

---

# 공공부문 가명정보 활용 체계 수립 방안

2024.06.05  
원주 한라대학교 AI정보보안학과  
김순석

---

[sskim@halla.ac.kr](mailto:sskim@halla.ac.kr)

# 목차

1. 가명정보의 개념 및 필요성
2. 가명정보 활용 체계의 구성요소
3. 공공부문에서의 가명정보 활용
4. 공공부문 가명정보 활용체계 수립 방안
5. 가명정보 활용 체계의 효과적 운영을 위한 방안
6. 요약 및 결론

# 1. 가명정보의 개념 및 필요성

## ❖ 가명정보란?

- (개인정보보호법 제2조제1호 다목) 추가 정보의 사용·결합 없이는 특정 개인을 알아볼 수 없는 정보

## ❖ 가명정보의 판단 기준

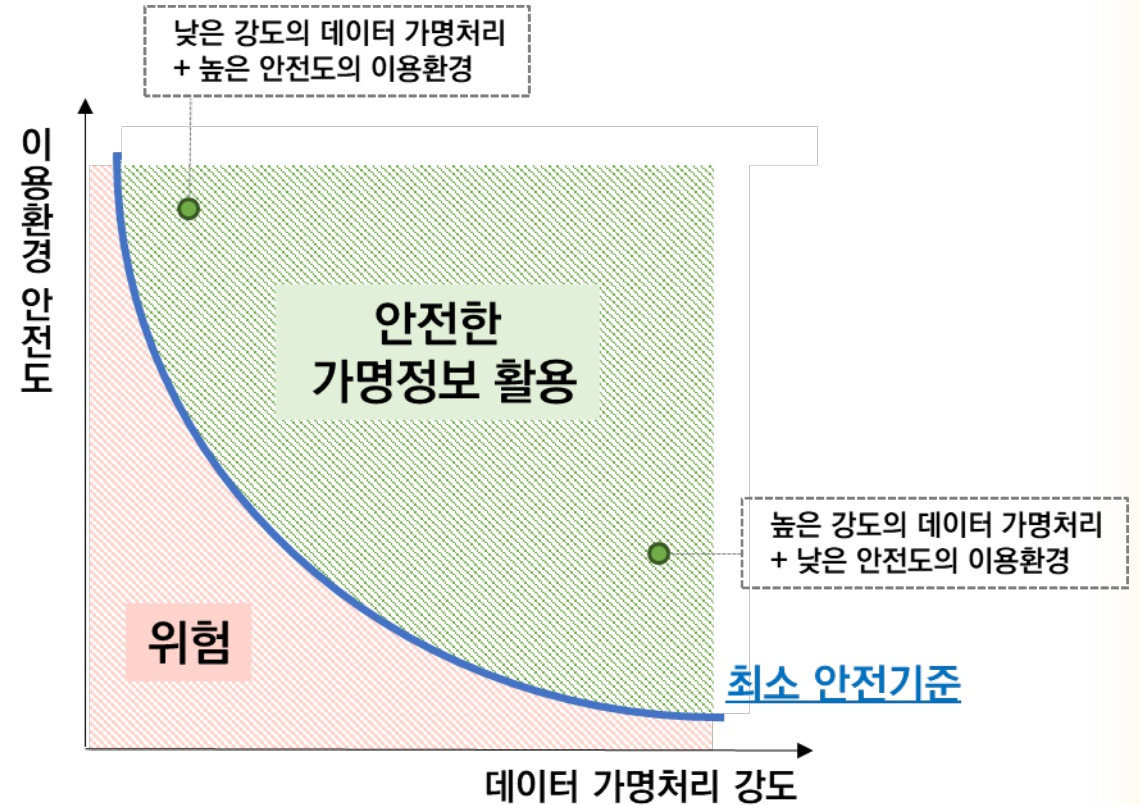
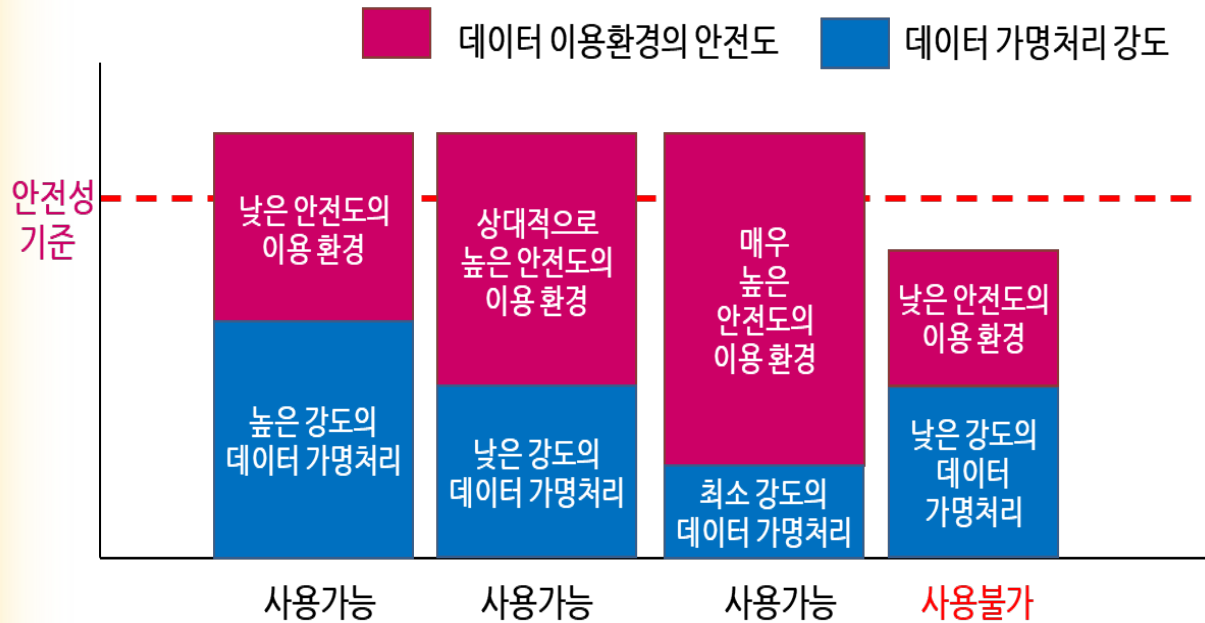
- (추가정보에 대한) 복원가능성 X & (특정 개인에 대한) 식별가능성 X
- 식별가능성 = 가명정보 자체 만으로 + 다른정보와의 결합을 통해서도

