

데이터 3법 이해

대한민국 법률체계

- 성문법
 - 헌법: 최상위 법
 - 법률: 국회의 의결을 거친
 - 명령: 시행령 -> 대통령령, 시행규칙 -> 부령, 총리령
 - 조례/규칙: 시도 조례, 시군구 조례
- 불문법
 - 관습법: 사회에서 인정되는 일정한 관습
 - 도덕적 요소 -> 노인에 대한 공경 등
 - 판례: 법원의 재판을 통해 형성됨
- 특별법 > 일반법
 - 특별법이 더 우선

법률 확인

- www.law.go.kr
- 국가법령정보센터 앱

정보보호/개인정보 관련 주요 법률체계

- 국가정보원 법
- 전자정부법 및 하위 규정
- 국가정보화 기본법 및 하위 규정
- 정보통신기반 보호법 및 하위 규정
 - 지침도 있음
 - 과기부 소관 지침, 행안부 기반 지침
 - 주요정보통신기반시설 취약점 분석평가기준
- 항공 보안법
- 정보통신망법 및 하위 규정
- 클라우드 발전 법
- 개인정보 보호법 및 하위 규정
- 전자서명법 및 하위 규정
- 신용정보보호법

데이터 관련 법

주무부처가 개인정보보호위원회, 과학기술정보통신부, 행정안전부, 산업통상자원부, 금융위원회로 되어 있는 '데이터' 관련 법이 무엇이 있는지 조사

- 데이터 3법
 - 개인정보 보호법
 - 신용정보의 이용 및 보호에 관한 법률
 - 정보통신망 이용촉진 및 정보보호 등에 관한 법률

- 데이터 이용 산업 활성화
- 과기부
 - 데이터 산업진흥 및 이용촉진에 관한 기본법
- 행안부
 - 공공데이터 제공 및 이용 활성화 법률
- 산업통상자원부
 - DX 관련 법

연도별

- ~2018 공공 데이터 전략 위원회
- 2020 가명정보처리 가이드라인
- 2021
 - 4차위 데이터특위 -> 산학연
 - 마이데이터 데이터 표준화 추진체계
 - EU GDPR 적정성 결정
 - 데이터산업법 제정
- 2022

데이터 산업진흥 및 이용 촉진에 관한 기본법 (데이터산업법)

- 데이터에 관하여 다른 법률에 특별한 규정이 있는 경우 제외 이 법으로 따름

산업 디저털 전환법

개인정보 관련 사항은 **개인정보 보호법**을 따름

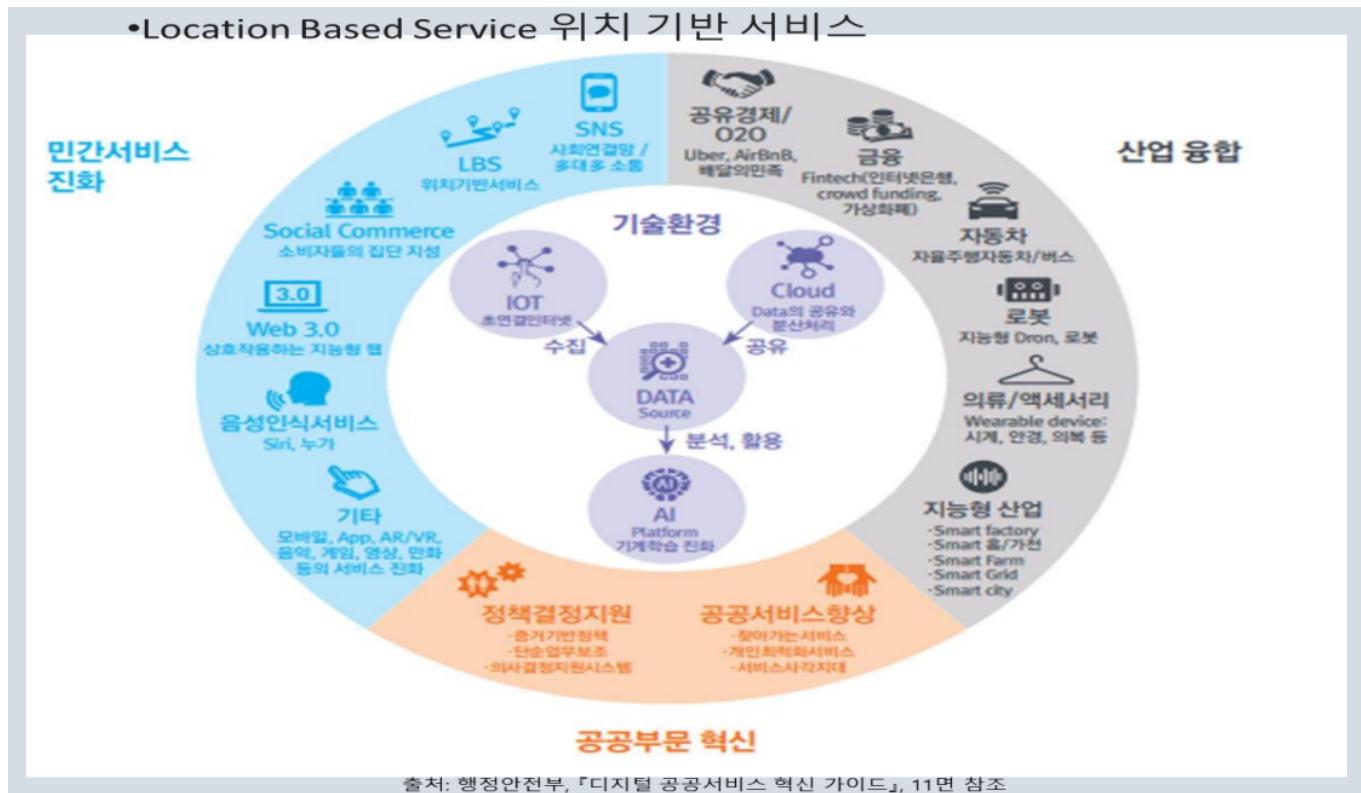
데이터 관련 법 제도 변화

1. 데이터 3법 개정
2. 데이터기반행정법 제정
3. 데이터산업법 제정
법을 보호하는 것이 아닌 산업화를 위해 개방
4. 산업디지털전환법 제정

데이터 3법 개정 (개인정보보호법 중심)

1. 법률
 2. 시행령 법이 나오고 각부처들이 실질적으로 시행해야 하는 것
 3. 고시 국가기관이나 공공 영역은 지켜야 함, 민간영역은 X
 4. 가이드 가이드를 중용하면 좋다는 것임, 안해도 처벌 X
- 비식별 조치 -> 가명정보 처리

개인정보보호법 개정 배경



전 세계적인 디지털 대전환 추세에 부합하는 법이 필요!

- 개인정보를 처리 및 관리하는 환경 마련 필요
 - 금융 -> 마이데이터
 - 위치기반 서비스
 - 가상화폐, Web3.0
 - AI 기반 -> 자율주행

정부서비스 모델

전통적 가치사슬 모델



- 기업이 공급-제조-유통 등 가치사슬 전 과정에 주체로 참여해 제품 또는 서비스를 생산
- 투자 대비 성과 극대화를 위한 규모의 경제에 초점 (공급자 중심 경제)
- 전통적 가치사슬 모델

- 디지털 플랫폼 비즈니스 모델



- 디지털 플랫폼을 기반으로 내외부 생산자가 제품, 서비스로 가치 제공
- 더 많은 참여가 가치를 증대시키고 다시 더 많은 참여를 유인하는 순환 생태계 (수요자 중심 경제)

개인정보보호법 개정 배경

- 데이터 생태계 조성과 개인정보 보호
- 기본 원칙
 - 민관과 함께 성장하는 혁신생태계 조성
 - 공공 데이터를 네거티브 원칙 하에 디지털 방식으로 전면 개방
 - 네거티브 규제: 법률로 금지된 것이 아니면 모두 허용하겠다
 - 정부는 인공지능, 데이터 기반으로 정책결정을 과학화
 - 마이데이터 -> 데이터와 서비스의 민관 공유를 위한 개방형 표준

디지털 대전환과 웹 환경 변화

- Web 1.0: 게시, 관리, 전송

- Web 2.0: 참여, 공유, 개방
 - 개인정보 처리자 중심
- Web 3.0: 정보주체 중심 (가명/익명정보 활용, 마이데이터)

데이터 보호

- 데이터 저장: 데이터 저장시 보호/보안 (DB암호화, 접근제어/인증) 법정책 존재
- 데이터 전송: 데이터 전송시 보호/보안 (HTTPS/VPN등 채널 암호화) 관련 법정책 존재
- 데이터 사용/활용 -> 최근에 활성화됨
 - 데이터 분석/공유/결합 및 이용 시 데이터 보호
 - 데이터3법과 같이 데이터를 활용하는 법정책이 생김
 - 이전까지는 보호하는 법들이 많았음

개인정보보호법 개정 배경

- 20년도 사회조사 -> 국민 54.7%가 우리사회가 개인정보 유출에 불안하다고 응답
 - n번방 사건 -> 공무원 계정으로 피해자 신상정보 유출
 - 자원봉사자명단 유출, 송파 살인 사건 등

개인정보 침해의 개념

개인정보 처리 전 과정에서 발생, 법적 근거 없는 개인정보 수집/이용/제공, 개인정보 유출/오남용 등을 모두 포괄

개인정보 보호법 목적

개인정보 처리 및 보호에 관한 사항을 정함으로써 개인의 자유와 권리를 보호하고, 나아가 개인의 존엄과 가치를 구현

- 제정 목적: 자연인에 관한 기본권과 자유
- 일반법 -> 짧음, 특별법 -> 제목이 길다 (~~에 관한 법률)
- GDPR: 개인 정보의 처리와 이전에 관한 사항을 규정한 EU 규칙
- 개인정보자기결정권의 제한?

정보주체의 실질적 권리 보장 vs 개인정보 활용

- 정보주체와 개인정보를 활용하는 사람 사이의 이견이 있음

개인정보보호법 만든 목적

- 데이터 경제 성장 견인
 - 마이데이터
- 개인정보 신뢰사회 구축
- 글로벌 스탠다드 선도
 - 국외이전 요건 다양화

개인정보보호법 개정 주요사항

1. 개인정보 보호체계 일원화
2. 데이터 이용 활성화

- 가명정보 도입

3. 개인정보처리자 책임성 강화

개인정보법 개정 주요 사항 (23.03.14)

- 사적 이용 제한

신기술 신산업 등 데이터 경제 성장 견인

- 개인정보 전송요구권 신설
- 이동형 영상정보처리기기 규정
- 온오프라인 규제 일원화
 - 오프라인에서 종이로 적는 것도 온라인과 동일하게 규제

글로벌 스탠다드에 부합하는 법 제도 정비

- EU의 GDPR을 주로 근거
- 담당자에 대한 형벌 중심에서 경제적인 재재 중심으로 전환
 - 담당자만 처벌하면 됐었음

디지털 신대에 적합한 국민의 적극적 권리 강화

- 개인정보 처리요건 정비
- 개인정보 처리방침 평가제도입
 - 처리방침의 적정성 여부, 알기 쉽게 작성했는지 여부
- 자동화된 결정 대응권
- 개인정보 분쟁조정
- 사적 목적 이용 금지

개인정보 보호법 개정 사항

- 영상정보처리기기

7. "영상정보처리기기"란 일정한 공간에 지속적으로 설치되어

어 사람 또는 사물의 영상 등을 촬영하거나 이를 유·무선망을 통하여 전송하는 장치로서 대통령령으로 정하는 장치를 말한다.

