발간등록번호
콜로키움 23-01

행정데이터 활용을 위한

가명정보처리의 이해

청년종합연구 II :
정책소외계층 청년 실태 및 정책개발

발표자 : 김윤중 수석(유피에스데이터)

일　시 : 2023년 4월 19일(수) 13:30~

장　소 : 한국청소년정책연구원 7층 중회의실
(온라인 병행)

※ 본 콜로키움은 한국인터넷진흥원에서 지원하는 '가명정보 활용 컨설팅' 과정의 일환임.

한국청소년정책연구원　개인정보보호위원회　KISA 한국인터넷진흥원　UPS DATA

한국청소년정책연구원

가명처리 컨설팅
사전교육자료

유피에스데이터㈜
김윤중 수석
010-4497-9915
yjkim@upsdata.info

개인정보보호위원회 KISA 한국인터넷진흥원 UPS DATA

가명처리 컨설팅 사전 교육

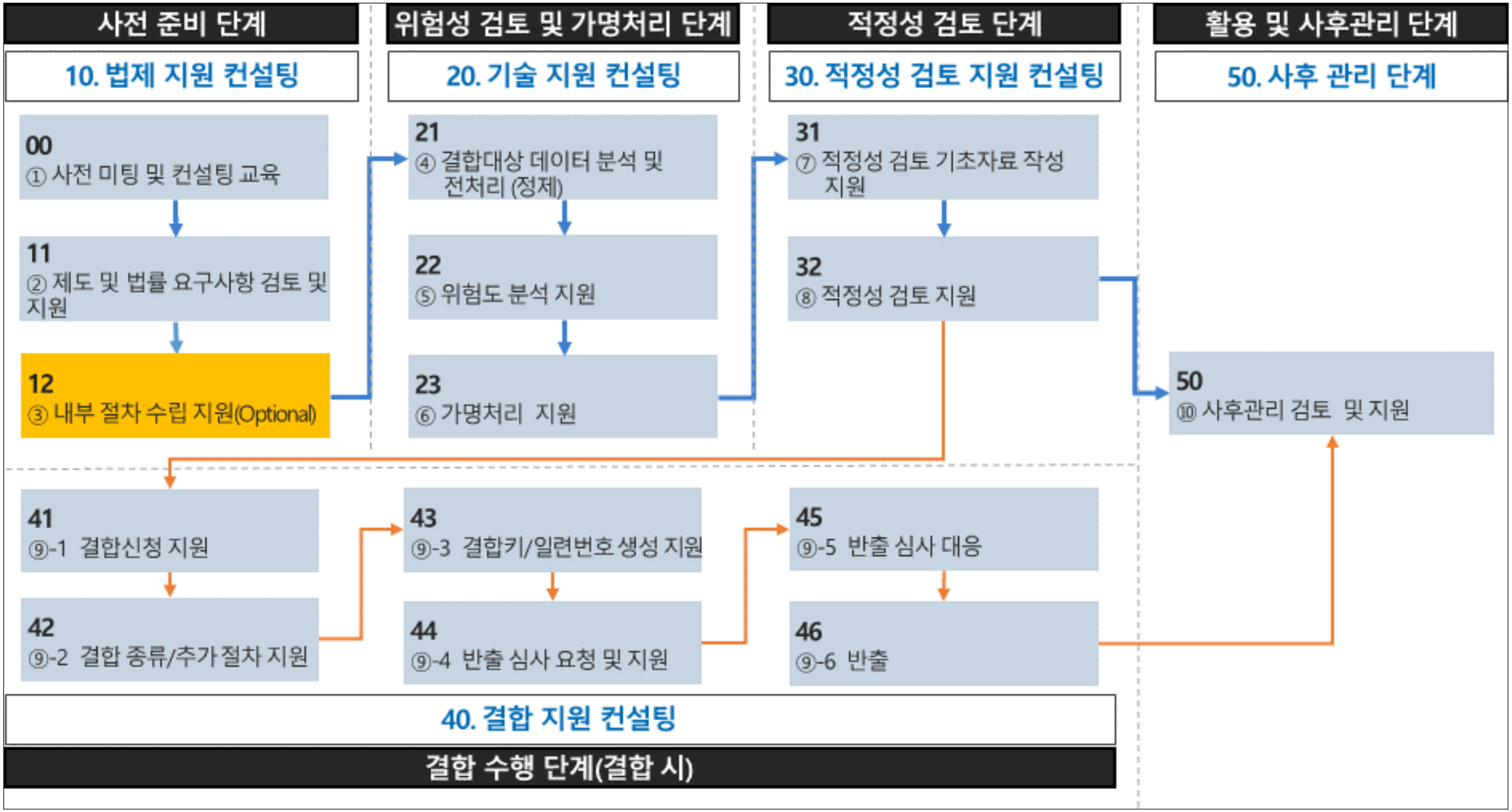
CONTENTS

컨설팅의 구성

1. 컨설팅 구성

1

컨설팅 종류 및 세부 수행 단계



단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

※ 각 컨설팅의 단계별 모듈은 요청기관의 컨설팅 요청 및 필요에 따라 상이하게 지원될 수 있음

1. 사전 미팅 및 컨설팅 신청		
결합 컨설팅 신청서 접수	컨설팅 기본 사항 교육 (인식 제고 교육 포함)	컨설팅 요구사항 협의
2. 제도 및 법률 요구사항 검토 및 지원		
관련 법률 1차 검토	관련 법제에 의한 처리 절차 안내	가명정보 내부 관리계획 작성 지원
적합성 검토	개인정보 처리 방침 수립	
3. 내부 처리 절차 수립 지원 (Optional)		
가명정보 운영 규정 수립	적정성 검토 위원회 운영 규정 수립	적정성 검토 위원회 운영 매뉴얼 제공
가명정보 처리 매뉴얼 수립 및 제공		
4. 결합 대상 데이터 분석 및 전처리		
데이터 구조 분석	개인정보 속성 분석	데이터 타입 분석
목적 컬럼 설정	데이터 분포 분석	기타 데이터 분석
5. 위험도 분석 지원		
데이터 활용 형태 위험성 검토	데이터 이용환경 위험성 검토	데이터 위험성 검토
위험성 검토 보고서 작성 지원		
6. 가명처리 지원		
가명처리 수준 정의	목적 달성 가능성 협의	가명 처리

단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

※ 각 컨설팅의 단계별 모듈은 요청기관의 컨설팅 요청 및 필요에 따라 상이하게 지원될 수 있음

7. 적정성 검토 기초자료 작성 지원	적정성 검토 대응 방안 수립	적정성 검토 관련 문서 검토	적정성 검토 기초자료 작성 지원
8. 적정성 검토 지원	적정성 검토 위원회 구성	적정성 검토 대응 지원	
9. 가명정보 결합	결합 신청 지원	추가 서비스 필요 여부 검토 (모의결합, 추출)	결합키 및 일련번호 생성
	결합키 관리 기관 전송	결합 대상정보 결합 전문기관 전송 지원	결합률 확인
	결합	모의 결합(필요시)	추출(필요시)
	추가 가명처리 필요 여부 확인	추가 가명처리 수행 지원	반출 심사 자료 작성 지원
	반출 신청서 작성 지원	안전성 확보조치 현황 확인	반출 심사 위원회 대응
10. 사후관리	사후관리 관련 교육 시행	재식별 모니터링 절차 및 방법 교육	컨설팅 종료 보고서 작성

가명 처리 관련 법령의 이해

1. **개인정보보호법의 가명처리**
2. 가명정보처리 가이드라인의 가명처리 절차

2

가명정보 처리 관련 법제


가명처리 관련 법제
금융분야
보건의료 분야
통계작성기관
교육기관
공공기관
그 외 분야
법률
특별법
일반법
신용정보의 이용 및 보호에 관한 법률
생명윤리 및 안전에 관한 법률
통계법
공공데이터법 등
그 외 분야
개인정보보호법
시행령
신용정보의이용및보호에관한법률시행령
개인정보보호법 시행령
고시
결합
안전성 확보조치
신용정보업 감독규정
가명정보의 결합 및 반출 등에 관한 고시
개인정보의 안전성 확보조치 기준
가이드라인
일반
분야별
금융분야 가명 익명처리 안내서
가명정보처리 가이드라인
보건의료데이터 활용 가이드라인
교육분야 가명, 익명 정보처리 가이드라인
공공분야 가명정보 제공 실무 안내서

가명처리에 대한 기본 조항

개인정보보호법 제3조

⑦ 개인정보처리자는 개인정보를 익명 또는 가명으로 처리하여도 개인정보 수집목적을 달성할 수 있는 경우 익명처리가 가능한 경우에는 익명에 의하여, 익명처리로 목적을 달성할 수 없는 경우에는 가명에 의하여 처리될 수 있도록 하여야 한다.

개인정보보호법 제15조 제3항

③ 개인정보처리자는 당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령으로 정하는 바에 따라 정보주체의 동의 없이 개인정보를 이용할 수 있다

개인정보보호법 제17조 제4항

④ 개인정보처리자는 당초 수집 목적과 합리적으로 관련된 범위에서 정보주체에게 불이익이 발생하는지 여부, 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부 등을 고려하여 대통령령으로 정하는 바에 따라 정보주체의 동의 없이 개인정보를 제공할 수 있다.

개인정보보호법 시행령 제14조 제2항

① 개인정보처리자는 법 제15조제3항 또는 제17조제4항에 따라 정보주체의 동의 없이 개인정보를 이용 또는 제공(이하 “개인정보의 추가적인 이용 또는 제공”이라 한다)하려는 경우에는 다음 각 호의 사항을 고려해야 한다.

1. 당초 수집 목적과 관련성이 있는지 여부
2. 개인정보를 수집한 정황 또는 처리 관행에 비추어 볼 때 개인정보의 추가적인 이용 또는 제공에 대한 예측 가능성이 있는지 여부
3. 정보주체의 이익을 부당하게 침해하는지 여부
4. 가명처리 또는 암호화 등 안전성 확보에 필요한 조치를 하였는지 여부

가명정보의 처리 등

개인정보보호법 제28조의2

① 개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다.

② 개인정보처리자는 제 1항에 따라 가명정보를 제3자에게 제공하는 경우에는 특정 개인을 알아보기 위하여 사용될 수 있는 정보를 포함해서는 아니된다.

통계작성 및
학술연구 등

제18조 2항 4호

상업적 목적의 통계 작성, 산학연구 및 민간연구 등의 **포함 여부 논란**

통계작성,
과학적 연구,
공익적 기록보존
등

제28조의2 제1항

新기술·제품 ·서비스 개발 등 **산업적 목적**을 포함하는 과학적 연구, 시장조사와 **상업 목적**의 통계작성도 **포함**

가명정보의 결합제한

개인정보보호법 제28조의3

① 제28조의2에도 불구하고 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 **서로 다른 개인정보처리자 간의 가명정보의 결합은 보호위원회 또는 관계 중앙행정기관의 장이 지정하는 전문기관이 수행**한다.

② 결합을 수행한 기관 외부로 결합된 정보를 반출하려는 개인정보처리자는 **가명정보 또는 제58조의2에 해당하는 정보로 처리한 뒤 전문기관의 장의 승인**을 받아야 한다.

③ 제1항에 따른 결합 절차와 방법, 전문기관의 지정과 지정 취소 기준 · 절차, 관리 · 감독, 제2항에 따른 반출 및 승인 기준 · 절차 등 필요한 사항은 대통령령으로 정한다.

기업 내부 데이터는 자체적으로 결합이 가능하며,
서로 다른 기업 간 데이터는 보안시설을 갖춘 전문기관 내에서 수행(1항)

결합된 데이터를 기관 외부로 반출할 경우 가명 또는 익명조치 후
전문기관의 승인을 거쳐 반출하도록 해 안전성을 높임(2항)

가명정보에 대한 안전조치 의무 등

개인정보보호법 제28조의4

① 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관 · 관리하는 등 해당 정보가 분실 · 도난 · 유출 · 위조 · 변조 또는 훼손되지 않도록 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적 · 관리적 및 물리적 조치를 하여야 한다. (가명정보의 안전성 확보조치)

② 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 경우 처리목적 등을 고려하여 가명정보의 처리 기간을 별도로 정할 수 있다.

③ 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자, 가명정보의 처리 기간(제2항에 따라 처리 기간을 별도로 정한 경우에 한한다) 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 관련 기록을 작성하여 보관하여야 하며, 가명정보를 파기한 경우에는 파기한 날부터 3년 이상 보관하여야 한다. (가명정보의 기록 및 파기 관리)

가명정보에 대한 안전조치 의무 등

개인정보보호법 시행령 제29조의5

① 개인정보처리자는 법 제28조의4제1항에 따라 가명정보 및 가명정보를 원래의 상태로 복원하기 위한 추가 정보(이하 이 조에서 "추가정보"라 한다)에 대하여 다음 각 호의 안전성 확보 조치를 해야 한다.

1. 제30조 또는 제48조의2에 따른 안전성 확보 조치
2. 가명정보와 추가정보의 분리 보관. 다만, 추가정보가 불필요한 경우에는 추가정보를 파기해야 한다.
3. 가명정보와 추가정보에 대한 접근 권한의 분리. 다만, 「소상공인기본법」 제2조에 따른 소상공인으로서 가명정보를 취급할 자를 추가로 둘 여력이 없는 경우 등 접근 권한의 분리가 어려운 정당한 사유가 있는 경우에는 업무 수행에 필요한 최소한의 접근 권한만 부여하고 접근 권한의 보유 현황을 기록으로 보관하는 등 접근 권한을 관리 · 통제해야 한다.

② 법 제28조의4제2항에서 "대통령령으로 정하는 사항"이란 다음 각 호의 사항을 말한다.

1. 가명정보 처리의 목적
2. 가명처리한 개인정보의 항목
3. 가명정보의 이용내역
4. 제3자 제공 시 제공받는 자
5. 그 밖에 가명정보의 처리 내용을 관리하기 위하여 보호위원회가 필요하다고 인정하여 고시하는 사항

개인정보의 안전성 확보 조치

개인정보보호법 시행령 제30조

① 개인정보처리자는 법 제29조에 따라 다음 각 호의 안전성 확보 조치를 하여야 한다.

1. 개인정보의 안전한 처리를 위한 내부 관리계획의 수립 · 시행
   → 이에 따라 가명정보의 내부 관리계획의 수립이 필요
2. 개인정보에 대한 접근 통제 및 접근 권한의 제한 조치
3. 개인정보를 안전하게 저장 · 전송할 수 있는 암호화 기술의 적용 또는 이에 상응하는 조치
4. 개인정보 침해사고 발생에 대응하기 위한 접속기록의 보관 및 위조 · 변조 방지를 위한 조치
5. 개인정보에 대한 보안프로그램의 설치 및 갱신
6. 개인정보의 안전한 보관을 위한 보관시설의 마련 또는 잠금장치의 설치 등 물리적 조치

② 보호위원회는 개인정보처리자가 제1항에 따른 안전성 확보 조치를 하도록 시스템을 구축하는 등 필요한 지원을 할 수 있다.

