

PRiVACY REPORT

2024 개인정보 이슈 심층 분석 보고서

개인정보 안심구역



| CONTENTS |

2024 Vol. 5

개인정보 안심구역

1. 개인정보 침해 최소화 및 유연한 개인정보 활용을 위한 개인정보
안심구역 제도 1
[윤수영/ 에이팩스프라이버시랩 대표]
2. PET 기술실증을 위한 개인정보 안심구역 활성화 방안 10
[오준형/ 서울여자대학교 조교수]
3. 제로트러스트 기반의 가명정보 처리 환경을 적용한 데이터 유용성
제고 방안 17
[박영수/ 윤법률사무소 변호사]

개인정보 침해 최소화 및 유연한 개인정보 활용을 위한 개인정보 안심구역 제도



윤수영

에이펙스프라이버시랩 대표

1. 들어가며

2020년 8월 데이터3법 시행에 따라 가명정보의 개념이 도입되고 개인정보처리자는 통계작성, 과학적 연구, 공공적 기록보존 등을 위한 목적으로 개인정보를 가명처리하여 활용할 수 있는 기반이 새롭게 마련되었다. 다만, 가명정보 처리 과정에서 데이터 품질이 저하되면 데이터 결합률도 떨어질 수 있어 원데이터를 사용할 때에 비해 데이터 분석 및 연구 활동에 한계가 존재한다.

한편, 개인정보의 보호 및 안전한 활용을 위해 개인정보보호위원회(이하 '개인정보위')는 2024~2026년 개인정보 보호 기본계획의 비전으로 '국민 신뢰 기반의 디지털 대전환 선도'를 선포하고 ① 데이터 경제시대 선도, ② 개인정보 안심사회 구현, ③ 글로벌 데이터 신질서 주도 등 3대 추진 전략 10대 과제를 수립한 바 있다. 이 중 첫 번째 전략인 '데이터 경제시대 선도'를 위한 '개인정보의 안전한 활용 촉진' 과제의 일환으로 '안전한 데이터 활용을 위한 법·제도 기반 조성'을 위해 환경적 안전성을 높임으로써 개인·가명정보를 보다 유연하게 활용할 수 있도록 하는 '개인정보 안심구역' 도입 계획을 밝힌 바 있다.^[1]

'개인정보 안심구역'이란, 유연한 개인·가명정보 활용, 개인정보 보호·활용기술 실증 등을 위해 개인정보위가 전문심의를 지원하고, 개인정보 처리에 필요한 안전한 이용환경을 지원하는 제도이다.^[2] 특히, 안심구역은 데이터를 장기간 보관·활용할 수 있으며, 보관된 데이터를 제3자가 재사용할 수 있다는 장점이 있다. 또한 가명정보 활용을 촉진하고, 데이터 연계 및 융합을 통해 고품질의 데이터를 제공하며, 안전하게 데이터를 활용할 수 있는 토대를 제공한다.

개인정보위는 결합전문기관, 가명정보 활용지원센터 등을 대상으로 2023년부터 시범운영 기관 신청을 받아왔으며, 2024년 10월 기준 통계청, 국립암센터, 한국도로공사, 한국사회보장정보원, (주)더존비즈온 등 5개소가 개인정보 안심구역으로 지정되어 있다. 개인정보위는 다양한 연구자와 기업이 안심구역을 통해 양질의 데이터를 보다 안전하게 활용할 수 있는 기반을 마련할 수 있을 것으로 기대하고 있으며 가명데이터의 이용 활성화에 따라 개인정보 안심구역의 지정은 더욱 확대될 전망이다. 이에, 개인정보 안심구역의 추진경과와 현황을 살펴보고 안심구역을 통한 보다 안전하고 활발한 개인정보 활용을 위한 방안을 모색하고자 한다.

2. 개인정보 안심구역 현황 분석

가. 개요

개인정보 안심구역의 기본 방향은 제로트러스트 원칙을 기반으로 데이터 활용 환경의 안전성을 높이고, 강화된 안전성에 비례하여 개인정보의 유연한 활용을 지원하는 것이다. 개인정보 보호 기본계획 및 안심구역 시범사업 지정 등을 통해 제시된 안심구역의 주요 기능으로는 안전한 환경 제공, 개인정보 유연한 활용, 개인정보보호 강화기술(PET, Privacy Enhancing Technology) 실증 지원 등이 있으며 관련 내용은 아래와 같다.

개인정보 안심구역 주요 기능

- ① 안전한 환경 제공 : 물리적 폐쇄환경, 개인정보 활용 전 과정 추적 관리, 멀티팩터 인증, 전문심의위원회 운영 분석결과물만 반출 등
 - ② 개인정보(가명정보)의 유연한 활용
 - (가명처리 수준 완화) 환경적 안전성 강화 수준에 비례하여, 가명처리 수준을 적정수준으로 완화하여 가명처리 후 분석 가능
 - (비정형데이터 활용) 비정형(영상, 이미지, 텍스트 등) 빅데이터에 대해 전문심의위가 검증한 가명처리 SW 적용 및 샘플링 검사 후 활용 허용
 - 불필요한 데이터 손실을 최소화시켜 양질의 데이터 확보·활용 가능
 - (가명정보 장기보관 및 재사용) AI 연구개발 등 지속적·반복적인 연구를 위해 가명정보의 장기관 보관 및 제3자 재사용* 가능
 - ※ 전제조건: ①「개인정보 보호법」에서 정한 목적 내 활용, ②원 보유기관과의 협의
 - ③전문심의위원회 심의, ④(필요시) 처리 목적·환경에 따른 추가 가명처리
 - AI 기술개발 등에 필요한 데이터 확보에 소요되는 시간·비용 대폭 절감 가능
 - ③ PET* 실증 지원
 - 기존 제도적용이 모호한 PET의 경우 전문심의위원회가 실증계획을 사전 검증하여 해당 PET를 적용한 개인정보 처리 허용
 - 프라이버시 우려, 모호한 규제적용 등으로 막혀있던 신기술 연구개발 가능
- * PET(Privacy Enhancing Technology): 동형암호, 합성데이터 등 프라이버시 향상 기술을 활용하고자 할 경우 가명·익명처리, 결합 등 제도적용이 모호하거나 불명확함

출처 : 저자 구성

나. 지정 평가 기준 및 절차

개인정보 안심구역의 지정은 개인정보위가 지정심사위원회(이하 '지정심사위')를 구성하고 지정심사위가 지정요건 충족 여부, 사업계획 구체성, 참여기관 적정성, 기대효과 등을 종합평가하여 지정한다. 지정심사위는 개인정보위가 기재부, 과기정통부, 복지부, 행안부, 금융위 등 관계부처로부터 2배수 추천을 받아 개인정보 보호, 데이터의 활용, 데이터 보안에 대한 전문가 5인으로 구성된다,

평가기준은 적합·부적합으로 평가되는 지정요건 및 정성평가되는 기관역량 종합심사로 구분된다. 유연한 개인정보 활용을 위한 환경적 안전성 구축 측면에서의 지정요건은 4인 이상의 담당조직 구성, 안심구역 운영을 위한 공간·시설·시스템 구축, 시스템·인증·이용환경·데이터 관리 등 보호조치 등이 있다. 기관역량에 대해서는 사업 추진 목표 및 추진 체계 등 사업 계획 구체성, 참여 인력 전문성 및 수요 창출 계획의 적절성 등 참여기관 적정성, 해당 분야 데이터 산업 발전 효과 및 기여도 등의 기대효과를 정성평가한다.