8. "과학적 연구"란 기술의 개발과 실증, 기초연구, 응용연구 및 민간 투자 연구 등 과학적 방법을 적용하는 연구를 말한다.

7. "고정형" 영상정보처리기기"란 일정한 공간에 설치되어 지속적으로 주기적으로

사람 또는 사물의 영상 등을 촬영하거나 이를 유·무선망을 통하여 전송하는 장치로서 대통령령으로 정하는 장치를 말한다.

7의2. "이동형 영상정보처리기기"란 사람이 신체에 착용 또는 휴대하거나 이동 가능한 물체에 부착 또는 거치(据置)하여 사람 또는 사물의 영상 등을 촬영하거나 이를 유·무선망을 통하여 전송하는 장치로서 대통령령으로 정하는 장치를 말한다.

8. "과학적 연구"란 기술의 개발과 실증, 기초연구, 응용연구 및 민간 투자 연구 등 과학적 방법을 적용하는 연구를 말한다.

- 이동형(드론)과 고정형을 분리

⑦ 정보주체는 제1항 및 제2항에
⑧ 제1항부터 제7항까지에서 규정
전송 요구의 거절 및 전송 중단의
[본조신설 2023. 3. 14.]
[시행일 미지정] 제35조의2

- 개인정보의 전송 요구권 (제35조의 2)
 - 시행 미지정
 - 기반이 잡힌 다음에 시행이 될 거 같음

시행일 (2024)

- 자동화된 결정에 대한 정보주체의 권리
- 개인정보관리 전문기관 등

개인정보보호위원회 격상

개인 정보 보호 관련 업무를 **개인정보보호위원회**로 통합 및 이관

- 원래는 행정안전부, 방송통신위원회, 개인정보보호위원회가 분산

데이터 이용 활성화

- 개인정보의 이용/제공 범위 확대
- 가명정보의 이용/제공 범위 확대
 - 통계작성, 과학적 연구, 공익적 기록보존에 확대됨
 - 가명정보 결합에 대한 법정 근거 명시
 - 다른 기업간 **가명정보 결합은 전문기관에서 수행**
 - 결합 정보를 반출할 경우 가명 및 익명 조치
- 결합정보 반출 승인 기준
 - 결합 목적과 반출 정보가 관련성이 있을 것
 - 특정 개인이 알아볼 수 없을 것
 - 반출 정보에 대한 안전조치

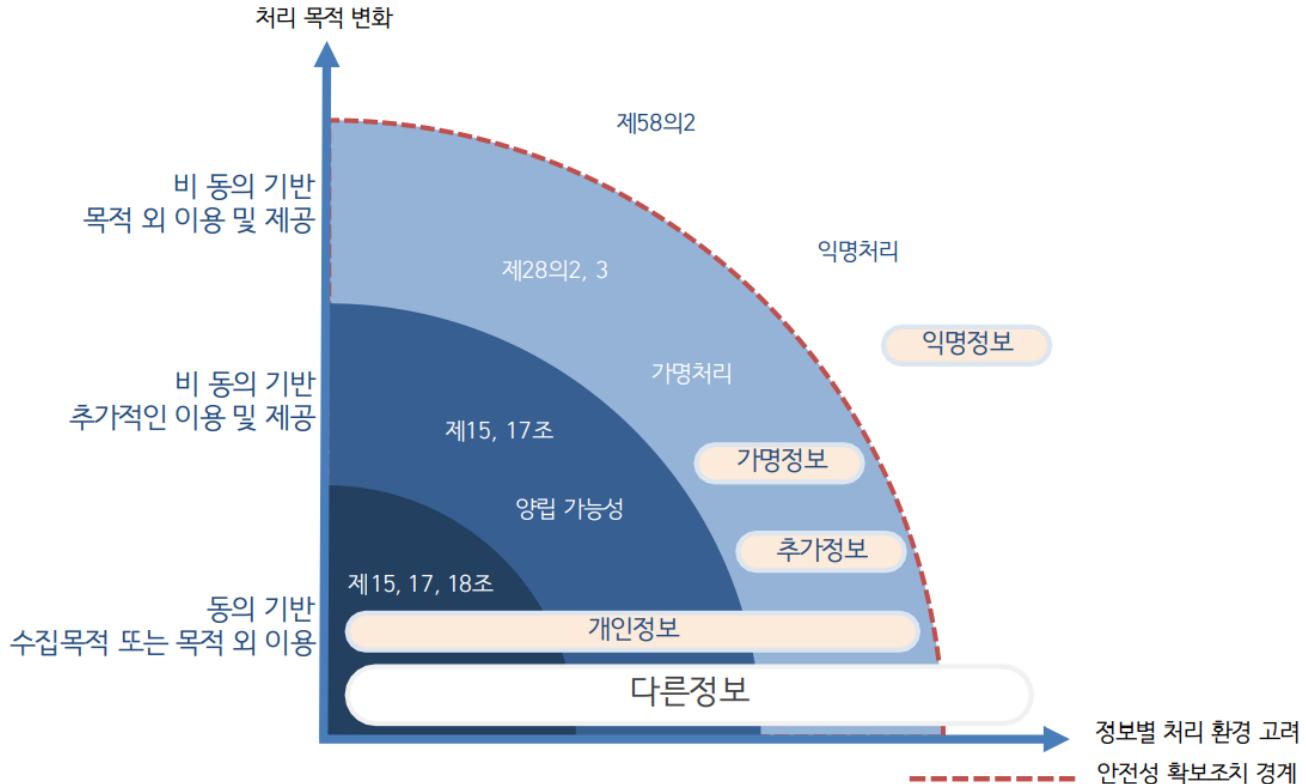
GDPR 적정성 결정 전후 비교

- 역외 이전 시 추가적 보호조치 불필요
- 준거 법률: GDPR, 개인정보보호법 -> 개인정보보호법 + 고시
- 감독기구 EU DPA + 개인정보보호위원회 -> 개인정보보호위원회
- 규제 준수 부담 낮아짐

가명정보의 정의

- 개인이 식별하 수 없도록 처리된 정보
- 가명처리된 정보는 개인정보처리자가 준수해야 할 요건 절차가 적용 X
- 가명처리된 정보가 다른 정보와 결합하여 특정 개인을 식별할 수 있으면 -> 개인정보로 취급

정보 정의



- 동의 기반 수집목적 또는 목적 외 이용
 - 개인정보, 다른정보
- 비동의 기반 추가적인 이용 및 제공
 - 가명정보
- 비동의 기반 목적 외 이용 및 제공

개인정보 판단 기준



제2조(정의) 이 법에서 사용하는 용어의 뜻은 다음과 같다. <개정 2014. 3. 24., 2020. 2. 4., 2023. 3. 14.>

1. “개인정보”란 살아 있는 개인에 관한 정보로서 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보

- 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보

— 중략 —

가. 개인을 알아볼 수 있는 정보

나. 쉽게 결합하여 알아볼 수 있는 정보

다. 가명처리한 가명정보

익명정보(58조의 2)

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

가. 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보

나. 해당 정보만으로는 특정 개인을 알아볼 수 없더라도 다른 정보와 쉽게 결합하여 알아볼 수 있는 정보. 이 경우 쉽게 결합할 수 있는지 여부는 다른 정보의 입수 가능성 등 개인을 알아보는 데 소요되는 시간, 비용, 기술 등을 합리적으로 고려하여야 한다.

다. 가명 또는 나름을 제1호의2에 따라 가명처리함으로써 원래의 상태로 복원하기 위한 추가 정보의 사용 · 결합 없이는 특정 개인을 알아볼 수 없는 정보(이하 “가명정보”라 한다)

1의2. “가명처리”란 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 등의 방법으로 추가 정보가 없이는 특정 개인을 알아볼 수 없도록 처리하는 것을 말한다.

- 개인정보 판단 기준 -> 쉽게 결합하여 알 수 있는지 여부
- 익명정보 -> 일반정보로 분류

개인정보 가명정보 익명정보

- 잘못된 가명정보는 식별이 가능하다
 - 마산에 국회의원 -> 식별 가능

잘못된 가명정보							
이름	생년월일	핸드폰	주소	직업	가족	예금평균잔액	대출액
		wd4e85d2z4qe	창원시 마산회원구	국회의원	배우자, 아들1, 딸1	4,567,900원	50,984,234원
삭제	삭제	암호화	범주화	개인식별가능			

- 제대로된 가명정보 -> 직업 삭제

가명정보							
이름	생년월일	핸드폰	주소	직업	가족	예금평균잔액	대출액
		wd4e85d2z4qe	창원시 마산회원구		배우자, 아들1, 딸1	4,567,900원	50,984,234원
삭제	삭제	암호화	범주화	삭제			

- 익명정보

익명정보							
이름	생년월일	핸드폰	주소	직업	가족	예금평균잔액	대출액
			창원시 마산회원구		가족3	400~500만원	4천~5천만원
삭제	삭제	암호화	범주화	삭제	일반화	범주화	범주화

- 범주화를 많이 한 것 (뭉뚱그린 것) -> 금액 400만원으로 범주화
- 통계 작성시 많이 활용

가명정보 vs 익명정보

- 취급자와 다른 정보에 의해 식별 가능성 존재

가명정보							
이름	생년월일	핸드폰	주소	직업	가족	예금평균잔액	대출액
		wd4e85d2z4qe	창원시 마산회원구		배우자, 아들1, 딸1	4,567,900원	50,984,234원

- 어떤 취급자와 다른 정보를 이용하더라도 식별 불가능

익명정보							
이름	생년월일	핸드폰	주소	직업	가족	예금평균잔액	대출액
			창원시 마산회원구		가족3	400~500만원	4천~5천만원
삭제	삭제	암호화	범주화	삭제	일반화	범주화	범주화

익명정보의 식별 가능성

- 통계에 의해서 개인정보가 유출될 수 있음
 - 특정할 수 있으면
- 빅데이터 개인정보 노출 사례
 - 유명인 개인정보 노출
 - 일론 머스크 비행기 경로 공개
- 익명정보도 어떻게 하면 재식별이 가능함

가명정보처리

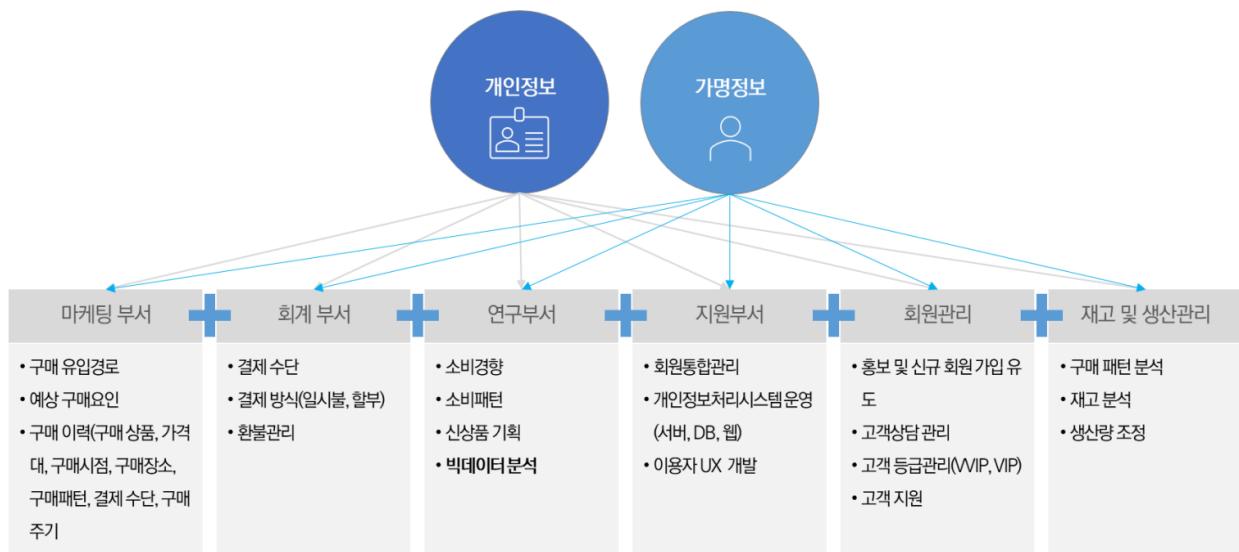
- 가명정보 처리 목적
통계작성, 과학적 연구 등 **정보주체의 동의 없이 가명 처리 가능**