## ❖ 가명정보의 조건

- (개인정보보호법 가명정보처리에 관한 특례 제28조의2~7)
- 개인정보처리자는 통계작성, 과학적 연구, 공익적 기록보존 등을 위하여 정보주체의 동의 없이 가명정보를 처리할 수 있다.
- 가명정보에 대한 특정 제3자 제공 가능, 결합(데이터)전문기관을 통한 서로 다른 개인정보처리자 간의 가명정보의 결합
- 개인정보로서 안전조치 의무, 재식별 금지 의무

# 1. 가명정보의 개념 및 필요성 - 계속

## ❖가명정보의 안전 활용 원칙 및 기준



### ※ 가명처리 기준

- 데이터만을 가지고 처리하는 것이 아니라 활용 목적, 이용 환경 등을 종합적으로 고려하여 유연하게 판단



# 1. 가명정보의 개념 및 필요성 - 계속

## ❖가명정보의 중요성

- (데이터 안전 활용) 가명정보는 개인정보를 보호하면서도 동시에 활용할 수 있는 중요한 개념
- 데이터 경제 활성화에 큰 도움을 줌
- 개인정보보호 강화에 큰 도움을 줌 -> PET 기술의 등장, 안전조치의 일환으로서의 가명정보

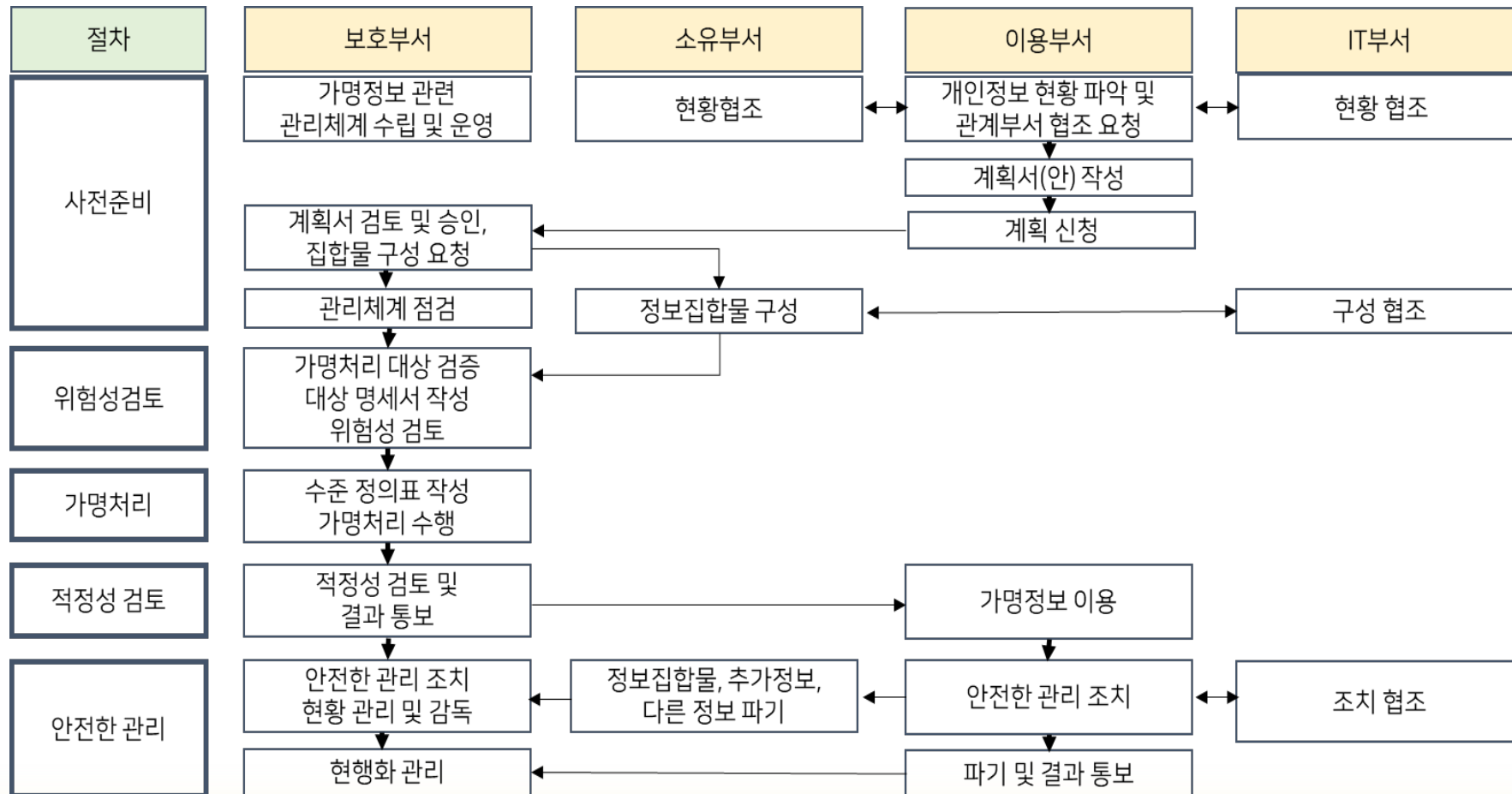
## 2. 가명정보 활용 체계의 구성 요소

### ❖가명처리 절차



## 2. 가명정보 활용 체계의 구성 요소 - 계속

### ❖가명처리 절차



## 2. 가명정보 활용 체계의 구성 요소 - 계속

### ❖가명정보의 안전조치 의무(개인정보보호법 제28조의4)

- 가명정보를 처리하는 경우에는 원래의 상태로 복원하기 위한 추가 정보를 별도로 분리하여 보관·관리하는 등 해당 정보가 분실·도난·유출·위조·변조 또는 훼손되지 않도록 안전성 확보에 필요한 기술적·관리적 및 물리적 조치를 하여야 함
- 처리목적 등을 고려하여 가명정보의 처리 기간을 별도로 정할 수 있음
- 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 보관



