의료 데이터 비식별화 기술 동향과 AI 기반 재식별에 관한 연구

서은경

연세대학교

seoek@yonsei.ac.kr

A Study on Trends of Medical Data De-identification Technology and AI-Based Re-Identification

Seo Eun Kyung

Yonsei Univ.

요 약

의료 데이터는 환자 진단, 치료 및 연구 등 다양한 분야에서 필수적인 자원으로 활용되지만, 민감한 개인정보를 포함하고 있어 안전한 데이터 활용이 중요한 과제로 부각되고 있다. 본 연구는 의료 데이터 비식별화 기술의 정의, 발전 과정 및 최신 동향을 정리하고, 실제 활용 사례를 통해 비식별화 기술의 적용 현황을 분석하였다. 또한, 인공지능(AI) 기술을 활용한 재식별 문제를 조명하고, 연구 결과 및 사례를 통해 비식별화만으로는 개인정보 보호가 충분치 않음을 확인하였다. 나아가, 재식별 위험에 대응하기 위한 기술적 방안과 법적·제도적 개선 방향을 검토하였다.

Ⅰ. 서 론

의료 데이터는 환자의 진단, 치료, 임상 시험 및 공중보건 연구 등 다양한 분야에서 필수적인 자원으로 활용되고 있다. 그러나 이러한 데이터에는 개인의 민감한 건강 정보가 포함되어 있어, 데이터 활용 과정에서 개인정보 보호 문제가 매우 중요하다. 특히, 전자건강기록(Electronic Health Records, EHR), 의료 영상, 유전체 정보 등은 높은 수준의 식별 가능성을 가지므로, 이를 안전하게 활용하기 위한 비식별화(de-identification) 기술이 필수적이다.[1]

최근 의료 분야에서는 인공지능(AI) 기반 분석 및 예측 모델이 급속히 확산되고 있으며, 이는 비식별화된 데이터의 활용을 촉진하는 동시에 새로운 보안 위협을 야기하고 있다. AI 기반 재식별(re-identification) 공격은 기존 비식별화 방법만으로는 충분히 대응하기 어렵다는 문제 제기가 나온 후, 이에 대한 연구가 활발히 진행되고 있다. 본 논문에서는 의료 데이터 비식별화 기술의 정의, 발전 과정과 최신 동향을 정리하고, 실제 활용 사례를 살펴본 후, AI 기반 재식별 문제와 이를 둘러싼 기술적·법적 대응 방안을 분석하고자 한다.

Ⅱ. 본론

1. 의료 데이터 비식별화 기술: 정의, 발전과정, 최신 동향

의료 데이터 비식별화는 개인을 직접 또는 간접적으로 식별할 수 있는 개인식별정보(Personal Identifiable Information, PII)를 제거하거나 변형하여, 데이터 활용 시 개인의 프라이버시를 보호하는 기술을 의미한다. 비식별화 기법은 크게 가명화(pseudonymization), 익명화(anonymization), 데이터 마스킹(data masking), 통계적 변형(statistical perturbation) 등으로 구분된다.[2]

초기에는 단순한 데이터 삭제나 이름·주민등록번호와 같은 직접 식별자의 제거를 중심으로 발전했으나, 최근에는 k-익명성(k-anonymity), l-다양성(l-diversity), t-근접성(t-closeness)과 같은 통계적 보호 기법과, 의료 영상 및 텍스트 데이터에 특화된 심층학습 기반 비식별화 기술이 등장하였다.[3] 특히, 의료 영상의 경우 단순한 픽셀 제거만으로는 충분하지 않아, AI 기반 알고리즘을 활용하여 얼굴 구조, MRI 스캔 내 고유 특징 등을 변형하거나 제거하는 방법이 활용되고 있다[4].

최근 동향으로는 AI를 활용한 자동 비식별화 기술과 클라우드 기반 데이터 비식별화 서비스가 주목받고 있다. 이러한 기술은 대규모 의료 데이터를 신속하고 정확하게 처리할 수 있으며, 데이터 활용의 효율성을 높이는 동시에 개인정보보호 수준을 유지할 수 있다는 장점이 있다.[5]

2. 의료 데이터 비식별화 기술 활용 사례

(1) PixelGuard

AWS(Amazon Web Services)의 PixelGuard는 AI 기반 의료 영상 비식별화 솔루션으로, CT·MRI·X-ray 이미지 내 개인 식별 요소를 자동으로 탐지하고 제거한다. PixelGuard는 기계학습 모델을 활용해 이미지 속 개인 정보를 식별하고, 이를 비식별화된 형태로 변환하여 연구용 데이터셋으로 제공한다.[6]