③ 제1항에 따른 안전성 확보 조치에 관한 세부 기준은 보호위원회가 정하여 고시한다.

→ 개인정보의 안전성 확보조치(2021-02) 고시에 안전성 화보조치에 대한 세부 기준이 명시되어 있음

→ 정보통신사업자는 개인정보의 기술적·관리적 보호조치 기준(2021-03)을 따라야 함

가명정보 처리 시 금지의무, 과징금 부과

개인정보보호법 제28조의5

① 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 자는 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 된다.

② 개인정보처리자는 제28조의2 또는 제28조의 3에 따라 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수 · 파기하여야 한다.

개인정보보호법 제64조의2 (과징금의 부과)

① 보호위원회는 다음 각 호의 어느 하나에 해당하는 경우에는 해당 개인정보처리자에게 전체 매출액의 100분의 3을 초과하지 아니하는 범위에서 과징금을 부과할 수 있다. 다만, 매출액이 없거나 매출액의 산정이 곤란한 경우로서 대통령령으로 정하는 경우에는 20억원을 초과하지 아니하는 범위에서 과징금을 부과할 수 있다.

6. 제28조의5제1항(제26조제8항에 따라 준용되는 경우를 포함한다)을 위반하여 특정 개인을 알아보기 위한 목적으로 정보를 처리한 경우

② 보호위원회는 제1항에 따른 과징금을 부과하려는 경우 전체 매출액에서 위반행위와 관련이 없는 매출액을 제외한 매출액을 기준으로 과징금을 산정한다.

③ 보호위원회는 제1항에 따른 과징금을 부과하려는 경우 개인정보처리자가 정당한 사유 없이 매출액 산정자료의 제출을 거부하거나 거짓의 자료를 제출한 경우에는 해당 개인정보처리자의 전체 매출액을 기준으로 산정하되 해당 개인정보처리자 및 비슷한 규모의 개인정보처리자의 개인정보 보유 규모, 재무제표 등 회계자료, 상품 · 용역의 가격 등 영업현황 자료에 근거하여 매출액을 추정할 수 있다.

가명정보 처리 시 금지의무, 과징금 부과

개인정보보호법 제71조(벌칙)

다음 각 호의 어느 하나에 해당하는 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처한다.

6. 제28조의3제1항(제26조제8항에 따라 준용되는 경우를 포함한다)을 위반하여 보호위원회 또는 관계 중앙행정기관의 장으로부터 전문기관으로 지정받지 아니하고 가명정보를 결합한 자

7. 제28조의3제2항(제26조제8항에 따라 준용되는 경우를 포함한다)을 위반하여 전문기관의 장의 승인을 받지 아니하고 결합을 수행한 기관 외부로 결합된 정보를 반출하거나 이를 제3자에게 제공한 자 및 그 사정을 알면서도 영리 또는 부정한 목적으로 결합된 정보를 제공받은 자

8. 제28조의5제1항(제26조제8항에 따라 준용되는 경우를 포함한다)을 위반하여 특정 개인을 알아보기 위한 목적으로 가명정보를 처리한 자

개인정보보호법 제75조(과태료)

① 다음 각 호의 어느 하나에 해당하는 자에게는 5천만원 이하의 과태료를 부과한다.

7의2. 제28조의5제2항을 위반하여 개인을 알아볼 수 있는 정보가 생성되었음에도 이용을 중지하지 아니하거나 이를 회수·파기하지 아니한 자

가명정보의 적용 예외

개인정보보호법 제28조의7

제28조의7(적용범위) 제28조의2 또는 제28조의3에 따라 처리된 가명정보는 제20조, 제20조의2, 제27조, 제34조제1항, 제35조, 제35조의2, 제36조 및 제37조를 적용하지 아니한다.

가명 처리 관련 법령의 이해

1. 개인정보보호법의 가명처리
2. **가명정보처리 가이드라인의 가명처리 절차**

2

가명정보 처리 가이드라인

- 목적
  - ✓ 빅데이터, AI 등 다양한 융·복합 산업에서의 데이터 이용 수요가 급증하는 가운데, 데이터 활용의 핵심인 가명정보 활용을 위한 법적 근거가 마련됨
  - ✓ 가명정보 활용에 필요한 가명정보 처리 목적, 처리 절차 및 방법, 안전조치에 관한 사항 등을 안내하여 안전한 데이터 활용 환경을 마련하고자 함

- 근거 및 범위
  - ✓ 개인정보보호법 제3장 제3절 '가명정보 처리에 관한 특례 ' 에 근거하여 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 가명정보의 처리이며 개인정보보호위원회와 소관 부처가 공동으로 발간한 개인정보의 가명정보 처리에 관한 분야별 가이드라인이 있는 경우에는 해당 분야의 가이드라인을 우선 적용함

- 주요 부처별 가이드라인
  - ✓ 개인정보보호위원회 : 가명정보 처리 가이드라인
  - ✓ 교육부 : 교육분야 가명·익명정보 처리 가이드라인
  - ✓ 보건복지부 : 보건의료 데이터 활용 가이드라인
  - ✓ 행정안전부 : 공공분야 가명정보 제공 실무안내서
  - ✓ 금융위원회 : 금융분야 가명·익명처리 안내서

가명정보 처리 가이드라인

가명처리 절차

- 개인정보의 가명처리는 ① 목적 설정 등 사전 준비, ② 처리 대상의 위험성 검토, ③ 가명처리 수행, ④ 적정성 검토 및 추가처리 ⑤ 안전한 관리 단계로 이루어 짐

1단계	2단계	3단계	4단계	5단계
목적 설정	위험성 검토	가명처리	적정성 검토 및 추가처리	안전한 관리
▪ 목적 설정 ▪ 적합성 검토 ▪ 필요서류 작성 ▪ 안전조치 수립 확인	▪ 데이터 식별 위험성 • 식별정보 포함 • 식별가능정보 포함 • 특이치 포함 • 재식별시 영향도 ▪ 처리환경식별위험성 • 활용 형태 • 처리 장소 • 처리 방법 →식별위험성검토결과서	▪ 가명처리 수준 정의 • 위험성 검토 결과를 바탕으로 각 컬럼별 가명처리 방법 및 수준을 정의 →가명처리수준정의표 ▪ 가명처리 • 가명처리 수준 정의표에 의해 가명처리 수행	▪ 적정성 검토 준비 • 기초자료 작성 • 기존 자료 취합 ▪ 적정성 검토 • 서류 및 절차 검토 • 위험성 검토 • 가명처리 검토 • 목적 달성가능성 검토 ▪ 추가 처리 • 적정하지 않은 경우 추가 처리 수행	▪ 재식별 금지 ▪ 재식별 모니터링 ▪ 안전조치 시행 ▪ 기록 작성 및 보관

1단계 – 목적 설정 개요

1단계 목적 설정

- 처리 목적 명확화 : 가명처리의 목적을 명확히 설정하고 가명정보 처리 목적의 적합성 검토 및 계약서, 개인정보 처리방침, 내부 관리계획 등 필요한 서류를 작성
- 적합성 검토 : 내부 승인이 존재할 경우 승인을 받기 위한 추가 설명자료 작성, 회의 개최 등을 진행할 수 있음
- 필요 서류 작성 : 가명정보의 제 3자 제공이나 가명처리 위탁 시 계약서의 작성이 필요

  ※ 특히 가명정보는 제공받는 순간 법적으로 내부 가명정보가 되어 이용기간의 제약이 없어지므로 제공하는 기관에서는 제공되는 가명정보의 위험성 등을 고려하여 계약서에 **삭제와 관련한 조항의 추가**가 필요할 수 있음

1단계 – 목적 설정 절차(1)

1단계 목적 설정

- **처리 목적 명확화** : 법률에서 허용하는 범위 내에서 가명정보 처리 목적 명확화
  - ✓ 법률에서 허용하는 범위는 통계작성, 과학적 연구, 공익적 기록보존 등에 한 함
  - ✓ (적절하지 않은 예시) 신제품 개발을 위한 과학적 연구 수행, 안전한 가명처리 방법을 적용하여 가명정보를 변환하도록 함

    ※ 목적이 구체적으로 명시되지 않아 적절하지 않음
  - ✓ (적절한 예시) OO제품의 성능 개선을 위해 개인별 OOO특성에 대한 설문조사를 토대로 개인별 특성과 성능 요인의 연관성에 대한 과학적 연구
  - ✓ 목적에 대한 증빙서류 필요(사업계획서 등)
- **적합성 검토** : 가명처리 적합성 검토(개인정보 보유부서 또는 전담부서)
  - ✓ 개인의 수집 목적 및 성격, 가명정보 활용 목적 등을 고려하여 가명처리 여부를 결정할 수 있음
  - ✓ 필요 시 심의위원회 구성 또는 외부전문가 평가 등을 통해 결정할 수 있음
  - ✓ 공공기관은 필수 절차임
  - ✓ 내부 승인 절차로 대체 가능

1단계 – 목적 설정 절차(2)

- **필요서류 작성**
  - ✓ 가명정보의 처리 또는 가명처리를 위탁(보호법 제26조에 따라 수행)하거나 가명정보를 제3자에게 제공하는 경우 필요에 따라 재식별 금지에 관한 사항, 기타 처리에 있어 유의해야 할 사항* 등을 포함한 계약서를 작성할 수 있음
    - ➢ (서류 예시) 제 3자 제공 계약서, 가명처리를 위한 위탁 계약서
    - ➢ (포함내용의 예시) 가명정보의 재제공 금지, 가명정보 재식별 금지, 가명정보의 안전성 확보조치, 가명정보의 처리기록 작성 및 보관, 가명정보의 파기, 재식별 시 손해배상 등
- **가명정보 처리를 위한 안전조치**
  - ✓ 개인정보 처리방침 수립·공개(보호법 제30조), 내부 관리계획 수립·시행(개인정보의 안전성 확보조치 기준 제4조, 개인정보의 기술적·관리적 보호조치 기준 제3조) 등 가명정보 처리에 앞서 이행하여야 할 사항 등 준비
  - ✓ 내부관리계획, 교육이수 증빙서류, 안전성확보조치를 위한 내부 문서 , 보안서약서 등
  - → 가명정보 처리에 관한 내부관리계획이 없는 경우, 계획 수립 필요
  - → 내부관리계획에 가명정보 포함 등 현행화 확인

2단계 – 위험성 검토 개요

- 가명처리 대상 데이터의 식별 위험성을 분석·평가하여 가명처리 방법 및 수준에 반영하기 위한 절차
- 식별 위험성은 1) 데이터의 식별 위험성과 2) 처리 환경의 식별 위험성으로 구분하여 검토

2단계 – 위험성 검토 절차(1)

데이터의 식별 위험성 검토

- 데이터의 위험성 검토는 가명처리 대상이 되는 정보에 식별 가능한 요소가 있는지를 파악하는 것

① 그 자체로 식별될 위험이 있는 항목 (식별정보)

→ 식별정보가 있는 경우 기본적으로 삭제하며 꼭 필요한 경우 대체기법을 적용

② 다른 항목과 결합을 통해 식별될 가능성이 있는 항목 (식별가능정보)

→ 준식별자의 경우 목적달성가능성을 함께 검토하여 처리 수준을 판단

→ 일반 속성의 경우 목적달성가능성과 필요 여부를 검토하여 삭제 또는 처리 수준을 판단

③ 그 밖에 특이정보 (특이정보 유무)

→ 특이치가 활용 목적에 필수적인지 검토, 필수적이라면 필요한 처리 수준을 판단, 필요하지 않으면 제거

④ 데이터 특성만으로 재식별 시 사회적 파장 등 영향도가 높은 항목 등이 있는지 검토 (재식별시 영향도)

→ 영향도가 높은 경우 식별 위험성 검토를 보수적으로 진행

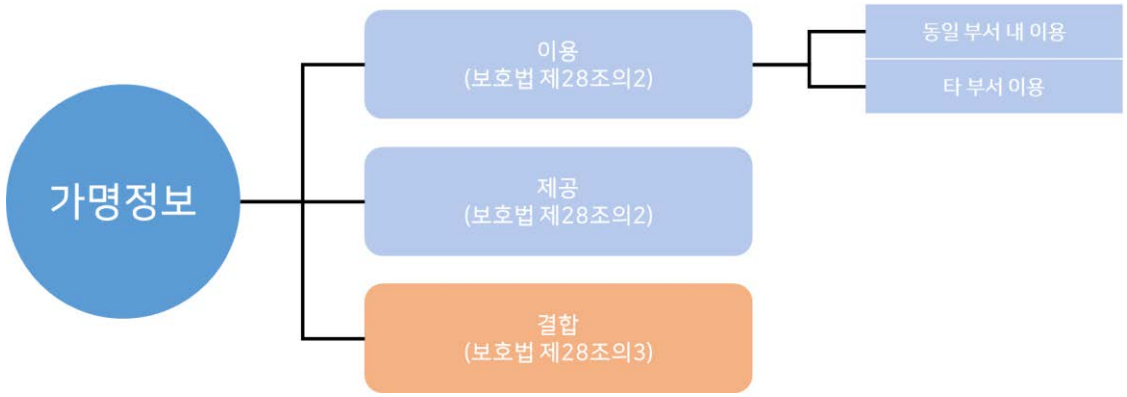
→ 데이터의 식별위험성 검토는 기본적인 가명처리 방안에 영향을 주며 이후 이용환경의 식별 위험성 검토 결과를 바탕으로 가명처리의 수준을 결정하게 됨

2단계 – 위험성 검토 절차(2)

처리 환경의 식별 위험성 검토

- 개인정보처리자는 가명정보 ① 활용 형태(이용·제공), ② 처리 장소, ③ 처리 방법(결합여부) 등 가명정보 처리 상황에 따라 발생할 수 있는 식별 위험성 검토

① 활용 형태 : 가명정보의 활용 형태는 이용, 제공, 결합으로 구분할 수 있음



a. 활용형태의 구분에 따라 보유하고 있는 다른 정보가 달라지게 되며 이에 따른 재식별 가능성도 다르게 나타남

처리 주체가 보유하고 있는 정보 <예시>

구분	처리주체	원본정보	추가정보	내부기관의 다른 정보[1)]	외부기관의 다른 정보[2)]	보유 경험 및 지식[3)]
내부	부서 내 처리자	○	○	○[1)]		○
	타 부서 처리자	○		○[1)]		○
외부	제3자				○[2)]	○