표 1 지정 절차 및 평가 기준

| 구분 | 주요 내용 |
|-----------------------------|--|
| 지정절차 | - 지정신청(지정신청자) → 신청서 검토 및 접수(개인정보위) → 지정심사(지정심사위) → 지정결정 통지 및 공고(개인정보위) |
| 지정요건 심사 (적합, 조건부적합, 부적합) | - (담당조직 및 정책) 4인 이상의 담당조직 구성 (관리책임자, 보안·운영·행정인력), 운영정책 및 개인정보보호 정책 수립, 교육계획 수립 등 - (공간·시설·시스템) 개인정보 안심구역 운영을 위한 공간·시설·시스템 구축(폐쇄망·망분리 환경, 멀티팩터 인증, 보안관제시스템 등) - (보호조치) 시스템, 인증, 이용환경, 데이터 관리 등 보호조치 수행 |
| 기관역량 심사 (정성평가) | - (사업계획 구체성) 사업 추진목표 및 업무 이해도, 내용의 명확성, 사업 추진체계, 기존 데이터 활용 사업과의 연계방안 등 - (참여기관 적정성) 참여인력의 전문성, 시설·시스템 구축의 적절성, 수요창출 계획의 적절성 및 실현가능성, 개인정보 안심구역 활용 수요자의 접근성, 사업 운영 계획의 적절성, 중장기 발전 방안 등 - (기대효과) 해당 분야 데이터 산업 발전 효과 및 기여도 |

다. 추진 경과

개인정보위는 결합전문기관, 데이터 안심구역, 가명정보 활용지원센터를 대상으로 2023년 4~5월 개인정보 안심구역 시범운영 대상기관 선정을 위한 신청을 받아 2023년 12월 ‘통계청’과 ‘국립암센터’를 안심구역 시범운영 대상기관으로 선정한 바 있다. ‘통계청’은 2024년 3월 현장실사를 거쳐 안심구역 운영기관으로 최종 지정되어 본격 운영을 시작하였고 ‘국립암센터’ 또한 7월부터 「보건의료 특화형 개인정보 안심구역」 본격 운영을 시작했다. 또한, 2024년 7월에는 한국도로공사, 한국사회보장정보원, 더존비즈온 등 3개 기관이 안심구역 시범운영기관으로 추가 지정되었다.

개인정보 안심구역 지정 주요 경과

- (개인정보위) 개인정보 안심구역 시범운영기관 공모 : '23. 11. 1 ~ 11. 30
- (통계청·국립암센터) 개인정보 안심구역 시범운영기관 최초 지정 : '23. 12월
- (통계청) 개인정보 안심구역 최종 지정 및 개소 : '24. 3월
- (국립암센터) 개인정보 안심구역 최종 지정 및 개소 : '24. 7월
- (개인정보위) 개인정보 안심구역 시범운영기관 지정 공모 : '24. 4. 23 ~ 5. 31
- (한국도로공사·한국사회보장정보원·더존비즈온) 안심구역 개인정보 시범운영기관 추가 지정 : '24. 7월

개인정보위는 이같은 안심구역을 2025년도에도 확대 추진하고, 관련 연구 용역 및 안심구역 성과 분석 등을 통해 기존 폐쇄망 시스템에서 클라우드 환경까지 확장할 수 있도록 개선하는 적용 방안을 마련할 예정이다.^[3]

개인정보위가 지정한 개인정보 안심구역을 통해 이용하고자 하는 이용신청자는 안심구역 지정기관에 이용신청서를 제출하면 안심구역 운영기관 검토 및 전문심의위원회의 이용심의를 통해 안심구역 이용 및 분석결과물 반출 신청이 가능하다. 안심구역 지정기관은 안심구역을 통한 데이터 활용결과를 기록·보관하고 홈페이지에 공개하여야 하며 전반적인 상세 이용 절차는 아래와 같다.

그림 1 개인정보 안심구역 이용절차^[4]

| | 이용신청자 | 개인정보 안심구역 운영기관 | 전문심의위원회 |
|-----------------|---|--|--|
| 1. 이용신청 | ① 개인정보 안심구역 이용신청 - 이용신청서 제출 (가명정보 이용 신청, 가명정보 재사용 신청, PET 실증 신청, 기타) - 개인정보 활용계획서 제출 (보완 요청시 서류 보완) | ② 신청 접수 - 서류 검토 및 미비서류 보완요청 | — |
| 2. 심의요청 | — | ③ 개인정보보호위원회에 전문심의위원회 구성 요청 - 전문심의위원회 심의계획서 작성 | ④ 전문심의위원회 구성 - 3인 이상 7인 이하 (심의요청일 기준 14일 이내 구성) |
| 3. 이용심의 | ⑧ 심의결과 확인 - 개인정보 안심구역 이용대상 · 범위 · 방법 확정 - 추가 프라이버시 보호조치 이행 | ⑤ 이용 심의 일정 · 절차 등 협의 및 조율 (전문심의위원회 운영) ⑦ 심의결과 안내 - 이용승인 시 이용현황 및 데이터 카탈로그를 운영 기관 홈페이지에 공개 | ⑥ 개인정보 안심구역 이용심의 - 개인정보 활용계획 사전검검 및 추가 프라이버시 보호 조치 권고 ※ 심의요청일 기준 30일 이내 심의완료 |
| 4. 개인정보 안심구역 이용 | ⑨ 개인정보 안심구역 이용 | ⑩ 이용 지원 · 관리 - 안심구역 인프라 기술지원, 개인정보 안심구역 이용에 대한 안전관리, 추가 프라이버시 보호조치 이행 확인 (필요시 전문심의위원회에 보고) - 가명처리 및 분석 지원 (가능한 경우) | — |
| 5. 반출심사 | ⑪ 분석결과물 반출 신청 | ⑫ 반출심사 일정 · 절차 등 협의 및 조율 ⑮ 반출승인 통보 및 분석결과물 반출 | ⑬ 분석결과물 반출심사 ⑭ 반출승인 ※ 심의요청일 기준 14일 이내 심사완료 |
| 6. 이용결과 관리 | — | ⑯ 활용결과 기록 · 보관 ⑰ 홈페이지에 활용결과 공개 | — |

출처: 한국사회보장정보원

라. 데이터안심구역과의 비교

과학기술정보통신부(이하 ‘과기정통부’) 또한 데이터산업법에 따라 기술적·물리적·관리적 보안대책 등 지정 기준이 충족되었음이 인정되는 경우, 데이터를 안전하게 분석·활용할 수 있는 데이터안심구역을 운영 중이다. 데이터안심구역 또한, 민감한 미개방데이터의 유출 우려를 해소하며 양질의 데이터 분석 수요를 충족할 수 있을 것으로 기대하고 있으며 2024년 10월 현재 강원도, 농림수산물교육문화정보원, 전라북도·국민연금공단, 한국데이터산업진흥원(서울센터), 한국도로공사, 한국전력공사, 한국지능정보사회진흥원, 한국데이터산업진흥원(대전센터) 등 총 9개소가 데이터안심구역으로 지정된 바 있다.^[5] 데이터안심구역 및 개인정보 안심구역은 그 취지 및 지정 절차가 유사한 면이 많으며 개인정보 안심구역의 지정 신청 대상으로 결합기관 외에도 데이터안심구역이 포함된다. 이에 한국도로공사는 데이터안심구역 및 개인정보 안심구역 모두로 지정되어 보다 효율적인 데이터 분석 및 활용이 가능할 것으로 전망된다.