## 2. 가명정보 활용 체계의 구성 요소 - 계속

### ❖가명정보의 활용 사례 - 2021년

#### | 가명정보 결합 우수사례 |

복지 노후소득보장 종합연구 보건복지부 사회보장위원회 사무국 **08**  
경진대회 수상작 1기시범사례

의료·건강 암환자의 합병증 및 만성질환 예측 연구 국립암센터 **10**  
경진대회 수상작 1기시범사례

의료·건강 폐암환자의 사망동향 및 사망예측 연구 국립암센터 **12**  
1기시범사례

의료·건강 맞춤형 산림치유 프로그램 분석 국립산림과학원 **14**  
1기시범사례

통신 불법스팸 실태연구 한국인터넷진흥원 **16**  
1기시범사례

유통·소비 지역별, 상품별, 상품별 소비패턴 분석 (주)케이티, 롯데멤버스 **18**  
1기시범사례

유통·소비 편의점 시 상품추천모델 구축 (주)이마트24 **20**  
경진대회 수상작

금융 수요자별 맞춤형 금융지원정책 수립 (주)국민은행기금 **22**  
경진대회 수상작

금융 중신용/신패일 고객 신용평가모형 개발 카카오뱅크 **24**  
경진대회 수상작

금융 코로나 전후 소비·투자 관점의 데이터 결합 및 분석 (주)미래에셋증권, (주)롯데카드 **26**  
경진대회 수상작

금융 금융소외계층을 위한 머신러닝 모형 개발 하나은행 **28**  
경진대회 수상작

통신·금융 통신정보를 활용한 사회적 취약층 대상 중금리 대출 지원 (주)케이티 **30**  
경진대회 수상작

교통 화물차 개인사업자 신용도 평가모델 연구 한국도로공사 **32**  
경진대회 수상작

#### | 가명정보 결합 아이디어 |

유통·소비 네트워크 회복을 통한 지역경제 활성화와 청년창업 지원 RE:DATA project **36**  
경진대회 수상작

복지 교통약자의 이동편의를 위한 대중교통 시설 점검 아이디어 도출 T-BTS **38**  
경진대회 수상작

복지 보건의료데이터 통합과 직업력 결합을 통한 국민직업건강정보체계 구축 최강이 이김 **40**  
경진대회 수상작

교육 선택과목별 학생 성취도 대조분석을 통한 효율적인 선택과목 운영 아이디어 도출 최형빈 **42**  
경진대회 수상작

## 2. 가명정보 활용 체계의 구성 요소

### ❖가명정보의 활용 사례 – 2022년, 2023년

2022 가명정보 활용 우수사례집

## CONTENTS

### 🔗 2022 가명정보 활용 우수사례 부문

복지 가명데이터 활용, 서울시 1인가구 정책 실효성 높인다 1인가구 분석팀 **06**  
대상

금융 비금융정보를 활용한 대안신용평가모형 개발 카카오펙크 신용리스크모델링팀 **08**  
대상

금융·복지 지역가입자의 주택금융부채공제를 위한 신용정보 결합 NHIS BIG DATA팀 **10**  
대상

저탄소에너지 친환경차 충전 인프라 수요예측 및 최적입지 선정 우결우팀 **12**  
대상 선도시례

문화·여가 맞춤형 가명정보로 철도협력여행사 코로나 극복 지원 한국철도공사 빅데이터팀 **14**  
대상

금융·복지 공무원의 기여금 성실 납부정보를 활용한 신용점수 향상 추진 공무원연금공단팀 **16**  
우수상

의료 차세대 염기서열 분석이 주요 암 진단과 치료에 미친 영향 분석 N to the G to the S팀 **18**  
우수상

건강 가명결합을 활용한 질병 및 재해 위험 예측모형 개발 한화생명 Data Analytics팀 **20**  
우수상

금융 코로나 상생 국민지원금의 경제적 효과 분석 신한카드 Data Biz 펠터팀 **22**  
우수상

금융 국가보훈대상자 신용실태 조사 국가보훈대상자 신용 실태 연구팀 **24**  
우수상

유통·소비 소상공인을 위한 맞춤형 고객정보 제공 지슬링 가이드팀 **26**  
우수상

문화·여가 게임&애니메이션 유저의 소비컨텐츠 및 특성 분석을 통한 비즈니스 인사이트 생성 넥스코리아팀 **28**  
우수상

금융 KBIZ-DATA 시장공급을 통한 소상공인 금융지원 중소기업중앙회 X NICE평가정보팀 **30**  
우수상

