



개인정보보호위원회, 비정형데이터 가명처리 관련 ‘가명 정보 처리 가이드라인’ 개정

2024.02.08

개인정보보호위원회는 2024년 2월 5일 개정된 ‘가명정보 처리 가이드라인’을 공개하였습니다. 이번 개정은 기존 가이드라인이 ‘정형데이터’에 대한 처리기준만 제시하고 있었던 한계를 보완하여, 이미지·영상·음성·텍스트 등 인공지능 기술개발의 핵심인 ‘비정형데이터’에 대한 가명처리 기준을 새롭게 마련하기 위한 것입니다.

※ 정형데이터와 비정형데이터의 차이점

구분	정형데이터	비정형데이터
정의	정해진 규칙에 맞게 구조화된 형식으로 존재하는 데이터	일정한 규격이나 정해진 형태가 없이 구조화되지 않은 데이터
특징	데이터 연산, 분석 등 데이터 처리 방식, 가명처리 기술·방법이 비교적 단순	연구목적·환경에 따라 데이터 처리방식 및 가명처리 기술·방법이 복잡·다양
예시	DB에 열과 행으로 저장된 테이블형식의 자료 등	사진·이미지, 비디오, 통화음성, 대화기록, 논문·보고서, 블로그 등

이번 개정을 통해 ‘가명정보 처리 가이드라인’에는 (i) 비정형데이터 가명처리 기준, (ii) 비정형데이터 가명처리 기술 및 예시 및 (iii) 비정형데이터 가명처리 시나리오 예시 등이 추가되었으며, 주요내용을 아래에서 더 자세히 살펴보도록 하겠습니다.

주요내용

1. 비정형데이터 가명처리 기준

- 데이터 처리목적·환경, 민감도 등을 종합적으로 고려하여 개인식별 위험성이 있는 정보를 판단하고, 합리적인 처리 방법·수준 설정 개인식별위험이 있는 정보가 비교적 명확히 구분되는 정형데이터와 달리 비정형데이터는 개인식별 가능 정보에 대한 판단이 상황에 따라 달라질 수 있음연구목적 달성에 필수적인 정보항목을 남기는 대신 그 외 정보항목에 대한 가명처리 수준을 높이거나, 다른 정보

및 소프트웨어(SW) 반입 제한 등 충분한 안전조치를 보완하여 활용

- 가명처리 기술의 한계 등을 보완하기 위해 사전 준비단계(연구 및 기술개발 기획)부터 위험성을 충실히 검토하고 적절한 안전조치를 수행 비정형데이터에 내재된 개인식별 위험 요인을 완벽하게 탐지하여 처리할 수 있는 기술이 아직 없음가명처리 기술의 한계를 보완하기 위해 다음 조치를 이행할 것을 권고

- ① 가명처리 기술의 적절성·신뢰성을 확인할 수 있는 근거 작성·보관
- ② 가명처리 기술 적용 이후, 처리 결과에 대해 자체적인 검수 수행
- ③ 가명처리 적정성 검토 과정에서 ①, ②를 포함하여 점검(외부전문가 과반 이상 참여 바람직)

- 데이터 복원기술 발달 등에 대응하여, 가명처리된 비정형데이터 활용 시 관련 시스템·SW의 접근·사용 제한 등 통제방안 마련 또한, AI 서비스 제공 과정에서도 개인식별 위험 등 정보주체 권익 침해 가능성을 지속 모니터링

2. 비정형데이터 가명처리 기술 및 예시

- 영상정보: 이미지 필터링 기술, 이미지 암호화, 얼굴 합성(프라이버시 보존형 데이터 마이닝), 인페인팅(Inpainting), AI 이용 영상정보 가명처리 등
- 음성정보: 음성정보 자체에 대한 가명처리(VT, VC, GMM매핑, DROPSY), 텍스트로 변환 후 가명처리 등
- 텍스트정보: 규칙기반 개인정보 삭제 혹은 마스킹, 스크리빙(Scrubbing), 정규표현식(Regular Expression), 주석 달기 (Annotation), AI기반 텍스트정보 가명처리, 텍스트를 테이블 형식으로 변환 등