- 주1) 원본정보와 추가정보를 제외한 개인정보처리자가 보유하고 있는 정보를 말함
- 주2) 외부이용기관이 보유하고 있는 정보를 말함
- 주3) 내·외부이용기관이 보유하고 있는 과거 유사 정보에 대한 수행 경험이나 지식 등을 말함

2단계 – 위험성 검토 절차(3)

처리 환경의 식별 위험성 검토

- 개인정보처리자는 가명정보 ① 활용 형태(이용·제공), ② 처리 장소, ③ 처리 방법(결합여부) 등 가명정보 처리 상황에 따라 발생할 수 있는 식별 위험성 검토

① 활용 형태 : 가명정보를 처리하는 처리자(또는 취급자)가 보유하고 있는 정보 또는 접근·입수 가능한 정보, 이용 범위 및 유형 등을 고려하여 식별가능한 항목이 있는지 검토, 단 처리자(또는 취급자)가 보유, 접근, 입수 가능한 모든 정보를 고려하여 식별가능성을 검토할 필요는 없으며 다음의 처리 장소, 처리 방법을 고려하여 식별 가능성 검토

② 처리 장소 : 가명정보가 해당 가명정보 외에 다른 정보의 접근·입수가 제한된 장소에서 처리되는지 검토, 장소는 물리적 분리 뿐 아니라 보안서약서, 계약서, 접근통제, 접근권한관리 등의 관리적, 기술적 통제를 통해 다른 정보와의 접근이 현실적으로 불가능한 경우 물리적 분리에 준하여 검토 (폐쇄환경(물리적, 기술적, 관리적)의 경우 다른 정보를 통한 식별가능성은 고려하지 않음)

③ 처리 방법

- 가명정보를 다른 정보와 연계 분석하는 경우 다른 정보와의 결합 후 식별가능한 항목이 있는지 검토
- 가명정보를 다른 정보와 내부 결합하는 경우 다른 정보와 결합 후 식별가능한 항목이 있는지 검토
- 가명정보를 반복 제공하는 경우 반복 제공을 통해 식별 위험이 높아지는 항목이 있는지 검토

2단계 – 위험성 검토 절차(4)

처리 환경의 식별 위험성 검토

- 식별위험성 검토 체크 리스트

  - 데이터의 위험성, 처리 환경의 식별위험성을 검토하여 식별위험성 검토 결과보고서를 작성하여야 하며 오른쪽의 체크리스트를 이용하여 이를 판단할 수 있음

구분		식별 위험성 검토 사항	
데이터	식별성	개인 식별 가능항 항목 여부	
		검토 항목	검토 결과
		(1) 식별이 가능한 단일항목의 정보가 있는가 * [항목설명] 직업에 개인의 식별성이 매우 높은 정보들이 포함되는 경우(장애인 여성 탁구 국가대표 감독, 지방자치단체장, 2급 이상의 공무원 등)	☐ 예 ☐ 아니오
		(2) 두 개 이상의 컬럼(항목)을 조합하여 식별가능성이 높아지는 정보가 있는가 * [항목설명] △ 동일 데이터셋 내 여러 이용 항목을 동일 목적으로 함께 분석함에 따라 식별가능성이 높아지는 경우(질병, 투약, 약품 등 연관있는 이용 항목을 종합적으로 분석하는 경우) △ 데이터셋 내 가족관계, 직책관계 등 계층적 특성을 가진 이용 항목이 포함되어 있어 개인 식별가능성이 높아질 수 있는 경우(회사 내 정보 분석 시 해당 데이터셋에 소수 직책이 포함되어 있는 경우) △ 시간, 위치, 행위 등 이용 항목을 함께 분석하는 경우	☐ 예 ☐ 아니오
		(3) 공개된 데이터와 결합·대조하여 식별가능성이 높아질 수 있는 이용 항목이 있는가 * [항목설명예를 들어 통계청의 인구 센서스 데이터를 사용하여 식별가능한 이용 항목이 있는지 검토	☐ 예 ☐ 아니오

<식별 위험성 검토 체크리스트 예시 (1)>

2단계 – 위험성 검토 절차(8)

구분		식별 위험성 검토 사항	
데이터	식별성	**개인 식별 가능항 항목 여부**	
		검토 항목	**검토 결과**
		(4) 데이터셋의 크기가 적어 식별이 가능할 우려가 있는가 * [항목설명] 예를 들어 연구대상 질병이 극희귀질병(국내 유병자가 200명 미만인 질병)이라 이 질병을 가진 대상이 한정된 인원이라 식별 가능성이 높아지는 경우가 있는지 검토	☐ 예 ☐ 아니오
		(5) 원본데이터 전체가 아닌 일부의 데이터를 처리하는 샘플링을 적용하지 않았는가 * [항목설명] 예를 들어 상품 구매이력 분석에서 특정 고객층의 일부를 샘플링해서 분석하는 것이 아니라 특정 고객층의 전체를 분석하는 경우인지 검토	☐ 예 ☐ 아니오
		(6) 시계열 성격을 가진 데이터가 포함되어 있는가 * [항목설명] 예를 들어 대학에서 학생들의 학점 데이터를 입학 때부터 졸업때까지의 모든 학점에 대해 분석하는 경우인지 검토	☐ 예 ☐ 아니오

구분		식별 위험성 검토 사항	
데이터	특이정보	**데이터 분포가 편중되어 있어 식별가능성이 있는 이용 항목 여부**	
		검토 항목	**검토 결과**
		(7) 연속적인 숫자형 데이터에서 데이터 값의 분포가 양 끝단의 정보 (분포 곡선에 따라 한쪽의 정보 포함)가 현저히 낮은 항목이 있는가	☐ 예 ☐ 아니오
		(8) 일반적인 문자형 데이터(비 연속적인 숫자형 데이터 및 코드형 데이터 포함)에서 특정 값으로 현저히 낮은 항목이 있는가	☐ 예 ☐ 아니오
	재식별 시 영향도	**재식별 시 정보주체에게 심각한 피해 또는 불이익을 초래할 수 있는 이용 항목 여부**	
		검토 항목	**검토 결과**
		(9) 사회통념상 차별받을 수 있는 정보 등으로 인해 정보주체가 피해 또는 불이익을 받을 수 있는 정보가 있는가	☐ 예 ☐ 아니오
		(10) 재식별로 인하여 받는 피해 또는 불이익의 정도와 규모가 상당히 클 수 있는 정보주체(대중적으로 유명한 사람 등)가 있는가	☐ 예 ☐ 아니오

2단계 – 위험성 검토 절차(8)

구분		식별 위험성 검토 사항	
처리환경	이용 및 제공	**가명정보 활용 형태 및 이용 기관의 개인정보 보호 수준 등을 고려하여 식별가능성이 있는 항목 여부**	
		검토 항목	**검토 결과**
		(11) 처리주체가 보유하고 있는 정보 또는 접근·입수 가능한 정보와 이용 범위 및 유형을 고려하여 식별가능한 항목이 있는가 * [항목설명] △ 시계열 분석 등을 위한 목적으로 가명정보를 반복 제공할 예정인 경우 반복 제공을 통해 식별 위험이 높아지는 항목이 있는 경우가 있는지 검토 △ 가명정보를 취급하는 자와 관련된 정보가 처리하는 데이터셋에 포함되어 있는 경우가 있는지 검토	☐ 예 ☐ 아니오
		(12) 추가정보를 삭제하지 않고 보관하고 있는가	☐ 예 ☐ 아니오
		(13) 가명정보 제공 시 제공받는 자의 개인정보 보호 수준 및 신뢰할 수 있는 인증을 받았는가(ISMS, ISMS-P, ISO 27001 등)	☐ 예 ☐ 아니오

구분		식별 위험성 검토 사항	
처리환경	처리장소	**가명정보가 관리적·기술적·물리적으로 안전한 장소에서 처리되는지 여부**	
		검토 항목	**검토 결과**
		(14) 가명정보 처리 시 다른 정보를 접근·입수할 수 있는 장소인가 • [항목설명] △ 누구나 접근 가능한 개방형 형태의 장소 및 네트워크인지 △ 내부인원만 출입할 수 있는 장소 및 네트워크가 아닌지 △ 가명정보 처리 관련 담당자만 접근할 수 있는 장소 및 네트워크가 아닌지 검토	☐ 예 ☐ 아니오
	다른 정보 결합	**가명정보를 다른 정보와의 결합하여 활용 시 식별가능성이 있는 항목 여부**	
		검토 항목	**검토 결과**
		(15) 다른 정보와의 연계 분석이 예정되어 있는가	☐ 예 ☐ 아니오
		(16) 처리주체가 보유하거나 접근·입수 가능한 정보 등 다른 정보와 연계 또는 결합하여 식별가능한 항목이 있는가	☐ 예 ☐ 아니오

가명정보 처리 가이드라인

2단계 – 위험성 검토 절차(9)

처리 환경의 식별 위험성 검토

- 식별 위험성 검토 결과보고서

  - 식별위험성 체크리스트 등을 통한 데이터의 식별위험성, 이용환경의 식별위험성을 종합하여 식별위험성 검토 결과 보고서를 작성

- 최종 검토의견은 외부 전문가에게 자문 및 작성을 요청할 수 있음

가명정보 활용목적		○ 본 기업은 전국적인 소매유통망을 가지고 있는 대형유통업체로 코로나 이전과 코로나 이후의 상품군별 판매 추이에 대한 통계학적 연구 분석을 통해 이후 코로나의 지속가능성이 높아짐에 따라 상품의 구매전략, 제품의 진열 위치 변경 등의 판매전략의 수립을 위해 데이터로 활용하기 위해 2019년 1월부터 12월까지의 주요 상품군별 판매액 정보와 2021년 1월부터 12월까지의 주요 상품군별 판매액 정보를 나이와 성별, 시군구 단위의 주소별로 비교하여 분석
가명처리 대상 데이터 항목		○ 고객ID, 나이, 주소, 성별, 2019년 1월~12월, 2021년 1월~12월까지의 여행용품, 식품류, 의류, 취미용품, 생활용품, 유아용품, 기타의 7개 범주의 구매금액의 월별 합계액, 월별 구매 총 금액, 월별 선호 제품군, 각 년도의 고객 등급, (전체 222개의 컬럼) ○ 전체 고객 800만명 중 25%를 무작위 샘플링하여 구성한 200만명에 대한 데이터
데이터 위험성	식별성 유무	○ '고객ID'는 개인식별정보임 ○ '나이', '주소', '성별'은 조합했을 때 개인의 식별이 가능한 개인식별 가능정보임
	특이정보 유무	○ 각 범주별 구매금액의 경우 특이정보로 인한 개인 식별성이 발생할 수 있음
	재식별시 영향도	○ 단순 고객의 구매데이터로 재식별 시 영향도는 크지 않을 것으로 판단됨

<식별 위험성 검토 결과보고서 예시>

2단계 – 위험성 검토 절차(9)

처리 환경의 식별 위험성 검토

- 식별 위험성 검토 결과보고서

처리 환경 위험성	이용 및 제공형태	ㅇ 내부 이용
	처리장소	ㅇ 가명정보는 인터넷과 원본 DB에 접근할 수 없는 차단된 별도의 분석 PC에서 분석 예정 ㅇ 분석PC가 있는 환경은 별도의 분석실로 엄격한 출입통제가 되어 있으며 출입 시 출입 관리대장을 기재
	다른 정보 결합가능성	ㅇ 가명처리 전 개인정보와 구매정보를 보유하고 있음
최종 검토의견*	ㅇ 해당 연구는 자사의 데이터를 자사의 내부에서 활용하는 것으로 데이터 자체 위험성과 처리 환경 위험성을 검토할 때 다음과 같은 조치가 필요함 - 결합 가능한 다른정보를 보유하고 있으나 처리 장소를 고려했을 때 결합 가능성은 매우 낮을 것으로 판단됨 - '고객ID'는 개인식별정보로 개인식별 가능성이 매우 높으며 이에 따라 연관관계가 없는 일련번호로 대체할 필요가 있음 - '나이', '주소', '성별'은 그대로 사용하는 경우 조합에 의한 개인식별 가능성이 있으며 이에 따라 다음과 같은 처리가 필요 · '나이' : 주 분석대상이 아닌 13세 미만의 경우 삭제처리가 필요하며 중학생과 고등학생은 하나로 묶어 처리하고 그 외의 나이에 대해서는 1살 단위로 제공하며 90세 이상의 나이에 대해서는 90세 이상으로 처리하는 것이 필요 · '주소' : 동단위와 상세 주소의 경우 통계목적에 필요하지 않기 때문에 삭제하며 시군구 단위의 주소까지만 사용하는 것이 필요 · '성별' : 성별은 분석목적에 필요하므로 그대로 사용 ㅇ 구매액 관련 정보들은 구매금액별 특이치를 검토하여 구매 금액에 대한 적절한 수준의 상단 코딩을 적용(19년 8월 취미용품 114,562,000원 --> 1억원 이상)해야 함 ㅇ 고객등급의 경우 식별성이 높은 VIP와 S를 하나로 묶어 식별성을 낮출 필요가 있음

<식별 위험성 검토 결과보고서 예시>

- 최종 검토의견은 외부 전문가에게 자문 및 작성을 요청할 수 있음

3단계 – 가명처리 개요

- 개인정보처리자는 '식별 위험성 검토 결과보고서 ' 를 기반으로 가명정보의 활용 목적 달성에 필요한 가명처리 방법 및 수준을 정의

<가명처리 수준 정의표 (예시)>

순번	항목명	처리수준	비고
1	소유자명	- 가명처리 (암호화: SHA2+Salt)	- 소유자명과 연락처는 추후 시계열 분석을 위해 가명처리 수행
2	연락처		
3	지번	- 가명처리(삭제)	- 세부 지번의 정보는 분석목적에 필요하지 않음
4	전세	- 기타기술 (라운딩: 만원 단위)	- 만원 단위의 금액만 분석목적에 필요
5	보증금		
6	월세		
7	주택구분	- 처리하지 않음 ※ 항목이 다수여서 작성이 어려운 경우 '별지'를 활용하여 목록만 제시	- 처리하지 않는 항목을 작성
8	시도		
9	시군구		
10	읍면동		
11	전용면적		
12	공급면적		

가명정보 처리 가이드라인

3단계 – 가명처리 절차(1)

가명처리

- ‘가명처리 수준 정의표’를 기반으로 가명처리를 수행
- 가명처리 단계에서 생성되는 추가정보는 **원칙적으로 파기**하고 필요한 경우 **가명정보와 분리하여 별도로 저장**

- 가명처리 (예시)