관광 고속도로 통행정보와 강원도 관광지 주차장 이용정보 결합 연구 강원도로의 힘팀 **32**  
우수상

### 🔗 2022 가명정보 활용 아이디어 부문

의료·건강 알코올 중독자의 진료정보 결합을 통한 질병 형태 분석 강현연구팀 **34**  
대상

건강 구매 식품과 건강의 상관관계 연구 삼카21총팀 **36**  
우수상

복지 재택근무제도 활성화를 통한 일·가정 양립 실현 및 고용 취약계층의 일자리 창출 박영주 **38**  
우수상

복지 서울복지실태조사와 행정데이터 결합 서울연구원팀 **40**  
우수상

관광 모두가 함께 즐기는 제주관광 무장애관광 분석팀 **42**  
우수상

의료·건강 생애주기에 따른 의료이용 실태분석 및 형평성 연구 건강연구221팀 **44**  
우수상 선도시례

### 「2023 가명정보 활용 우수사례 · 아이디어 경진대회」 수상작 내역

구분	상격	시상 주체	작품명	팀명 (참여기관)
우수사례 부문	대상 (5)	개인정보보호위원회 위원장상	가명정보의 결합적 활용을 통한 교통약자 보호구역 지정 분석	양평군
		금융위원회 위원장상	가명정보 결합물을 활용한 Gig 등급모형 개발	나이스지니데이터(주)
		보건복지부 장관상	한방병원에 내원한 요추추간판탈출증 환자에서 한방의료이용에 따른 차이 분석	부천자생 (부천자생한방병원)
		과학기술정보통신부 장관상	제주 한달살이 분석	카름누리 (제주관광공사 통계팀 SKI)
		국세청 청장상	코로나19 위험도 순위점수 조회 서비스	국민건강보험공단 빅데이터연구부
	우수상 (11)	국민건강보험공단 이사장상	응급실 뺑뺑이를 방지하기 위한 소아 응급실 방문 AI예측	HIRA AI 리딩팀 (건강보험심사평가원)
		건강보험심사평가원 원장상	암환자에서의 딥러닝 기반 관상동맥 석회화 점수 유용성 검증 및 심혈관계 사건 위험도 분석 연구	세브란스병원 (연세대, 중앙대)
		금융결제원 원장상	대안정보를 활용한 소상공인 특화 신용평가모형 개발	카카오뱅크 신용리스크모델링팀
		금융보안원 원장상	지역화폐 혁신 운영을 위한 마이크로 소비행태 분석	CAMELLIA SEVEN (부경대, 부산시, 부산TP, 코리아세븐)
		롯데정보통신 대표상	에너지이용 소외계층의 현황 파악을 위한 '에너지 사용량 데이터 분석'	한국에너지재단 에너지복지연구센터 (한국에너지재단 등 12)
		LG CNS 대표상	가스 및 통신데이터 결합을 활용한 취약계층 생활패턴 분석	Projectmakers (경동스기스, SKT, 통계청)
		더존비즈온 대표상	동반 질환에 따른 중상이상 교통사고의 위험성 규명을 위한 '의료정보-중상이상 교통사고' 자료 결합	서울대학교병원
		한국사회보장정보원 원장상	정책소외계층 청년 실태 및 정책개발·시설토소 자립준비청년 생활실태 조사	한국청소년정책연구원 청년종합연구팀
		한국데이터산업진흥원 원장상	내일채움공제 사업별 가입자 및 미가입자간 장기재직효과 분석	중소벤처기업진흥공단 성과보상처
		한국신용정보원 원장상	신정원의 신용정보와 보험정보를 활용한 자동차 사고예측 모형	고려대 캡스톤 연구팀 (고려대 경영대학원)
		한국지능정보사회진흥원 원장상	암환자 병원임상정보-국가보건의료 융합 데이터베이스 플랫폼 개발을 통한 장단기 예측, 비용-효과평가 및 암경험자 질병 발생 예측모델 개발	서울대병원 암정보과제 연구팀
	대상 (1)	개인정보보호위원회 위원장상	노인을 위한 버스정류장 시설 개선 우선순위 선정	운영공동체 (성균관대학교)
		한국인터넷진흥원 원장상	광명시민 취약계층 안전 확보를 위한 지능형 CCTV의 효율적인 설치 위치 선정을 위한 연구	정보통신과민만세 (광명시)
			민간 및 공영주차장 기반 체납차량 우선 단속지역 분석	신세릭스 (대구디지털혁신진흥원)
			건강보험 외국인 부정수급 및 원정치료(공수치료) 분석방안 제시	강민석 (한국수자원공사)
			MZ세대 소비 행태에 따른 퇴사예측 모형	청년희망을 위한 데이터 네티지의 시작 (한국능률협회컨설팅, 한국산업단지공단)
아이디어 부문	우수상 (5)		살기 좋은 지역을 만들기 위한 데이터 기반 접근 방향	정도범, 유화선 (한국과학기술정보연구원)
				<b>10</b>

