
보험사 VOC 데이터의 안전한 활용을 위한 가명처리 기술 및 적용 전략 연구

 최우연

목차

1. 연구 배경 및 목표
2. AI 기반 가명처리 시스템의 연구 가치
3. VOC 데이터 수집 및 데이터 전처리 전략
4. 가명처리 모델링 – NER 활용
5. 가명처리 모델링 예시 – NER 활용
6. 가명처리 모델링 – LLM 활용
7. 가명처리 매커니즘 및 매핑 전략
8. 의학 개체명 인식 및 하이브리드 검출 전략
9. 의료 데이터 가명처리 대상 및 평가 기법
10. 의료 데이터 가명처리 성능 향상 기법
11. 가명처리 통합 파이프라인 설계
12. 운영 및 보안, 테스트 전략
13. 연구결과 및 기대효과

1. 연구 배경 및 목표

연구 배경

- 금융권 데이터 3법(개인정보보호법·정보통신망법·신용정보법) 강화로 AI 서비스 도입에 제약 발생
- 보험사 VOC 등의 데이터 활용이 사업 방향 설정에 핵심이나, 법적 제약이 상존
- 보험사 VOC 데이터에는 의학정보 등 민감정보가 포함되어 특화된 가명처리방식이 필요함
- 온프레미스 환경에서 운용할 수 있는 AI 기반 가명처리 서비스가 필요
- 본 연구의 목적은 가명처리 기술을 활용해 법무팀의 동의를 얻고, 안전하게 데이터를 활용할 수 있는 토대를 마련하고자 함

연구 목표

- 보험사 VOC 데이터 내 민감정보에 대해 AI 기반 가명처리 기술을 적용하고, 법무팀의 사전 동의 체계를 구축하여, 데이터의 안전한 내부 활용 기반을 마련하는 것.
- 가명처리 대상
 - 직접식별자*: 이름, 주민등록번호, 휴대전화, 집전화, 이메일 주소, 고객 ID/회원번호, 계좌·카드번호
 - 준식별자*: 상세 주소(시/군/구 이상), 생년월일 또는 나이, 직업, 직급, 보험상품 가입 이력, 거래, 청구 내역
 - 민감정보*: 진단명, 처방·투약 정보, 검사 결과, 입·퇴원 일자, 수술·치료 이력
- 처리 방식
 - 배치 가명처리 방식 채택
 - 일정 주기(예: 매시간, 매일, 매주)로 누적된 데이터를 한꺼번에 처리
 - 대량 데이터에 최적화된 일괄 작업(Job)으로 높은 처리량 달성
 - 향후 실시간 가명처리 방식 활용 고려

"AI 기반 가명처리서비스를 통해 데이터 활용과 개인정보 보호 간 균형 달성을 추구"

- 직접식별자 : 단일 정보만으로 특정 개인을 유일하게 식별할 수 있는 정보
- 준식별자 : 단독으로는 개인 식별이 어렵지만 다른 정보와 결합 시 특정 개인을 식별할 수 있는 정보
- 민감정보 : 개인의 사생활, 신체, 정신, 사회적 특성을 내포해 노출될 경우 피해가 클 수 있는 정보

2. AI 기반 가명처리 시스템의 연구 가치

주요 상용 솔루션 비교

- Private AI, Lingvanex, IRI FieldShield, IBM Guardium 등 다양한 온프레미스 기반 가명처리 솔루션 존재.
- Private AI: 고도화된 NLP 기술을 통해 비정형 텍스트 기반 VOC 데이터 가명처리에 매우 적합.
- Lingvanex: 주로 다국어 번역에 강점. 가명처리는 부가적 기능으로 보험사 VOC 용도에는 보완 솔루션 필요.
- IRI FieldShield: DB 기반 정형 데이터 가명처리에 특화.
- IBM Guardium: 데이터 보호·감사·정책 관리에 중점. 보안 거버넌스와 연계한 가명처리 체계 구축에 강점