식별정보 / 식별가능정보

소유자 명	연락처	주택 구분	시도	시군구	읍면동	지번	전세 (천원)	보증금 (천원)	월세 (천원)	전용 면적	공급 면적
김철수	090-1234-5678	아파트	서울 특별시	동작구	사당동	1388-4	-	25,000	750	104.00	84.00
이영희	090-2468-3579	오피스텔	대전 광역시	서구	둔산동	656	81,250	-	-	56.45	24.32
박민호	090-9876-5432	아파트	부산 광역시	해운대구	우동	111-13	125,000	-	-	100.00	84.00

(소유자 명, 연락처) +Salt 암호화

삭제 / 라운딩

(가명처리)

ID	주택 구분	시도	시군구	읍면동	전세 (천원)	보증금 (천원)	월세 (천원)	전용 면적	공급 면적
wd4e85D2C1qe89rwqe	아파트	서울 특별시	동작구	사당동	-	25,000	800	104.00	84.00
r5w1e2SXzi4wd64qwz	오피스텔	대전 광역시	서구	둔산동	81,300	-	-	56.45	24.32
ghe6W15Z5ax40e24jx	아파트	부산 광역시	해운대구	우동	125,000	-	-	100.00	84.00

4단계 – 적정성 검토 및 추가처리 개요

- 가명처리 결과에 대해 적절한 수준으로 가명처리가 이루어졌는지, 재식별 가능성은 없는지 등에 대한 최종적인 판단 절차를 수행(데이터의 분포, 내용 등을 고려하여 특이정보까지 감안한 판단 필요)

→ 각 분야별 가이드라인에 적정성 검토의 절차 및 방법에 대해 정의되어 있는 경우 그 가이드라인을 따라야 함 (보건의료분야, 교육분야)

- 가명처리 결과가 가명정보 활용 목적 달성에 적합하지 못하거나 가명처리 수준이 부족한 경우 추가적인 가명처리가 필요

- 적정성검토 항목별 필요서류 예시
  - ✓ 사용목적 : 사업계획서
  - ✓ 이용환경(개인정보보호 수준, 보유정보 등) : 별도 서류
  - ✓ 가명처리 대상 정보집합물 및 가명정보 명세: 가명처리 정보집합물 명세서
  - ✓ 개인정보 항목별 적용 가명처리 기법 : 식별 위험성 검토 결과보고서, 가명처리 수준 정의표
  - ✓ 가명처리 수준 기준 : 가명처리 수준 정의표
  - ✓ 기타 프라이버시 보호모델 적용 여부 등이 기초자료에 포함

4단계 – 적정성 검토 및 추가처리 개요

- (1. 필요서류 및 위험성 검토)
  - ✓ 필요서류 내용, 데이터 식별 위험성, 처리 환경의 식별 위험성 등 판단 항목을 누락 없이 검토하였는지 확인
    ※ 사전준비 단계에서 필요서류가 법/제도 목적에 적법하게 작성되었는 지와 가명처리 단계에서 체크리스트 및 결과보고서 기반으로 위험성 판단 항목을 누락 없이 검토하였는지 여부
- (2. 가명처리 방법 및 수준의 적정성 검토)
  - ✓ 가명처리 단계에서 위험성 검토 결과를 반영하여 가명처리 방법 및 수준을 적정하게 정의하였는지 확인
- (3. 가명처리의 적정성 검토)
  - ✓ 정의한 가명처리 방법 및 수준에 따라 실제 가명처리를 수행하였는지 확인
  - ✓ 가명정보 항목 전체를 검토하여 가명처리가 제대로 되었는지 확인
    ※ 특히 대용량 정보의 경우 중간에 처리되지 않은 부분이 있을 수 있으므로, 가능한 가명정보 항목 전체 확인 필요
- (4. 처리 목적 달성 가능성 검토)
  - ✓ 가명처리 정보가 당초 가명정보 처리 목적을 달성할 수 있는지 여부 검토
    ※ 목적 달성에 필요한 최소한의 항목으로 처리되었는지와 처리된 정보가 당초 목적을 달성하기에 적정한지 판단

4단계 – 적정성 검토 및 추가처리 개요

추가 가명처리

- 적정성 검토 결과 가명처리가 적정하지 않은 경우 가명처리를 다시 수행하거나 부분적으로 추가 가명처리를 수행함
- 명확한 추가처리를 조건으로 한 경우는 제한적으로 추가처리 적정하지 않은 경우는 다시 심의위원회 구성할 수 있음
- (특이정보처리 부분) 위험도를 바탕으로 가명처리한 경우에도 '특이정보'를 통해 개인식별이 가능한 경우 추가처리
- (특이정보 예시)

사례1) 국회의원 같이 특정 지역에서 소수만 존재하는 직업의 경우 지역구 국회의원 명단 등을 통해 개인이 식별될 수 있음

※ (가명처리 예시) 특정 지역을 인접 지역과 병합* 하거나, 직업을 일반화(정치인)

* 국가통계기관의 경우 세부 지역단위 통계 시 2천명이 되지 않는 경우 인접 지역에 병합

사례2) 호텔, 렌터카 등 여행업종에서 보유중인 최고급 객실이용정보, 특정 차량이용정보는 개인(공인 등)이 SNS등 온라인에 공개하는 정보와 결합되어 개인이 식별될 수 있음

※ (가명처리 예시) 특정 차량(슈퍼카)의 이름을 일반화(스포츠카)하여 게시하거나, 호텔 최고급 객실정보를 일반객실 정보 대체

5단계 – 안전한 관리 개요

- 가명정보처리자는 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 안됨
- 가명정보 처리 과정에서 개인식별 가능성이 증가하는지 여부 등을 지속적으로 모니터링
- 특정 개인이 식별되는 경우 즉시 처리 중지, 회수, 파기 등의 재식별 위험을 제거하기 위한 조치 수행
- 검토결과 적정으로 판단된 가명정보에 대해 관련 법령에 따라 기술적·관리적·물리적 안전조치 이행

5단계 – 안전한 관리 절차(1)

- 적정성 검토 이후 생성된 가명정보는 법에 따라 기술적·관리적·물리적 안전조치 등 사후관리를 이행하여야 함

  (보호법 제28조의4)

제28조의4(가명정보에 대한 안전조치의무 등) ① 개인정보처리자는 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관·관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 대통령령으로 정하는 바에 따라 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 한다.

② 개인정보처리자는 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 관련 기록을 작성하여 보관하여야 한다.

5단계 – 안전한 관리 절차(2)

1. 재식별 금지 및 모니터링

✓ 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니되며(보호법 제28조의5 제1항), 가명정보를 처리과정에서 우연히 특정 개인이 식별되는 경우 처리중지, 회수, 파기 등과 같이 위험을 제거하기 위한 적절한 조치를 즉시 수행하여야 함(보호법 제28조의5 제2항)

제28조의5(가명정보 처리 시 금지의무 등) ① 누구든지 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니된다.

② 개인정보처리자는 가명정보를 처리하고자 하는 경우에는 가명정보의 처리 목적, 제3자 제공 시 제공받는 자 등 가명정보의 처리 내용을 관리하기 위하여 대통령령으로 정하는 사항에 대한 관련 기록을 작성하여 보관하여야 한다.

✓ 개인정보처리자는 가명정보 처리 과정에서 특정 개인이 식별될 위험이 있는지 여부를 지속적으로 모니터링 하는 등 가명정보를 안전하게 처리하여야 함

5단계 – 안전한 관리 절차(3)

2. 안전조치 시행

✓ 개인정보처리자는 사전준비 단계에서 수립한 내부 관리계획에 따라 가명정보를 안전하게 관리하여야 함

3. 가명정보 처리 관련 기록 작성 및 보관

✓ 개인정보처리자는 가명정보의 처리 목적, 개인정보 항목, 이용내역, 제3자 제공 시 제공받는 자를 작성하여 보관하여야 함

결합의 종류

내부결합

- 내부결합
  - ✓ 1단계(사전준비)에서 결합대상 정보 간 결합키로 활용될 공통속성(항목)과 결합알고리즘(암호종류+salt포함)을 선정
  - ✓ 서로 다른 개인정보처리자가 보유한 정보 간의 결합은 보호위 또는 관계부처 장관이 지정한 결합전문기관을 통해 처리

결합의 종류

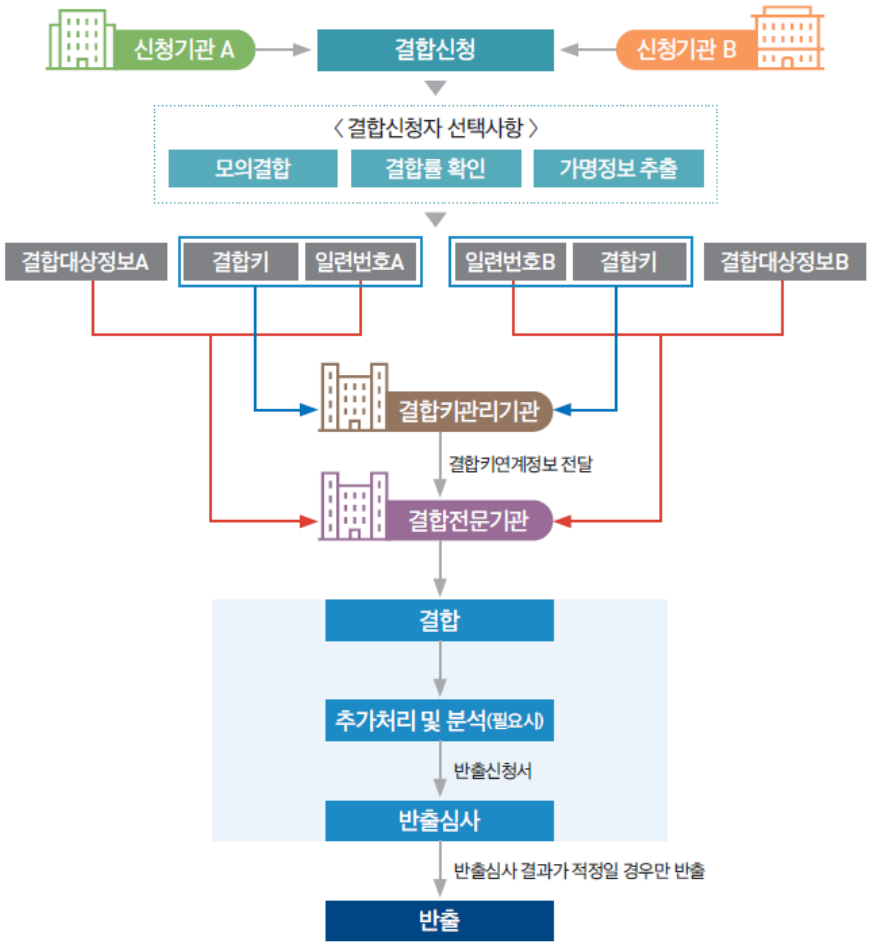
전문기관 결합

- 전문기관결합
  - ✓ 개인정보처리자는 결합전문기관을 통해 통계작성, 과학적 연구, 공익적 기록보존 등을 위한 가명정보 결합 가능
    - ➢ 서로 다른 개인정보처리자가 보유한 가명정보를 결합하여 활용하고자 하는 경우에는 개인정보위 또는 관계 중앙행정기관의 장이 지정한 결합전문기관을 통하여 수행하여야 함 (개인정보보호법 제28조의3 제1항)

※ 가명정보의 결합은 원칙적으로 서로 다른 결합신청자가 결합키를 동일하게 가지고 있는 것에 대해서만 수행. 단, 결합 목적에 따라 결합되지 않는 정보의 분석*이 필요한 경우 개인정보위의 검토 및 승인 후 활용 가능

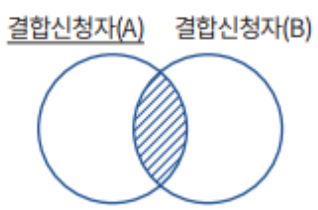
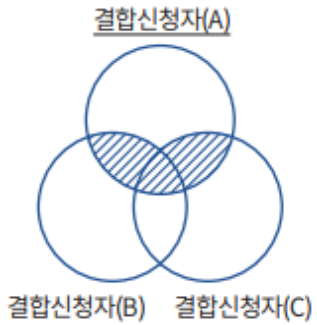
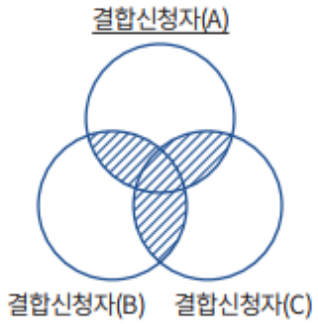
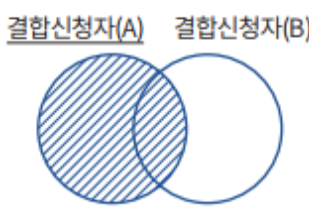
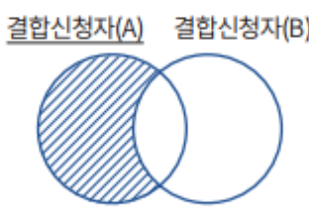
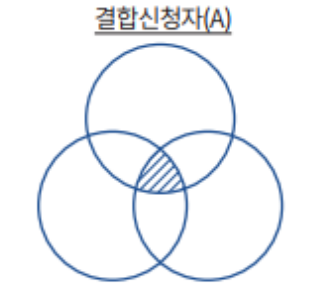
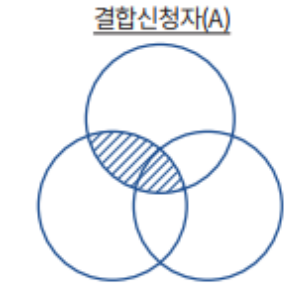
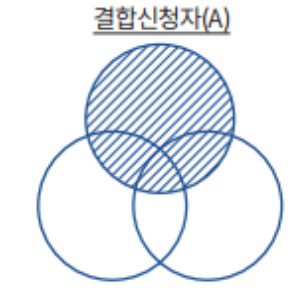
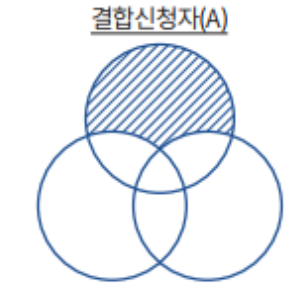
* 결합신청자(A)의 결합되지 않는 정보는 결합신청자(A)만 분석할 수 있음