표 2 데이터안심구역 및 개인정보 안심구역 비교

| 구분 | 데이터안심구역 | 개인정보 안심구역 |
|-------|--|--|
| 법적 근거 | 데이터산업법 데이터안심구역 지정 및 운영 관한 지침 | 개인정보 안심구역 운영 가이드 |
| 주요 경과 | ’22년 9월 | ’23년 11월 |
| 지정요건 | <ul style="list-style-type: none"> ■ (인력구성 등) 운영책임자, 보안책임자를 포함 4명 이상 조직구성, 보안대책이 구비될 수 있는 건물·시설 (인터넷기반자원공유(클라우드) 포함) ■ (보안대책 등) 안심구역 운영에 관련한 기술적·물리적·관리적 보안대책 마련, 운영계획, 이용절차·데이터 반입·반출 정책 등 마련 | <ul style="list-style-type: none"> ■ (담당조직 및 정책) 4인 이상의 담당조직 구성 (관리책임자, 보안·운영·행정인력), 운영정책 및 개인정보보호 정책 수립, 교육계획 수립 등 ■ (공간·시설·시스템) 개인정보 안심구역 운영을 위한 공간·시설·시스템 구축 (폐쇄망·망분리 환경, 멀티팩터 인증, 보안관제시스템 등) ■ (보호조치) 시스템, 인증, 이용환경, 데이터 관리 등 보호조치 수행 |
| 지정절차 | <ul style="list-style-type: none"> ■ 서류제출(신청기관) → 안심구역 자문단 심사(서면 및 현장심사 등) → 지정(과기정통부) | <ul style="list-style-type: none"> ■ 지정신청(지정신청자) → 신청서 검토 및 접수(개인정보위) → 지정심사(지정심사위) → 지정결정(개인정보위) → 통지 및 공고(개인정보위) |
| 지정현황 | 강원도, 농림수산물교육문화정보원, 전라북도·국민연금공단, 한국데이터산업진흥원(서울센터), 한국도로공사, 한국전력공사, 한국지능정보 사회진흥원, 한국데이터산업진흥원(대전센터) 등 총 9개소 | 통계청, 국립암센터 (개소 준비 중) 한국도로공사, 한국사회보장정보원, ㈜더존비즈온 등 총 5개소 |

3. 기대효과 및 전망

개인정보위가 요구하는 안전성이 강화된 데이터 처리 환경을 갖추어 개인정보 안심구역으로 지정되면 ▲가명처리 수준의 적정 완화 ▲다양한 결합키를 활용한 결합률 제고 ▲지속적·반복적 연구를 위한 가명정보의 장기 보관 및 제3자 재사용 허용 ▲빅데이터(영상·이미지 등) 표본(샘플링) 검사 ▲개인정보 보호 강화기술(PET) 실증 등의 지원이 이루어진다.^[6] 이에, 개인정보 안심구역 지정기관들은 안심구역을 통해 양질의 데이터 분석과 결합이 이루어져 더 정확하고 생산적인 데이터 활용이 가능할 것으로 기대되고 있다.

| 안심구역 주요 기대효과 | |
|------------------|-----------------|
| 기존 | 개인정보 안심구역 도입 후 |
| 단기 활용(통상 6개월~1년) | 장기 활용(5년+추가연장) |
| 일회성 활용 후 파기 | 재사용 제3자 활용 가능 |
| 가명처리 적정성 전수 검사 | 빅데이터 샘플링 검사 허용 |
| 결합키 사용 제한 | 결합키 다양화(결합률 향상) |

이와 같은 기대효과는 지정기관별 특성 및 데이터 처리 분야에 따라 다양할 것으로 전망된다. 우선 최초로 개인정보 안심구역으로 지정된 통계청은 통계목적고유번호 등 다양한 결합키 활용한 데이터 결합 및 데이터 장기 활용에 따른 시계열 분석 등 지속적 반복적 연구 측면의 효과가 클 것으로 기대된다. 통계청은 이미 전국 110개 기관에서 300여 종의 행정자료를 수집해 66종의 새로운 통계 데이터를 생산하고 있으며 개인정보 보호법에 의거한 결합전문기관의 지위 및 신용정보법상 데이터전문기관의 지위를 가지고 있어 통계 데이터 결합을 통한 다양한 연구 활동이 가능하다.

그러나, 가명정보 기반의 데이터 결합제도의 경우 까다로운 절차 및 낮은 결합률로 데이터의 정확성이 많이 떨어지는 한계가 있어 3개 기관의 가명정보 데이터 결합 시 낮게는 30%, 높아도 78%, 89% 수준에 불과하였다고 한다.^[7] 반면, 국가통계 특화형 개인정보 안심구역은 주민등록번호를 SHA256 모듈을 통해 일방향으로 암호화한 데이터인 통계목적 고유번호를 보유한 통계작성기관들끼리 정보를 공유하여, 그 결합키로 데이터를 결합하도록 만든 체계이다. 이 결합키를 활용한 결합 방식은 훨씬 높은 결합률을 제공하며, 안전성과 신뢰성도 높을 것으로 기대된다. 이를 위해 기존 통계데이터센터 분석실 내에 개인정보 안심구역 분석공간과 심사공간을 마련하고 지정PC를 두었으며 모든 시스템·네트워크와 분리된 안전한 폐쇄망 환경을 구축하여 안전한 데이터 처리 환경을 구축하였다.^[8]

또한, 보건의료계에서는 개인식별 위험이 크다는 이유로 연구자체가 불가능하거나 희귀암, 민감상병 관련 연구에서는 과도한 가명처리로 데이터 품질이 훼손돼 연구결과를 신뢰할 수 없는 경우가 많았으나, 개인정보 안심구역에서 안전한 데이터 처리 환경에서 이러한 연구를 제대로 이루어질 수 있을 것으로 기대되고 있다.

2024년 7월부터 「보건의료 특화형 개인정보 안심구역」을 운영하는 국립암센터는 희귀암, 민감상병*, 유전체 데이터 분야 등의 연구에 있어, 기존 연구에서의 가명처리 데이터의 품질 저하로 인한 연구 수행 불가 혹은 연구 결과 신뢰성 저하 문제점들을 안심구역을 통해 해소할 수 있을 것으로 전망하고 있다. 안심구역 내에서는 데이터의 품질이 일정수준 이상 유지될 수 있어 데이터 품질을 과도하게 훼손하지 않고도 다양한 연구가 가능할 것으로 기대되기 때문이다. 또한, 가명처리 데이터의 적정성 검사를 위해선 원칙상 전체 데이터의 전수검사가 필요하지만, 안심구역에서는 표본(샘플링) 검사가 가능하기 때문에 머신러닝 등 의료 영상·이미지 빅데이터를 활용하는 데 소요되는 시간과 비용도 절감할 수 있을 것으로 기대된다.^[9]

* 민감상병이란 환자의 부상, 질병, 질환 등의 건강정보이며 이 중 정보주체의 인권 및 생활보호에 중대한 피해를 야기할 수 있는 상병으로 정신질환, 희귀질환, 확대 및 낙태 관련 상병 등이 해당

한국사회보장정보원(이하 '사보원') 또한 데이터 기반 정책 연구 활성화와, 패널 데이터의 장기 활용 및 다른 연구자의 재사용 허용으로 다른 데이터와의 결합 등 새로운 관점의 연구가 가능해져 보다 많은 연구자들이 가명정보를 창의적으로 활용할 수 있을 것으로 기대하고 있다. 특히, AI 분야에 있어서도 사보원이 운영 중인 복지사각지대 발굴 AI모델 개선을 안심구역에서 진행할 수 있게 되어 위기가구 발굴을 위한 보다 정교한 AI 모델 개선이 가능할 것으로 예상되며 복지 분야 AI 적용 확대를 위한 거점 역할을 수행할 수 있도록 할 예정이다.^[10]

앞서 살펴보았듯이 개인정보 안심구역을 통한 데이터 활용결과를 기록·보관하고 홈페이지에 공개하도록 되어 있어, 다양한 기업 및 연구자들이 향후 공개될 안심구역 데이터 활용결과들을 참고하여 안심구역을 통한 활발한 데이터 분석이 가능할 것으로 기대된다.