상용솔루션 대비 AI연구의 차별점

항목	상용솔루션	AI 연구기반
데이터 유형	정형 데이터 위주 분석	비정형 VOC 데이터에 최적화
탐지 방식	고정된 룰 기반 (규칙 기반) 탐지	문맥 이해를 바탕으로 한 AI 기반 탐지 가능
범용성	다양한 산업에 맞춤 불가	보험 특화 모델로 도메인 전문성 확보
설명력	수동 검토 필요, 설명력 부족	법무팀과 협업 가능한 설명형 AI 처리 제공
인프라 환경	솔루션 종속적인 제한된 운영 환경	클라우드 및 온프레미스 환경 모두에 유연하게 적용 가능

기대 효과

1. 데이터 활용성 증대
 - 기존 상용 솔루션의 대용량 처리 한계 및 수동 컬럼 설정의 번거로움을 해소하여, 보다 효율적이고 실용적인 데이터 활용 환경을 조성.
2. AI 기반 자동화 향상
 - AI 도입을 통해 가명처리 자동화 수준을 향상시킴으로써, 수작업 부담을 줄이고 정확도와 일관성을 강화.
3. 확장 가능한 워크플로우 구축
 - 유연하고 확장성 높은 처리 구조를 마련하여, 다양한 데이터 환경 및 요구사항에 효과적으로 대응.
4. 금융·보험 특화 기술 확보
 - 도메인 지식 기반의 AI 학습을 통해 금융/보험 산업에 최적화된 가명처리 기술 개발 가능, 관련 산업에서의 실질적 적용성과 경쟁력 제고

3. VOC 데이터 수집 및 데이터 전처리 전략

데이터 수집

- ✓ AI-HUB 제공 민원(콜센터) 질의응답 음성 데이터 활용
 - o <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=dat&dataSetSn=98>
- ✓ 금융·보험 도메인 중심 데이터 활용

데이터출처	수집데이터 종류
금융/보험업계	잔고 및 거래내역 조회
	이체,출금,대출서비스
	상품 가입 및 해지
	사고 및 보상 문의

- ✓ 생성형 AI를 통해 개인정보·의료정보가 포함된 합성 데이터 추가 생성

문제 및 해결전략

- ✓ 음성 데이터를 Whisper 모델로 텍스트 변환 후, 오타자 교정과 문장 경계 재정의 진행
 - Whisper 단계: 음성 특징을 뽑아내고 텍스트 예측
 - 후처리 단계: 사람이 읽기·분석하기 편하도록 "맞춤법·구두점"을 다듬고 "문장 단위"를 확실히 구분
- ✓ 임베딩* 및 클러스터링 기법으로 화자 분리 및 타임스탬프 유지하여 정확성 확보
 - 임베딩: 목소리를 수치로 표현
 - 클러스터링: 비슷한 수치를 가진 목소리끼리 묶기
 - 타임스탬프: 발화 시작·종료 시간을 함께 기록