가명정보처리 개정사항

- 가명정보 제공
제 3자에게 제공하는 경우 **개인을 특정하는 정보**를 포함해서는 안됨
- 가명정보 결합
 - 서로 다른 개인정보처리자 간의 결합시 -> **전문기관을 통해서만** 결합 가능
- 개인정보의 파기 (개정 후 정보)
 - 가명정보의 처리 기간 경과 등 개인정보가 불필요하게 될 때는 파기해야 함
 - 다른 법령이 있을 때는 파기하지 않음
- 가명정보에 대한 안전조치의무 (개정후)
 - 가명정보의 처리 기간을 별도로 정할 수 있음
 - 가명정보를 파기한 경우 파기한 날부터 3년 이상 보관해야 함
- 가명정보처리기 금지의무 등 (개정후)
 - 가명정보를 처리하는 자는 특정 개인을 알아내기 위해 가명정보를 처리하면 안됨
- 적용범위
 - 제20조의 이용 제공 내역통지, 제34조의 2는 적용범위에서 빠짐

개인정보 이용 및 활용 원칙 변화

- 전통적 활용
 - 마케팅 부서
 - 구매 이력, 패턴
 - 회계 부서
 - 결제
 - 등등
- 가명정보의 활용



- 가명처리만 하지 동일함

- 다른 산업군의 가명처리한 가명정보를 결합하여, 새로운 Insight 형성, 서비스 개발에 사용



가명정보 활용 우수 사례

구분	등급	시상주체	내용	소속(팀명)
활용사례 부분(13)	대상(4)	보건복지부 장관상	암환자의 합병증 및 만성질환 예측연구	국립암센터
		과학기술정보통신부 장관상	편의점 AI 상품추천모델 구축	(주)이마트24
		개인정보보호위원회 위원장상	수요자별 맞춤형 금융지원정책 수립	(주)국민행복기금
		금융위원회 위원장상	중신용/싼파일 고객 신용평가모형 개발	카카오뱅크
	우수상(5)	국민건강보험공단 이사장상	노후소득보장 종합연구	보건복지부 사회보장위원회 사무국
		금융결제원 원장상	코로나 전후 소비투자 관점의 데이터 결합 및 분석	(주)미래에셋증권, (주)롯데카드
		금융보안원 원장상	금융소외계층을 위한 머신러닝 모형 개발	하나은행
		신용정보원 원장상	통신정보를 활용한 사회적 취약층 대상 중금리 대출 지원	(주)케이티
		한국지능정보사회진흥원 원장상	화물차 개인사업자 신용도 평가모델 연구	한국도로공사
	시범사례(4)	폐암환자의 사망동향 및 사망예측 연구	국립암센터	
		맞춤형 산림치유 프로그램 분석	국립산림과학원	
		불법스팸 실태연구	한국인터넷진흥원	
		지역별, 상권별, 상품별 소비패턴 분석	(주)케이티, 롯데멤버스	

사례 - 노후소득보장 종합연구

여러 부처의 데이터를 결합해서

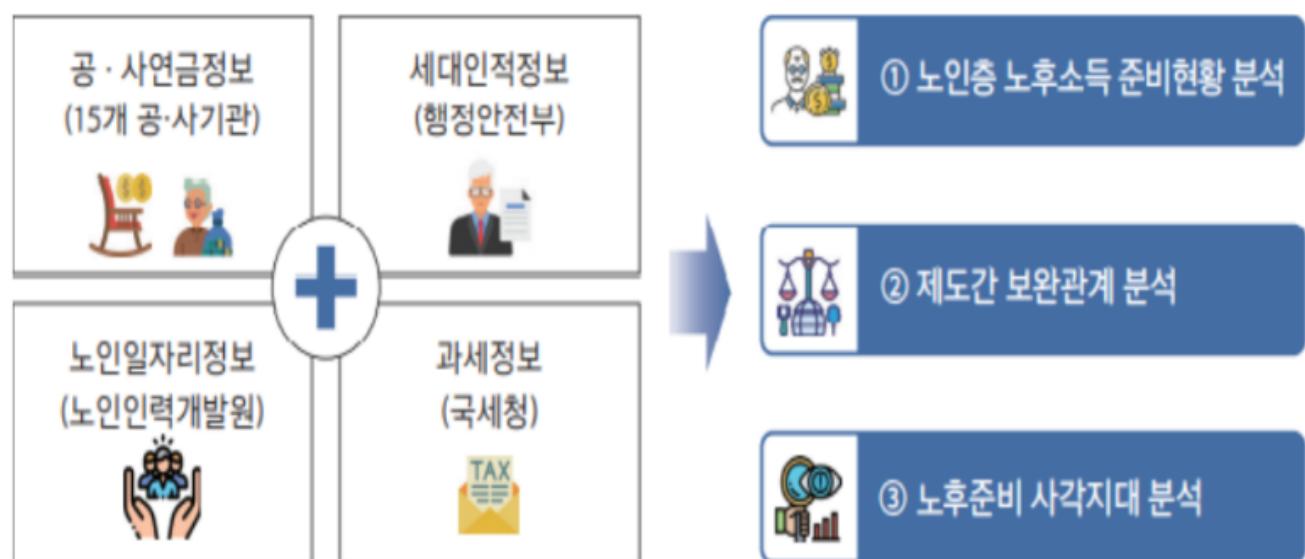
1. 결합데이터

분류	데이터보유기관	데이터 개요
공공	행정안전부	세대·인적 정보
	국세청	과제정보
	국민연금공단	국민연금 정보
	한국사회보장정보원	기초연금 정보
	공무원연금공단	공무원연금 정보
	국군재정관리단	군인연금 정보
	국방전산정보원	
	별정우체국연금관리단	별정우체국연금 정보
	사립학교교직원연금공단	사립학교교직원연금 정보
	한국주택금융공사	주택연금 정보
	한국농어촌공사	농시연금 정보
민간	한국노인인력개발원	노인일자리 정보
	국민·기업·농협·신한·우리·하나은행	퇴직연금 정보

분석결과를 도출함

2. 분석결과 및 향후 계획

- 65세 이상 노인층의 연금수급률은 47%, 연평균 수급액 710만원으로 분석됨
- 65세 이상 남성의 연금수급률 및 평균 연간수급액이 여성 대비 약 2배로 나타남
- 연령이 높아질수록, 연금수급액이 낮을수록 기초연금 수급률이 증가함
- 해당 분석결과를 기반으로 노후소득보장제도 간 보완관계를 분석하고, 저소득층 노인의 소득보장실태 파악 및 사각지대 분석 등 보다 심층적인 분석을 수행할 예정임



사례 - 친환경차 충전 최적입지 선정

성남시 정보 + 티맵모빌리티



사례 - 폐암환자의 사망동향 및 사망예측 연구

접근이 어려운 개인정보 사례



- 국립암센터: 폐암 환자의 진단, 검사, 병리, 수술, 항암제 등 임상 정보(02년~19년, 2만명)
- 보험공단: 폐암 환자의 건강검진, 상병내역, 요양기관 등 요양 급여 청구 정보(11년~19년, 2만명)
- 통계청: 사망연월일 및 사망원인 103항목 분류 정보(04년~19년, 423만명)

사례 - 비금융정보를 활용한 대안신용평가모형 개발

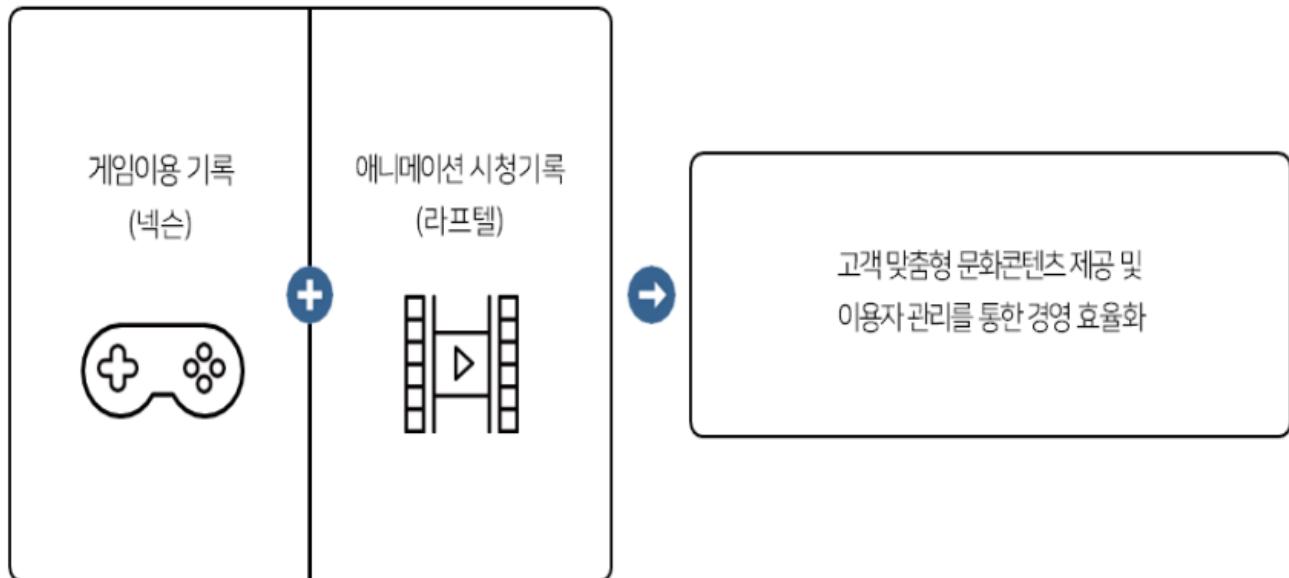
1. 결합데이터



데이터보유기관	데이터 개요	레코드 수
카카오뱅크	카카오뱅크 대출정보, 은행이용실적 정보	약 3천 7백만 건
코리아크레딧뷰로	카카오뱅크 조회 고객, 전국민 CB데이터	약 1억 2천만 건
NICE평가정보	카카오뱅크 조회 고객, 전국민 CB데이터	약 1억 6백만 건
카카오모빌리티	카카오택시, 대리 사용자 호출 및 결제 현황	약 3억 건
카카오페이	카카오페이 이용자 사용성 지표	약 2억 5천만 건
카카오	결합키별 카카오 선물하기 이용관련 데이터	약 4억 2천 2백만 건
금융결제원	자동이체 요약정보	약 3억 5천 1백만 건
다날	다날 소액결제 스코어 및 요약항목	약 5백만 건
롯데멤버스	롯데멤버스 통합등급 및 요약항목	약 4천 9백만 건
교보문고	교보문고 통합등급 및 요약항목	약 1천 2백만 건
LG유플러스	통신 이용정보	약 1억 2천 3백만 건

사례 - 게임 & 애니메이션 유저의 소비컨텐츠 분석을 통한 비즈니스 인사이트 생성

1. 결합데이터



데이터 보유기관	데이터 개요	레코드 수
넥슨코리아	온라인/모바일 게임 이용 기록	약 108억 건
라프텔	애니메이션 시청 기록	약 8억 건

실습 - 자료조사

- 가명정보 활용 사례 조사
- 특정 기관의 개인정보 처리방침에서 가명정보 처리 기록 공개 확인
<https://privacy.sktelecom.com/view.do?ctg=policy&name=policy>

15. 가명정보의 처리

SK텔레콤은 수집한 개인정보를 특정 개인을 알아볼 수 없도록 가명처리하여 통계 작성, 과학적 연구, 공익적 기록 보존 등을 위해 활용할 수 있습니다. 가명정보 처리 및 위탁, 제3자 제공에 대한 내용은 아래와 같습니다.