4. 맺음말

개인정보 안심구역 제도는 개인정보 보호와 데이터 활용 사이의 균형을 유지하며, 다양한 연구와 혁신을 가능하게 만드는 중요한 기반을 제공한다. 아직은 시행 초기이나 보호와 활용의 비례성에 입각하여 공공기관을 중심으로 한 초기 안심구역 운영이 안정적으로 자리 잡아가면서, 앞으로 더 많은 민간 기업의 참여를 유도하는 것이 중요할 것이다. 이를 위해서는 민간 기업이 안심구역을 통해 데이터를 안전하게 활용할 수 있도록 법적·제도적 지원이 필요하다. 나아가 개인정보 안심구역 제도의 활성화를 위해서는 보다 유연하고 효율적으로 운영할 수 있도록 절차 간소화와 시스템을 개선하는 것 또한 고려할만한 사항이다. 이를 통해 공공과 민간 모두가 데이터 활용을 통해 혁신적 성과를 창출하고, 개인정보 보호라는 기본 원칙을 유지하면서도 데이터 경제를 활성화하는 데 기여할 수 있을 것이다.

주석

- [1] 개인정보보호위원회(‘23.6.28), 개인정보 보호 기본계획(2024-2026), (URL: <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS061&mCode=C010010000&nttId=8998>)
- [2] 개인정보보호위원회, 개인정보 안심구역 시범운영기관 지정 공모 계획(‘24.4.23) (URL: <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=10096>)
- [3] 개인정보보호위원회(‘24.6.21), 제5기 공공데이터전략위원회 4차 회의 의안번호 제21호, 「가명정보 활용 활성화 방안」
- [4] 한국사회보장정보원(사업소개)정책사업지원)개인정보 안심구역, (URL: <https://www.ssis.or.kr/lay1/S1T1972C1973/contents.do>)
- [5] 과학기술정보통신부(‘24.1.4), “과기정통부, 데이터안심구역 신규 지정” (URL: https://www.msit.go.kr/bbs/view.do?jsessionid=h_hwT1TjfwEtAAHj-NKT95Lw-HAz4d5e6SvT-blT.AP_msit_1?sCode=user&mPid=238&mId=113&bbsSeqNo=94&nttSeqNo=3183922)
- [6] 개인정보보호위원회(‘24.4.23), “개인정보위, 「개인정보 안심구역」 신규 기관 공모” (URL: <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS074&mCode=C020010000&nttId=10096>)
- [7] 보안뉴스(‘24.3.11) “[현장탐방] ‘개인정보 안심구역’ 시범인증기관, 통계청을 가다” (URL: <https://m.boannews.com/html/detail.html?idx=127592>)
- [8] 통계청(‘24.3.28), “이제는 개인정보 안심구역을 통해 안전하게 통계데이터를 연계하고 분석하세요” (URL: https://kostat.go.kr/board.es?mid=a10301010000&bid=246&act=view&list_no=430165&tag=&nPage=1&ref_bid=203,204,205,206,207,210,211,11109,11113,11814,213,215,214,11860,11695,216,218,219,220,10820,11815,11895,11816,208,245,222,223,225,226,227,228,229,230,11321,232,233,234,12029,10920,11469,11470,11817,236,237,11471,238,240,241,11865,243,244,11893,11898,12031,11825,246&keyField=T&keyWord=&bodo_b_type=all)
- [9] 개인정보보호위원회(‘24.7.31), “개인정보 안심구역 개소식” (URL: <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=BS211&mCode=C040020000&nttId=10401>)

- [10] 한국사회보장정보원('24. 7. 3), “한국사회보장정보원, ‘개인정보 안심구역’
시범운영기관 지정” (URL:
[https://www.ssis.or.kr/lay1/bbs/S1T1910C1976/B/101/view.do?mode=vi
ew&article_seq=125962&cpage=1&rows=10&condition=&keyword=](https://www.ssis.or.kr/lay1/bbs/S1T1910C1976/B/101/view.do?mode=view&article_seq=125962&cpage=1&rows=10&condition=&keyword=))

PET 기술실증을 위한 개인정보 안심구역 활성화 방안



오준형

서울여자대학교 조교수

1. 배경

디지털 대전환 시대에 데이터는 산업 발전과 가치 창출의 핵심 경쟁력이자 촉매제 역할을 담당하고 있다. 이에 따라 주요 국가들은 데이터 패권 경쟁에서 우위를 차지하기 위해 데이터 전략 자산 확보 및 데이터 신산업 육성을 국가적 어젠다로 추진 중이다. 미국은 2019년 12월 연방데이터 전략(Federal Data Strategy)에 이어 2021년 9월 기업데이터 전략(Enterprise Data Strategy)을, EU는 2020년 2월 EU 데이터 전략(European Data Strategy)을 수립했다. 마찬가지로 중국은 2021년 빅데이터 산업 발전계획(2021-2025)을, 영국은 2020년 9월 국가데이터전략(National Data Strategy)에 이어 2021년 9월부터 인공지능 전략(National AI Strategy)을 펼쳐나가고 있다.^[1]

이러한 세계적인 흐름 속에서 가장 중요한 데이터 중 하나로 떠오르는 것은 개인정보이다. 개인정보의 효과적인 활용은 국가 데이터 전략의 성공 여부를 좌우하는 핵심 요소이다. 하지만 현실적으로는 가명정보 및 개인정보의 활용에 대한 여러 제약이 존재한다. 글로벌 IT 시장 분석 및 컨설팅사 IDC(International Data Corporation)의 조사에 따르면 전 세계에서 생성되는 데이터의 70~75%는 개인정보에 해당하지만^[2], 재식별 우려와 컴플라이언스 부담 등으로 인해 데이터 제공과 활용이 쉽지 않다. 특히 AI 연구개발 및 데이터 학습에 필요한 데이터 확보와 활용 면에서 이러한 제약은 더욱 부각된다.^[3]

가명정보는 처리 목적, 방법, 환경 등이 사전에 구체적으로 설정되어야 하며^[4], AI 연구개발 및 학습을 위한 장기간의 데이터 축적이 어려운 상황이다.^[5] 또한 프라이버시 향상을 위한 다양한 기술인 PET(Privacy Enhancing Technology, 프라이버시 강화 기술)가 발전하고 있지만, 이를 실제 현장에서 적용하고 활용할 수 있는 법적, 제도적 장치가 미흡하다. 동형암호, 합성데이터, 차분 프라이버시 기술 등이 대표적인 예로, 가명정보 제도를 적용해야 하는지, 익명정보로 자유롭게 쓸 수 있는지에 대한 불명확함이 존재한다.^[6]

이러한 제약을 해결하고 개인정보 및 가명정보의 안전한 활용을 도모하기 위해서는 환경적 통제장치가 필요하다. 이에 따라 정부 차원의 제도적 지원과 인프라 구축이 시급히 요구되고 있다. 개인정보 안심구역은 이러한 맥락에서 등장한 개념으로, 개인정보 보호와 데이터 활용의 혁신을 동시에 이룰 수 있는 제도로 주목받고 있다.

2. 개인정보 안심구역

개인정보 안심구역은 데이터 처리 환경의 안전성을 높여 가명정보의 활용을 활성화하기 위해 도입된 제도로써, 기존의 데이터 활용에서 제한적이었던 부분들을 해소하고 보다 안전한 환경에서 데이터의 분석 및 활용이 가능하도록 지원한다.^[7] 이를 위해 개인정보보호위원회는 2024년 4월 23일부터 ‘개인정보 안심구역 시범운영기관’ 지정 공모를 추진하였으며, 이러한 개인정보 안심구역은 제로 트러스트(Zero Trust) 기반의 보안모델과 데이터 처리과정에 대한 사전·사후적 통제를 바탕으로 다양한 데이터 처리에 대한 안전성을 보장하고 있다.^[8]

제로 트러스트 보안 모델은 “아무것도 신뢰하지 않는다”는 원칙을 기초로 내부 사용자에 대해서도 무조건적인 신뢰를 두지 않으며, 데이터 처리 과정 전반에 대한 검증을 통해 보안 수준을 강화하는 것을 목표로 한다. 구체적으로는 4인 이상의 담당조직, 오프라인 폐쇄망 분석 환경^[9], 멀티팩터 인증^[10], 데이터 외부 반출 금지 등으로 데이터 처리의 안전성을 확보한다. 이러한 환경적 안전성은 제로 트러스트 보안 모델의 원칙에 따라 데이터의 안전한 활용을 보장하며, 이를 통해 다양한 데이터 처리에 대한 제약이 완화되고 기존에 접근이 어려웠던 데이터의 활용이 용이해진다. 특히, 개인정보 안심구역의 운영을 통해 아래와 같은 기능들이 가능해진다.