• 임베딩 : 문자나 음성 등 비정형 데이터를 수치 벡터로 변환해 기계가 이해할 수 있도록 표현하는 방법

4. 가명처리 모델링 - NER 활용

1

데이터 라벨링

- 이름, 전화번호, 의료정보 등 PII 대상 라벨링 도구로 B-ENTITY/I-ENTITY 태그 부여

2

모델 세팅 및 Fine-Tuning

- KoBERT와 Transformer 기반* NER 모델*로 라벨링 데이터 학습, 한국어 VOC 특화

3

추론 및 성능평가

- Precision, Recall, F1 지표로 토큰·엔티티* 단위 성능 검증 수행

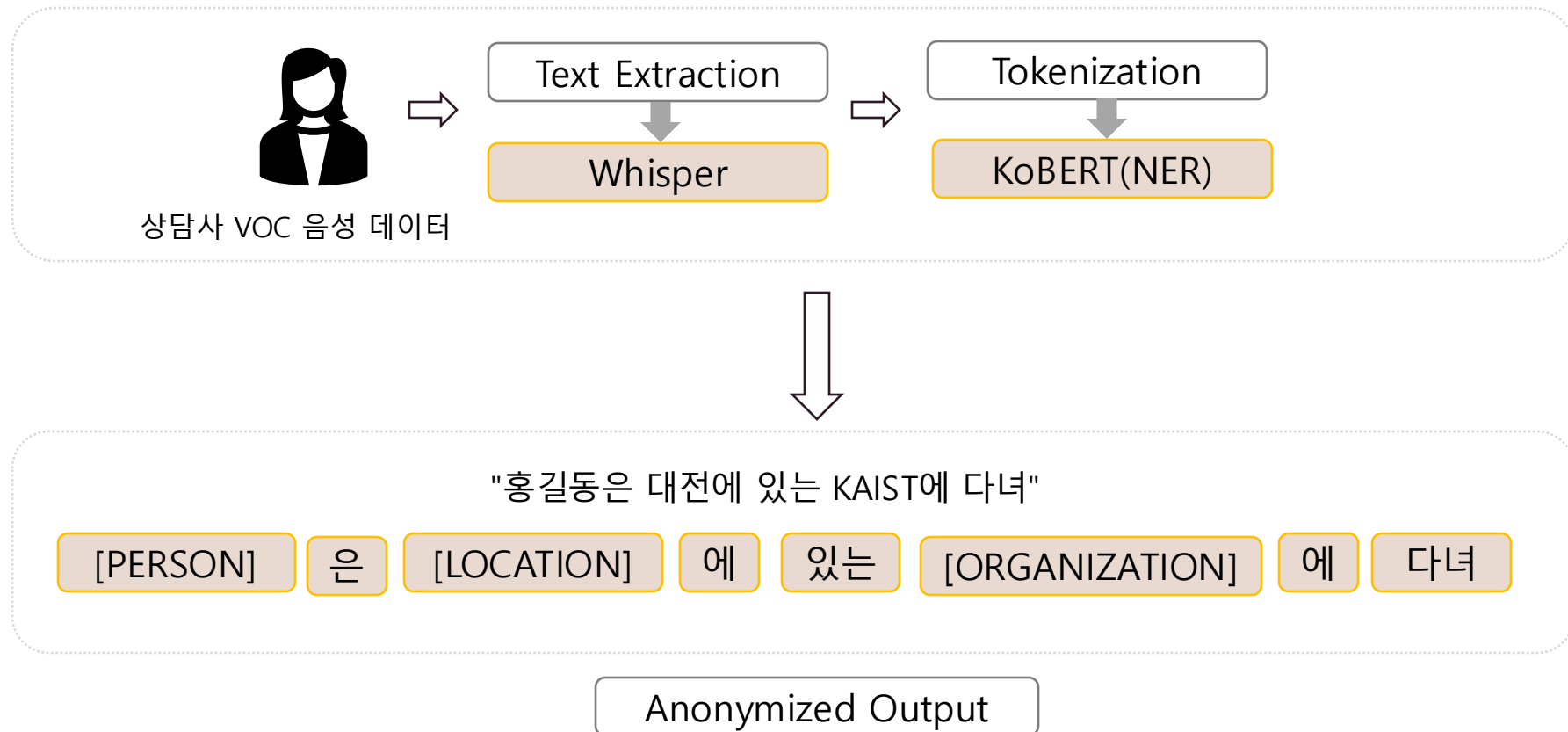
- B-ENTITY: 해당 개인정보 항목의 시작 단어에 붙임
 - 예: [B-ENTITY 홍길동 I-ENTITY] → "홍길동"이라는 엔티티의 첫 토큰
- I-ENTITY: 같은 엔티티의 두 번째 이후 단어에 붙임
 - 예: [B-ENTITY 급성골수성백혈병 I-ENTITY]