신한 카드가 인증을 하는 전문기관이겠음

가명처리/제공목적	정보의 종류, 항목	보유/이용기간	제공받는 자	위탁 업무
문화/예술행사 관람자 특성 분석	2022년 고객 사용 현황 (연령/성별, 추정 거주지, 요금제/업체, 라이프스타일/관심도 관련 추정 정보, 방문지)	목적 달성을 후 1개월 내	예술경영지원센터	가명처리 (SK C&C)
				가명정보 결합 (신한카드)

건강보험공단 개인정보처리방침 https://www.nhis.or.kr/nhis/etc/privacy_policy.do SK와 비교해서 명확

하게 적혀있지 않음 그리고 보유기간을 준영구로 함

12 가명정보 처리에 관한 사항



① 공단은 통계작성, 과학적 연구, 공익적 기록보전 등의 목적으로 가명정보를 처리하고 있습니다. 공단이 처리하는 가명정보에 관한 사항은 다음과 같습니다.

가명정보 처리 업무명	처리목적	처리항목	보유 및 이용기간
국민건강정보DB제공	통계작성, 과학적 연구, 공익적 기록보전	자격 및 보험료, 진료상세, 건강검진, 요양기관, 장기요양 정보 등	준영구

3. 조사된 자료로 아래 항목 구성 아래 요소가 빠지면 법규 위반임

- 가명정보 처리의 목적
- 가명처리한 개인정보의 항목
- 가명정보의 이용내역
- 제3자 제공 시 제공받는 자

관련 직무 혼선

데이터 관련 업무는 빅데이터팀에서 수행했으나 데이터 3법 개정으로 혼선

AS-IS



TO-BE



가명정보 관리체계 수립

1. 개인정보 보호지침
2. 절차
3. 메뉴얼 & 양식

가명처리 절차

1. 사전준비
2. 월미성 검토
3. 가명처리 절차
4. 적정성 검토
5. 안전한 관리

법정 요구사항 준수를 위한 절차 마련

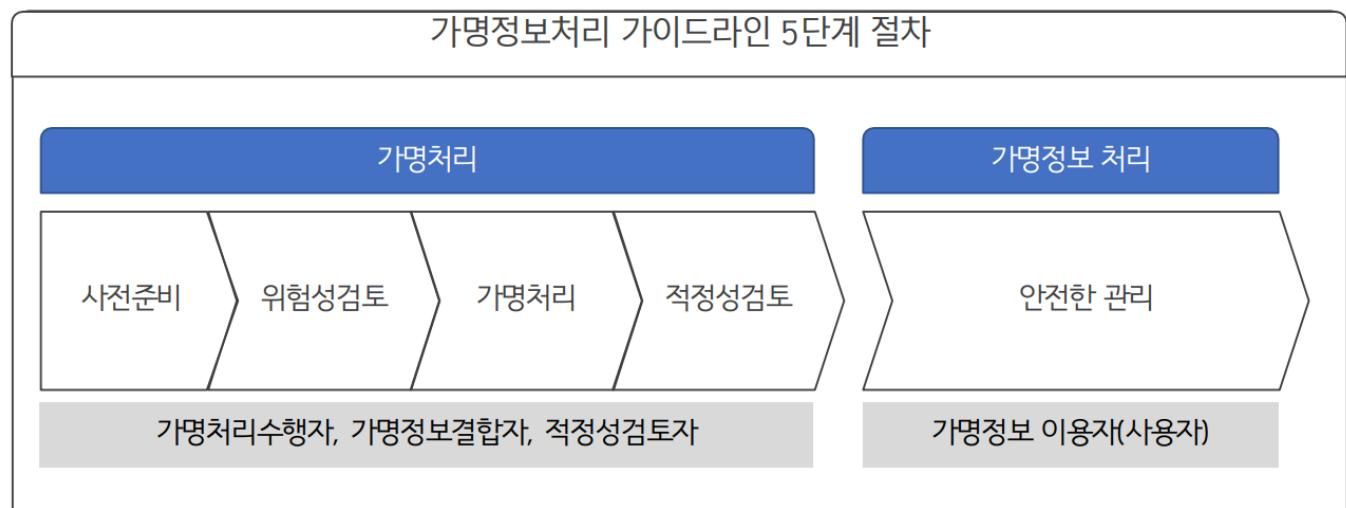
- 컴플라이언스 요구사항
 - 법령, 지침 및 고시, 가인드라인, 특화 제도

가명정보 관련 제도 및 법률

연번	구분	내용	소관부처
1	법률	개인정보보호법	개인정보보호위원회
2		신용정보의 이용 및 보호에 관한 법률	금융위원회
3	시행령	개인정보 보호법 시행령	개인정보보호위원회
4		신용정보의 이용 및 보호에 관한 법률 시행령	금융위원회
5	고시	가명정보의 결합 및 반출 등에 관한 고시	개인정보보호위원회
6		공공기관의 가명정보 결합 및 반출 등에 관한 고시	개인정보보호위원회
7		신용정보업감독규정	금융위원회
8	가이드라인	가명정보 처리 가이드라인	개인정보보호위원회
9		보건의료 데이터 활용 가이드라인	보건복지부
10		교육분야 가명·익명정보 처리 가이드라인	교육부
11		공공분야 가명정보 제공 실무안내서	행정안전부
12		금융분야 가명·익명처리 안내서	금융위원회

- 가이드라인 -> 현업에서 참고할 만한 것

가명처리와 가명정보의 처리 차이점



- 가명처리 -> 개인정보의 일부를 삭제하거나 일부 또는 전부를 대체하는 과정
- 가명정보 처리 -> 가명처리를 통해 생성된 가명정보를 활용하는 행위
- 그니깐 가명처리는 개인정보를 비식별화 하는 과정, "가명정보 처리"는 이를 필요로 하는 것

사전준비 단계

- 목적 설정 -> 가명정보를 왜 만들려고 하는지 목적을 명확히 설정
- 가명처리 대상 선정
- 처리 목적 적합성 검토

- 가명정보처리를 위한 안전조치 이행
- 필요 서류 작성

사전준비 단계 - 가명정보 이용 제공 신청서

접수번호	ZZANG-20220522001		접수일	2022년 5월 22일
신청자	조직/부서명	A 유동회사 구매부		
	담당자 직위	과장	담당자 성명	이유동
	전화번호	02-1234-5678	이메일 주소	gamyongzzang@utong.com
처리 목적	<input checked="" type="checkbox"/> 등계작성 <input type="checkbox"/> 과학적 연구 <input type="checkbox"/> 공익적 기록보존 코로나 이전과 코로나 이후의 상품군별 판매 추이에 대한 통계 작성을 위한 가명정보의 제공			
	<input type="checkbox"/> 내부이용 <input checked="" type="checkbox"/> 제3자 제공 <input type="checkbox"/> 결합전문기관을 통한 결합			
활용 형태	<input type="checkbox"/> 동일 개인정보처리자 <input checked="" type="checkbox"/> 제3자(제공받는 자) 제공			
이용 주체	<input type="checkbox"/> 내부 <input checked="" type="checkbox"/> 외부			
처리 환경	<input type="checkbox"/> 폐쇄 환경 <input type="checkbox"/> 제한 없음			
처리 장소	<input checked="" type="checkbox"/> 1회 제공 <input type="checkbox"/> 시계열 분석 등을 위한 반복 제공(회 예정)			
반복 제공 여부	<input type="checkbox"/> 온라인 <input checked="" type="checkbox"/> 오프라인			
제공 방법				
제공 받는 자	(내부) <input checked="" type="checkbox"/> 동일부서 <input type="checkbox"/> 타부서 (외부) <input type="checkbox"/> 제3자, 결합전문기관 제공			

- 목적, 활용 형태, 등등이 적혀짐

신청 명세	가명처리 대상 명칭	자사의 고객 구매정보
	데이터 내역	구매부에서 보관하고 있는 자사 고객 중 회원가입이 2019년 1월 이전인 고객의 구매정보 중 2019년 1월 ~ 2019년 12월, 2021년 1월 ~ 2021년 12월 까지의 구매정보데이터 전체 222컬럼, 전체 대상인원 820만명 중 50% 무작위 샘플링한 410만명의 정보
	가명정보 이용 기간	2022년 4월 1일 ~ 2022년 10월 31일(7개월)

위와 같이 가명정보 이용·제공을 신청합니다.

2022년 5월 22일

신청인(부서장명)

이유동(서명 또는 인)

첨부서류	1. 개인정보 유형 분류표
	2. 활용데이터 요구 수준표(선택)
3. 데이터 처리 위탁 계약서(위탁 계약의 경우) / 추후 제출	
4. 데이터 제공 계약서(제3자 제공의 경우) / 추후 제출	

- 신청 명세 -> 비식별 데이터가 있겠음
- 이용기간 -> 이용기간도 특정해야 함

위험성검토 단계

위험성 검토 주요 내용 구분	주요 내용 설명
대상선정	[1단계. 사전준비]에서 설정한 목적을 달성하기 위해 필요한 항목을 개인정보파일에서 선정함 ※ 가명처리 대상 항목 선정 시 가명정보처리 목적 달성을 위한 최소 항목으로 해야 함
위험성검토	위험성 검토는 가명처리 대상 데이터의 식별 위험성을 분석·평가하여 가명처리 방법 및 수준에 반영하기 위한 절차이며, 식별 위험성은 1) 데이터의 식별 위험성과 2) 처리 환경의 식별 위험성으로 구분하여 검토해야 함

- 대상선정: 불필요한 데이터를 넣지 말아야 겠음
- 위험성검토: 식별 위험성을 분석 및 평가를 함
 - 데이터 자체 식별 위험성 요소: 재식별 시 영향도, 식별정보 등
 - 처리 환경 식별 위험성 요소: 물리적 요소

데이터 자체 식별 위험성 검토 내용

1. 식별정보: 특정 개인과 직접적으로 연결되는 정보
2. 식별기능정보: 다른 항목과 결합되는 경우에 식별 가능성이 높아지는 항목
3. 특이정보: 전체데이터에 식별 가능성을 가지는 고유값
4. 재식별시영향도: 재식별로 불이익이 될 수 있음

처리환경 식별 위험성 검토 내용

처리환경 식별 위험성 요소	요소 설명
활용형태	<ul style="list-style-type: none"> ▪ 내부이용 또는 외부제공하는 경우 처리자가 보유하거나 접근입수가 가능한 정보, 이용 범위 및 유형 등 ▪ 보안서약서, 계약서 등을 통해 파악 가능한 범위 정보를 고려하여 식별 가능성을 검토 가능
처리장소	<ul style="list-style-type: none"> ▪ 해당 가명정보 외에 다른 정보의 접근입수가 제한된 장소에서 처리되는지 여부 ▪ 보안서약서, 계약서 등으로 내외부 활용이 제한된 경우 폐쇄 환경에 준하여 검토 가능
처리방법	<ul style="list-style-type: none"> ▪ 가명정보를 다른 정보와 연계 분석 내부 결합하는 경우 결합 후 식별 가능 한 항목 ▪ 가명정보 반복 제공 시 식별 위험이 높아지는 항목