- 가명처리 수준 완화
- 다양한 결합 키 활용으로 결합률 상승
- 시계열 분석 등 지속적 반복적 연구를 위한 가명정보의 장기간(5년) 보관 및 제3자 재사용 허용
- 영상·이미지 등 빅데이터에 대한 표본(샘플링) 검사 실시
- 동형암호 등 개인정보보호 강화기술(PET) 실증

개인정보 안심구역은 양질의 가명정보를 환경적 안전성을 갖춘 구역에서 처리할 수 있어 제한됐던 다양한 데이터 처리가 가능해진다는 점이 핵심이다. 다양한 데이터 처리의 유연성과 안전성을 동시에 추구하며, 가명정보의 활용을 촉진하는 역할을 수행한다.

개인정보 안심구역 지정 및 운영 사례

개인정보보호위원회는 다양한 연구자와 기업이 양질의 데이터를 안전하게 활용할 수 있도록 기반을 마련하기 위해 안심구역을 도입하였다. 다양한 분야의 기관들이 개인정보 안심구역으로 지정되어 운영 중이거나 개소를 준비 중에 있다.

표 1 개인정보 안심구역 지정 및 운영 사례

| 기관명 | 주요 분야 | 데이터 제공 방식 | 데이터 종류 |
|-----------------|-------|----------------------------------|---|
| 통계청 | 국가통계 | 국가통계 특화형 안심구역을 통해 고품질 데이터 | - |
| 국립암센터 | 보건의료 | 비정형(영상, 이미지, 음성, 텍스트, 유전체 등) 데이터 | 희귀암, 민감상병 데이터 등 기존 가명처리 수준으로 연구 목적 달성이 어려웠던 연구 지원 |
| 한국도로공사 (심사중) | 국토교통 | 도로, 버스, 철도, 시설, 자율주행 데이터 | - |
| 한국사회보장정보원 (심사중) | 사회복지 | 사회복지 분야에서 정책 수립에 필요한 데이터 | 국민 맞춤형 복지를 실현하는 사회복지 분야 정책 수립에 필요한 양질의 데이터를 생산 예정 |
| 더존비즈온 (심사중) | 정밀의료 | 의료 빅데이터 | 민간 가명정보 결합전문기관 중 최초 개인정보 안심구역 시범운영기관으로 지정 |

3. 개인정보 안심구역 활용 및 활성화 방안

사례분석: 과기정통부 데이터 안심구역 운영 사례

과학기술정보통신부에서는 데이터 안심구역의 활성화를 위해 아래와 같은 활동을 진행하고 있다. 이런 사례들은 개인정보 안심구역의 활성화에 참고할 수 있을 것이다.

표 2 과기정통부의 데이터 안심구역 활성화 방안 및 주요 활동

| 활성화 방안 | 활동명 | 내용 |
|---------------------|---|--|
| 데이터 분석 및 활용 교육 프로그램 | 2024 데이터 안심구역 분석캠프, 2023 데이터 안심구역 분석캠프 | 서울/대전센터에서 데이터 활용 및 분석을 위한 캠프 진행 |
| | 데이터 안심구역 분석캠프 미개방 물류데이터 활용과정, 데이터 안심구역 분석캠프 미개방 소비데이터 활용과정 | 안심구역에 적재되어 있는 데이터를 중심으로 데이터 기반 문제해결 및 의사결정 등 미니 프로젝트를 통한 데이터 활용 경험 제공 |
| | 「데이터 안심구역 분석캠프」 AtoZ 상반기 과정 | 데이터 안심구역의 분석 환경 및 데이터를 소개하고 안심구역에서 제공하는 분석도구 설명 및 미개방데이터 탐색 등 맛보기 실습 진행 |
| | 「데이터 안심구역 분석캠프」파이썬 과정(초급), 파이썬 야간과정(초급), 대전센터 「데이터 안심구역 분석캠프」 R과정(초급) | 데이터 안심구역의 분석 환경 및 데이터를 소개하고 범용 프로그래밍 언어(Python) 혹은 패키지(R)를 활용한 데이터 분석 기초 및 방법 안내 |

| | | |
|------------------------|-------------------------------------|---|
| | 2022년 데이터분석입문과정 교육 (1차~6차) | 데이터 안심구역의 미개방 데이터를 활용한 데이터 분석 기초 교육으로 실무 중심의 데이터 분석 인재 양성을 위한 프로그램 |
| 데이터 활용 경진대회 및 사례 공모 | 데이터 안심구역 활용 공동 경진대회 (2024, 2023) | 데이터 안심구역 내 데이터 활용 우수 사례 발굴 및 데이터 활용 저변 확대를 위한 경진대회 개최 |
| | 2024 데이터 안심구역 활용 사례 모집 | 데이터 안심구역 및 미개방데이터 활용 우수 사례 발굴을 위한 공모전 개최 |
| | 데이터 활용 아이디어 및 시각화 경진대회 | 데이터 활용 아이디어 발굴 및 시각화 경진대회를 통해 창의적인 데이터 활용 아이디어와 기술적 방안 마련 |
| | 데이터 활용 아이디어 공모 (2022) | 미개방 데이터를 활용한 블루오션 발굴을 목적으로 데이터 분석 및 활용 아이디어를 모집하는 공모전 개최 |
| 설명회 | 2024 데이터 유통 활용 제도 지역설명회 | 데이터 산업진흥 및 이용 촉진을 위한 설명회 개최, 데이터 유통·활용 제도 및 지원 사업 소개 |
| 기업 지원 | 24년 데이터 안심구역 활용 지원 수요기업 모집 | 미개방 데이터 및 고성능 분석환경 활용 지원, 데이터 안심구역 활용 사업화 컨설팅 지원 |

위 내용 외로도 데이터 안심구역에서는 ‘대한민국 4차 산업혁명 2020 페스티벌’이나 ‘데이터 활용 아이디어 및 시각화 경진대회’ 등 다양한 활동을 통해 데이터 이용자와 개발자들이 개인정보 보호를 준수하면서 데이터를 활용할 수 있도록 지원하고 있다. 이러한 접근 방식을 개인정보 안심구역에도 적용하여, 실무적인 교육 프로그램과 참여의 장을 마련함으로써 데이터 이용자와 개발자들이 개인정보 보호의 중요성을 인식하고 실천할 수 있도록 할 수 있다.

개인정보 안심구역의 활성화 방안

a) 개인정보 안심구역 내 데이터 활용 가이드라인 및 표준 절차 수립

분야별, 목적별 데이터 활용 가이드라인을 수립하고 표준화된 절차를 마련한다면 데이터의 안전한 결합과 활용을 위한 표준화된 프로세스를 구축할 수 있을 것이다. 각 기관이 개인정보 보호법 및 관련 지침을 준수하여 데이터를 활용할 수 있도록 지원하는 것이다. 특히, 개인정보의 가명 처리, 안전한 데이터 결합 및 분석 절차에 대한 구체적인 가이드라인을 제공함으로써 데이터 활용의 일관성을 높이고, 안전한 데이터 유통 환경을 조성할 수 있다.

b) 개인정보 안심구역 내 데이터 활용 및 보호 인식 제고를 위한 세미나 및 홍보 캠페인 추진

데이터 활용과 개인정보 보호에 관심 있는 다양한 이해관계자들을 대상으로 정기적인 세미나,

포럼 등을 개최하여 최신 동향, 사례, 기술 등을 공유한다. 이를 통해 업계 간 협력을 강화하고, 데이터를 안전하고 효과적으로 활용하기 위한 지식과 네트워크를 구축할 수 있다.