## 2. 가명정보 활용 체계의 구성 요소

### ❖가명정보의 결합선도 사례 – 2021년(1기), 2022년(2기)

분야	1기 가명정보 결합 시범사례
의료+인구	① 폐암 환자의 사망동향 및 사망예측 연구 (국립암센터) (21.5.28. 성과발표)
	[결합정보] 암 환자정보(암센터), 암 환자 진료정보(국민건강보험공단), 사망정보(통계청)
	[분석결과] 폐암 환자의 5년 이내 사망률 77.4%, 5년 이상 생존 폐암환자의 22.2%가 암 이외 원인으로 사망, 이 중 24.8%가 심뇌혈관 질환이 원인인 것으로 분석
의료+인구	② 암 환자의 장기합병증 및 만성질환 예측 연구 (국립암센터) (21.6.4. 성과발표)
	[결합정보] 약 20만 명의 암 환자 임상정보(암센터, 병원), 암 환자 진료정보(국민건강보험공단)
	[분석결과] 6대 암환자가 일반 환자군에 비해 주요 질병 발생빈도가 높음(심부전 81%, 심근경색 50%, 뇌졸중 25%, 골절 47%, 당뇨병 35%), 결합데이터의 AI학습을 통해 암예방 예측 모델 개발 예정
통신+유통	③ 불법스팸 실태연구 (한국인터넷진흥원) (21.6.25. 성과발표)
	[결합정보] 스팸신고정보(한국인터넷진흥원), 고객정보(SK텔레콤)
	[분석결과] 남성(64.4%)이 여성(35.6%)보다 스팸신고가 많음, 남성은 주로 주식정보 스팸에 주로 노출되며, 여성은 대출카드 스팸에 노출됨, 50대 남성에 대한 도박대출 스팸 발송이 가장 많음
통신+유통	④ 지역별, 상권별, 상품별 소비패턴 분석 (KT, 롯데멤버스) (21.7.16. 성과발표)
	[결합정보] 이동통신 고객정보(KT), 유통소비정보(롯데멤버스)
	[분석결과] 1인 가구는 의류패션잡화·가전·디지털기기 구매 비중 높음, 구매금액 상위 고객은 30대, 개인별 관심도가 높은 품목에 대한 구매 비중이 2배 이상 높음
레저+건강	⑤ 맞춤형 산림치유프로그램 분석 (국립산림과학원) (21.12.27. 성과발표)
	[결합정보] 의료 건강정보(국민건강보험공단), 운동 활동정보(주비글)
	[분석결과] 등산 빈도·거리가 건강 증진 및 의료비 감소에 높은 상관관계를 보임, 연령 높아질수록 등산 비율 증가, 뛰기 비율 감소, 체중·체질량·허리둘레 감소에 등산 횟수·거리가 큰 영향
소득+복지	⑥ 노후소득보장 종합연구 (보건복지부 사회보장위원회) (21.12.30. 성과발표)
	[결합정보] 약 340만 명 대상 소득정보(국세청), 세대인적정보(행안부), 개인연금정보(은행사 등)
	[분석결과] 노인층 연금수급률 47%(남성 66%, 여성 33%), 연간수급액 710만원(남성 861만원, 여성 489만원), 소득수준 높을수록 공적연금 가입률 증가, 연령이 높아질수록 기초연금수급률 증가
금융+보훈	⑦ 국가보훈대상자 신용 실태연구 (국가보훈처) (결과 비공개)
	[결합정보] 보훈대상자 정보(국가보훈처), 보훈대상자 신용정보(신용정보원)

#### < 2기 가명정보 결합 4대 중점 선도사례 >

분야	선도사례	정보보유기관 및 결합정보	이용기관	결합전문기관
소득 불평등	①건강보험 급여에 따른 소득분배 영향 분석	(건보공단) 건강보험 가입현황 및 보험료·급여 정보 등 (통계청) 가계금융복지조사 정보	가천대학교	통계청 (개인정보위 지정)
의료 형평성	②생애주기에 따른 의료이용 실태분석 및 형평성 비교	(한국노동연구원)노동패널정보 (건강보험심사평가원) 진료자료 (건보공단) 검진·장기요양자료 등	심사평가원 서울대학교	건보공단 (복지부 지정)
장애인 복지	③장애인 복지정책 효과 분석	(건보공단) 국민건강보험 정보 (사회보장원) 장애인 연금정보 등	서울대학교	건보공단 (복지부 지정)
저탄소 경제	④친환경차 충전 인프라 수요예측 및 최적입지 선정	(성남시) 성남시 차량등록정보 (티맵모빌리티) Tmap 운행이력정보	성남시 티맵모빌리티	SK(주) C&C (과기정통부 지정)

## 2. 가명정보 활용 체계의 구성 요소

### ❖가명정보의 결합선도 사례 – 2023년(3기)

**참고 1** 3기 가명정보 결합 선도사례 총괄 현황

분야	선도사례 과제	연구기관	정보 제공기관 및 결합대상 정보
복지 (노인) +경제	①고위험 취약노인 예측 모형개발·고령자 지원 정책 효과 분석	KDI 국제정책대 학원	(복지부) 기초연금, 노인돌봄 등 (고용부) 고용보험 취득·상실정보 등 (건보공단) 보험자격 등
인구 +사회	②인구감소지역 지원을 위한 생활인구 산정 및 활성화 방안	행안부, 통계청	(행안부) 주민등록 (법무부) 외국인등록 (통신사) 이동통신
복지 (여성) +경제	③일·가정 양립 정책이 여성 근로자의 경제활동·출산에 미치는 영향 분석	서울대	(고용부) 육아휴직 관련 등 (건보공단) 보험자격 등
의료 +건강	④실업이 의료서비스 이용과 국민 건강에 끼치는 영향 분석	고려대	(고용부) 고용보험 취득·상실정보 등 (건보공단) 보험자격 등
사회 +복지	⑤공공부조 제도 확대가 대상자의 고용·건강·의료 이용에 미치는 영향 분석	강남대	(고용부) 고용보험 취득·상실정보 등 (복지부) 기초생활보장 수급 (건보공단) 보험자격 등