- 활용 형태: 내부이용 또는 외부제공하는 경우 처리자가 보유하거나 이용 범위
- 처리 장소: 제한된 장소에서 처리되는지 여부
- 처리 방법: 가명정보를 반복하고 제공 시 식별 위험이 높아지는 항목

위험성검토 - 처리환경

- 가명정보를 제공받은 자가 신뢰할 만한 인증을 받았는가?
 - ISMS, ISMS-P, ISO 27000
- 처리목적
 - 가명정보 처리 목적을 선정하고 명확히 하였는가

식별 위험성 검토 결과보고서

가명정보 활용목적	<ul style="list-style-type: none"> B동계업체는 코로나 이전과 코로나 이후의 상품군별 판매 추이에 대한 통계작성을 위해 당사의 구매정보 데이터 중 2019년 1월부터 12월까지의 주요 상품군별 판매액 정보와 2021년 1월부터 12월까지의 주요 상품군별 판매액 정보를 나이와 성별, 시군구 단위의 주소별 데이터를 분석 	
가명처리 대상 데이터 항목	<ul style="list-style-type: none"> 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 820만명 중 50%를 무작위 샘플링하여 구성한 410만명에 대한 데이터 	
데이터 위험성	식별성 유무	<ul style="list-style-type: none"> '고객 ID'는 개인식별정보임 '나이', '주소', '성별'은 조합했을 때 개인의 식별이 가능한 개인식별 가능정보임
	특이정보 유무	<ul style="list-style-type: none"> 각 범주별 구매금액의 경우 특이정보로 인한 개인 식별성이 발생할 수 있음
	재식별시 영향도	<ul style="list-style-type: none"> 단순 고객의 구매데이터로 재식별 시 영향도는 크지 않을 것으로 판단됨
처리 환경 검토	이용 및 제공 형태	<ul style="list-style-type: none"> 제3자 제공 <ul style="list-style-type: none"> -데이터 제공 계약을 체결하여 데이터를 제공 -데이터 제공 계약에는 재제공 금지, 목적 달성 후 삭제, 재식별 금지 및 재식별 시 조치에 관한 사항들이 포함되어 있음 B동계업체는 개인정보(가명정보)처리시스템에 대한 ISMS-P인증을 취득하고 있음
	처리 장소	<ul style="list-style-type: none"> B동계업체에서 가명정보는 인터넷에 접근할 수 없는 차단된 별도의 분석 PC에서 분석 예정 분석PC가 있는 환경은 별도의 분석실로 내부적인 출입통제를 적용하는 것으로 파악됨
	다른 정보와의 결합 가능성	<ul style="list-style-type: none"> B동계업체는 다양한 통계를 생성하는 업체로 유사 업종에 대한 통계정보 등 결합 가능성이 있는 정보를 보유하고 있음

- 사전준비에 있는 것들을 검토를 함
- 데이터 위험성
 - 식별성 유무, 특이정보 유무, 재식별시 영향도
- 처리 환경 검토
 - 이용 및 제공 형태
 - 제3자 제공
 - 처리 장소
 - 차단된 환경 -> 해킹 위협 X
 - 다른 정보와의 결합 가능성
 - 유사 업종에 대한 통계정보가 결합될 위험성이 있음
 - 하지만 처리장소가 차단된 환경이니 갠춘
- 최종 검토의견
 - 처리 장소와 개인정보 보호 수준을 검토할 때 결합에 대한 시도는 거의 없을듯
 - 개인정보 보호 수준 -> ISMS-P니 수준이 높따!
 - 고객 ID는 개인 식별 가능성 ○ 매우 높으므로 salt값이 포함된 해시처리 기법 적용 필요
 - 주소, 나이, 주소, 구매액, 고객등급 -> 범주화 필요성 보임
 - 성별은 분석 목적으로 그대로 사용

가명처리 단계

개인정보처리자는 식별 위험성 검토 결과를 기반으로 가명정보의 활용 목적 달성을 위한 가명처리 방법 및 수준을 정하여 항목별 가명처리 계획을 설정함

개인정보처리자는 '항목별 가명처리계획'을 기반으로 가명처리를 수행함

가명처리 과정에서 생성되는 추가정보는 원칙적으로 파기하고 필요한 경우 가명정보와 분리하여 별도로 저장하여야함

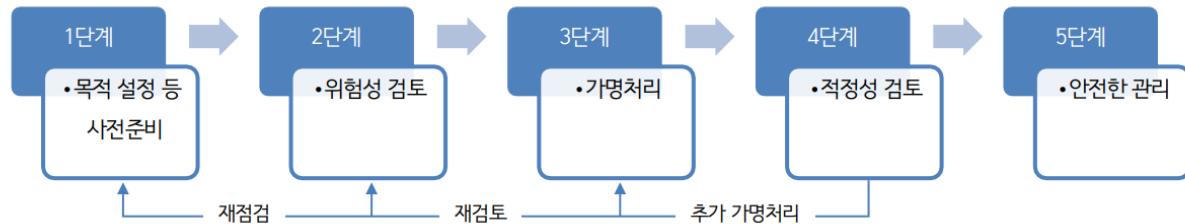
1. 식별 위험성 검토 결과를 기반으로 항목별 가명처리 계획 설정
2. 항목별 가명처리계획을 기반으로 가명처리
3. 가명처리 과정에서 생성되는 추가정보는 원칙적으로 파기 (필요한 경우 가명처리와 분리 보관)

항목별 가명처리계획 작성 예시

- 개인식별가능정보 -> 합쳐질때 식별 가능

적정성 검토

- 1,2,3 단계의 가명처리에 대한 적정성을 최종 검토



❖ 적정성 검토는 ① 필요서류, ② 처리 목적 적합성, ③ 식별 위험성, ④ 가명처리 방법 및 수준의 적정성, ⑤ 가명처리의 적정성, ⑥ 처리 목적 달성 가능성 단계로 검토가 이루어짐

❖ 적정성 검토 시 위원장을 선정하여 절차에 따라 검토를 진행할 수 있도록 하고, 종합적인 내용과 각 검토위원의 의견을 고려한 최종검토결과 및 종합검토의견을 개인정보처리자에게 제출함

적정성 검토 - 예시 1

단계	검토내용	부적정 시 조치사항
① 필요서류	사전준비 단계에서 필요서류가 법·제도 목적에 적합하게 작성되었는지 검토 * 처리 위탁 및 제3자 제공에 대한 계약서, 이용환경에 대한 보호수준을 확인할 수 있는 서류 등	해당 자료 보완 작성 등 재점검
	가명정보 이용 제공 신청서*, 가명정보 안전조치 의무이행 확약서 * 처리 위탁 및 제3자 제공에 대한 계약서, 이용환경에 대한 보호수준을 확인할 수 있는 서류 등	
② 목적 적합성 등	보호법에서 정한 가명정보 처리 목적에 해당하는지, 처리목적을 구체적으로 설정하였는지 검토	목적 명확화 및 재설정
	가명처리 및 결합 목적 증빙자료(가이드라인 참고5)	
③ 식별 위험성에 대한 결과 적정성	가명처리 단계에서 데이터의 자체 식별 위험성, 처리 환경의 식별 위험성 등 판단 항목을 누락 없이 검토하였는지 식별 위험성 검토 점검표 및 결과보고서 기반으로 검토	식별위험성 재검토, 결과보고서 보완 등
	개인정보 유형 분류표, 활용데이터 요구 수준표, 식별 위험성 검토 결과보고서, 가명정보 안전조치 의무이행 확약서	
④ 항목별 가명처리 계획의 적정성	가명처리 단계에서 위험성 검토 결과를 반영하여 항목별 가명처리 방법 및 수준을 적정하게 계획하였는지 확인	항목별 가명처리계획 보완
	식별 위험성 검토 결과보고서, 항목별 가명처리계획	
⑤ 가명처리 결과에 대한 적정성	계획한 가명처리 방법 및 수준에 따라 실제 가명처리를 수행하였는지 확인 ※ 특히 대용량 정보의 경우 중간에 처리되지 않은 부분이 있을 수 있으므로 가능한 가명정보 항목 전체를 확인 필요	가명처리가 적정하지 않은 경우 가명처리를 다시 수행하거나 부분적으로 추가 가명처리를 수행
	가명정보 처리 기초자료 명세서 ※ 원본데이터 특징, 유형, 분포 등 가명정보의 생성 및 활용 등과 관련하여 설명할 수 있는 기초자료	
⑥ 처리 결과에 대한 목적 달성 가능성	가명처리된 정보가 당초 가명정보 처리 목적을 달성할 수 있는지 여부 검토	항목별 가명처리계획 보완, 추가 가명처리 등
	가명처리 및 결합 목적 증빙자료(가이드라인 참고5)	

안전한 관리

- 재식별 금지 및 재식별 간으성 모니터링
 - 특정 개인을 알아보기 위한 목적으로 가명정보 처리 X
- 안전조치 시행
 - 개인정보처리자는 사전준비 단계에서 수립한 내부 관리계획에 따라 가명정보를 안전하게 처리
- 가명정보 처리 관련 기록 작성 및 보관
 - 개인정보처리자는 가명정보의 처리목적, 개인정보항목, 이용내역, 제3자제공시제공 받는자를 작성하여 보관하여야 함

개인정보 데이터 이해

개인정보 속성 분류

1. 고유식별자
2. 준식별자
3. 민감정보
4. 일반정보

개인정보 자료 일반적 자료 형태

- 자료
 - 범주속성 -> 자료의 범위를 지정, 요약속성과 결합하여 활용
 - 요약속성 -> 통계수치 자료
- 메타데이터 -> 자료의 생산규칙

자료 형태로 데이터 분석

- 범주형 -> 이산형/범주형
 - 명목형 -> 각 범주간 순위가 없음, 성별/인종/혈액형
 - 순위변수/척도 => 순위의 의미를 가짐, 성적/학력/경제수준
- 연속형
 - 구간변수/척도 -> 범주 사이에 등간성이 있음, 덧셈가능 곱셈불가능, 온도
 - 비율변수/척도 -> 범주 사이에 등간성 있음, 덧셈곱셈 둘다가능, 키/몸무게/나이/시간

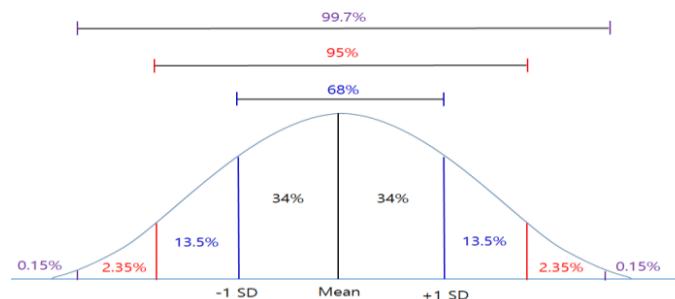
특이정보 관찰 방법

- k-익명성 - 동질 그룹 분석
- Outlier
 - 3 시그마 규칙 -> 정규분포에서 3시그마 벗어난 것
 - 사분위수 분석 -> 박스플롯 그래프를 이용
- 도수분포표 -> 빈도분석, 막대 그래프 드을 이용하여 빈도가 적은 데이터 관찰