뿐만 아니라 개인정보 안심구역에서의 안전한 데이터 결합과 활용에 대한 홍보 캠페인 및 홍보 자료를 제작하여 일반 시민과 기업의 인식을 높이고, 개인정보 보호와 데이터 활용이 조화롭게 이루어질 수 있도록 정책적 지원과 법적 기반을 강화할 수 있다.

c) 우수 사례 공모전 개최

개인정보 안심구역 내 데이터 활용에 대한 실증적 활용을 장려하기 위해 안심구역 내에서 기술을 적용한 실제 데이터 활용 및 결합 프로젝트를 추진한다. 이를 통해 안전한 데이터 활용의 효과를 시범적으로 보여줄 수 있다. 또한 우수 활용 사례에 대한 공모전을 개최하여 PET 기술의 혁신적인 활용 사례를 발굴하고 포상함으로써 기술 도입 및 활용에 대한 동기를 부여할 수 있다. 이러한 사례는 교육 및 홍보 자료로 활용되어 개인정보 안심구역의 활용성과 필요성을 알리는데 기여할 것이다.

d) 아이디어톤 및 해커톤 개최

개인정보 안심구역에서 안전하게 가명 처리된 실제 데이터셋을 효과적으로 활용할 수 있는 아이디어를 발굴하는 아이디어톤이나 데이터셋을 활용하여 PET 기술 적용 및 데이터 분석을 주제로 해커톤을 개최한다. 이를 통해 PET 기술을 효과적으로 적용할 수 있는 새로운 기술과 아이디어를 발굴하고, 우수한 아이디어에 대해서는 사업화 지원을 통해 실제 서비스로 구현될 수 있도록 돕는다. 데이터 활용 방안을 모색하고 혁신적인 접근을 제안하여 안심구역의 가치를 높이고 참여자들이 개인정보 보호를 준수하면서도 데이터를 활용할 수 있는 구체적 방안을 탐색할 수 있을 것이다.

4. 결론

개인정보 안심구역은 데이터 활용과 개인정보 보호의 균형을 맞추고, 안전하고 효율적인 데이터 결합 및 분석 환경을 제공함으로써 다양한 분야에서 데이터 활용의 가치를 제고할 수 있는 핵심 제도이다. 안심구역의 운영을 통해 기존 가명정보 처리에서 발생했던 데이터 훼손 문제를 최소화하고, 보다 고품질의 데이터를 안전하게 활용할 수 있는 기반을 조성하고 있다. 이러한 노력은 데이터 기반의 산업 발전과 혁신 기술 도입을 가속화하고, AI 연구개발을 비롯한 다양한 분야에서 데이터 활용 역량을 향상시키는 데 기여할 것으로 기대된다.

또한, 개인정보 안심구역의 활성화를 위해 추진되는 다양한 교육 프로그램, 경진대회, 사례 공모, 설명회, 기업 지원 등의 활동은 데이터 이용자와 개발자들이 개인정보 보호를 준수하면서

안전하게 데이터를 활용할 수 있도록 실무적인 역량을 강화하고 참여의 장을 마련하고 있다. 향후 개인정보 안심구역의 지속적인 활성화를 위해서는 분야별 가이드라인 및 표준 절차 수립, 홍보 활동, 다양한 데이터 활용 아이디어의 발굴과 사업화 지원 등이 필요하다. 특히 PET 기술의 활용과 실증은 개인정보 보호와 데이터 활용의 공존을 가능하게 하는 핵심 방안으로서 주목할 필요가 있으며, 이를 통해 개인정보 안심구역이 사회 전반에 안전한 데이터 활용을 촉진하는 역할을 수행하도록 노력해야 할 것이다.

주석

- [1] 보안뉴스, "AI 활용 개인정보 유출사고 빈번, 보안대책 마련 시급", 2024.9.28.
(URL: <https://m.boannews.com/html/detail.html?idx=130283>)
- [2] AIT타임스, "전 세계 연간 생성되는 디지털 데이터, '25년까지 '16년 대비 10배 규모 163조 기가바이트로 증가", 2017.11.15.
(URL: <https://www.aitimes.kr/news/articleView.html?idxno=10993>)
- [3] 메디게이트뉴스, "AI 신약개발 강국 핵심은 양질의 데이터...바이오 데이터 활용 촉진법 도입," 2023.9.26.
(URL: <https://www.medigatenews.com/news/3339728471>)
- [4] UK Research and Innovation (UKRI), "GDPR: Identifiability, Anonymisation, and Pseudonymisation," UKRI,
(URL: <https://www.ukri.org/wp-content/uploads/2021/11/MRC-291121-GDPR-Identifiability-Anonymisation-Pseudonymisation.pdf>).
- [5] 아이뉴스24, "개인정보 활용 확대방안 발표...데이터 신경제 선도 목표," 2021.07.21.
(URL: <https://www.inews24.com/view/1378392>).
- [6] OECD, "Emerging privacy-enhancing technologies: Current regulatory and policy approaches," OECD Digital Economy Papers, No. 351, OECD Publishing, Paris, 2023, <https://doi.org/10.1787/bf121be4-en>.
- [7] 개인정보보호위원회, "가명정보 활용 확산으로 데이터 신경제 선도 추진 - 관계부처 합동, 「가명정보 활용 확대방안」 발표," 2023.07.21.
(URL: <https://www.korea.kr/docViewer/skin/doc.html?fn=d985835e84f1637c2caef2196f7b43d5&rs=/docViewer/result/2023.07/21/d985835e84f1637c2caef2196f7b43d5>).
- [8] 개인정보보호위원회, "개인정보위, '개인정보 안심구역' 신규 기관 공모," 2024.04.23.
(URL: <https://www.pipc.go.kr/np/cop/bbs/selectBoardArticle.do?bbsId=B5074&mCode=C020010000&nttId=10096>).
- [9] 특정 환경에 설치된 폐쇄성 네트워크 공간. 연구자는 플랫폼에서 지정한 분석환경의 컴퓨터를 이용하여 데이터 분석 가능함
- [10] 사용자가 서비스에 로그인할 때 신원 확인을 위해 다양한 인증 요소를 사용하는 보안 기술

제로트러스트 기반의 가명정보 처리 환경을 적용한 데이터 유용성 제고 방안



박영수

윤법률사무소 변호사

1. 들어가며

가명정보는 추가정보 없이는 특정 개인을 알아볼 수 없는 정보로, 가명처리 수준은 데이터의 식별 위험성과 처리환경의 식별 위험성을 종합하여 결정된다. 즉, 가명정보는 절대적 개념이 아닌 상대적 개념으로 가명정보의 처리 수준은 처리환경 수준에 따라 달라질 수 있기 때문에 가명정보의 처리가 다른 정보의 접근 및 이용이 제한되는 폐쇄환경에서 이루어지는 경우에는 폐쇄환경이 아닌 장소에서 처리되는 가명정보에 비해 상대적으로 데이터 유용성이 높은 가명처리를 수행하는 것이 가능하다.

이에, 가명정보 적정성을 평가하는 경우 가명정보 처리를 하는 자가 이용 목적 달성을 위해 데이터 유용성이 높은 가명정보를 요청하는 경우에는 인터넷이 차단된 폐쇄환경에서 가명정보를 처리하는 것을 조건으로 가명처리 수준을 완화하여 왔다. 하지만, AI 개발, 공동 연구 등을 위해서는 클라우드 환경에서의 가명정보 처리가 수반되어야 하는 경우가 많아 폐쇄환경 내에서 가명정보를 처리하도록 하는 것은 현재 데이터 산업의 현실을 반영하지 못한다는 비판을 받아왔다. 이에, 클라우드 환경에서 가명정보를 처리하는 경우에도 데이터 유용성이 높은 가명처리가 이루어질 수 있도록 하기 위해 제로트러스트 기반의 가명정보 처리에 대한 논의가 시작되게 되었는데, 아래서는 제로트러스트의 개념과 기본원리 및 핵심 요소를 토대로 가명정보 처리환경에 어떻게 적용할 수 있는 지에 대해 검토해보고자 한다.