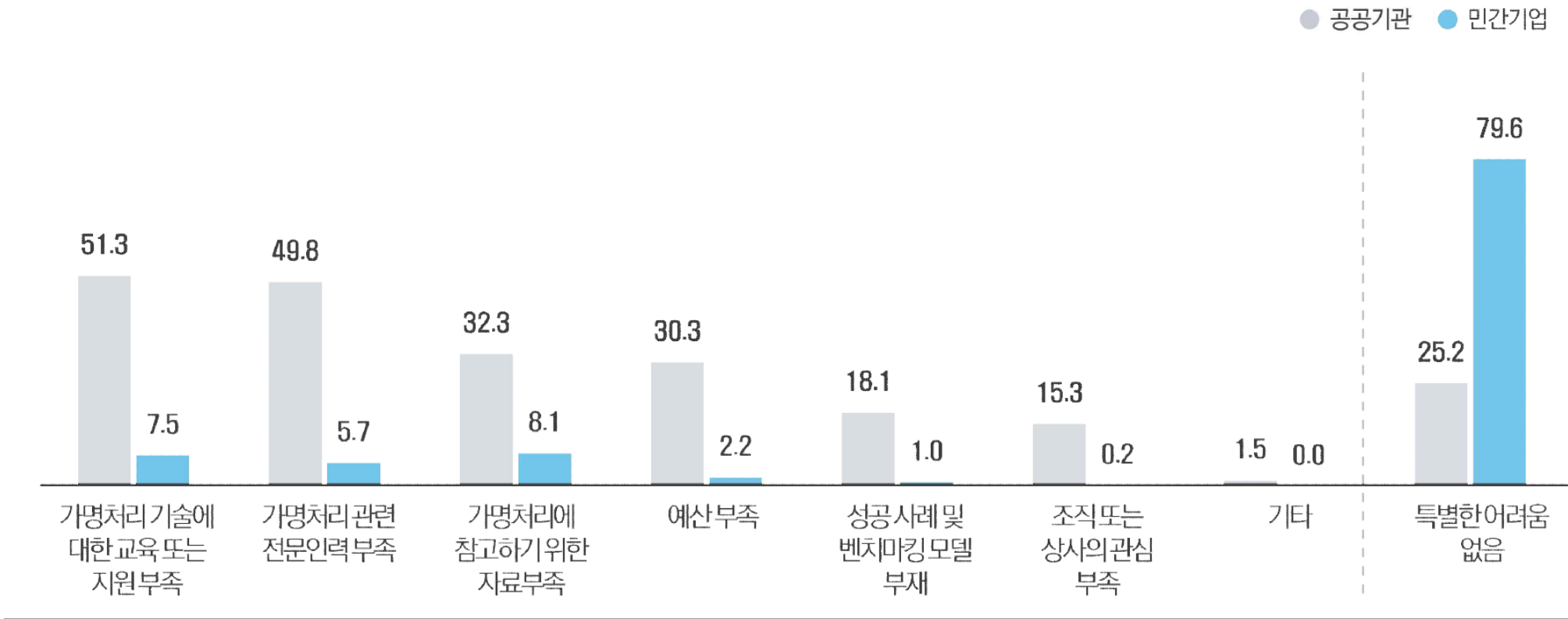
복지 (아동) +건강	⑥복지서비스 이용 아동과 청년의 서비스 이용실태 및 효과분석	아동권리 보장원	(권리보장원) 아동 신체 마음·건강 등 (사회보장원) 급여 서비스 내역 등
교육 +취업	⑦졸업 후 경력경로에 기반한 맞춤형 학생역량 시스템 개발	군산대	(군산대) 졸업생 학적 등 (건보공단) 건강보험 정보
지역 소비 +경제	⑧부산 지역화폐 운영 정책개선을 위한 이용행태 분석	부산시, 부경대	(부산시) 지역화폐 승인내역 (롯데멤버스) 구매이력
지역 의료 +건강	⑨코로나19 환자치료 지역격차 분석	강원도	(질병청) 코로나19 역학조사 (통계청) 사망원인 정보 (강원도) 코로나19 역학조사 (심사평가원) 진료·처방정보 등