〈가명정보 결합·반출 절차〉



결합의 종류

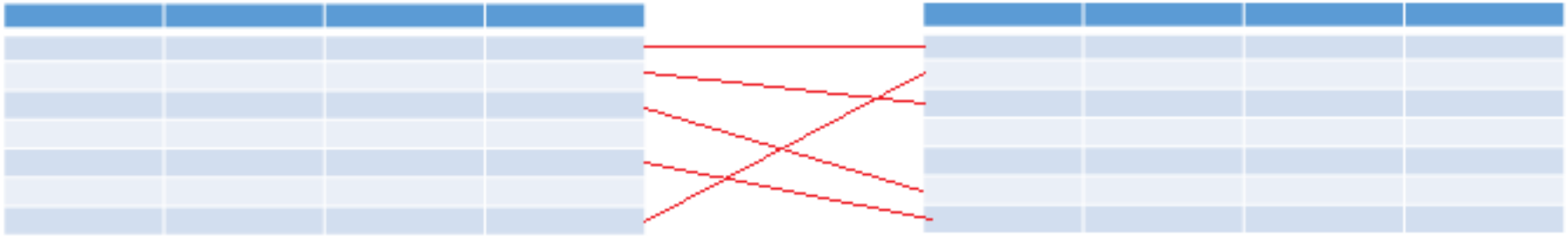
결합의 유형

<table>
<tr><th colspan="3">공통결합
(INNER JOIN)</th><th>확대결합
(OUTER JOIN)</th><th>잔여결합
(ANTI-INNER JOIN)</th></tr>
<tr><th>단일
(INNER SINGLE)</th><th>다중
(INNER MULTI)</th><th>완전
(INNER FULL)</th><th>단일
(OUTER SINGLE)</th><th>단일
(ANTI-INNER SINGLE)</th></tr>
<tr>
<td> 
</td>
<td> 
</td>
<td rowspan="2"> 
</td>
<td> 
</td>
<td> 
</td>
</tr>
<tr>
<td> 
</td>
<td> 
</td>
<td> 
</td>
<td> 
</td>
</tr>
</table>

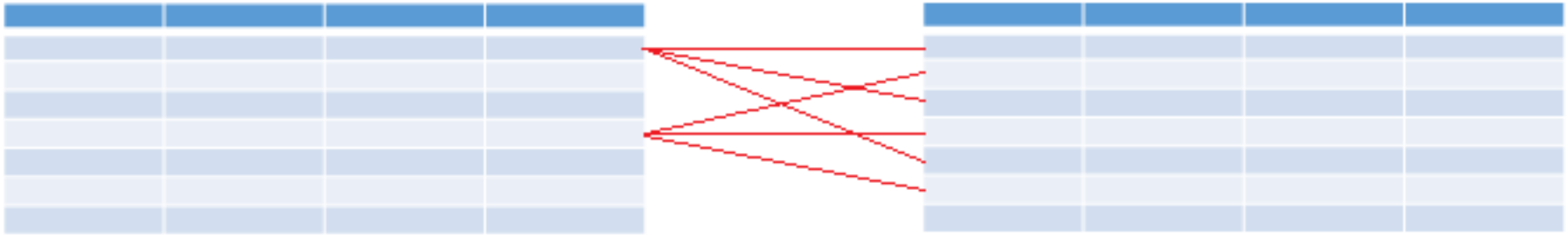
결합의 종류

결합의 기술적 종류

① 1:1 **결합** - 두개의 결합 대상 테이블에서 한줄과 한줄을 결합하는 방식



② 1:N **결합** - 두개의 결합 대상 테이블에서 한줄과 다른 테이블의 여러줄을 결합하는 방식



※ 기존의 '개인정보 비식별 조치 가이드라인'의 전문기관에서는 1:1 **결합**만을 시행하였으며 N:N결합은 지원할 수 없음

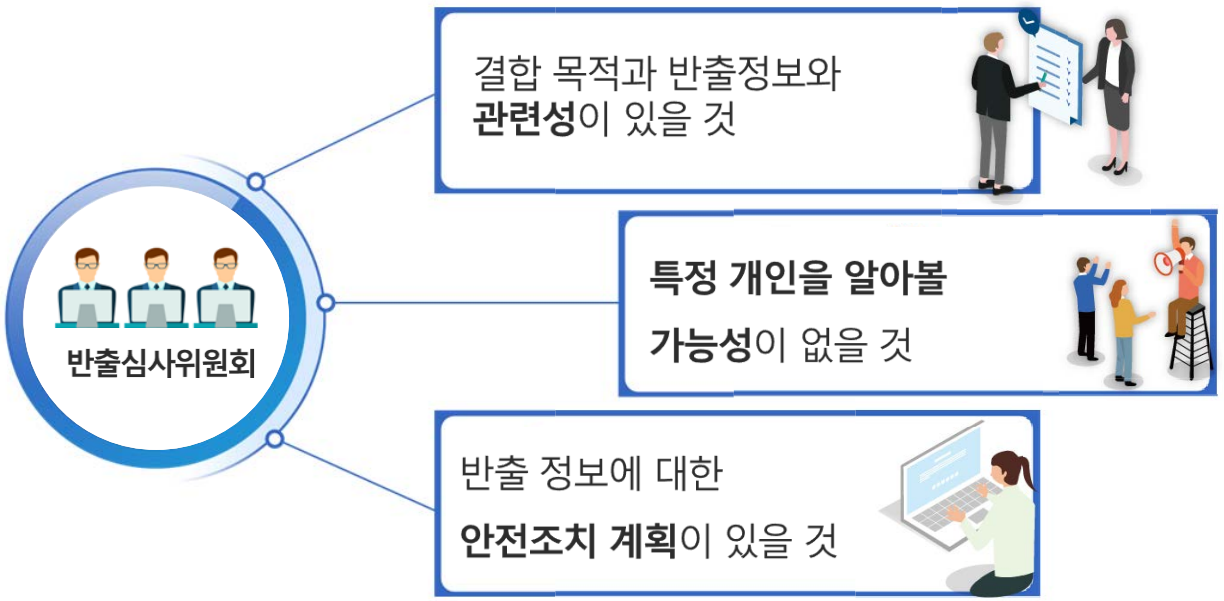
결합정보의 반출

반출 심사 시 고려사항

- **반출 심사 시 고려사항**

- **개인정보처리자 간 가명정보의 결합 및 반출 등(개인정보보호법 시행령 제29조의3 제4항)**

④ 결합전문기관은 다음 각 호의 기준을 충족하는 경우에는 법 제28조의3제2항에 따른 반출을 승인해야 한다.
이 경우 결합전문기관은 결합된 정보의 반출을 승인하기 위하여 반출심사위원회를 구성해야 한다.



가명정보의 반복결합

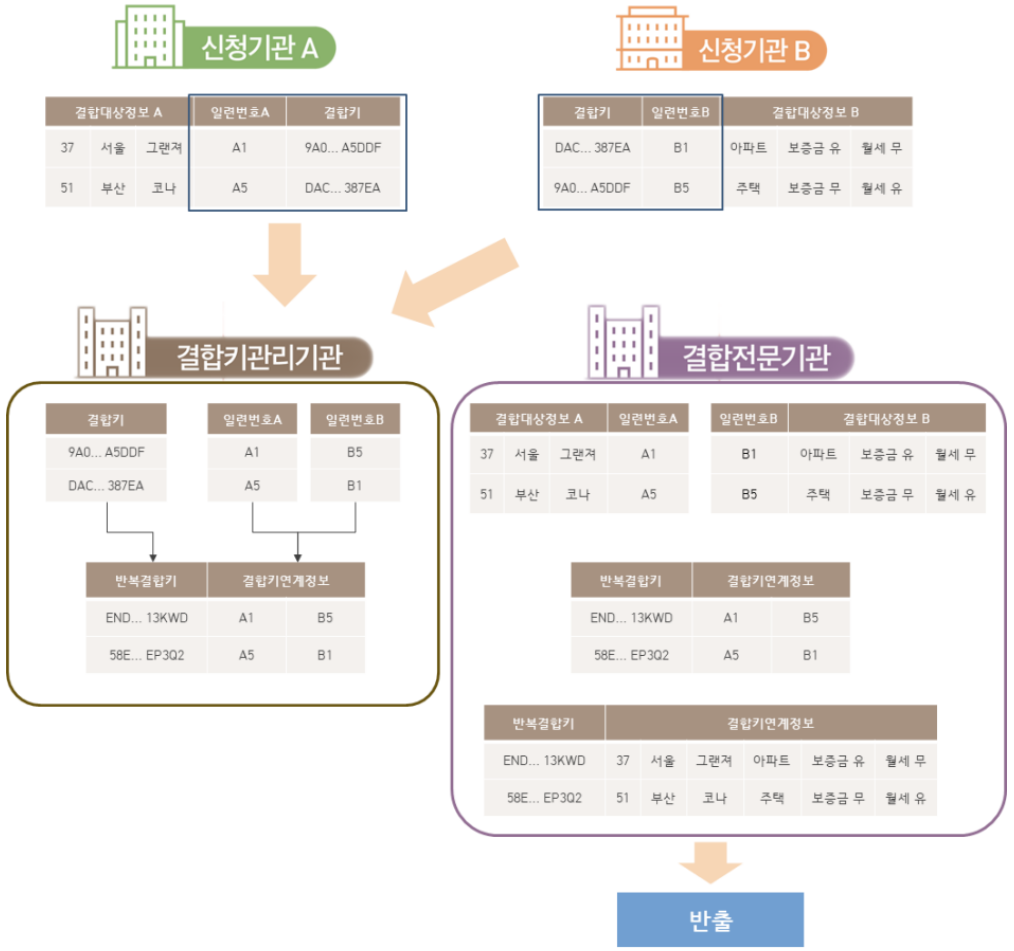
가명정보의 반복결합

- 시계열 분석* 등을 목적으로 가명정보를 결합할 때에는 동일한 서로 다른 개인정보처리자 간의 가명정보를 지속적·반복적으로 반복하며 결합 할 수 있음

  * 시계열 분석 : 어떤 현상에 대하여 과거에서부터 현재까지의 시간에 흐름에 따라 기록된 데이터를 바탕으로 미래의 변화에 대한 추세를 분석하는 방법

  ✓ 반복결합이 필요한 경우, 결합신청 시 반복결합을 선택하여 신청

  ※ 반복결합의 경우, 반출정보에 반복적인 분석을 위해 필요한 키(반복결합키)가 추가 포함됨



가명처리의 이해

1. **관련 용어 정리**
2. 가명정보, 익명정보

03

개인정보의 정의

개인정보보호법 제2조제1호

“개인정보”란 살아있는 개인에 관한 정보로서 다음 각 목의 어느 하나에 해당하는 정보를 말한다.

가. 성명, 주민등록 번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보
나. 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다.
다. 가목 또는 나목을 제 1호의 2에 따라 가명처리함으로써 원래의 상태로 복원하기 위한 추가정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보(이하 “가명정보”라 한다)

개인정보보호법 제2조제1호의2(신설)

“가명처리”란 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말한다.

가명정보의 개념

개인정보보호법 제2조제1호의2(신설)

- 가명처리 과정에서 개인정보의 전부 또는 일부를 대체하는데 이용된 수단이나 방식(알고리즘 증), 가명정보와의 비교·대조 등을 통해 가명처리 된 개인정보 부분을 복원할 수 있는 정보를 말한다. (개인정보보호법 해설서 13P)
- 특정 개인을 식별할 수 있다는 점에서 다른 정보와 동일하나 가명정보를 원래의 개인정보로 복원할 수 있는(일부 복원이 불가능한 경우 포함)정보라는 점에서 다른 정보와 다른 특징을 가짐(개인정보보호법 해설서 13P)

