

K-ICT 빅데이터센터
개인정보 가명·익명조치기술 전문교육

NIA 한국정보화진흥원

개인정보 가명· 익명처리 국내· 외 동향

1일차 | 기본1

김 승 환

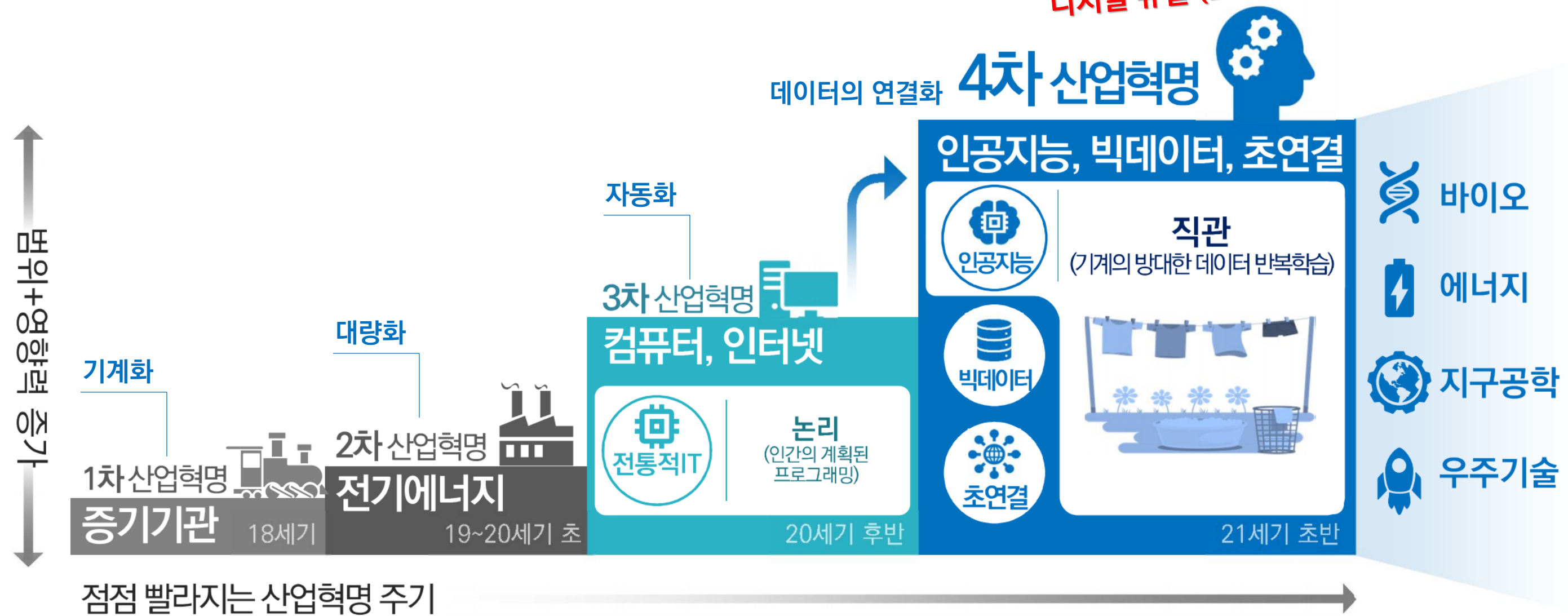
인하대학교 데이터 사이언스학과

swkim4610@inha.ac.kr

이 노트는 라곰 소프트 오형섭 이사 강의자료를 편집
하였음을 밝혀 둡니다.

인공지능, 빅데이터 등 디지털 기술로 촉발되는 초연결 기반의 지능화 혁명

디지털 뉴딜 (DNA 생태계 강화)



※ (출처) 대통령직속 4차산업혁명위원회, “4차산업혁명위원회”