### 3. 공공부문에서의 가명정보 활용

그림 11-27. 개인정보 가명처리 애로사항(복수응답)

(단위 : %, 기준 : 가명처리 경험자, 검토 중인 응답자)



2021년 개인정보보호 실태조사, 개인정보위



# 4. 공공부문 가명정보 활용 체계 수립 방안

## ❖체계 수립을 위한 기본 원칙

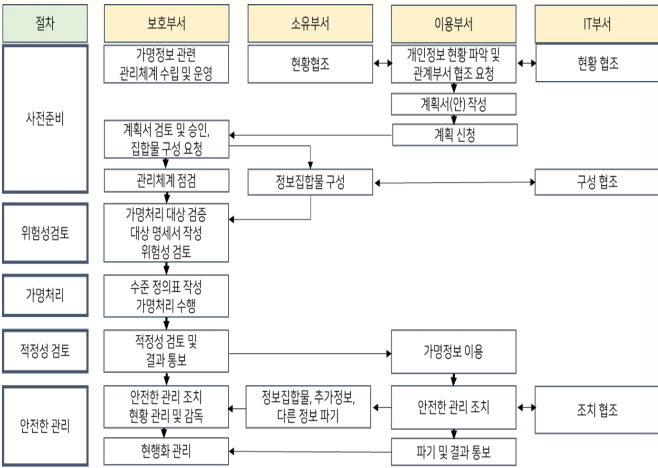
### ➢ 데이터 거버넌스

- ✓ 기관 내 가명정보 관리체계를 수립,  
운영할 담당 부서 지정, 교육 필요

### ➢ 데이터 컴플라이언스

- ✓ 개인정보보호법, 신용정보법
- ✓ 시행령, 고시, 가이드라인...
- ✓ 내부관리계획 수립,
- ✓ 개인정보처리방침 고지

공공부문 가명정보 활용 체계 수립 방안



### 가명정보 관련 제도 현황에 대한 참고

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

## 4. 공공부문 가명정보 활용 체계 수립 방안

### ❖ 체계 수립을 위한 기본 원칙

#### ➤ 가명처리를 위한 기술 인프라

- ✓ 가명처리 시스템, 가명정보 및 추가정보 보관을 위한 서버 구성 등 HW
- ✓ 가명처리를 위한 SW
- ✓ HW 및 SW 운용 인력 전문 교육

## 5. 가명정보 활용 체계의 효과적 운영을 위한 방안

### ❖가명정보의 품질 관리 방안

- 핵심 이해관계자들과의 상시 소통
- 가명정보 처리의 주 목적은 분석과 활용에 있음. 따라서 분석과 활용을 위한 데이터 품질이 중요
  - ✓ 이를 위해 가명정보를 이용할 부서나 혹은 기관과의 협의를 통해 데이터 품질 관리가 필요
  - ✓ 가명처리 절차에서는 3단계(항목별 가명처리 수준 정의표 생성 이후 실제 데이터 가명처리 이전)에 해당

## 5. 가명정보 활용 체계의 효과적 운영을 위한 방안

### ❖ 가명정보 활용 모니터링 및 평가

- 내부관리계획의 경우 1년에 한번씩 이행 실태를 점검
- 가명정보 활용 또한 최소 1년에 한번 정도는 안전성에 대한 모니터링 필요
  - ✓ 기술은 계속 발전하기 때문임

### ❖ 가명정보 재식별 방지 대책

- 가명정보의 재식별 위험 최소화 방안
  - ✓ 보호법 제28조의4 가명정보에 대한 안전조치 의무 이행
  - ✓ 보호법 제28조의5 가명정보 처리시 금지 의무 이행
    - 특정 개인을 알아보기위한 목적으로 가명정보를 처리해서는 안됨
    - 가명정보 처리 과정에서 특정 개인을 알아볼 수 있는 정보 생성시 즉시 처리를 중지하고 지체없이 회수, 파기

## 6. 요약 및 결론

### ❖ 공공부문 가명정보 활용 체계 수립 방안

#### ➤ 체계 수립을 위한 기본 원칙

- ✓ 데이터 거버넌스, 컴플라이언스, 기술 인프라 마련 필요

### ❖ 가명정보 활용 체계의 효과적 운영을 위한 방안

#### ➤ 가명정보의 품질관리

#### ➤ 가명정보 활용 모니터링 및 평가

#### ➤ 가명정보 재식별 방지 대책 – 가명정보의 재식별 위험 최소화 방안

- ✓ 보호법 제28조의4~5 – 가명정보에 대한 안전조치 의무, 가명정보 처리 시 금지의무 이행 필요



# 감사합니다.

---