- KoBERT 기반 모델
 - 한국어에 최적화된 사전학습 BERT 모델(KoBERT)을 활용하여 자연어 이해 성능 확보
- Transformer 기반 개체명 인식(NER)
 - 셀프어텐션 구조*를 활용해 문장에서 이름, 전화번호 등 주요 개체를 자동 식별
- 라벨링 데이터 기반 학습
 - 사람이 지정한 예시(예: "홍길동 → PERSON")를 바탕으로 모델이 개체 구분 기준을 학습
- 보험사 VOC 특화 튜닝
 - 콜센터·챗봇 대화문에서 자주 사용되는 어투와 표현을 반영하여 도메인 맞춤형 모델로 고도화

"KoBERT 기반의 Transformer NER 모델을 활용해, 비정형 VOC 데이터에서도 고도화된 가명식별 자동화 구축"

• Transformer 기반 : 입력 간의 관계를 파악하기 위해 자기주의 메커니즘(Self-Attention)을 활용하는 딥러닝 모델 구조
• 셀프어텐션 구조 : 입력 문장 내 단어들 간의 관계를 동적으로 계산해, 중요한 정보에 더 집중할 수 있게 하는 메커니즘
• NER 모델 : 문장에서 사람, 기관, 날짜 등 특정 개체명을 자동으로 식별하고 분류하는 자연어처리 모델
• 토큰 : 문장을 구성하는 가장 작은 단위로, 보통 단어나 형태소 수준의 분절된 요소, 엔티티 : 문장에서 의미 있는 이름이나 개념으로, 사람, 장소, 조직 등 특정 범주로 분류되는 개체

5. 가명처리 모델링 예시 - NER 활용



6. 가명처리 모델링 - LLM 활용

1

데이터 라벨링

- Zero-Shot / Few-Shot 방식 (라벨링 불필요 또는 최소화)
- 도메인 특화 튜닝 (필요시 소규모 라벨링)

2

모델 세팅 및 Fine-Tuning

- Llama 2 (Meta): 7B~70B 파라미터, 로컬 GPU 서버에 설치 가능
- GPT-J / GPT-NeoX: 오픈소스, 비교적 경량화된 버전 확보 쉬움
- Mistral 7B: 최근 공개된 고성능·경량 모델

3

추론 및 성능평가

- Precision, Recall, F1 지표로 토큰·엔티티 단위 성능 검증 수행

- 프롬프트 역할 지정
- 예시
 - 당신은 PII 추출 전문가입니다... JSON 형식으로 뽑아주세요

- 구체적 지시 + 출력 포맷
- 예시

다음 문장에서 개인정보(PII)를 전부 찾아, JSON 배열로 출력하세요.

[INPUT]

"홍길동 고객님의 연락처는 010-1234-5678, 주민번호는 123456-1234567 입니다."

[OUTPUT 예시]

```
[  
  {"entity": "홍길동", "type": "PERSON"},  
  {"entity": "010-1234-5678", "type": "PHONE"},  
  {"entity": "123456-1234567", "type": "ID_NUMBER"}  
]
```

"오프라인 LLM 모델을 활용해 비정형 VOC 데이터에서도 고도화된 가명식별 자동화 구축"

7. 가명처리 메커니즘 및 매핑 전략

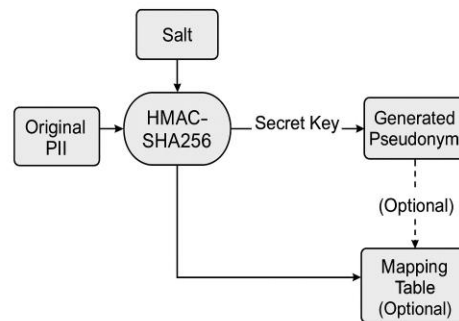
시스템 환경과 보안 요구에 맞춰 적합한 가명화 방법 적용

동적 매핑 테이블 방식

- 순차 가명 할당
 - "사용자A", "사용자B" 형태로 PII별 가명 지정
- 일관성 유지
 - 기존 매핑값 재사용으로 동일 대상 간 가명 일치
- 재식별 가능성 고려
 - 분석·법무 협업 등 재식별 필요한 환경에 적합
- 가독성 우수
 - 사람이 인식하기 쉬운 형태로 가명 설정 가능
- 제한사항
 - 테이블 유출 시 보안 리스크 존재, 관리 부담 발생