(2) Shaip

Shaip의 비식별화 기술은 개인정보를 보호하면서 데이터 활용을 가능하게 하는 솔루션이다. 환자 기록, 진료 노트 등 다양한 데이터 유형을 처리할 수 있다. 이름, 주소, 주민등록번호 같은 직접 식별자와 생년월일, 직업 등 간접 식별자를 제거하거나 일반화하며, 마스킹, 익명화, 토큰화, 가명화 등 다양한 방법을 사용한다. 특히 의료 데이터에서 미국 건강 보험 양도 및 책임에 관한 법(HIPAA) 규정을 준수하며, 자동화된 처리와 전문가 검수를 결합해 정확성을 높이고 있다. 이를 통해 연구, AI 학습, 데이터 공유 등에서 개인정보를 안전하게 보호하면서 데이터를 활용하는 데 기여하고 있다.[7]

(3) Google Healthcare API

Google Healthcare API는 FHIR(Fast Healthcare Interoperability

Resources) 기반 의료 데이터를 대상으로, 다양한 비식별화 기능을 제공하며, 데이터 마스킹, 제거, 일반화 등의 기법을 통합하여 안전한 데이터 공유를 지원하고 있다.[8]

이러한 서비스들은 의료 연구의 효율성을 높이는 동시에 데이터 프라이버시 보호를 강화하는 사례로 주목받고 있으며, 특히 다기관 연구, AI 모델 학습용 데이터셋 구축 등에서 필수적인 기반 기술로 활용되고 있다.

3. 의료 데이터 비식별화 기술과 AI 기반 재식별 문제

(1) AI 기반 재식별 문제: 연구결과 및 사례

비식별화된 의료 데이터라도, AI 기술을 활용하면 개인 식별이 가능해지는 사례가 보고되고 있다. Rocher 등(2019)은 생성 모델을 이용해 불완전한 데이터셋에서도 높은 성공률로 개인을 재식별할 수 있음을 실험적으로 입증하였다.[9] 또한, 최근 연구에서는 k-익명성 기반 데이터에서도 AI를 활용한 combinatorial refinement attack 기법을 통해 재식별 공격이 가능함이 보고되었다.[10]

의료 영상 분야에서도 유사한 위험이 존재한다. 예를 들어, MRI 영상의 경우 AI가 미세한 해부학적 특징을 학습하여 개인을 재식별할 수 있으며, 이는 단순 비식별화만으로는 방어가 어렵다.[11] 이처럼 AI 기반 재식별 문제는 비식별화 기술의 한계를 드러내며, 데이터 보호 정책과 기술적 대응의 필요성을 동시에 부각시킨다.[12]

(2) AI 기반 재식별 문제에 대한 기술적 차원에서의 대응

재식별 공격에 대응하기 위해 최근에는 다중 보호 계층(Multi-layered protection) 접근이 제안된다. 이는 데이터 전처리 단계에서 익명화, 통계적 변형, GAN(Generative Adversarial Network) 기반 합성 데이터 생성 등을 결합하는 방식이다.[13] 또한, 동적 데이터 마스킹과 접근 제어를 결합하여, 데이터 분석자는 필요한 최소한의 정보만 접근하도록 제한함으로써 재식별 위험을 낮출 수 있다.[14]

또한, AI 기반 재식별 가능성을 평가하고 사전에 위험도를 산정하는 프레임워크 개발이 필요하다. 예를 들어, pseudo-identification risk 평가, 민감도 기반 데이터 분할, 그리고 합성 데이터 생성 기법은 실제 연구 환경에서 재식별 가능성을 줄이는 데 기여할 수 있다.[15]

4. 비식별화된 의료 데이터의 재식별 문제와 개인정보보호법

비식별화 데이터의 재식별은 개인의 민감한 건강 정보가 외부로 유출될 수 있다는 점에서 심각한 개인정보 침해 문제를 야기한다. 현행 개인정보보호법과 가명정보 처리 가이드라인은 데이터 비식별화를 통해 개인정보를 보호하도록 규정하고 있으나, AI 기반 재식별 문제를 충분히 고려하지 못하고 있다.[2] 따라서, 재식별 위험 평가를 의무화하고, 데이터 활용 목적과 범위를 명확히 제한하며, 비식별화 과정에 대한 검증 절차를 강화하는 법적·제도적 보완이 필요하다. 특히, 임상 연구에서 다기관 데이터를 공유할 때에는 재식별 위험을 최소화하기 위한 기술적 조치와 법적 준수사항을 병행해야 한다.[16]