수치형 관찰 방법

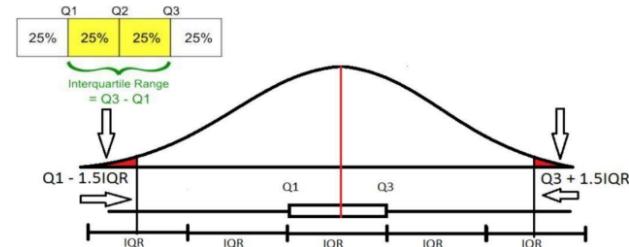
수치형 관찰 방법

❖ 3시그마 규칙(68-95-99.7 규칙) 정규분포형태의 데이터 분포에서 3시그마를 벗어나는 데이터는 특이치로 판단

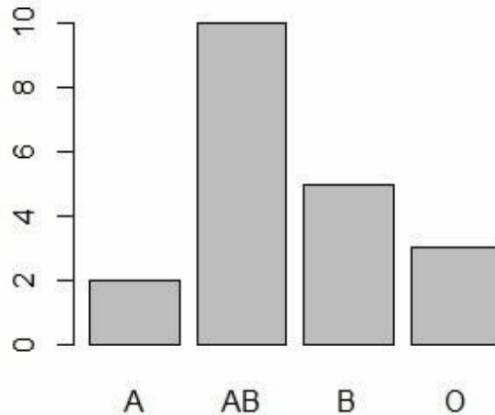


❖ 1.5*IQR(Inter-Quartile range) 방법 Q1에서 하단으로 1.5*IQR을 벗어난 값과 Q3에서 상단으로 1.5*IQR을 벗어난 값을 특이치로 판단

$$IQR = Q3 - Q1$$



범주형 관찰 방법



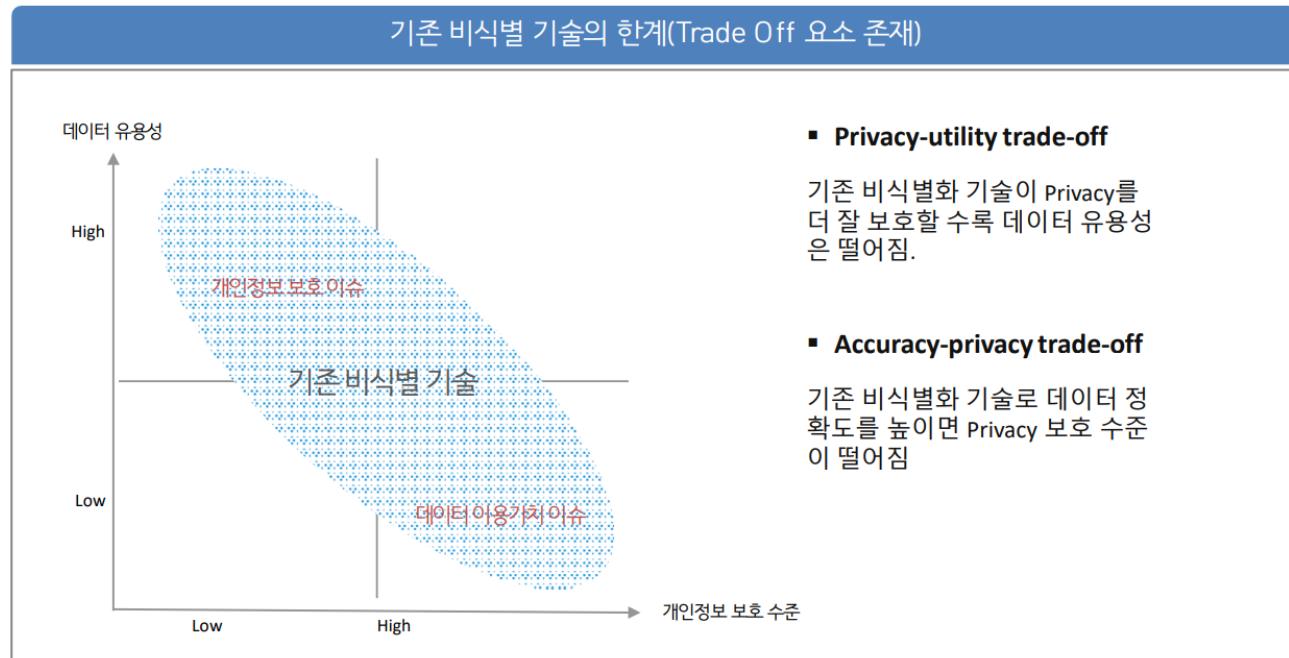
· 지역, 직업			· 나이	
주소	직업	빈도	나이(세)	빈도
경기	국회의원	5	10~20	4
경기	국회의원	5	20~30	11
강원	국회의원	1	30~40	21
경기	국회의원	5	40~50	18
경기	국회의원	5	50~60	5
경기	국회의원	5	60~70	1

※ 지역에 대한 도수분포(빈도)를 이용하여 특이정보 검토

※ 특정 나이에 도수분포(빈도)를 측정하여 특이정보 검토

가명처리 기수 이슈

가명처리 기술 이슈



- 개인정보보호 수준을 높이면 -> 데이터 이용가치가 떨어짐
- 수준을 낮추면 -> 개인정보보호이슈

가명처리 기법

- 통계 도구 기법 -> 데이터 전체 구조를 변경 (표본 추출, 총계 처리 등의 기법)
- 암호화 도구 -> 암호화 기법을 사용
- 삭제 기술 -> 레코드, 속성값 등 데이터 값을 제거하는 기술
 - 마스킹, 레코드 삭제, 컬럼 삭제 등
- 가명화 기술 -> 원래 식별자를 찾을 수 없게 대체식별자를 사용
 - 해시함수, 매팅테이블 등등
- 해부화 -> 기존 테이블을 식별자 테이블과 유용한 속성 테이블로 분할
- 일반화 기술 -> 범주화, 라운딩
 - 1억 4천 2백 50만원 -> 1억 4천만원
- 무작위화 기술 -> 속성의 값을 수정하여 원래의 값과 다르게 변형

- 잡음추가, 순열, 부분총계

가명처리 기법1

ISO/IEC 20889 기준 분류	기법 설명
통계 도구	<ul style="list-style-type: none"> ■ 데이터의 전체 구조를 변경하는 통계적 성질을 가진 기법, 비식별화 기술의 효율성을 높이기 위해 사용 ■ 표본 추출, 총계 처리 등의 기법이 포함됨
암호화 도구	<ul style="list-style-type: none"> ■ 비식별화 기술의 효율을 향상시키는 보안 대책을 시행 및 비식별화 도구의 일부로서 사용 ■ 결정성 암호화, 순서보존 암호화, 형태보존 암호화, 동형 암호화, 동형 비밀분산 등의 기법이 포함됨
삭제 기술	<ul style="list-style-type: none"> ■ 레코드, 속성값, 또는 데이터셋에서 선택된 데이터 값을 제거하는 기술 ■ 마스킹, 레코드 삭제, 컬럼 삭제 등의 기법이 포함됨
가명화 기술	<ul style="list-style-type: none"> ■ 데이터 주체의 식별자를 각 데이터 주체에 대해 특별 생성된 간접식별자로 대체하는 비식별 기술의 범주 ■ 해시함수, 매핑테이블, 임의 생성 등의 기법이 포함됨

가명처리 기법2

ISO/IEC 20889 기준 분류	기법 설명
해부화	<ul style="list-style-type: none"> ■ 기존 테이블을 식별자가 포함된 테이블과 유용한 속성을 포함하는 테이블로 분할하여 데이터로부터 식별자의 관계를 분리시키는 비식별화 기술로 각 테이블은 서로 다른 접근 권한을 부여해야 함, 해부화는 데이터를 변경시키지 않고 구조만을 변경시키는 기법임
일반화 기술	<ul style="list-style-type: none"> ■ 범주화로도 불림, 특정한 값을 상위의 속성으로 대체하는 비식별 기술 ■ 라운딩, 랜덤라운딩, 제어라운딩, 상하단 코딩, 로컬일반화 등의 기법이 포함됨
무작위화 기술	<ul style="list-style-type: none"> ■ 속성의 값을 수정하여 원래의 값과 다르게 변형되도록 하는 비식별 기술 추론 시도의 효율성을 감소시키는 기술이며 Perturbation이라고도 함 ■ 잡음추가, 순열, 부분총계 등의 기법이 포함됨

액셀을 이용한 가명처리 기법

1. 통계처리

1. 통계 처리

구분	설명	함수
최대값	범위 중 가장 큰 값	MAX
최빈값	빈도수가 높은 값	MODE.SNGL
중앙값	값들을 크기의 순서대로 정렬했을 때 가장 중앙에 위치하는 값	MEDIAN
평균값	여러 수의 합을 수의 개수로 나눈 값	AVERAGE
최소값	범위 중 가장 작은 값	MIN
백분위수값	데이터 배열 또는 범위에서 k값에 해당하는 백분위수를 나타내는 숫자	PERCENTILE
라운드	지정 실수 자릿수로 올림 또는 내림	ROUND
샘플링	지정된 범위에서 임의 값 생성	RANDBETWEEN, INDEX
몫구하기	지정된 숫자를 나누어 몫 구하기	QUOTIENT

2. 문자 처리

2. 문자 처리

구분	설명	함수
문자열합치기	여러 문자열을 하나로 합치기	CONCATENATE
왼쪽문자반환	텍스트 값에서 맨 왼쪽에서 지정된 범위까지 문자열 반환	LEFT
오른쪽문자반환	텍스트 값에서 맨 오른쪽에서 지정된 범위까지 문자열 반환	RIGHT
문자추출	텍스트 문자열의 지정한 위치로부터 지정한 개수의 문자를 표시	MID
문자수확인	텍스트 문자열의 문자 개수를 반환	LEN
문자반복	지정한 횟수만큼 문자 반복	REPT
형식변환	숫자 표시 형식을 지정하고 텍스트로 변환	TEXT
문자찾기	텍스트 값에서 다른 텍스트 값을 검색	FIND
문자열치환	문자열을 지정된 문자로 변환처리	REPLACE, SUBSTITUTE

3. 사용자 정의

3. 사용자 정의

구분	설명	함수
단방향 암호화	동일한 입력값은 동일한 결과를 반환하고 결과값으로 입력값 확인 불가	octoSHA256(변환할 텍스트) octoSHA512(변환할 텍스트)
구분자 분리	구분자에 의해 분리하고 수준별로 조합한 결과 가져오기	octoSplit(대상 텍스트, 구분자, 반환 수준 정의)
자리수 분리	텍스트 문자열의 지정한 위치로부터 지정한 개수의 문자를 표시	octoMid(대상 텍스트, 반환 시작 자리번호, 반환 문자수)
빈도 처리	지정된 빈도수 조건 대상 데이터 처리	octoFreq(빈도범위, 측정값, 대체값, 빈도기준 수)
텍스트 변경	변환 대상 텍스트의 일부라도 일치하면 전체 텍스트 변경 처리	octoFindAndReplaceAll(대상값, 검색값, 변환값)
샘플링추출	지정된 영역에서 정해진 수만큼 샘플링 추출	octoSample(대상범위, 추출수)

4. 기타

4. 기타

구분	설명	함수
비교	조건이 참일 때, 거짓일 때 분기 처리	IF
	수식이 오류로 평가하는 경우 지정한 값을 반환	IFERROR
찾기	데이터를 찾아서 원하는 데이터를 가져오기	VLOOKUP
	영역에서 지정된 행 정보 값을 반환	INDEX
	칼럼 순서 반환	COLUMN
	지정된 영역에서 지정된 값을 가지는 셀의 개수	COUNTIF