해시 기반 처리 방식

- 동일성 보장
 - 동일 PII → 동일 토큰 생성 (HMAC-SHA256 + salt)
- 보안성 강화
 - 단방향 해시로 역추적 방지, 재식별 위험 최소화
- 클라우드 전송 최적
 - 민감정보 유출 방지를 위한 강력한 보호 수단
- 제한사항
 - 사람이 이해하기 쉬운 가명 생성 어려움
 - 데이터 구조·패턴 유지에는 한계



- Original PII와 Salt를 함께 HMAC-SHA256 함수에 입력
- Secret Key를 사용해 해시 토큰(Generated Pseudonym) 생성
- (Optional) 매핑 테이블에 원본과 가명 관계 저장

[해시 기반 처리 방식 예시]

8. 의학 개체명 인식 및 하이브리드 검출 전략

VOC 내 의료정보 정밀 식별을 위한 하이브리드 개체명 인식 전략

사전 구축

- 의료 도메인 특화 개체 사전(SNOMED-CT, ICD-10, LOINC) 제작 및 희귀병 리스트 확보
 - SNOMED-CT: 전 세계적으로 가장 방대한 임상용어(임상 개념)가 모여 있는 표준화된 용어집
 - ICD-10: WHO가 정의한 질병 분류 코드 체계
 - LOINC: 실험실 검사 및 임상측정 항목(검사 코드) 사전

하이브리드 검출

- 정규표현식과 모델 기반 접근 병행하여 날짜·수치·진단명 등 추출 정확도 극대화
- 표준 날짜 형식과 수치 단위를 처리하는 정규표현식 룰과, 문맥 정보를 학습한 딥러닝 모델 기반 NER을 병행 적용.

NER 모델 학습

- KoBERT/ClinicalBERT 기반 fine-tuning으로 의료 문서 내 개체명 정밀 검출
 - ClinicalBERT(의료 코퍼스*로 재학습된 KoBERT)는 진단서·처방전·병리보고서 등 실제 의료 문서를 바탕으로 사전학습되어, 약물명·질병명·검사명 등 의료 개체의 분포와 문맥을 효과적으로 이해

AI 접근법의 장점

- 의료 문맥 이해 기반 추출 정확도 향상
 - 딥러닝 기반 NER 모델은 문장 내 의미 흐름과 상관관계를 반영하여 단순 키워드 탐지보다 정밀한 개체 식별 가능
- 도메인 적응형 확장성 확보
 - ClinicalBERT 등 사전학습된 의료 특화 모델 활용으로 진단명, 약물명, 검사명 등 다양한 의학 개체에 대한 대응력 강화

• 의료 코퍼스: 의학 분야에서 수집된 텍스트 데이터 집합으로, 의료용어, 진료기록, 논문 등으로 구성되어 NLP 모델 학습에 활용

9. 의료 데이터 가명처리 대상 및 평가 기법

- ✓ 의학 PII는 식별 위험도에 따라 분류되며, 유형별 맞춤 비식별화와 평가 기준 적용을 통해 데이터 보호와 활용 간 균형을 실현해야 함

의학 PII 분류

• 직접식별자: 이름, 주민번호 등

• 준식별자: 희귀 질환명, 검사 수치, 입원 날짜

• 민감속성: 정신건강, 유전질환 정보



평가 기법

• k-익명성*: 개인 식별 위험 최소화 그룹화

• t-근접성*: 검사 수치 분포 노이즈 추가

• l-다양성*: 그룹 내 속성 다양성 보장

- K-익명성 : 어떤 사람을 알아보기 어렵게, 최소 K명과 구분되지 않도록 정보를 바꾸는 개인정보 보호 방법
- T-근접성 : 같은 그룹 내 민감한 정보의 분포가 전체 데이터와 비슷하도록 만들어, 특정 정보가 드러나지 않게 하는 익명화 기법
- l-다양성 : 동일한 그룹 안에 다양한 민감 정보가 포함되도록 해, 특정 값을 유추하기 어렵게 만드는 개인정보 보호 기법

