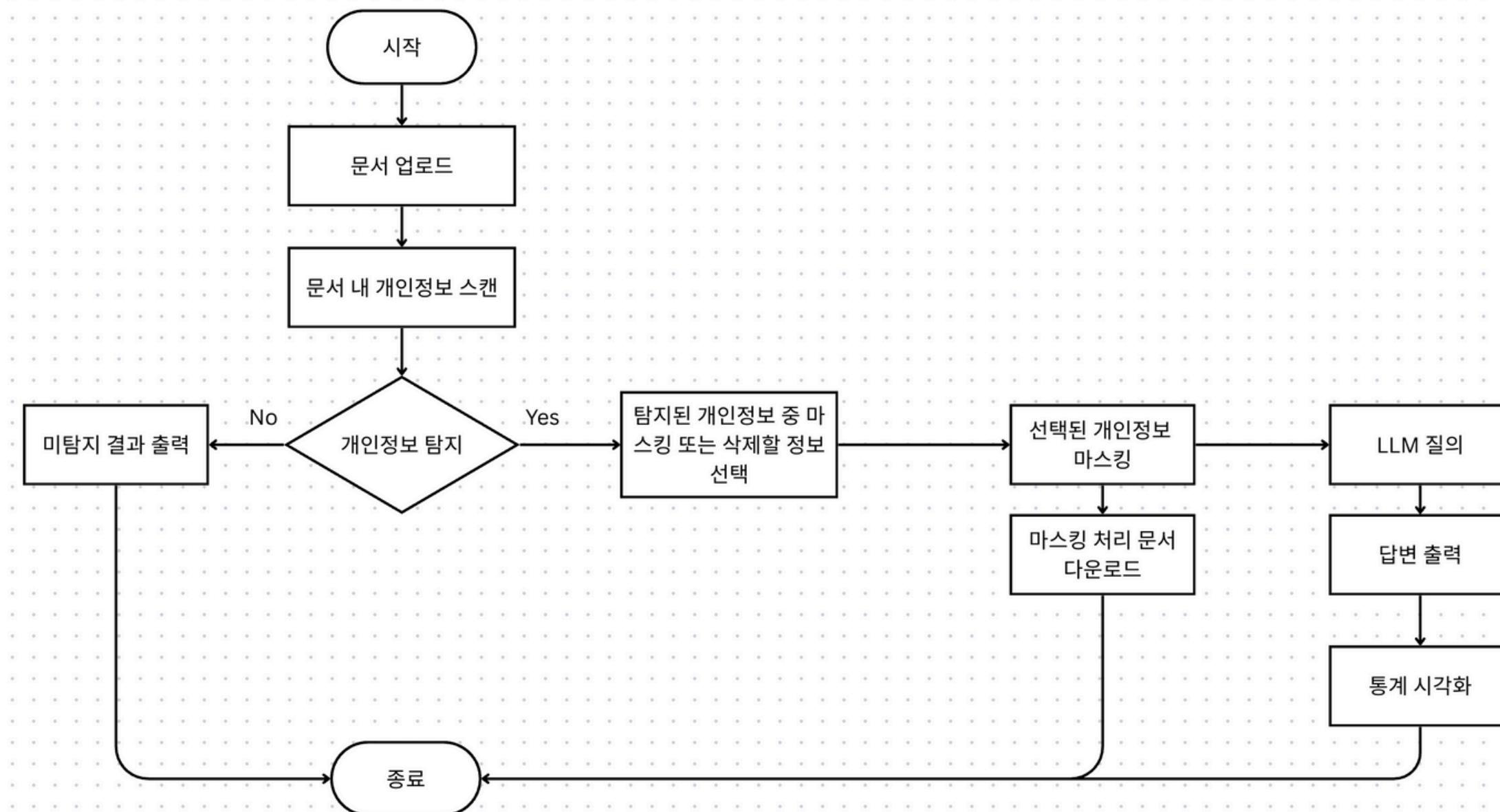
- 코드 디버깅 및 리팩토링
- 데이터 추가 정제 · 통합 테스트
- 발표 자료 초안 제작