III. 결론

의료 데이터의 비식별화는 개인정보 보호와 연구 활용의 균형을 맞추는 핵심 기술로 자리 잡았다. AI 기술의 발전은 비식별화 효율성을 높였으나, AI 기반 재식별 문제는 여전히 해결해야 할 과제로 남아 있다. 따라서 기술적 보호, 위험 평가, 법제도적 보완을 통합적으로 고려한 접근이 필요하다. 향후 연구는 AI 기반 재식별 가능성을 사전에 평가하고, 다중 보호 계층과 합성 데이터 활용 등 안전한 데이터 공유 방법론을 개발하는 방향으로 진행되어야 할 것이다.

참 고 문 헌

[1] 보건산업브리프 Vol.268. "의료데이터 활용을 위한 개인정보 비식별화 기술 및 프로그램 동향." 2019.

[2] 개인정보보호위원회. "가명정보 처리 가이드라인(2024.2.)." 2024.

[3] 정영수. "신뢰 기반의 인공지능을 위한 개인정보 보호 제도개선 방안 연구." 성균관대학교 일반대학원, 박사학위논문. 2024.

[4] Faustini, P., McIver, A., Sullivan, R., Dras, M. “De-identification of clinical data: A systematic review of free text, image and tabular data approaches.” International journal of medical informatics vol. 208 (2026): 106225. doi:10.1016/j.ijmedinf.2025.106225

[5] Gunay, M., Keles, B., Hizlan, R. “LLMs-in-the-Loop Part 2: Expert Small AI Models for Anonymization and De-identification of PHI Across Multiple Languages.” arXiv, 2024, arXiv:2412.10918. doi.:10.48550/arXiv.2412.10918

[6] Amazon Web Services(AWS). "PixelGuard: Advancing Healthcare Data Privacy through AI-Driven De-Identification System for Medical Imaging Research." 2025. (https://aws.amazon.com/ko/blogs/publicsector/pixelguard-advancing-healthcare-data-privacy-through-ai-driven-de-identification-system-for-medical-imaging-research/).

[7] Shaip. "Data De-identification." (https://ko.shaip.com/offerings/data-deidentification/).

[8] Google Cloud. "Healthcare API: De-identification Concepts." 2026. (https://docs.cloud.google.com/healthcare-api/docs/concepts/de-identification?hl=ko).

[9] Rocher, L., Hendrickx, J. M., and de Montjoye, Y. A. "Estimating the success of re-identifications in incomplete datasets using generative models." Nat Commun 10, 3069. 2019. doi:10.1038/s41467-019-10933-3

[10] Wang, Z., Khatibi, E., Firouzi, F., Mousavi, S. R., Chakrabarty, K., and Rahmani, A. M. “Linkage Attacks Expose Identity Risks in Public ECG Data Sharing.” arXiv, 2025, arXiv:2508.15850.

[11] Sarkar, A. R., Chuang, Y.-S., Mohammed, N., and Jiang, X. “De-identification is not enough: a comparison between de-identified and synthetic clinical notes.” Sci Rep 14, 29669. 2024. doi:10.1038/s41598-024-81170-y.

[12] Hallaj, S. et al. “Navigating open data sharing and privacy in the age of clinical AI research: from reidentification to pseudo-reidentification.” EClinicalMedicine vol. 91 103729. 29 Dec. 2025. doi:10.1016/j.eclinm.2025.103729.

[13] Morris, J. X., Campion, T. R., Nutheti, S. L., Peng, Y., Raj, A., Zabih, R., and Cole, C. L. “DIRI: Adversarial Patient Reidentification with Large Language Models for Evaluating Clinical Text Anonymization.” arXiv, 2024, arXiv:2410.17035.

[14] Aguelal, H., and Palmieri, P. “De-Anonymization of Health Data: A Survey of Practical Attacks, Vulnerabilities and Challenges.” International Conference on Information Systems Security and Privacy. 2025;2:595-605. doi: 10.5220/0013274200003899

[15] Chhillar, S., Righi, M. K., Sutter, R. E., and Kornaropoulos, E. M. “Exposing Privacy Risks in Anonymizing Clinical Data: Combinatorial Refinement Attacks on k-Anonymity Without Auxiliary Information.” arXiv, 2025, arXiv:2509.03350.

[16] Madhusudhanan, S., and Jose, A. C. “Privacy preservation techniques through data lifecycle: A comprehensive literature survey.” Computers & Security 155, 2025, doi:10.1016/j.cose.2025.104473.