<원본정보(예시)>

성명
김희선
권율
강수지
이순신
박은하
장동건
정우성
이황
오동구

<가명정보(예시)>

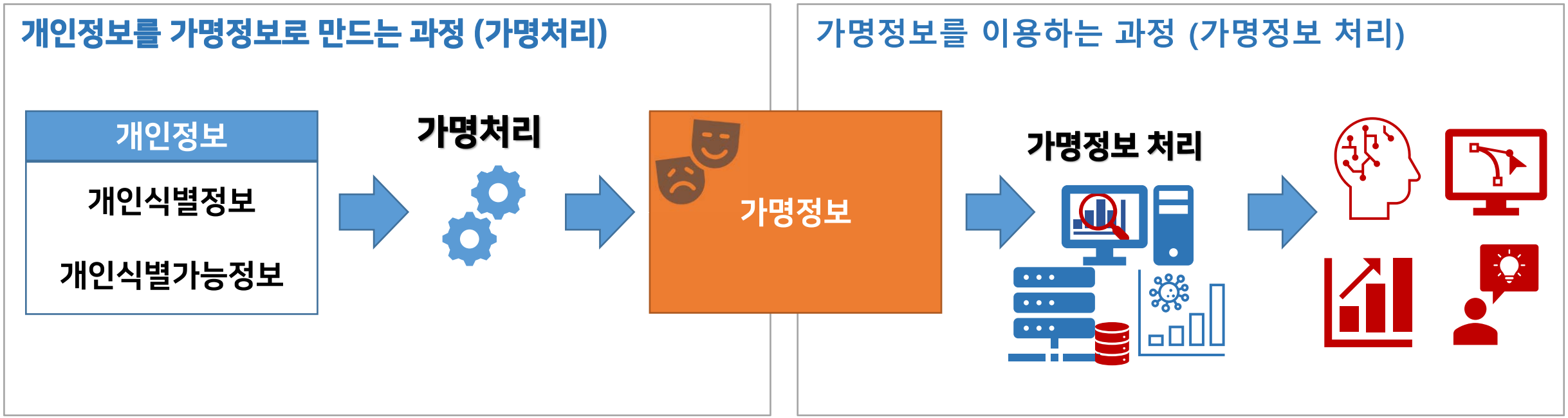
가명
최수지
권민준
강하늘
김라희
박민지
최재영
박상희
강윤희
육동희

<매핑테이블> 추가정보

원본 성명	가명
김희선	최수지
권율	권민준
강수지	강하늘
이순신	김라희
박은하	박민지
장동건	최재영
정우성	박상희
이황	강윤희
오동구	육동희

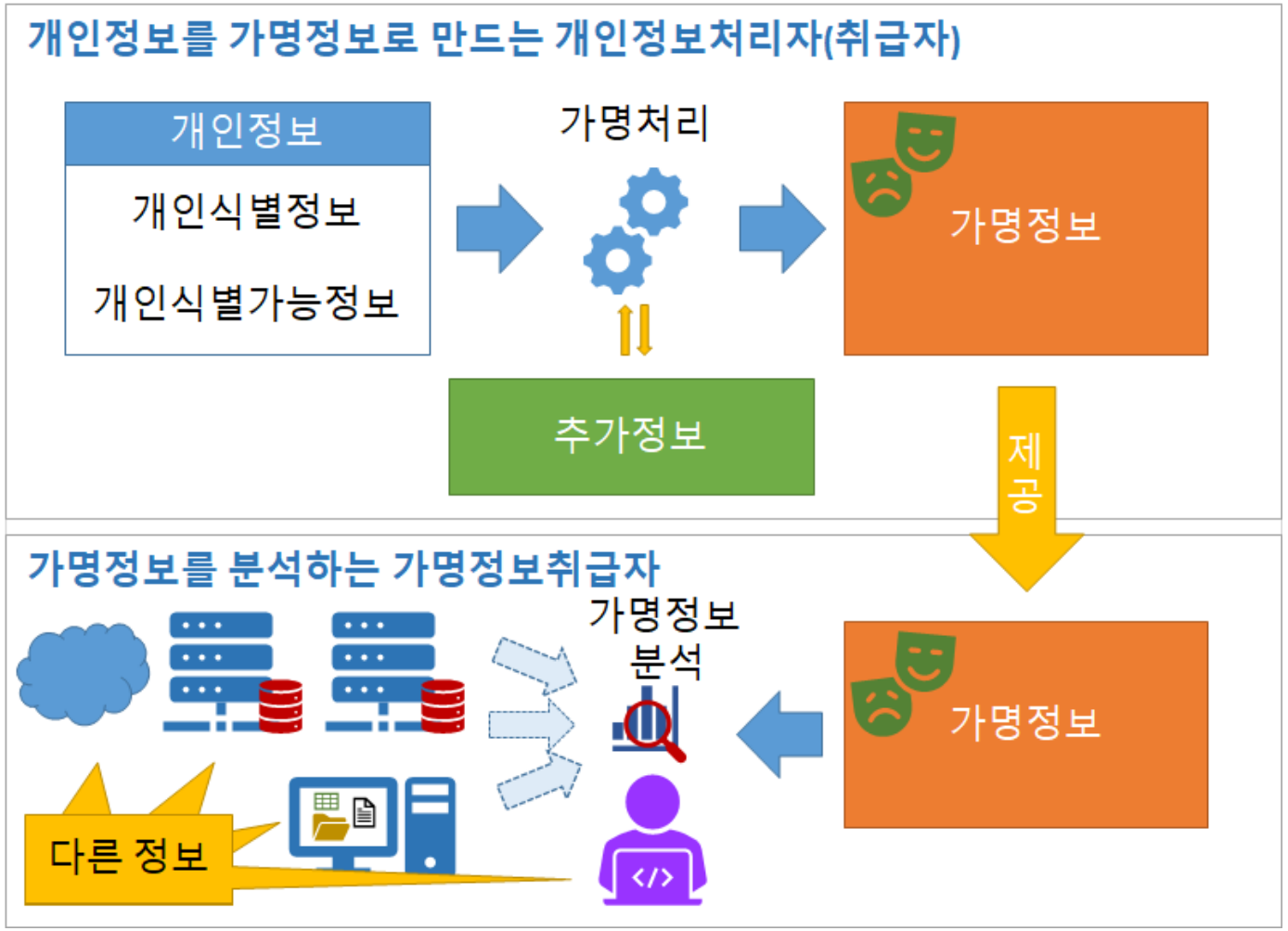
가명처리와 가명정보 처리

- 가명처리는 개인정보를 가명정보로 만드는 과정
- 가명정보 처리는 가명정보를 활용하여 가공, 분석, 이용, 제공 등의 행위를 하는 것



다른 정보의 개념

- 가명정보처리자가 가명정보를 처리하는 환경에서 합법적으로 접근, 입수할 수 있는 모든 정보
  - ✓ 가명처리의 적절성을 판단하기 위한 중요 요건 중의 하나로 적절한 수준의 가명처리가 되지 않는 경우 다른 정보를 통해 특정 개인을 식별할 수 있는 위험이 있음



익명정보의 개념

익명정보

- 개인정보보호법 제 58조의2(적용제외) 이 법은 시간, 비용, 기술 등을 합리적으로 고려할 때 다른 정보를 사용하여도 더 이상 개인을 알아볼 수 없는 정보에는 적용하지 아니한다

- 신용정보의 이용 및 보호에 관한 법률 제2조(정의) “익명처리 " 란 더 이상 특정 개인인 신용정보주체를 알아볼 수 없도록 개인신용정보를 처리하는 것

✓ 기술적인 익명정보에 대한 정의는 없으며 개념적 정의만 있음

✓ 신용정보법에서는 실제 기술적인 수준에 대해 금융위에서 판단기준을 마련했음

- ➢ 개인정보 비식별 조치 가이드라인의 기준과 유사
- ➢ 적정성 검토위원의 검토에 의해 익명을 판단하도록 하고 있음

개인정보의 분류

비식별 관점의 개인정보의 분류

대분류	소분류	개념 및 정의	비고
개인정보	개인식별정보	살아 있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보(가목)	개인정보보호법 제2조 제1호
	개인식별가능정보	해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 것(나목)	

법적 분류	기술적 분류			
구분(출처)	ISO/IEC 20889 (2018년 11월)		정부부처합동 개인정보 비식별 조치 가이드라인 (2016년 6월)	가명정보 처리 가이드라인 (2020년 9월)
대상 범위	운영 환경(전반)	분석 대상 데이터셋	구분 없음	구분 없음
개인식별정보	직접식별자	유일식별자	식별자	개인식별정보
개인식별가능정보	간접식별자	준식별자	속성자 (준식별자+민감정보)	개인식별가능정보

개인정보의 분류(가명, 익명처리의 기술적 분류)

법률	명칭	정의	정의 원문
데이터 운영환경을 고려한 기준	직접 식별자 (Direct identifier)	특정 운영 환경에서 데이터 주체를 고유하게 **식별**할 수 있도록 해주는 속성	attribute that alone enables unique identification of a data principal within a specific operational context
	간접 식별자 (Indirect identifier)	데이터셋에 포함되어 있거나 외부에 속한 속성과 함께 특정 운영 환경에서 데이터 주체의 고유 **식별**을 가능하게 하는 속성	attribute that, together with other attributes that may be in the dataset or external to it, enables unique identification of a data principal within a specific operational context
하나의 데이터셋 기준	고유 식별자 (Unique identifier)	데이터셋에서 데이터 주체를 독립적으로 **선정(singles out)**해 내는 데이터셋에서의 속성	attribute in a dataset that alone singles out a data principal in the dataset
	준식별자 (Quasi-identifier)	데이터셋에서 데이터셋에 포함된 다른 속성과 함께 고려할 때, 데이터 주체를 **선정(singles out)**해 내는 속성	attribute in a dataset that, when considered in conjunction with other attributes in the dataset, singles out a data principal

- **운영환경(Operational context)** : 데이터를 실제 사용하는 환경을 지칭하는 용어로 데이터가 사용되는 환경에서 보유한 모든 다른 데이터와 제3자 또는 잠재적인 공격자가 소유하고 있거나 공개 도메인 상에 존재하는 정보들을 감안한 개념임
- **식별(Identification)** : 개인정보보호법 상의 개인에 대한 식별은 특정 개인정보주체를 정확하게 알아보는 것을 의미함
- **선정(Single out)** : 주어진 데이터 주체를 고유 식별하기 위해 알려진 특성집합들을 관찰하여 데이터셋 내에 해당 데이터 주체에 속한 레코드를 격리(isolation)하는 행위를 말함

개인정보의 분류

프라이버시 침해 관점에서의 개인정보분류

민감정보

- 식별자를 통해 신원이 공개되는 경우 개인에게 가볍게는 수치심으로부터 심각하게는 사회적 차별, 경제적 피해를 끼칠 수 있는 모든 정보
- 개인정보보호법의 민감정보와는 다른 개념이며, 개인정보보호법의 민감정보는 아래의 법률적 민감정보에 해당함
- **민감정보의 분류 : 개인이 스스로 공개되기를 원하지 않는 정보***

※ Database Anonymization Techniques with Focus on Uncertainty and Multi-Sensitive Attributes, B.K.Tripathy, 2013

① **법률적 민감정보** : 유전정보와 같이 법령에 의해 사용이 제한된 정보 (법률 준수 여부)

② **식별성 희귀정보** : 희귀난치성질환(상병코드), 희귀의약품(투약코드), 우리나라에서 가장 높은 급여를 받는 사람 등과 같이 정보 자체의 일반적인 특성은 식별성을 가지지 못하지만 해당되는 개인들이 유일하거나 매우 희귀하여 개인의 식별이 가능하도록 할 수 있는 정보 (재식별 위험)

③ **낙인성 정보** : 종교, 사채 금액, 입양아 여부, 특정 병명(AIDS, 한센병 등) 등 공개 시 개인에게 사회적·경제적으로 심각한 피해를 줄 수 있는 정보 (영향도 위험)

④ **기타 민감한 정보** : 기타 소득, 재산 상태 등 민감한 정보 (영향도 위험)

- 가명처리 관점에서 민감정보가 있는 경우 위의 위험성을 검토하여 추가적인 처리가 필요한지 검토하여야 함
- 교육정보는 산업군의 민감정보에 포함되며 이에 따라 다른 산업군에 비해 좀 더 강한 처리가 필요(보건의료, 복지정보와 유사)

개인정보의 분류

데이터 분석 관점에서의 개인정보분류

분석 시 활용되는 개인정보(TA, Target Attribute)

- 데이터 분석에 개인정보가 활용되는 경우 다음과 같이 두가지의 분류로 나눠질 수 있음

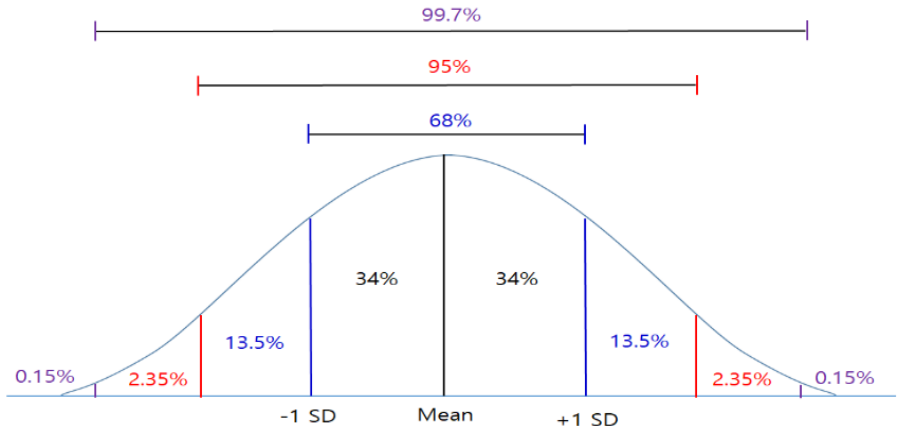
1. 분석 대상을 정의하는 컬럼
   - ✓ 분석 대상을 분류하거나 분석 대상의 통계학적 특성을 나타내는 정보
   - ✓ 대부분의 경우 준식별자에 포함되며 일부 직접 식별자가 포함될 수 있다.
     분석의 목적에 적합한 정도의 가명화, 익명화의 적용이 필요
2. 분석 목적을 달성하기 위한 컬럼
   - ✓ 분석의 직접적인 목적이 되는 컬럼, 이 컬럼에 강도가 높은 비식별을 적용하는 경우 분석의 목적달성이 어려울 수 있음
   - ✓ **매우 식별성이 높아 분석의 정확도 차원의 목표를 수정하지 않고는 가명화, 익명화가 불가능한 경우 데이터 상황에 대한 추가적인 통제가 필요함**
   - ✓ **비식별 시 개인정보의 분류 방법 중 가장 중요한 분류 방법으로 이 개념을 포함하지 않는 경우 데이터의 유용도가 매우 낮아질 수 있음**

특이치의 개념

특이치의 검출 방법

3시그마 규칙(68-95-99.7 규칙)

- 정규분포형태의 데이터 분포에서 3시그마를 벗어나는 데이터는 특이치로 판단



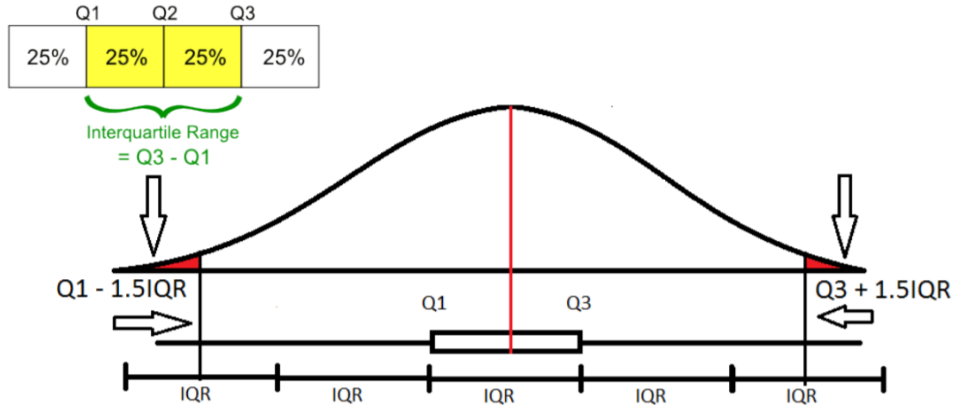
캡핑(Capping Method)

- 데이터 분포에서 5~95%의 범위를 벗어난 값을 특이치로 판단 (2시그마를 기준으로 하는 것과 유사함)

15*IQR(inter-Quartile range) 방법

- Q1에서 하단으로 1.5IQR을 벗어난 값과 Q3에서 상단으로 1.5IQR을 벗어난 값을 특이치로 판단

IQR = Q3 – Q1



가명처리의 이해

1. 관련 용어 정리
2. **가명정보, 익명정보**

3

가명처리 예시

원본데이터

고객번호	NAME	SEX	AGE	ADDRESS	BIRTHDAY	MOBILE
KD384108281	김수미	F	51	서울 성동구 성수동1가 169번지	5월 24일	910-3045-6144
KD650878117	김애란	F	50	서울 강서구 화곡7동 367번지	12월 24일	910-9659-4043
KD491709404	이경아	F	53	경기 가평군 설악면 412-22번지	7월 24일	910-9830-6422
KD102984041	김정애	F	58	경기 의정부시 녹양동 449-45번지	5월 23일	910-3891-3865
KD768757023	김찬수	M	49	서울 서대문구 연희동 213-33번지	3월 23일	910-4632-2821
KD685952071	정순기	M	43	경기 고양시 주엽1동 549번지	12월 30일	910-1142-2843
KD575291753	장용원	M	46	경기 성남시 수내동 116-86번지	1월 15일	910-9716-9910
KD174298911	이의수	F	41	경기 고양시 대화동 86번지	10월 9일	910-9628-3123
KD011531953	천계순	F	52	경기 고양시 탄현동 548번지	6월 28일	910-6376-5929
KD146335971	박찬옥	F	52	경북 포항시 연일읍 377-86번지	1월 5일	910-8460-5807
KD998246845	이금자	F	51	서울 강서구 방화3동 467-5번지	7월 3일	910-9567-7433
KD036003472	김인숙	F	57	경남 창원시 진동면 200-37번지	12월 10일	910-2976-6394