DAY 4

▶ 통합 및 최종 점검

- 전체 모듈 연동 및 통합 테스트
- 시연 영상 제작
- 발표 자료 및 최종 산출물 정리

04 시스템 아키텍처



05 데이터 전처리

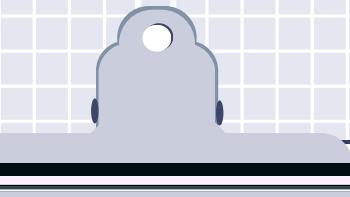


```
id,pii_type,source_path,source_type,container,row,col,header,bbox,text
39313bee8fe67e4c,address,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,1
b77621a587b369d2,driver_license,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,2
ce004541540214ae,email,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,3
0fb780ff754ce9ed,name,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,4,0
b271d7095e611b29,phone,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,5
a348550ca1c7954b,rrn,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,1,6,주
32181e8b7379be5e,address,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,1
1164a62d0347158c,driver_license,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,2
22be19b112135043,email,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,3
c87475f06249610,name,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,4,0
620aa6a27a2493b2,phone,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,5
d497305f0b54e9a2,rrn,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,2,6,주
dbda4b86469124c4,address,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,1
2db85ce637cfdf74,driver_license,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,2
a5d51079a6d41391,email,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,3
c8802eb86eaae9ed,name,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,4,0
87ccf82b001e0d8d,phone,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,5
e64f3952c54bd4ac,rrn,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,3,6,주
e42f2548478455ea,address,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,1
2ecc847bf43ed84e,driver_license,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,2
030c6075576e1f04,email,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,3
1dc3ee11845672be,name,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,4,0
7e820ee441754619,phone,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,5
a417ffab13981d56,rrn,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,4,6,주
32a730b4682466b8,address,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,5,1
40f221e2c8b27ecf,driver_license,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,5,2
ch5c07602-17-71,rrn,C:\수업\modul_proj2_v\data\test_data_v3.xlsx,Worksheet,5,3
```



```
address,driver_license,email,name,phone,rrn
강원도 춘천시 **** * **** ***,11-42-6*****-**,xg*****@hotmail.com,서*우,***_**
광주시 북구 **** * **** ***,13-96-4*****-**,v2*****@gmail.com,박*영,***_**-53
경기도 수원시 *** * **** ***,23-92-5*****-**,3f*****@naver.com,신*윤,***_**-**
경기도 성남시 *** * **** ***,22-28-7*****-**,pz****@gmail.com,조*진,***_**-9393,
부산시 해운대구 **** * **** ***,12-97-4*****-**,36*****@hotmail.com,오*수,***_*
충청북도 청주시 **** * **** ***,27-95-8*****-**,wr*****@yahoo.com,한*진,***_**
제주시 **** * **** ***,23-84-8*****-**,4f****@gmail.com,이*은,***_**-2885,62071
충청북도 청주시 **** * **** ***,12-51-4*****-**,to*****@naver.com,황*린,***_**
대전시 서구 **** * **** ***,21-39-2*****-**,iw*****@daum.net,한*호,***_**-15
제주시 **** * **** ***,27-42-3*****-**,m8*****@daum.net,윤*윤,***_**-4780,611
강원도 춘천시 **** * **** ***,26-13-1*****-**,eq****@hotmail.com,윤*수,***_**-2
제주시 *** * **** ***,13-40-6*****-**,40*****@hotmail.com,박*영,***_**-2148
대구시 수성구 *** * **** ***,14-84-7*****-**,p4*****@naver.com,오*린,***_**-642
충청북도 청주시 **** * **** ***,18-17-3*****-**,hy*****@naver.com,정*진,***_**-*
경기도 성남시 *** * **** ***,24-23-4*****-**,o6*****@yahoo.com,한*은,***_**-8
부산시 해운대구 **** * **** ***,26-86-7*****-**,i8*****@gmail.com,오*늘,***_**
대전시 서구 **** * **** ***,13-98-5*****-**,xi*****@daum.net,정*민,***_**-86
강원도 춘천시 **** * **** ***,19-69-1*****-**,w6*****@naver.com,신*수,***_**
인천시 연수구 *** * **** ***,20-58-9*****-**,rv*****@naver.com,박*영,***_**-5
서울시 강남구 *** * **** ***,27-19-1*****-**,al*****@daum.net,이*린,***_**
대구시 수성구 *** * **** ***,15-10-9*****-**,0u*****@daum.net,이*연,***_**-687
경기도 고양시 *** * **** ***,12-45-4*****-**,3z*****@naver.com,조*은,***_** Col 5
부산시 남구 *** * **** ***,12-64-3*****-**,e6*****@naver.com,서*린,***_**-8999,
```