개인정보 가명·익명처리 국내외 동향

1. 데이터 경제 활성화 동향

■ 데이터 경제(Data Economy)의 개념

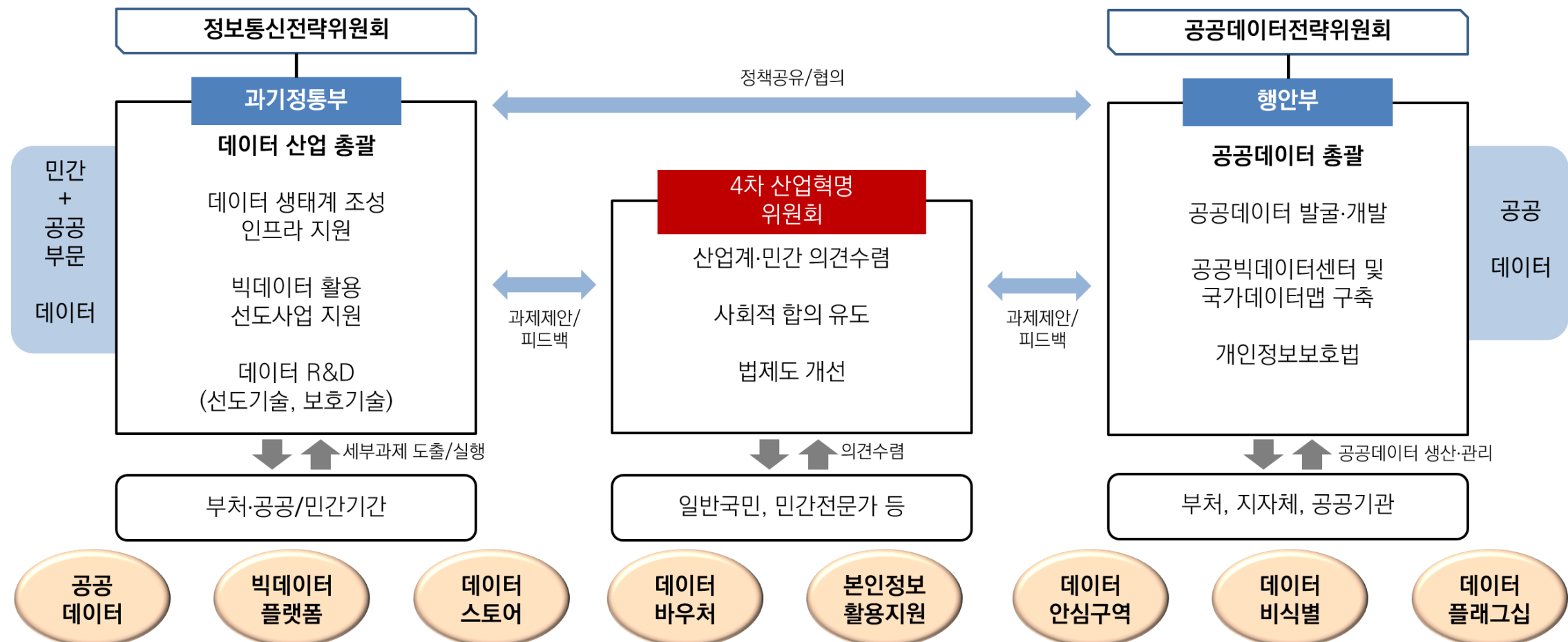
가트너 2011	<ul style="list-style-type: none"> 응용 프로그램, SW, HW 경제가 아닌 빅데이터, 오픈/연결데이터 등 데이터로 파생되는 경제가 경쟁 우위를 이끌어가는 시대
EU 집행위원회(EC) 2014	<ul style="list-style-type: none"> 데이터를 다루는 구성원이 만들어내고 있는 생태계, 데이터의 생성·수집·저장·처리·분배·전달 등을 모두 포괄하는 개념
MIT 2016	<ul style="list-style-type: none"> 데이터 자본은 재화·서비스를 생산하는데 필요한 저장된 정보로, 기존의 물리적 자산처럼 장기적인 경제적 가치를 보유
IBM 2016	<ul style="list-style-type: none"> 데이터를 내·외부적으로 가치를 창출하는데 사용하는 것을 의미하며, 이러한 현상은 이용 가능한 데이터와 데이터 기반의 의사결정이 증가하면서 기업들 사이에서 더 많은 데이터가 교환됨으로써 발생
Digital Reality 2018	<ul style="list-style-type: none"> 조직이나 비즈니스의 방대한 데이터를 저장·검색·분석해서 생성되는 금융이나 경제적 가치
한국정보화진흥원 2018	<ul style="list-style-type: none"> 모든 데이터가 활용하기 쉽게 자유롭게 흘러 타 산업 발전의 촉매 역할을 하면서, 혁신적 비즈니스와 서비스를 창출하는 경제

※ (출처) 한국정보화진흥원, “데이터 경제의 부상과 사회경제적 영향”

■ 데이터 경제 활성화 전략과 데이터 경제 활성화 추진체계

주요내용

- ① 데이터경제 활성화를 위한 산업 육성
 - 빅데이터 및 인공지능 등에 2019년 1조원 투자
 - 정보주체 중심 데이터 활용(MyData) 사업, 빅데이터센터 구축, 중소벤처 데이터 구매·가공 바우처 지원
- ② 데이터 활용 관련 규제혁신
 - 가명정보 개념의 도입, 익명정보는 개인정보보호 대상에서 배제, 데이터 결합 법적 근거 마련



■ 데이터 경제 활성화를 위한 빅데이터 플랫폼 및 센터 현황

분야		과제명(약칭)	플랫폼 주관기관	센터 구성
금융		금융 빅데이터 플랫폼	비씨카드	10개 (다음소프트 등)
환경		환경 비즈니스 빅데이터 플랫폼	수자원공사	10개 (그린에코스 등)
문화		문화체육관광 빅데이터 플랫폼	문화정보원	11개 (야놀자 등)
교통		DIAMOND-E 빅데이터 플랫폼	교통연구원	10개 (아이나비 등)
헬스케어		암 빅데이터 플랫폼	국립암센터	10개 (삼성서울병원 등)
유통소비		유통·물류 빅데이터 플랫폼	매일방송	10개 (식신 등)
통신		통신 빅데이터 플랫폼	케이티	14개 (두잉랩 등)
기타	중소기업	중소-중견기업 빅데이터 유통 플랫폼	더존비즈온	10개 (인크루트 등)
	지역경제	지역경제 빅데이터 플랫폼	경기도청	8개 (코나아이 등)
	산림	산림 어메니티 빅데이터 플랫폼	임업진흥원	10개 (한국산림복지진흥원 등)