10. 의료 데이터 가명처리 성능 향상 기법

K-익명성

- 추출 방법
 - 준식별자(QI)를 기준으로 그룹핑
 - 예: (성별, 연령대, 우편번호) 같은 조합별로 레코드 묶기
 - 그룹별 크기(레코드 개수) 확인
- K-익명성을 높이는 방안
 - 일반화: 연령→'20대', '30대' 같이 범주 넓히기
 - 삭제: 드물게 나타나는 QI 조합 일부 레코드 제거
 - QI 속성 축소: (성별, 연령, 우편번호)→(성별, 연령) 등 차원 낮추기

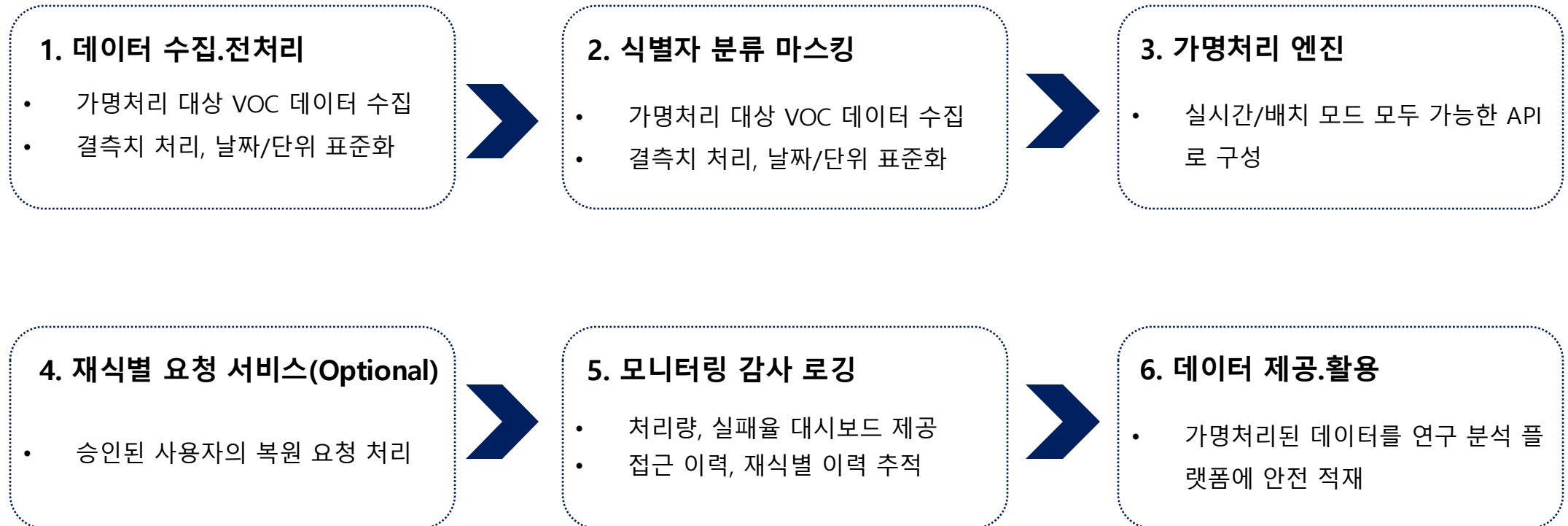
T-근접성

- 추출 방법
 - 그룹 내 민감속성 분포 산출
 - 예: 그룹 A의 혈압 수치 분포
 - 전체 데이터 민감속성 분포 산출
 - 거리 계산 즉, 그룹 분포와 전체 분포 간 최대 거리
- T-근접성을 높이는 방안
 - 노이즈 추가: 그룹 내 값들을 전체 분포에 가깝게 조정
 - 예: 각 검사치에 소량의 랜덤 노이즈를 더해 분포 차이 최소화
 - 버킷화: 검사항목 연속값을 구간별 카테고리로 묶어 분포 단순화
 - 그룹별 대표값(mean)으로 치환