05 주요 기능



개인정보 탐지

문서를 업로드하면 주민등록번호,
운전면허번호 등 주요 개인정보 탐지

커스텀 마스킹

탐지된 개인정보 중에서 사용자가
원하는 항목만 선택해 마스킹 가능

대시보드

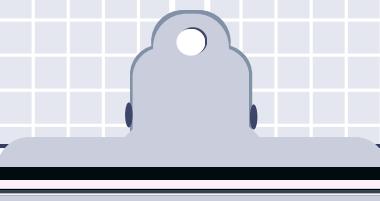
탐지 현황 · LLM 기반 챗봇 등
한눈에 확인할 수 있는 대시보드



차트 시각화

개인정보를 통해 비즈니스 인사이트를
얻을 수 있는 통계 자료 제시

05 주요 기능



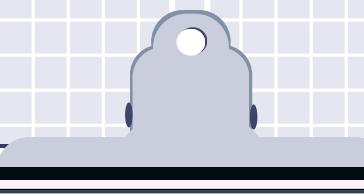
식별정보

개인정보 항목	설명	마스킹 예시
성명	성명 중 성을 제외한 나머지	홍**
	성명 중 이름의 첫 번째 글자 이상	홍*동
영문 성명	영문 성명 중 앞 4자리 철자 노출	SEO *****
휴대폰번호	전화번호 또는 휴대폰 뒤 4자리	010-1234-****
주민등록번호	뒤에서부터 6자리	971212-2*****
여권번호	뒤에서부터 4자리	12345****
주소	도로명 이하의 건물번호 및 상세주소의 숫자 건물번호, 상세주소(동/층/호)	인천시 미추홀구 주안로 *, ****동, ****호
이메일 주소	ID 중 앞 2자리를 제외한 나머지	sy*****@naver.com
CI(온라인 주민등록번호)	88자리의 숫자 중 앞 7자리 노출	cDQ12ci***** ~ *****

고급정보

개인정보 항목	설명	마스킹 예시
카드번호	카드번호 중 7~12번째 숫자 또는 9~12번째 숫자	9430-20*****-239 9430-2000-****-239
국제카드 업계 표준(PCI-DSS)준용		9430-*****-****-239*
카드 유효기간	연/월을 모두 마스킹	YY//MM > **/**
온라인 회원 ID	ID중 앞 2자리를 제외한 나머지	sy*****
계좌번호	뒤에서부터 5자리 변환 (요구사항에 따른 변경 가능)	430-20-1***** 454222-01-1*****
I-PIN	뒤에서부터 5자리 변환	123*****
고객의 IP 주소	앞에서 부터 3자리 변환	***.8.7.12

05 주요 기능



● 개인정보 마스킹 선택

이름 주민번호

연락처 여권번호

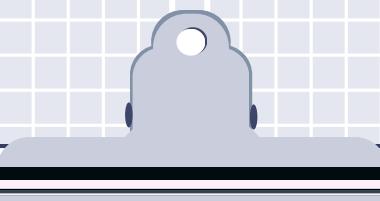
주소 면허번호

이메일 카드번호



김*경	R8254****	0-06-8*****-660311-1*****	자치시 일산동구 ***** *
서*우	R3406****	16-03-2*****-990414-1*****	대구광역시 흥덕구 **** *
권*희	R3241****	16-01-5*****-780408-1*****	서울특별시 연수구 **** *
고*호	R3506****	28-15-5*****-770403-2*****	충청북도 일산동구 **** *
문*우	M8383****	13-05-1*****-540910-2*****	광주광역시 달서구 **** *
조*윤	R2613****	14-04-8*****-120903-3*****	경상북도 서초구 **** ***
박*경	R7462****	19-18-8*****-060128-3*****	충청북도 동래구 **** ***
송*원	M5193****	20-01-5*****-210715-4*****	세종특별자치시 수성구 *
강*진	R1952****	23-03-7*****-661027-2*****	서울특별시 연수구 **** *
고*수	M8454****	21-18-6*****-250805-3*****	대전광역시 해운대구 ***