직업코드	주거형태	결혼유무	연카드사용액	추정연소득
12101	단독주택	미혼	8,359,972	52,249,823
11200	아파트	미혼	18,570,935	29,953,120
820000	빌라 및 다세대	기혼-한사람이상재혼	37,852,287	122,104,150
231300	아파트	기혼-모두초혼	2,545,077,452	21,404,897,510
231400	단독주택	기혼-한사람이상재혼	41,639,864	74,356,899
231500	아파트	기혼-모두초혼	5,012,338	8,353,896
232100	단독주택	기혼-모두초혼	1,062,197	8,851,641
232900	아파트	미혼	5,775,319	8,493,116
232901	오피스텔	기혼-모두초혼	754,410,647	48,328,677
232902	빌라 및 다세대	기혼-모두초혼	14,925,034	27,638,952
232903	단독주택	기혼-한사람이상재혼	98,626,257	124,843,363
232904	아파트	기혼-모두초혼	1,684,751	5,809,487

가명처리 예시

가명처리된 데이터

SN	SEX	AGE	ADDRESS	BIRTHDAY
0678399980cc2fbdfc863c09f7cdbdc65899627135bf0c88a441d343fef0f0bd	F	51	서울 성동구 성수동1가	5월 24일
49f22907ab16cef4c9bbec43fda6e435cf527ee98e074ebcb41f889d234c79c7	F	50	서울 강서구 화곡7동	12월 24일
3966d6a5a82c1e9a901f5d6817a7e6757f72da85d220bd4e1857b7adee69be19	F	53	경기 가평군 설악면	7월 24일
7b22d7c2a0afbd86762ebed5ae0a76a95e08f19296643bc3f6b403c2119d7842	F	58	경기 의정부시 녹양동	5월 23일
1ab6b981f05a734a2c66c8cd330c9f031abf4fa81cbf6ffc070ea293a45719dc	M	49	서울 서대문구 연희동	3월 23일
52f36203643666c8778d8d6ca5ca71af9225e0871cda219e7f2f39c70fce9fa6	M	43	경기 고양시 주엽1동	12월 30일
a644f4597b0b7f9aa61edf1be786a4752ee144fe88ca94c0dfd903fae13831ed	M	46	경기 성남시 수내동	1월 15일
6be62d8b0af1a0d6874dc46a29347a25270c1f15c0fcc9a3a411f450278e4810	F	41	경기 고양시 대화동	10월 9일
9ca941895ea6781b143328c6433b223c45eaba04ca8998ab48be54c87e1a00fd	F	52	경기 고양시 탄현동	6월 28일
effd0082a235075888b952a3ed3d511b8426df6e3a8de9fa90ce61ba63025793	F	52	경북 포항시 연일읍	1월 5일
bca8333fed238de5484fe64c0ca95df4eabdc9f1c55c7439c2cfa0ae3ee17bc0	F	51	서울 강서구 방화3동	7월 3일
fa82e65f7ffbdb92c6e50b5a038d8aca9cf392dcf95a177e61a9a3ba8f7a4a0f	F	57	경남 창원시 진동면	12월 10일

직업코드	주거형태	결혼유무	연카드사용액	추정연소득
12101	단독주택	미혼	8,359,972	52,249,823
11200	아파트	미혼	18,570,935	29,953,120
820000	빌라 및 다세대	기혼-한사람이상재혼	37,852,287	122,104,150
231300	아파트	기혼-모두초혼	2억 이상	5억 이상
231400	단독주택	기혼-한사람이상재혼	41,639,864	74,356,899
231500	아파트	기혼-모두초혼	5,012,338	8,353,896
232100	단독주택	기혼-모두초혼	1,062,197	8,851,641
232900	아파트	미혼	5,775,319	8,493,116
232901	오피스텔	기혼-모두초혼	754,410,647	48,328,677
232902	빌라 및 다세대	기혼-모두초혼	14,925,034	27,638,952
232903	단독주택	기혼-한사람이상재혼	98,626,257	124,843,363
232904	아파트	기혼-모두초혼	1,684,751	5,809,487

가명처리 예시

가명처리에 사용된 방법

구분	고객번호	NAME	SEX	AGE
원본정보	KD384108281	김수미	F	51
가명처리방법	SHA 512를 적용한 일방향 암호화	컬럼 삭제	그대로 사용	그대로 사용
가명처리결과	0678399980cc2fbdfc863c09f7cdbdc65899627135bf0c88a441d343fef0f0bd		F	51

구분	ADDRESS	BIRTHDAY	MOBILE
원본정보	서울 성동구 성수동1가 169번지	5월 24일	910-3045-6144
가명처리방법	상세 번지 삭제, 동단위로 범주화	그대로 사용	컬럼 삭제
가명처리결과	서울 성동구 성수동1가	5월 24일	

구분	직업코드	주거형태	결혼유무	연카드사용액	추정연소득
원본정보	12101	단독주택	미혼	2,545,077,452	21,404,897,510
가명처리방법	그대로 사용	그대로 사용	그대로 사용	상단 코딩, 경계값 치환	상단 코딩, 경계값 치환
가명처리결과	12101	단독주택	미혼	2억 이상	5억 이상

가명처리 기준

가명처리의 기준 1. 식별가능성

① 식별 가능성

- ✓ 추가 정보가 없는 상태(다른 정보가 있는 상태)에서 가명정보의 분석(다른 정보를 포함한)을 통한 특정개인에 대한 식별 가능성
- ✓ 가명처리 시 가명정보 자체만으로 특정 개인을 알아볼 수 있는 지와 추가정보 또는 다른 정보의 결합가능성을 고려할 필요가 있음 (가명정보 처리 가이드라인(p.9))
- ✓ 식별이란?
  - ➢ 연봉 8억 7천을 받는 서울 서대문구 연희동의 52세 남자 → 일반적으로는 식별이 아님
  - ➢ 연봉 8억 7천을 받는 서울 서대문구 연희동의 52세 남자 → **동일 직장의 사람이 분석하는 경우는 어떨까?**
  - ➢ **아래 예시는 대한은행의 고객정보를 대한은행에서 분석하는 것으로 직장이 동일한 사람이 분석대상에 포함된 경우 급여의 특이치로 인해 식별가능성이 발생하는 경우를 나타냄**

고객번호	NAME	SEX	AGE	ADDRESS	직장명	연소득
KD384108281	김수미	F	51	서울 성동구 성수동1가 169번지	KISA	76,542,156
KD650878117	김애란	F	50	서울 강서구 화곡7동 367번지	삼성전자	97,248,512
KD491709404	이경아	F	33	경기 가평군 설악면 412-22번지	대신건설	32,158,245
KD102984041	김정애	F	48	경기 의정부시 녹양동 449-45번지	우영정보	51,248,185
KD768757023	김찬수	M	52	서울 서대문구 연희동 213-33번지	대한은행	874,256,148

일년번호	SEX	AGE	ADDRESS	직장명	연소득
136	F	51	서울 성동구 성수동1가	KISA	76,542,156
748	F	50	서울 강서구 화곡7동	삼성전자	97,248,512
1356	F	33	경기 가평군 설악면	대신건설	32,158,245
2823	F	48	경기 의정부시 녹양동	우영정보	51,248,185
4175	M	52	서울 서대문구 연희동	대한은행	874,256,148

또는 분석가가 국세청 사람으로 전국민의 연소득을 알 수 있다면?

- ➢ 가명처리의 식별가능성은 **가명정보의 가명처리 수준**과 이용환경의 **다양한 요소를 모두 고려해야 함**

가명처리 기준

가명처리의 기준 1. 식별가능성

- 데이터의 이용환경 요소(미국: Context, 영국 : Context와 Data를 포함하여 Data situation(데이터 상황))
  - ✓ 에이전트: 데이터 흐름의 모든 시점에서 데이터에 대해 행위를 하고 상호작용할 수 있는 사람 및 단체
  - ✓ 거버넌스 프로세스 : 에이전트와 데이터의 관계가 관리되는 방식, 데이터의 관리 및 보호에 대한 정책, 개인정보 보호수준, 데이터 제공에 대한 계약, 규범과 관행을 통한 비공식적인 행동(리스크 회피 정책의 보유 여부, 개인정보보호를 우선시 하는 문화 등)
  - ✓ 인프라 : 데이터가 흐르고 데이터 환경을 형성하도록 허용하는 구조와 설비, 보안 인프라를 포함하나 보안 인프라보다 광범위한 사회적이고 경계적인 구조가 포함
  - ✓ 다른 정보 : 활용을 하기 원하는 데이터에 연결되어 재식별을 가능하게 만들 수 있는 모든 정보, 다음과 같은 4가지 세부 범주로 구분
    - 에이전트가 가지고 있는 개인지식(분석가의 개인지식, 주요 분석 분야 등)
    - 공개적으로 사용 가능한 데이터(Data.go.kr 등)
    - **업무를 위해 사용하는 제한된 엑세스 환경의 모든 데이터**
    - 기타 유사한 데이터의 공개 여부

- 이용환경 요소에 따른 가명처리
  - ✓ 정보의 이용환경 요소에 대한 통제가 가능한 경우 데이터의 유용성을 최대한 살리기 위해 위험분석을 통해 통제를 통한 잔여위험을 파악하여 이를 기준으로 개인정보(데이터)를 처리
  - ✓ **데이터의 이용환경의 위험성**과 **데이터의 위험성**을 검토하여 가명처리 수준을 정의

가명처리와 익명처리의 차이점

해외 각국의 표준에서의 이용환경에 대한 통제의 고려

- 다양한 국가의 비식별 데이터 활용 모델에서 이용환경에 대한 통제를 고려하고 있음

구분	UKAN의 Access Model	IPCO의 Release Model	NIST의 Data Sharing Models
완전 공개모델	Open Access	Publicly	The Release and Forget Model
제약적인 공개모델	Delivered access	semi-publicly	The Data Use Agreement(DUA) Model
한정된 장소 사용 모델	On-site safe settings	non-publicly	The Enclave Model
기타 모델	Virtual Access	없음	The Synthetic Data with Verification Model

- 가명처리는 이용환경에 대한 환경적 통제를 고려하고 있음
- 가명처리의 적정성 수준은 결국 이러한 이용환경에 대한 통제를 포함한 위험기반의 판단이 필요함
- 가명 처리된 정보는 이용환경에 대한 통제를 고려하여 가명처리 기준에 적합해야 함
  - 식별 가능성 : 추가 정보를 사용하지 않은 상태에서의 가명정보의 분석을 통한 개인에 대한 식별 가능성
  - 복원 가능성 : 추가 정보를 사용하지 않은 상태에서 다시 원래의 정보로의 복원 가능성

가명처리 기준

가명처리의 기준 2가지

② 복원 가능성(가명정보에 대한 기준)

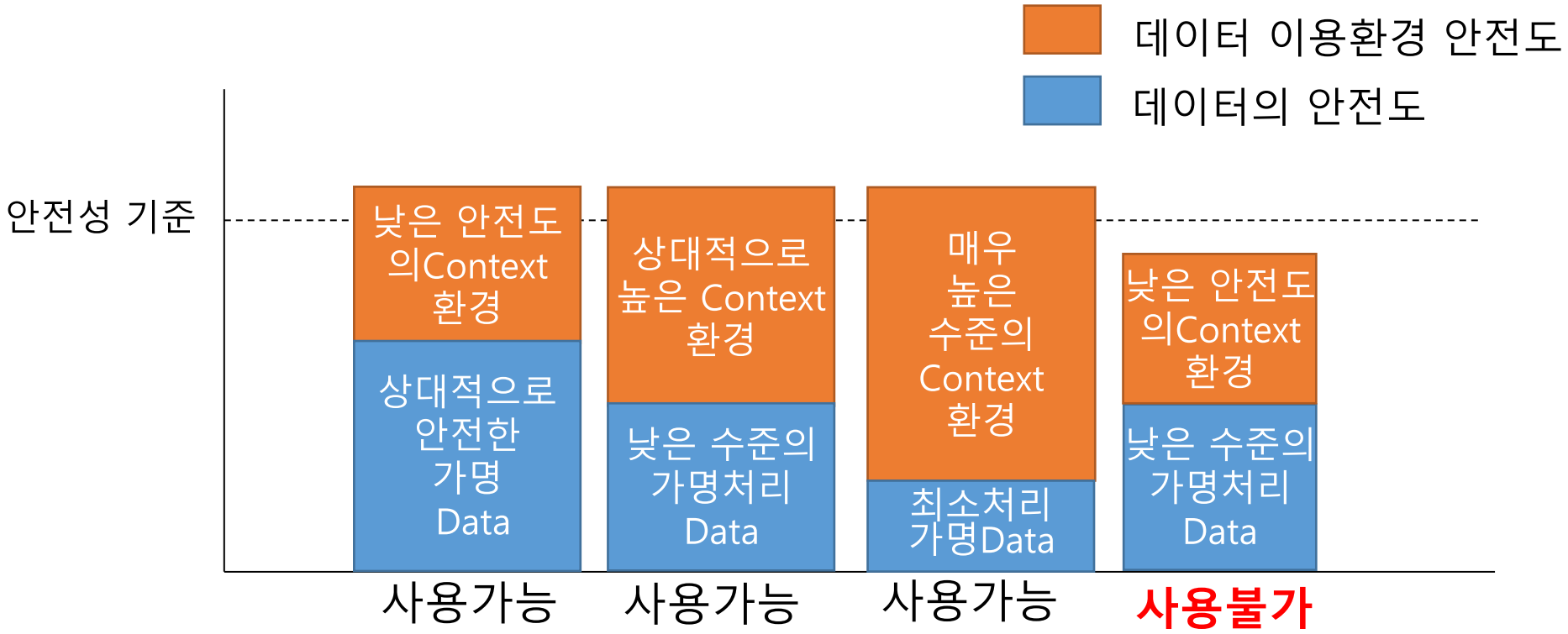
- ✓ 추가 정보(다른 정보는 있을 수 있는 상태)가 없는 상태에서 다시 원래 정보(원본개인정보)로의 복원 가능성
- ✓ 오른쪽의 예시와 같이 주민등록번호에 암호화를 적용하였으나 암호화의 강도가 낮아 다시 원본을 알아낼 수 있을 가능성 (De-anonymizing South Korean Resident Registration Numbers Shared in Prescription Data, 2015, Latanya Sweeney and Ji Su Yoo)
- ✓ 또는 암호화 알고리즘 자체의 문제로 원래의 값이 노출될 가능성
- ✓ 한국인터넷진흥원의 암호 알고리즘 및 키 길이 이용 안내서(2018년)에 따라 강력한 암호화 방식의 사용이 필요