3. 비정형데이터 가명처리 시나리오 예시

- 의료데이터, CCTV 영상, 음성 대화·상담 정보를 활용하는 바이오·헬스, 인공지능, 데이터 솔루션 개발 등 다양한 분야에서의 가명 정보 활용 전 과정을 상세히 안내 주요 분야별 비정형데이터 가명처리 시나리오 예시

분야	비정형데이터	예시
의료	이미지, 영상, 텍스트	유방암·골밀도 감소 여부 진단 AI 개발
의료	이미지	구강질환 진단 AI 개발
의료	이미지, 영상	안면골 골절 진단 AI 개발
교통	이미지, 영상	자율주행차 주행 시 비정상 상황 인지 AI 개발
교통	이미지	고속도로 다인승전용차 단속 AI 개발
대화·검색	텍스트	한국어 대화가 가능한 AI 챗봇 제작
대화·검색	음성, 텍스트	콜센터 직원 교육용 가상상담 시나리오 생성 AI 개발

시사점

- AI 기술과 컴퓨팅 자원의 발달로 '비정형데이터'에 대한 활용수요는 폭발적으로 증가하였음에도 기업, 연구기관 등은 비정형데이터의 적합한 가명처리 방법이나 수준을 정확히 알 수 없었습니다.
- 이번 개정 '가명정보 처리 가이드라인'에 포함된 비정형데이터 가명처리 사례 내지 시나리오는 단순 참고용으로 각 사업자의 데이터 활용 분야·상황에 맞게 가명처리 방법·수준 등을 자유롭게 적용할 수 있는데, 가이드라인에서 제시된 사례 등을 통해 비정형데이터를 활용하는 다양한 신기술 영역에서 현장의 불확실성을 어느정도 해소할 수 있을 것으로 보입니다.
- 이번 가명처리 가이드라인 개정 외에도 개인정보보호위원회는 「가명정보 활용 확대방안」(23.7)을 발표하는 등 안전한 데이터 활용을 진흥하고자 하는바, 추후에도 이와 관련한 정부의 주요 동향들을 지속적으로 파악할 필요가 있습니다.

About Shin & Kim's ICT Group

법무법인(유) 세종은 개인정보 분야에 차별화된 전문성과 인적 네트워크(윤종인 전 개인정보보호위원회 위원장, 김영호 전 행정안전부 차관, 최재유 전 과학기술정보통신부 차관 등)를 보유하고 있으며, 기업들을 위하여 개인정보보호법과 GDPR을 비롯한 국내외 개인정보 규제, 개인정보 유출사건 대응, 개인정보보호 컴플라이언스 체계 수립 등 개인정보보호에 관한 전문적인 자문을 제공하고 있습니다. 특히, 데이터 3법 개정 및 하위법령 제정, 관련 제도개선에 있어 민간영역에서 주도적인 역할을 수행한 바 있으며, 가명정보, 데이터 활용, ICT 산업 전반에 대한 규제 동향 파악 및 대관, 입법컨설팅, 규제영향력 분석과 기업의 전략 수립 등에 대한 법률자문을 제공하고 있으므로, 보다 전문적인 내용이나 궁금하신 사항이 있으면 언제든지 연락 주시기 바랍니다.

관련구성원

강신욱

대표변호사

02-316-4059

sokang@shinkim.com

안정호

변호사

02-316-2891

jhahn@shinkim.com

장준영

변호사

02-316-4985

jyojang@shinkim.com

윤호상

변호사

02-316-2584

hsyoon@shinkim.com

강지현

변호사

정영운

변호사

02-316-1518
jhykang@shinkim.com

02-316-1822
ygjeong@shinkim.com

Copyright SHIN & KIM LLC. All rights reserved.