가명정보 가이드라인

3단계 가명처리

(원본정보)

소유자명	연락처	주택구분	법정동코드	시도	시군구	읍면동	지번	건물명	전세(천원)	보증금(천원)	월세(천원)	전용면적	공급면적
김철수	090-1234-5678	아파트	2635010700	서울특별시	동작구	사당동	1388-4	한글아파트	-	25,000	750	104.00	84.00
이영희	090-2468-3579	오피스텔	3611011000	대전광역시	서구	둔산동	656	나주시티오	81,250	-	-	56.45	24.32
박민호	090-9876-5432	아파트	4311410100	부산광역시	해운대구	우동	111-13	세종아파트	125,000	-	-	100.00	84.00

선정

선정

선정

- 대상선정 -> 법정동코드, 건물명 뺄
- 위험성 검토
 - 지번을 삭제 조치 -> 등기부열람 가능
 - 소유자명 연락처는 암호화

식별정보				식별가능정보									
소유자명	연락처	주택구분	시도	시군구	읍면동	지번	전세(천원)	보증금(천원)	월세(천원)	전용면적	공급면적		
김철수	090-1234-5678	아파트	서울특별시	동작구	사당동	1388-4	-	25,000	750	104.00	84.00		
이영희	090-2468-3579	오피스텔	대전광역시	서구	둔산동	656	81,250	-	-	56.45	24.32		
박민호	090-9876-5432	아파트	부산광역시	해운대구	우동	111-13	125,000	-	-	100.00	84.00		

(소유자명, 연락처)+Salt값
암호화

삭제 라운딩

- 식별정보는 salt값 + 암호화
- 식별가능정보 -> 삭제 또는 라운딩

(가명처리)

ID	주택구분	시도	시군구	읍면동	전세(천원)	보증금(천원)	월세(천원)	전용면적	공급면적
wd4e85D2C1qe89rwqe	아파트	서울특별시	동작구	사당동	-	25,000	800	104.00	84.00
r5w1e2SXzi4wd64qwz	오피스텔	대전광역시	서구	둔산동	81,300	-	-	56.45	24.32
ghe6W15Z5ax4Qe24jx	아파트	부산광역시	해운대구	우동	125,000	-	-	100.00	84.00

안전한 관리

1. 재식별 금지 및 재식별 가능성 모니터링

1. 재식별 금지 및 재식별 가능성 모니터링

- 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 자는 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 되며(보호법 제28조의5 제1항), 가명정보 처리 과정에서 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 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 자는 특정 개인을 알아보기 위한 목적으로 가명정보를 처리해서는 아니 된다.
② 개인정보처리자는 제28조의2 또는 제28조의3에 따라 가명정보를 처리하는 과정에서 특정 개인을 알아볼 수 있는 정보가 생성된 경우에는 즉시 해당 정보의 처리를 중지하고, 지체 없이 회수·파기하여야 한다.

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

※ 가명처리 기술의 취약점으로 인한 재식별 가능성 및 다른 정보와 결합 시 재식별가능성이 있는 새로운 공개데이터의 발생 여부

가명처리를 해도 가명정보가 합쳐져서 정보유출이 될 수 있음

2. 비정형데이터 가명처리 기준

IT 신기술과 개인정보보호

생체정보와 개인정보보호

- 생체정보 정의
 - 지문, 얼굴, 흥채, 정맥, 음성 등 개인의 신체적, 생리적, 행동적 특징
 - 특정 개인을 식별 가능
- 생체인식정보
 - 생체정보 중 특정 개인을 인증 식별할 목적으로 처리
 - 생체 인식 원본 정보 -> 개인을 인증 또는 식별할 목적
 - 생체 인식 특징 정보 -> 원본정보로부터 특징점을 추출하는 등 기술적 수단에 의해 생성되는 정보

개인정보

생체정보도 개인정보임

- 개인정보보호법 시행령 제18조 (민감정보의 범위)
 - 기술적 수단을 통해 성성한 정보

생체정보 수집단계 유의사항

원본정보 수집·이용 동의			
항목	목적	보유·이용기간	동의여부
얼굴정보 (원본정보)	[필수*] 이용자 식별 및 본인 인증	• 특징정보 생성 시 까지	<input type="checkbox"/> 동의함 <input type="checkbox"/> 동의안함
	[선택*] 얼굴 인식 알고리즘 연구	• 수집일로부터 1년	<input type="checkbox"/> 동의함 <input type="checkbox"/> 동의안함
※ 위의 개인정보 처리에 대한 동의를 거부할 권리가 있습니다. 그러나 필수 항목의 동의를 거부할 경우 얼굴정보를 이용한 로그인 기능을 이용할 수 없습니다.			
민감정보(특징정보) 처리 동의			
항목	목적	보유·이용기간	동의여부
얼굴정보 (특징정보)	이용자 식별 및 본인 인증	• 회원탈퇴 시 까지	<input type="checkbox"/> 동의함 <input type="checkbox"/> 동의안함
※ 위의 민감정보 처리에 대한 동의를 거부할 권리가 있습니다. 그러나 동의를 거부할 경우 얼굴정보를 이용한 로그인 기능을 이용할 수 없습니다.			

- 생체인식정보 수집시 개인정보 수집 이용 동의서 징구
- 원본정보 보관이 필요한 경우 -> 목적 및 보유기간 구분하여 안내
- 정보통신망을 통해 생체인식정보를 전송하는 경우 암호화 후 전송
 - SSL/TLS, VPN 적용

생체정보 이용단계 유의사항

- 생체인식정보는 인사, 연구, 개발 등 다른 목적으로 활용 불가
 - 안면 인식 목적 얼굴사진 -> 인사관리, 신분증 발급에 활용 X
- 이용자가 생체인식 정보 제공과 이용을 결정할 수 있도록 열람, 정정, 삭제 등의 통제수단 제공
- 생체인식 정보 수집, 입력은 단말에서 처리
 - 19.8 -> 인터넷에 노출된 서버에서 암호화되지 않은 이름, 비밀번호, 지문, 안면 정보 발견

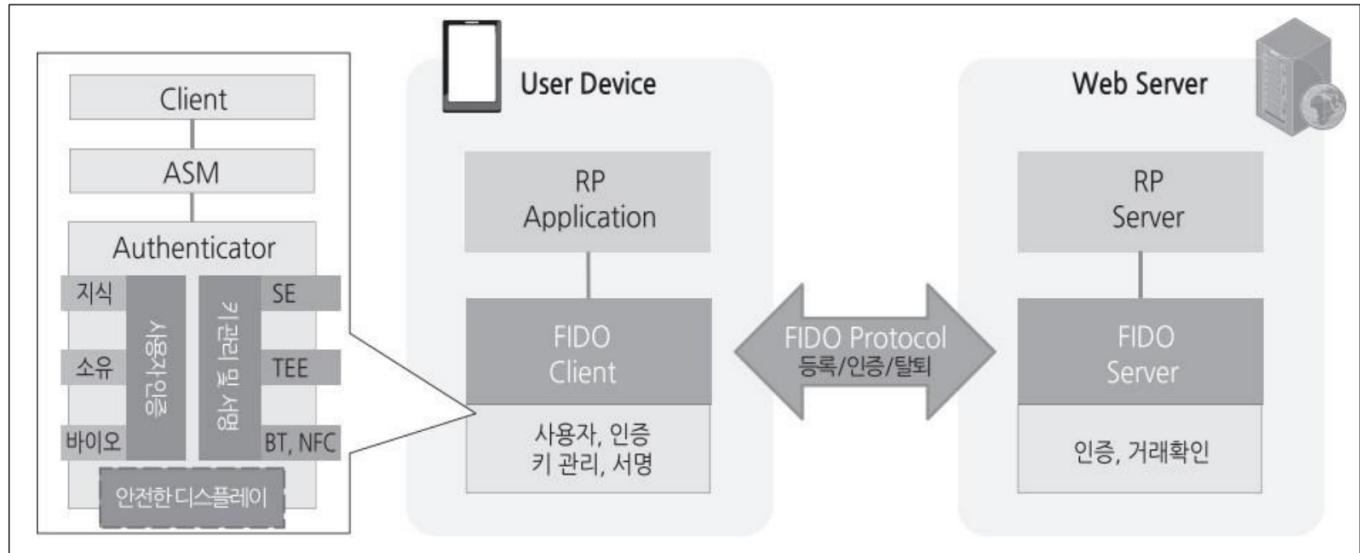
생체정보 보관/파기 단계 유의사항

- 생체인식 정보 저장 시 암호화 하여 저장
- 암호화에 사용하는 암호키 생성, 이용 보관, 배포, 파기 등에 대한 정차 수립
- 보유기간 경과 또는 처리목적 달성 시 생체인식정보를 지체없이 파기
- 원본정보를 보관해야 하는 경우 다른 개인정보와 분리하여 별도 저장
 - 물리적으로 분리
 - 이용자가 식별되지 않도록 공통 식별자는 임의 값 활용

생체인증 FIDO

- 아이디 비번 조합 대신, 지문, 홍채 등을 활용한 새로운 인증 시스템
- 개인정보는 단말에 저장
- 예시
 - 삼성페이 -> 인증방법을 지문으로 사용

FIDO 유니버설 인증 프레임워크 아키텍쳐



유저 디바이스에서 인증을 함

생체인증 유출 및 오남용 사례

실리콘 지문, 지문 유출 등등

클라우드 개인정보보호

클라우드 서비스 유형

- IaaS
- SaaS
- PaaS
- FaaS

클라우드 보안

저비용으로 많이 이용했지만, 개인정보 보호 문제 대두

클라우드 종료

- private 클라우드 -> 특정회사 전용, 확장성 떨어짐, 초기투자비용 높음
- public 클라우드 -> 적은 비용, 보안성 낮음, CSP에 대한 의존도
- Hybrid 클라우드 -> 보안성은 private, 나머지는 public

CSP MSP

- CSP (Cloud Service Provider)

- AWS, Azure, GCP
- 국내는 KT, 네이버, NHN 클라우드
- MSP
 - 메가존, 베스핀글로벌, 삼성 SDS, LG CNS 등

클라우드 서비스 시장 전망

- 국내 퍼블릭 클라우드 서비스 시장이 전년대비 17.8% 성장하며, 2022년 3조 226억원의 규모를 기록할 것으로 전망
- 클라우드 서비스 시장은 5년간 연평균 15.5% 기록하였으며, 2026년 5조 1,010억원의 규모를 기록할 것으로 전망
- IaaS 시장은 정부 클라우드 전환 정책으로 국내 클라우드 사업자들의 시장 진출 및 데이터센터 확장 움직임이 이어질 것으로 예상
- PaaS 시장의 경우 기업들이 산업별로 RPA, AI, 데이터베이스 기반의 PaaS 서비스에 대한 수요 확대로 플랫폼을 도입하는 움직임이 확산될 것으로 예상
- SaaS 시장은 비대면 업무 환경을 지원하는 VDI와 네트워크 보안강화를 위한 신규 수요가 지속적으로 발생하고 있어 기업 예산 변동성이 높게 나타날 것으로 전망



전반적으로 시장이 확대되는 추세

클라우드 사례

- NC 소프트 -> VMWARE가 프라이빗 클라우드 구축
- BNK금융그룹 뉴타닉스 위에 그룹사 프라이빗 클라우드 구현
- 국가관세망 -> 클라우드 전환 (유니파스)