I-다양성

- 추출 방법
 - 각 QI 그룹 내 민감속성 distinct 개수 계산
 - 예: 그룹 A의 진단명 5개
 - 최소 다양성 추출
- I-다양성을 높이는 방안
 - 그룹 크기 조정: 그룹 병합 또는 세분화로 민감속성 고유값 늘리기
 - 동의어·유의어 반영: "고혈압", "HTN" 같은 동의어를 하나로 묶거나 분리
 - 추가 속성 활용: 검사치·처방약 정보 등을 민감속성 식별변수로 함께 고려

11. 가명처리 통합 파이프라인 설계



12. 운영 및 보안, 테스트 전략

운영 유의사항	Fine tuning 방안	테스트 및 검증
<ul style="list-style-type: none"> ➤ 법무·윤리 심의위와 가명처리 리스트 사전 합의 ➤ 모델 및 룰셋 변경 이력관리, 모니터링과 피드백 체계 구축 ➤ 역할 기반 접근 제어(RBAC)* 혹은 속성 기반 접근 제어(ABAC)* 등의 책임 분리 및 접근 통제 <ul style="list-style-type: none"> ○ 관리자/개발자/검토자 등 역할 구분에 따른 권한 체계 명시 ➤ 보안성을 높이기 위해 시스템 로그 수집 및 정기 검토 체계 구축 <ul style="list-style-type: none"> ○ 접근 이력 및 모델 업데이트 로그 저장 및 주기적 검토 	<ul style="list-style-type: none"> ➤ 병목 원인 파악 <ul style="list-style-type: none"> • CPU vs I/O vs 네트워크 병목 구분 → 멀티스레드·병렬화 적용 ➤ 정확도 개선 <ul style="list-style-type: none"> • 정규표현식 룰 보강: FP/FN 사례 수집 후 룰 업데이트 • 모델 재학습: 문제 사례 포함한 추가 라벨링으로 NER 모델 Fine-tuning ➤ 자동화 보고서 <ul style="list-style-type: none"> • 매 실험마다 결과 수집·시각화 ➤ 데이터 품질 관리 <ul style="list-style-type: none"> • 오타/누락 데이터에 대한 정기적 리뷰 및 재라벨링 프로세스 구축 ➤ 사전 학습 적용 <ul style="list-style-type: none"> • 도메인 적합도를 고려한 이미 학습된 모델을 활용하는 Transfer Learning 전략 활용 	<ul style="list-style-type: none"> ➤ Precision, Recall, F1 등 정량적 성능 평가 ➤ 법무팀 및 현업 리뷰 통한 정성적 평가 ➤ 상용 솔루션 대비 처리량 및 정확도 벤치마크 ➤ 상용 솔루션과 구축 모델 성능을 비교할 수 있는 평가용 파이프라인 별도 구축 ➤ 처리 실패 시 리스크가 큰 항목에 대한 우선 순위 테스트 전략 적용

- 역할 기반 접근제어(RBAC) : 사용자에게 부여된 역할에 따라 시스템 자원 접근 권한을 관리하는 보안 방식
- 속성 기반 접근제어(ABAC) : 사용자, 자원, 환경 등의 다양한 속성을 기준으로 정교하게 접근 권한을 제어하는 보안 방식

13. 연구 결과 및 기대 효과

연구 요약

- 금융/보험사 데이터 특화 가명처리 자동화 파이프라인 구축 및 법적 검증 방안 마련

차기 연구 방향

- 다국어 지원, 실시간 스트리밍 처리 기술 도입, 강화학습 기반 가명처리 고도화

기대 효과

- 데이터 3법 준수 하에 AI 서비스 확대, 금융/보험사 고객 맞춤형 서비스 품질 향상

Thank you

AI 기반 가명처리 기술을 활용하여 보험사 VOC(고객의 소리) 데이터를 안전하게 처리하고 활용하는 방안제시