05 주요 기능



규칙:

반드시 제공된 통계값만 바탕으로 답변해라. 새로운 추측은 하지 마라.
주민번호: `rrn_gender_counts` / `rrn_full_birth_year_counts` 사용.
운전면허번호: `dl_prefix_counts` 사용.
이메일: `email_domain_top10` 사용.
전화번호: `phone_last_4_counts` 사용.
성씨: `family_name_top10` 사용.
주소: `address_region_top10` 사용.
불필요한 설명은 줄이고, 통계 수치와 간단한 해설만 제시하라.

"""

```
def prepare_planner_prompt(user_query: str, available_columns: list, known_regions: list) -> str:  
    cols = ", ".join(available_columns)  
    regs = ", ".join(sorted({r for r in known_regions if r})) or "(없음)"  
    return f"""
```

당신은 '질문을 집계 계획(JSON)'으로 변환하는 플래너다.

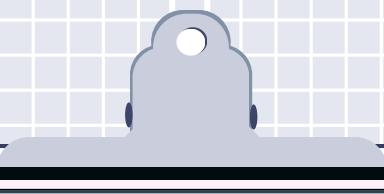
사용자 질문: "{user_query}"

데이터프레임 컬럼: [{cols}]
`region_first`(주소 1토큰) 가능한 값(예시): [{regs}]

규칙:

질문에 위 지역 값(예: 서울, 부산, 경기 등)이 포함되어 있으면 `filters`에 `["column": "region_first", "op": "eq", "value": "해당지역"]`을 반드시 추가하라.
두 가지 의도가 같이 오면(예: 남녀 통계 + 나이별 남녀) 더 구체적인 의도(나이별 남녀)를 우선한다.
아래 스키마로만 JSON 응답. 해설 금지. 코드블록(```json)으로 감싸도 좋다.

05 주요 기능



Deploy ⋮

PII Masking Dashboard

안전하게 개인정보를 처리하고 Gemini와 함께 인사이트를 비즈니스에 활용해 보세요.

안전한 개인정보 분포 분석

마스킹 처리가 완료된 후 시각화 자료가 표시됩니다.

Gemini와 함께 다양한 인사이트를 얻어보세요!

Gemini에게 물어보기

분석 실행

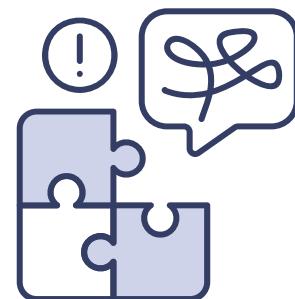
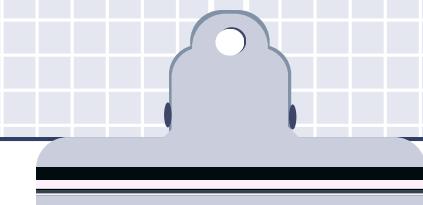
파일 처리 시작

파일 다운로드

가이드

- 사이드바에서 파일을 업로드하세요.
- 보호가 필요한 개인정보 항목을 선택한 뒤 '파일 처리 시작' 버튼을 눌러주세요.