홀수자리		짝수자리	
a	1	f	0
b	2	g	9
c	3	h	8
d	4	i	7
e	5	j	6
f	6	k	5
g	7	l	4
h	8	m	3
i	9	n	2
j	0	o	1

fgjnjganjohng → 690209-1201827

가명처리와 익명처리의 차이점

가명처리의 환경의 위험성과 데이터의 가명처리 수준과의 상관관계



✓ 국내에서는 이러한 이용환경에 대해 법과 제도로 규정하고 있음(개인정보보호법, 시행령, 고시 등, 2장 참조)

익명처리 예시

익명처리 전 데이터

고객번호	NAME	SEX	AGE	ADDRESS	BIRTHDAY	MOBILE
KD384108281	김수미	F	51	서울 성동구 성수동1가 169번지	5월 24일	910-3045-6144
KD650878117	김애란	F	50	서울 강서구 화곡7동 367번지	12월 24일	910-9659-4043
KD491709404	이경아	F	53	경기 가평군 설악면 412-22번지	7월 24일	910-9830-6422
KD102984041	김정애	F	58	경기 의정부시 녹양동 449-45번지	5월 23일	910-3891-3865
KD768757023	김찬수	M	99	서울 서대문구 연희동 213-33번지	3월 23일	910-4632-2821
KD685952071	정순기	M	93	경기 고양시 주엽1동 549번지	12월 30일	910-1142-2843
KD575291753	장용원	M	46	경기 성남시 수내동 116-86번지	1월 15일	910-9716-9910
KD174298911	이의수	F	41	경기 고양시 대화동 86번지	10월 9일	910-9628-3123
KD011531953	천계순	F	52	경기 고양시 탄현동 548번지	6월 28일	910-6376-5929
KD146335971	박찬옥	F	52	경북 포항시 연일읍 377-86번지	1월 5일	910-8460-5807
KD998246845	이금자	F	51	서울 강서구 방화3동 467-5번지	7월 3일	910-9567-7433
KD036003472	김인숙	F	57	경남 창원시 진동면 200-37번지	12월 10일	910-2976-6394

직업코드	주거형태	결혼유무	연카드사용액	추정연소득
12101	단독주택	미혼	8,359,972	52,249,823
11200	아파트	미혼	18,570,935	29,953,120
820000	빌라 및 다세대	기혼-한사람이상재혼	37,852,287	122,104,150
231300	아파트	기혼-모두초혼	2,545,077,452	21,404,897,510
231400	단독주택	기혼-한사람이상재혼	41,639,864	74,356,899
231500	아파트	기혼-모두초혼	5,012,338	8,353,896
232100	단독주택	기혼-모두초혼	1,062,197	8,851,641
232900	아파트	미혼	5,775,319	8,493,116
232901	오피스텔	기혼-모두초혼	754,410,647	48,328,677
232902	빌라 및 다세대	기혼-모두초혼	14,925,034	27,638,952
232903	단독주택	기혼-한사람이상재혼	98,626,257	124,843,363
232904	아파트	기혼-모두초혼	1,684,751	5,809,487

익명처리 예시

익명처리 후 데이터

SN	SEX	AGE	ADDRESS	직업	주거형태	연카드사용액	추정연소득
1	F	50	제주	서비스업	오피스텔	오백만원 이상 이천만원 미만	6천만원 미만
2	F	50	경기	서비스업	단독주택	오백만원 이상 이천만원 미만	삼천만원 미만
3	F	50	서울	서비스업	아파트	이천만원 이상 오천만원 미만	일억원 이상
4	F	50	경기	기능직	단독주택	1억원 이상 3억원 미만	오억원 이상
5	M	80이상	서울	사무직	아파트	이천만원 이상 오천만원 미만	팔천만원 미만
6	M	80이상	경기	기능직	아파트	오백만원 이상 이천만원 미만	일천만원 미만
7	F	40	경기	전문직	아파트	일백만원 이상 이백만원 미만	일천만원 미만
8	F	40	인천	전문직	아파트	오백만원 이상 이천만원 미만	일천만원 미만
9	F	50	인천	기타	아파트	1억원 이상 3억원 미만	오천만원 미만
10	F	50	서울	판매직	아파트	오백만원 이상 이천만원 미만	삼천만원 미만
11	F	50	경기	기타	아파트	오천만원 이상 1억원 미만	이억원 미만
12	F	50	서울	단순노무	아파트	일백만원 이상 이백만원 미만	일천만원 미만

익명처리 기준

익명처리 기준 3가지

① 특정 가능성(Single out)

✓ 데이터 주체를 고유 식별하기 위해 데이터 셋의 일련의 특성들(characteristics)을 관찰하여 개인에 속한 레코드를 격리(Isolation)해 낼 가능성, 준식별자가 유일한 레코드나 나오면 안됨

우편번호가 124**인 사람은 한사람으로 다른 사람과는 달리 유일한 정보를 보유하고 있어 다른 정보와의 격리가 가능하며 이에 따라 사전지식공격, 검사시나리오 등에 취약하여 식별가능성이 높아질 수 있음

우편번호	나이	성별	질병
123**	30	M	전립선염
123**	30	M	상기도염
123**	20	M	고혈압
123**	20	M	심부전증
123**	30	M	전립선염
124**	20	M	갑상선암
125**	30	M	고혈압
125**	30	M	고혈압
125**	30	M	당뇨
125**	20	M	상기도염
125**	20	M	전립선염
125**	20	M	상기도염

익명처리 기준

익명처리 기준 3가지

② 연결 가능성(Linking)

✓ 동일한 데이터 주체 혹은 데이터 주체 그룹과 관련된 레코드를 별도의 데이터 셋(다른 정보)에 연결하여 개인을 식별할 가능성

처방기록데이터

우편번호	생년월일	성별	질병	처방
12345	790101	남	전립선염	A12
23456	790102	여	고혈압	A11
34567	790103	남	전립선염	B12
45678	790104	여	고혈압	B14
56789	790105	남	당뇨	A13
67890	790106	여	HIV	B13

입당기록데이터

우편번호	생년월일	성별	정치성향	등록일
12345	690101	남	공산주의	120304
23456	690102	여	민주주의	120304
34567	690103	남	공산주의	120304
45678	690104	여	민주주의	120304
56789	790105	남	민주주의	120304
67890	790106	여	공산주의	120304

보유하고 있는 두 정보를 결합하면

56789에 거주하는 생년월일이 790105인 남성이 당뇨로 A13을 처방받고, 민주주의임을 유추 가능

또한 67890에 거주하는 생년월일이 790106 여성은 HIV에 감염, B13을 처방받고, 공산주의 성향을 가짐을 유추 가능

익명처리 기준

익명처리 기준 3가지

③ 추론 가능성(Inference)

✓ 무시할 수 없는 확률로 다른 속성 집합의 값(다른 정보 또는 분석자의 개인 지식)에서 속성의 값을 추론하여 개인을 식별할 가능성

왜 더 엄격한 기준이 필요한가?

데이터 자체의 재식별 위험이 높아 **123**에 살고있는 40-45세, 키 160-169, 몸무게 61-66인 남성**은 **고혈압** 또는 **전립선염**을 앓고 있는 것을 재식별 하는 경우

우편번호	나이	성별	키	몸무게	질병
123**	[40, 45]	M	164	61	전립선염
123**	[40, 45]	M	166	63	전립선염
123**	[40, 45]	M	165	62	고혈압
123**	[40, 45]	M	167	64	고혈압
123**	[40, 45]	M	168	65	고혈압
123**	[40, 45]	M	169	66	고혈압
123**	[40, 45]	F	167	64	상기도염
123**	[40, 45]	F	162	67	갑상선암
123**	[40, 45]	F	168	65	고혈압
123**	[40, 45]	F	160	62	빈혈
123**	[40, 45]	F	169	66	고지혈증
123**	[40, 45]	F	161	65	부정맥

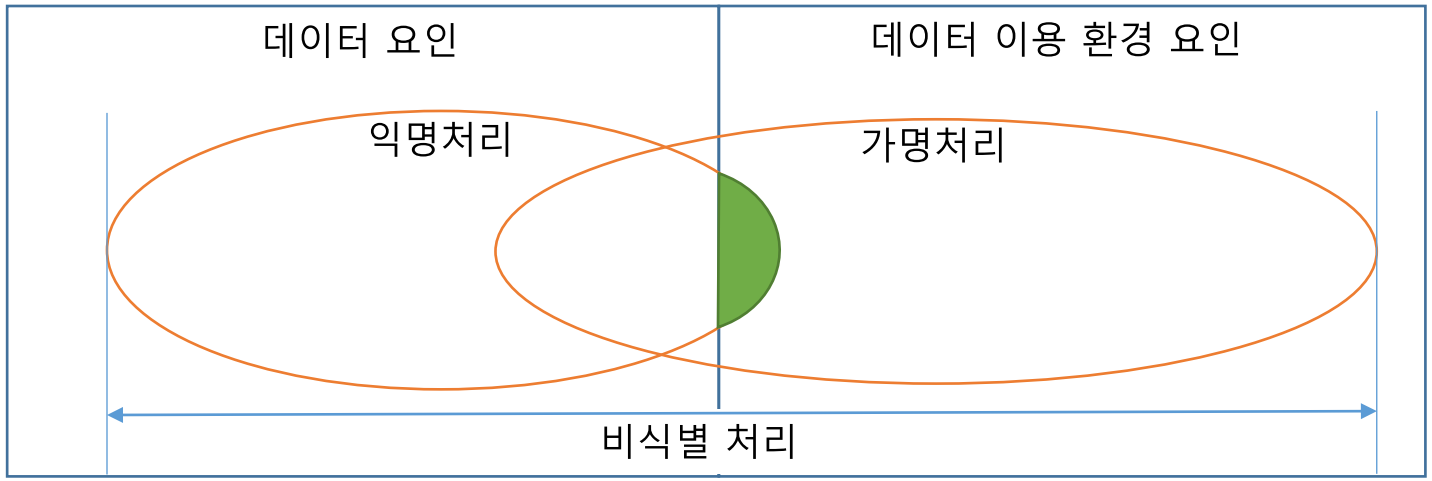
가명처리와 익명처리의 차이점

익명처리

-익명처리

- ✓ **이용환경에 대한 통제가 불가능**하여 **데이터의 위험에 따라 데이터를 비식별 처리**
- ✓ 데이터 이용환경의 통제가 가능한 경우 이용환경의 안전도에 따라 처리 수준을 줄일 수 있음

-국내에서는 법적으로는 허용되지 않음, 별도의 계약을 통해서만 가능(연두색 부분)
(문제 발생 시 법적 책임이 따름)



가명처리와 익명처리의 차이점

익명처리

가명처리

- 데이터의 이용 환경에 대한 제어를 통해 재식별 위험을 통제
- 데이터의 처리 수준에 대해서는 데이터 이용환경에 대한 위험분석과 데이터 자체의 위험을 분석하여 적정 수준을 결정
- 재식별 위험은 실제 특정인의 식별에 대한 위험을 기준으로 판단
- 이에 따라 적정성 검토는 이용환경에 대한 통제와 이에 따른 데이터의 처리 수준을 검토하는 형식으로 적용

익명처리

- 데이터 자체의 익명 처리를 통해 재식별 위험을 통제
- 재식별 위험은 식별이 아닌 특정되는 것을 기준으로 검토해야 하며 이에 따라 식별가능성보다 훨씬 가혹한 조건을 요구
- 개인정보보호법 상 익명처리는 더 이상 개인정보가 아니게 됨으로 완전공개 수준의 처리를 요함
- 이용환경을 통한 경감은 별도의 계약을 통해서만 가능 (문제 발생 시 법적 책임이 따름)

단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

단계	세부 절차	신청기관	UPS	비고
1. 사전 미팅 및 컨설팅 신청	결합 컨설팅 신청서 접수			
	컨설팅 기본 사항 교육			
	컨설팅 요구사항 협의			
2. 제도 및 법률 요구사항 검토 및 지원	관련 법률 1차 검토			
	관련 법제에 의한 처리 절차 안내			
	가명정보 내부 관리계획 수립			
	적합성 검토			
	개인정보 처리 방침 수립			
3. 내부 처리 절차 수립 지원 (Optional)	가명정보 운영 규정 수립			
	적정성 검토 위원회 운영규정 수립			
	적정성 검토 위원회 운영 매뉴얼 수립			
	가명정보 처리 매뉴얼 수립			

단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

단계	세부 절차	신청기관	UPS	비고
4. 결합 대상 데이터 분석 및 전처리	데이터 구조 분석			
	개인정보 속성 분석			
	데이터 타입 분석			
	목적 컬럼 설정			
	데이터 분포 분석			
	기타 데이터 분석			
5. 위험도 분석	데이터 활용 형태 위험성 검토			
	데이터 이용환경 위험성 검토			
	데이터 위험성 검토			
	위험성 검토 보고서 작성 지원			
6. 가명처리 지원	가명처리 수준 정의			
	목적 달성 가능성 협의			
	가명처리			
7. 적정성 검토 기초자료 작성 지원	적정성 검토 위원회 운영 매뉴얼 수립			
	가명정보 처리 매뉴얼 수립			

단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

단계	세부 절차	신청기관	UPS	비고
7. 적정성 검토 기초자료 작성	적정성 검토 대응 방안 수립			
	적정성 검토 관련 문서 검토			
	적정성 검토 기초자료 작성			
8. 적정성 검토	적정성 검토 위원회 구성			
	적정성 검토 대응			
9. 가명정보 결합	결합 신청			
	추가 서비스 필요 여부 검토(모의결합, 추출)			
	결합키 및 일련번호 생성			
	결합키 관리기관 전송			
	결합대상정보 결합 전문기관 전송			
	결합			
	추가 가명처리 수행 지원(필요시)			
	반출 심사 자료 작성 지원			
	반출 신청서 작성			
	반출심사위원회 대응			

단계별 컨설팅 상세 업무

단계별 컨설팅 상세 업무

단계	세부 절차	신청기관	UPS	비고
10. 사후관리	사후관리 교육 시행			
	재식별 모니터링 절차 및 방법 교육			
	컨설팅 종료 보고서 작성			

행정데이터 활용을 위한

가명정보처리의 이해

청년종합연구 II :

정책소외계층 청년 실태 및 정책개발

한국청소년정책연구원 개인정보보호위원회 KISA 한국인터넷진흥원 UPS DATA