2. 제로트러스트 개념과 기본원리

제로트러스트란 “절대 신뢰하지 말고, 항상 검증하라(Never Trust, Always Verify)”라는 문구로 대표되는 보안 모델이다. 이러한 제로트러스트 보안 모델은 코로나19 이후 재택·원격 근무가 확산되면서 기업의 업무 환경이 내부 시스템 기반의 데이터 저장·이용 방식에서 클라우드 기반의 외부 시스템 방식으로 전환됨에 따라 주목받게 되었다. 기존에는 데이터를 기업 내부 시스템에 저장하고 이용하였기 때문에 네트워크 위치를 기반으로 높은 수준의 신뢰성을 부여하는 경계성 보안 모델을 주로 적용하였으나, 클라우드 환경에서는 기업 데이터를 외부에 저장하고 원격으로 접근하는 방식이 일반화되면서 신뢰할 수 있는 주체를 식별하기가

어려워졌다. 이에 따라 기존 경계성 보안 모델의 한계를 극복하고자 ‘누구도 신뢰하지 않는다’는 원칙을 기반으로 하는 제로트러스트 보안 모델의 도입이 필요하게 된 것이다.

미국 바이든 정부는 2021년 5월 “국가 사이버 보안 개선을 위한 행정명령(Executive Order 14028- Improving the Nation’s Cybersecurity)”에서 “연방 정부의 사이버보안 현대화”를 위해 제로트러스트 아키텍처*를 도입하도록 하였고, 일본, 영국 등 여러 나라, 정부, 글로벌 기업에서도 제로트러스트 기반의 보안 체계를 도입하려는 움직임을 보이는 등 제로트러스트 아키텍처는 사이버보안의 새로운 패러다임으로 급부상하고 있다. 이에 발맞추어 우리나라에서도 제로트러스트 도입 계획을 발표하며, 2023년 과학기술정보통신부는 제로트러스트 가이드라인 1.0을 발간한 바 있다. 제로트러스트 가이드라인 1.0에서는 아래 [표 1], [표 2]와 같이 제로트러스트 6가지 기본 원리와 핵심 요소를 정의하고 있는 바, 아래서는 위 내용을 기반으로 데이터 유용성을 높일 수 있는 가명처리 환경 구축 방안을 제시해 보고자 한다.

* 제로트러스트 아키텍처란 제로트러스트 개념을 활용하여 기업 내부의 네트워크, 시스템 및 리소스(데이터, 응용, 서버, 하드웨어, 네트워크, 클라우드 서비스 등)(이하 “리소스”)를 보호할 수 있는 보안 구조를 말함(제로트러스트 가이드라인 1.0. 2023. 6. 28p).

표 1 제로트러스트 기본 원리

※ 제로트러스트 가이드라인 1.0, 2023. 6., 32p

| 구분 | 주요 내용 |
|----|---|
| 가 | 모든 종류의 접근에 대해 신뢰하지 않을 것(명시적인 신뢰 확인 후 리소스 접근 허용) |
| 나 | 일관되고 중앙 집중적인 정책 관리 및 접근제어 결정, 실행 필요 |
| 다 | 사용자, 기기에 대한 관리 및 강력한 인증 |
| 라 | 리소스 분류 및 관리를 통한 세밀한 접근제어(최소 권한 부여) |
| 마 | 논리 경계 생성 및 세션 단위 접근 허용, 통신 보호 기술 적용 |
| 바 | 모든 상태에 대한 모니터링, 로그 및 이를 통한 신뢰성 지속성 검증, 제어 |

표 2 제로트러스트 구현에 따르는 핵심 요소 및 목표

※ 제로트러스트 가이드라인 1.0, 2023. 6., 90p~102p

| 구분 | 핵심요소 | 목표 |
|----|------------|---|
| 1 | 식별자·신원 | 기업망에 접근하는 사용자는 기업에서 관리하는 식별정보를 이용하여 업무에 활용하는 응용에 접근한다. 피싱 등 다양한 공격에 강한 다중 인증(MFA, Multi-Factor Authentication) 기법을 도입함으로써, 더욱 정교한 형태의 온라인 공격으로부터 사용자를 보호할 수 있어야 한다. |
| 2 | 기기 및 엔드포인트 | 기업은 업무용으로 인가되어 동작하는 모든 기기 목록을 관리하고, 해당 장치에서 발생하는 사고를 예방, 감지 및 대응할 수 있어야 한다. |
| 3 | 네트워크 | 기업망 내에서 모든 DNS, HTTP 트래픽을 암호화하며, 보안 경계를 격리된 환경으로(소프트웨어적으로) 분할하는 계획을 실행해야 한다. |
| 4 | 시스템 | 기업은 보유하고 있는 서버(Server) 시스템에 대해서 권한 사용자(예: 시스템 관리자 등) 레벨로 접속하는 경우 서버 시스템의 주요 파일 접근제어, 주요 사용 명령어 통제, 영역별 접근제어 등이 통제되고 관리됨으로써 외부 해커의 주요 서버 시스템 공격에 대해 대비하여야 한다. |
| 5 | 응용 및 워크로드 | 기업은 모든 응용이 인터넷에 연결되어 있다고 간주하고, 정기적으로 응용 프로그램을 엄격하게 시험하며, 외부 취약성 보고서를 참고해야 한다. |
| 6 | 데이터 | 기업은 완전한 데이터 분류를 사용하는 보호 기능을 배포하기 위해 명확하고 공유된 경로를 따른다. 기관은 클라우드 보안 서비스 및 도구를 활용하여 민감한 데이터를 검색, 분류 및 보호하고 전사적 로깅 및 정보 공유를 구현해야 한다. |

3. 데이터 유용성 향상을 위한 제로트러스트 기반의 가명정보 처리 환경 구축 방안

제로트러스트 기반의 가명정보 처리 환경을 논의하는 이유는 폐쇄환경이 아닌 곳에서도 데이터 유용성이 높은 데이터를 처리할 수 있도록 하는 데 그 목적이 있다. 따라서, 제로트러스트 기반의 가명정보 처리 환경을 정의하기에 앞서 폐쇄환경에서 처리하는 가명정보에 대하여 가명처리 수준을 낮게 설정할 수 있는 요인이 무엇인지 우선적으로 살펴보고자 한다.

가명정보는 추가 정보 없이는 특정 개인을 알아볼 수 없는 정보로 가명처리는 식별정보, 식별가능정보, 특이정보[표 3]를 대상으로 한다. 이중 식별정보는 그 자체로 개인을 식별할 수 있는 정보에 해당하므로 1차적 가명처리 대상이 되며, 식별가능정보와 특이정보는 그 자체로는 개인의 식별이 어렵지만 다른 정보와 결합하는 경우 식별가능성을 가지기 때문에 처리환경을 고려하여 처리 여부와 방법을 결정하게 된다. 즉, 처리환경에서는 가명정보를 처리하는 자가 가명정보 외의 다른 정보를 이용하여 가명정보를 재식별할 위험이 있는지에 대해 검토하게 되는데 폐쇄환경에서 가명정보를 처리하는 경우 가명정보를 처리하는 자의 행위 및 다른 정보의 이용을 통제할 수 있어 다른 정보와 결합 가능성과 그로 인한 식별가능성이 없게 되므로 식별가능정보와 특이 정보에 대한 처리 수준을 낮게 설정하는 것이 가능해진다. 따라서, 제로트러스트 기반의 가명정보 처리 환경을 정의하는 경우에도 위와 같은 폐쇄환경의 특성을 반영하는 것이 요구된다.