06 주요 문제점 및 개선 방향



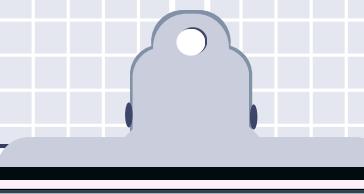
문제점

- 마스킹 처리 수준 정의
- 데이터 민감도 기준
- 다량 데이터 처리 한계
- Gemini 무료 API 한계

해결방안

- PDF, Word 등 다양한 확장자 지원
- 빠르고 정확한 통계 산출
- 지원 카탈로그 8개 외 추가 카탈로그 지원

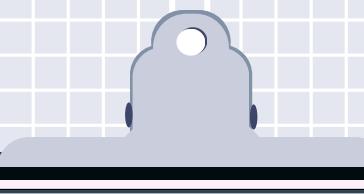
09 시연영상



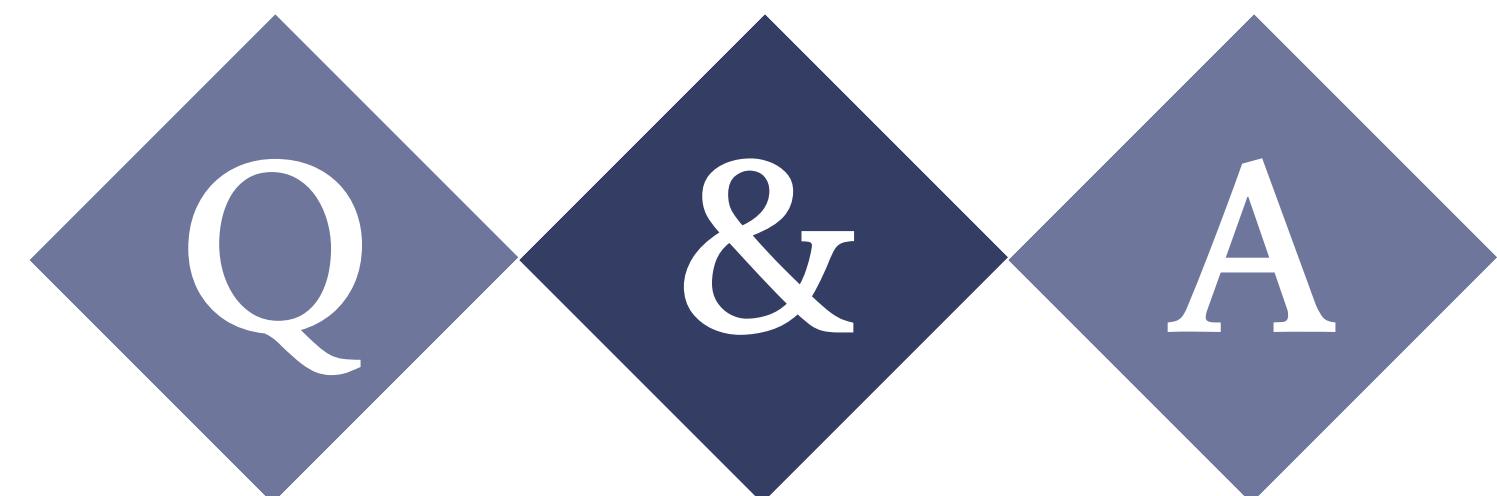
The screenshot shows a web-based PII Masking Dashboard application. The interface includes:

- 파일 업로드:** A section for uploading files, with a "Drag and drop file here" area and a "Browse files" button.
- 개인정보 마스킹 선택:** A list of checkboxes for selecting data types to mask, including 이름, 주민번호, 연락처, 여권번호, 주소, 면허번호, 이메일, and 카드번호.
- 마스킹 처리 시작:** A button to start the masking process.
- 파일 다운로드:** A button to download the masked file.
- 가이드:** A section with instructions:
 - 사이드바에서 파일을 업로드하세요.
 - 보호가 필요한 개인정보 항목을 선택한 뒤 '파일 처리 시작' 버튼을 눌러주세요.
- PII Masking Dashboard:** The main title and a small icon.
- 안전한 개인정보 분포 분석:** A section showing analysis results.
- Gemini와 함께 다양한 인사이트를 얻어보세요!** A call-to-action section with a search bar and a "분석 실행" button.

The browser status bar at the bottom shows the temperature (34°C), battery level (부분적으로 맑음), and system time (오후 5:09 2025-08-21).



10 Q & A



궁금한 사항을 자유롭게 문의해 주세요.



감사합니다

THANK YOU