클라우드 보호조치 - 접근 권한 관리

- 애플리케이션 접근 권한
- DBMS 접근 권한
- 서버 운영체제 접근권한

접근 권한 관리 고려사항

- 클라우드 관리콘솔에서 접근권한을 차등 부여하는 기능을 제공하고 있는지 검토
 - AWS IAM
- PaaS를 사용하면 서비스 별 접근 권한 부여 최소화
- 외부 스토리지 등 기타 저장 장치에 대한 접근권한 부여 최소화
- 권한 부여, 변경, 말소 처리 이력을 저장하는 기능 있는지 확인
 - cloudwatch

접근 통제 고려사항

어떻게 접근을 통제할 것인지

클라우드 안정성 확보 방안

역할별 계정 관리	역할별 계정 관리
네트워크 격리 및 접근통제	논리적으로 격리된 네트워크 적용 및 접근통제 정책 적용
전용회선 또는 IPSec VPN 연계	망분리 요건을 만족하도록 네트워크 연동 구성
암호화 키 관리 적용	중요정보 암호화 및 키 관리 서비스 적용
가상환경보안솔루션/ 패치	내부 보안 정책에 맞게 보안 정책 커스터마이징 적용
보안관제 서비스	보안 관제 및 침입대응, 취약점 점검 주기적 실시 보고
감사로그 수집	계정 활동 로그 자동으로 수집, 기록, 분석 기능 제공
보안교육 수행	보안 책임감 및 보안의식 고취를 위한 주기적 교육 실시

- 전용회선 또는 VPN 사용하는지 -> 클라우드는 원격지이므로

클라우드 보안 위협

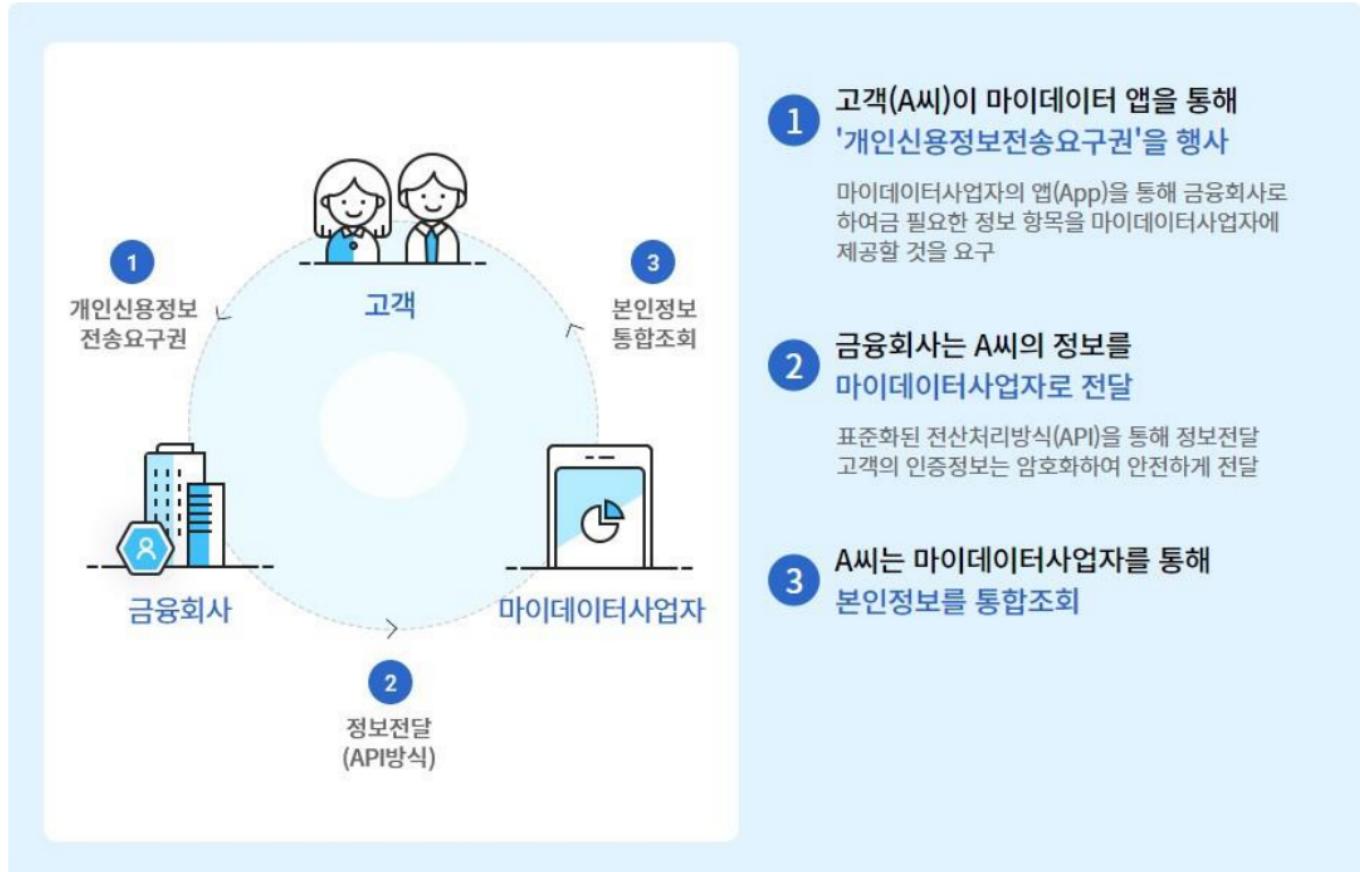
- 중요 데이터는 암호화 하자
- 계정 관리 미흡 또는 과도한 권한 부여
- 멀티 클라우드 -> 여러 클라우드 사용하면 가시성 확보 여려움
- 사용자 실수와 설정 오류
 - 장기간 동일 비밀번호 사용, 잡속 키 관리 미흡

마이데이터

다수 정보시스템에 분산되어 있는 개인의 정보에 대한 권리를 보장, 동의에 따라 본인 데이터를 개방 활용하는 것

- 개인정보 이동권 개인정보처리자가 처리하고 있는 개인정보 혹은 개인데이터를 정보주체가 그 사본을 넘겨받아 재사용
- A은행의 계좌정보를 -> C은행에서 볼수 있도록 땡김
- 개인정보 자기결정권 스스로 정보의 유통, 언제 공개할 건지 정하는 권리

마이데이터 서비스



개인이 금융데이터를 금융회사로 전달

- 넘겨주는 규격이 통일된 것임

마이데이터의 스크래핑 vs API

- 스크래핑 -> 사업자가 고객의 인증정보로 정보를 자동으로 추출
 - 과도한 수집 문제
- API -> 고객이 원하는 정보만 선택하여 정보 전송 요구 가능

마이데이터 사례 국내

금융	의료
KB국민은행 자산포트폴리오 소비자패턴 진단 금융습관/목표챌린지	신한카드 오늘의 자산 리포트 지출 분석/관리 금융 캘린더
토스 자산/계좌 통합조회 소비 내역 마이데이터 알림	뱅크샐러드 DTC유전자검사 발병률 예측 습관 등 건강증진 활동
나이스평가정보 신용상태 컨설팅 신용점수 올리기 대출현황/한도예측	나의 건강기록 진료 이력 건강검진 이력 투약 이력
메디팡팡 비대면 진료 진료 이력 건강검진 이력	비헬씨 처방전 이력 건강검진 이력 맞춤형 건강리포트

마이데이터 해외 사례

구분	미국 'Blue Button'	영국 'Midata'	프랑스 'Mes Infos'	핀란드 'MyData'
목적	• 개인 온라인 의료 기록 접근, 정보 공유 확대	• (개인) 현명한 의사결정 지원 • (기업) 비즈니스 혁신 창출 • (국가) 경제 성장 추구	• (개인) 개인 데이터 통합 지점 확보 • (기업) 개인정보 제공으로 기업/기관 간 가치 제고 • (국가) 기술적, 법적 문제 측정	• (개인) 디지털 인권 강화 • (기업) 새로운 사업 기회 제공, 데이터 기반 서비스 발전
주체	정부(10~175)	정부+민간(11.4~)	비영리기관+공공·민간(12.1~)	비영리기관+대학(15~)
분야	의료	에너지, 통신, 금융 집중	에너지, 통신, 보험 집중	특정 없음
참여 대상	• 정보제공자 : 병원, 약국, 공공기관, 보험사 등 • 서비스 이용자: 전 국민	Visa, Google, British Gas, Lloyds TSB, O2 등 (총 26개 기업)	비영리기관, 민간기업, 클라우드 제공기업, 실험대상자(개인)	비영리기관, 주요 대학
수혜 대상	참전용사 → 모든 환자	개인(소비자)	개인, 기업, 데이터 활용자	개인, 기업
현황	정보 업데이트 종료(17.5)	법적 규제 시도, Midata Innovation Lab 설립(13.7~)	실험 완료 후 파일럿 프로젝트 진행	프로젝트 진행
결과	1억 5천만 명이 의료 정보 이용 (16.3 기준)	금융서비스 분야 데이터 제공 강제 시행(18.1~)	조직에 대한 신뢰성 향상에 기여	헬스케어, 금융 등의 분야에서 프로젝트 수행

마이데이터 보안 위협 및 대책

마이데이터 사업자의 보안 취약점을 이용하여 데이터베이스, 서버 등 IT 인프라 대상으로 사이버 공격 시도 증가

마이데이터 사업자는 정기적인 취약점 진단을 통해 알려진 보안 취약점으로 인한 공격 대비

공격자에 의해 유출된 개인정보가 암호화되지 않아 데이터 재식별을 통한 개인정보 유출

개인정보 전송 및 저장 시 안전한 암호화 적용, 비식별 처리

마이데이터 서비스 사칭을 통해 이용자가 가짜 웹사이트나 모바일 앱에서 인증정보 입력 시 정보 탈취

이용자 대상 서비스 사칭 웹 사이트, 모바일 앱 주의 안내 및 이용자 보안수칙 안내

- 마이데이터 사업자의 보안 취약점을 이용하여 사이버 공격 시도 증가
 - 데이터베이스, 서버 등 IT 인프라 대상
- 유출된 개인정보가 암호화되지 않음 -> 데이터 재식별을 통한 개인정보 유출
- 마이데이터 서비스 사칭을 통한 인증정보 갈취

마이데이터 보호조치

- 관리적 보안
 - 관리 감독 철저 -> 금융감독원, 금융보안원등의 감사
 - ISMS-P 인증심사 등 준수, 공공기관은 PIA 법적 준수 추가 고려
 - 각 참여기업 및 기관의 관련 규정, 계획 수립 철저
- 기술적 보안
 - SOAR기반의 모니터링
 - Security Orchestration & Automation Response-> 차세대 보안 시스템 검토
 - 표준화된 Open API 적용
 - 벤더종속이나 업데이터 어려움 API 사용 지향

- 사용자 SIDE 보안 강화
- 개인정보보호
 - 개인정보 수집 및 제공 제한 -> 최소한의 정보만 수집
 - PPLM, PPDM 기술 활용
 - 인공지능 기술활용
 - 시스템 장애 및 SW오류 최소화

마이데이터 보안 사고 사례

- 어떤은행의 개인정보 유출 사고
- CI 정보 오류 -> 본인의 정보가 아닌 타인의 카드 사용내역, 대출내역 등이 노출
- 카드사들의 무리한 마이데이터 사업 추진으로 개인정보유출
- 마이데이터는 주로 데이터 전송 시 사고가 많이 발생함

복습 사항

- 개인정보법 개정 사유
 - 가명처리가 무엇인지
 - 가명처리 주의할 점, 강조할 점
- 마이데이터