표 3 식별정보, 식별가능정보, 특이정보

※ 가명정보 처리 가이드라인, 2024. 2., 16p

| 구분 | 주요 내용 |
|--------|---|
| 식별정보 | <ul style="list-style-type: none"> · 특정 개인과 직접적으로 연결되는 정보 · (예시) 성명, 고유 식별정보, (개인)휴대전화번호, (개인)전자우편주소, 의료기록번호, 건강보험번호 등 |
| 식별가능정보 | <ul style="list-style-type: none"> · 다른 항목과 결합하는 경우 식별가능성이 높아지는 항목 · (예시) 성별, 연령, 거주 지역, 국적, 직업, 위치정보 등 해당정보를 처리하는 자를 기준으로 판단 |
| 특이 정보 | <ul style="list-style-type: none"> · 전체 데이터에 식별가능성을 가지는 고유(희소)값, 편중된 분포를 가지는 단일·다중이용항목 · (예시) 희귀성씨 등 특이한 값, 국내 최고령 등 극단값, 특정 데이터 분석집단에서 희소한 값 등 |

앞서 살펴본 제로트러스트 기본 원리와 폐쇄환경의 특성을 고려할 때 제로트러스트 기반 가명정보 처리 환경은 아래 5가지 사항을 구현하는 것이 필요해 보인다.

가. 모든 종류의 접근을 신뢰하지 않고, 최소한의 접근 권한을 부여할 것

가명정보를 처리하는 시스템 등에 대해서는 원칙적으로 접근을 제한하여야 하며, 인증을 거친 처리자 및 단말기에 대해서만 접근을 허용하도록 하여야 한다. 또한, 가명처리, 분석, 관리 등 각 업무별로 접근·이용이 필요한 정보 등이 다르므로 이를 고려하여 필요 최소한의 인력에게 접근 권한을 부여하는 것이 필요하다.

나. 처리자 및 단말기에 대한 관리 및 강력한 인증

가명정보를 처리하는 시스템에 접근할 수 있는 단말기를 지정하여 목록을 관리하여야 한다. 또한, 비인가자 접근을 방지하기 위해 등록된 처리자에 대하여 멀티팩터 인증, 생체인증 등 접근 권한자를 확인할 수 있는 강력한 인증수단을 적용하여야 한다.

다. 모니터링 및 점검

가명정보 처리에 대한 로그를 기록·관리·점검하여야 한다. 또한, 재식별 행위, 유출행위를 방지할 수 있도록 PC 화면모니터링시스템을 수행하여 이상행위를 모니터링하여 신뢰성을 지속적으로 점검·관리하여야 한다. 또한, 이상행위 발생 시 접근 세션을 종료시키는 등 조치를 통해 이상행위 발견에 대한 조치 방안을 마련하여야 한다.

라. 다른 정보의 접근·이용 차단(클라우드 이용 시 외부 인터넷 차단)

제로트러스트 기반 가명정보 처리 환경에서 다른 정보와의 결합을 통한 식별가능성을 방지하기 위해 폐쇄환경과 같이 다른 정보의 접근·이용을 제한하는 방안을 마련하는 것이 필요하다. 만약, 클라우드 환경을 통한 가명정보 처리 환경을 구축하는 경우라면 클라우드 접속 시 클라우드 이용을 위한 인터넷 접속 외에 외부 인터넷 접속을 차단하여 다른 정보의 접근·이용을 통제하는 등 조치를 수행하여야 한다.

마. 가명정보 유출, 오·남용 방지(가명정보 송·수신 및 저장시 암호화 및 키관리)

처리 환경을 고려하여 가명처리 수준을 낮춘 가명정보는 외부로 유출되는 경우 재식별 위험에 노출될 수 있다. 따라서, 데이터 유·노출 시 비인가자의 이용을 방지할 수 있도록 송·수신 및 저장 시 암호화하여야 한다. 만약, 가명정보를 클라우드에 저장하여 처리하는 경우에는 암호키를 클라우드 외부에 보관하는 것이 필요하다.

4. 제로트러스트 기반의 가명정보 처리 환경을 활용한 가명정보 적정성 평가

가명처리 수준은 데이터 식별 위험성과 처리 환경 위험성을 종합하여 결정하기 때문에 데이터 유용성을 높이기 위해서는 처리 환경의 위험성을 낮추어야 한다. 이에, 가명정보 적정성 평가를 요청한 자가 가명정보 처리 목적 달성을 위해 유용성이 높은 데이터를 요청하는 경우 지금까지는 인터넷이 차단된 폐쇄환경에서 가명정보 처리하는 것을 조건으로 가명처리 수준을 완화해 왔다. 따라서, 지금까지 클라우드 환경에서 가명정보를 처리하는 경우에는 가명처리 수준을 완화하는 것에 제약이 있어 왔는 바, 클라우드 환경에서는 제로트러스트 기반의 가명정보 처리환경을 검토하여 가명처리 수준 완화의 조건으로 적용해 볼 수 있을 것으로 생각된다.

5. 결어

2020년 8월 데이터 3법 개정 시에는 가명정보 제도 시행 초기임을 감안하여 "안전성"에 무게를 둔 가명정보 제도를 도입하였다. 제도 시행 이후 4년의 시간 동안 가명정보 활용 활성화를 위한 노력은 계속되고 있지만, 가명정보 제도를 유연화하는 것에 대한 시도는 이루어지지 못했다. 가명정보 이용 활성화를 위해서는 무엇보다 가명정보 이용 편의성 및 접근성을 높이는 것이 필요하다. 아직까지 데이터 유용성을 유지하면서도 가명정보 처리환경의 유연성을 높이기 위한 제로트러스트 기반의 가명정보 처리는 본격적으로 논의되지 않은 상황이지만, 제로트러스트 기반의 가명정보 처리 논의는 이러한 가명정보 이용 편의성 및 접근성을 높이는 첫걸음이 될

수 있어 보인다. 본 보고서를 시작으로 가명정보 제도를 유연화하는 논의가 구체적으로 시작되어 “활용성”까지도 갖춘 가명정보 2.0 제도가 마련되길 기대해 본다.

참고 문헌

과학기술정보통신부, 한국인터넷진흥원, 제로트러스트포럼, 제로트러스트 가이드라인 1.0, 2023.06

개인정보보호위원회, 가명정보 처리 가이드라인, 2024.02.

개인정보보호위원회, 개인정보 보호 법령 및 지침, 고시 해설서, 2020. 12.

John Kindervag (Forrester), “Build Security Into Your Network's Dna: The Zero Trust Network Architecture”, 2010.11

John Kindervag (Forrester), “No More Chewy Centers: Introducing The Zero Trust Model Of Information Security”, 2010.09

NIST Sp 800-207, “Zero Trust Architecture”, 2020.08

Privacy Report

2024 개인정보 이슈

심층 분석 보고서

『Privacy Report 2024 개인정보 이슈 심층 분석 보고서』는
디지털·정보보호 관련 글로벌 트렌드
및 주요 이슈를 분석하여
정책 자료로 활용하기 위해 한국인터넷진흥원에서
기획, 발간하는 심층 보고서입니다.

한국인터넷진흥원의 승인 없이 본 보고서의 무단전재나
복제를 금하며, 인용 출처 『Privacy Report 2024
개인정보 이슈 심층 분석 보고서』를 밝혀주시기 바랍니다.

본 보고서의 내용은
한국인터넷진흥원의 공식 견해가 아님을 알려드립니다.

발행

발행일 2024년 11월
발행처 한국인터넷진흥원 개인정보제도팀
전라남도 나주시 진흥길 9
Tel : 061-820-1231

PRIVACY REPORT

2024 개인정보 이슈 심층 분석 보고서

2024 Vol. 5

PRiVACY
REPOR[T