1. 데이터 경제 활성화 동향

빅데이터 vs 인공지능

- 인공지능이 과거 보다 좋은 결과를 창출하려면 막대한 빅데이터를 학습하여야 함
- 데이터는 원유이고, 머신러닝 기법은 원유정제 기술임
- 결국, 원유를 보유하는 것이 경쟁력임



인공지능(AI) Artificial Intelligence

1. 데이터 경제 활성화 동향

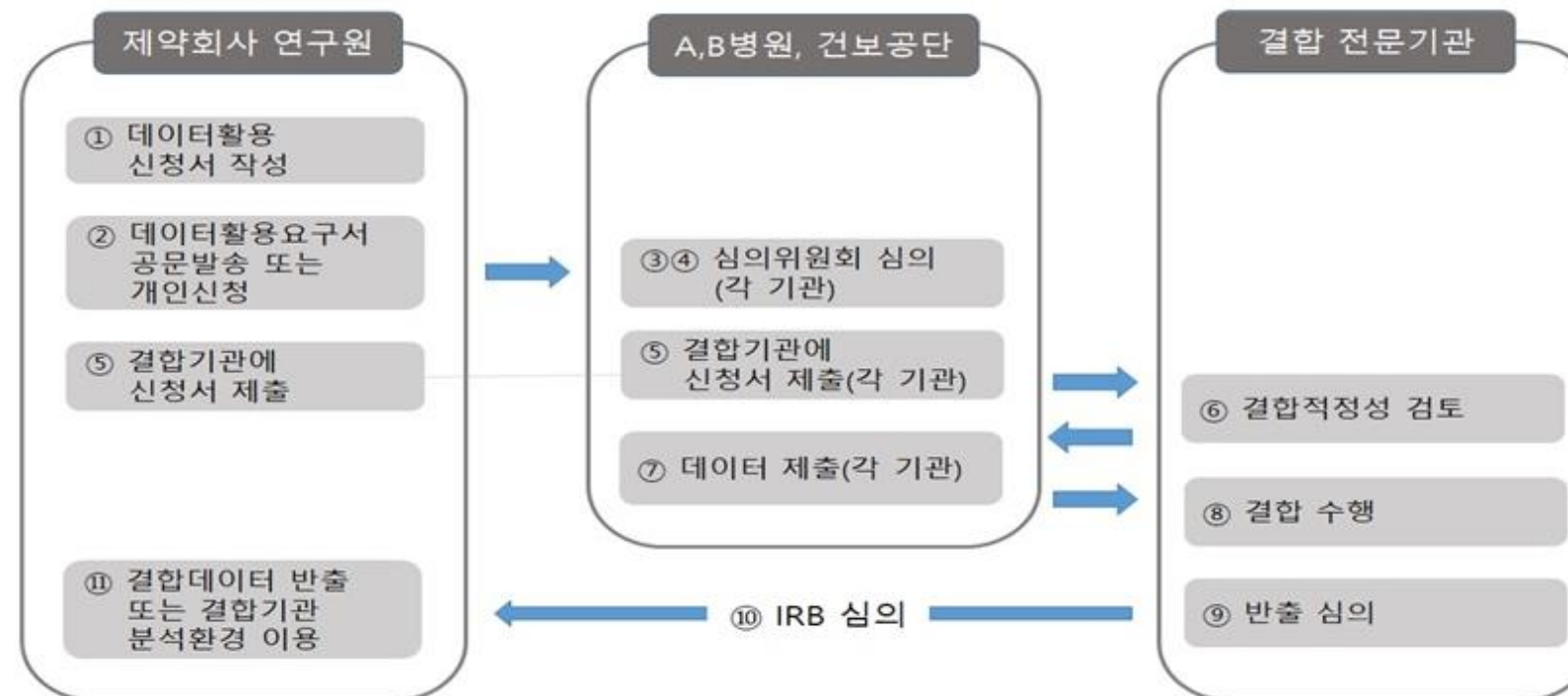
■ 빅데이터 vs 인공지능

- 과거에는 고객에 대한 정보를 카드 매출, 점원의 경험등을 이용해 얻었지만, Untact 시대에는 고객이 온라인에서 발생하는 막대한 빅데이터에서 고객에 정보를 얻을 수 밖에 없음



■ 빅데이터 vs 인공지능

- 사람들이 발생하는 데이터(교통, 의료, SNS 등)는 기업, 병원, 국가기관 등에 분산되어 있고 사용목적이 개인정보 보호법에 의해 제한되어 있음
- 개인정보를 통한 사익을 취하지 않는 공공목적, 연구, 제품개발 등도 동시에 제한되어 정보활용에 걸림돌이 되고 있어 이에 대한 규제를 완화하였음
- 또한, 기관별로 분절되어 있는 데이터를 결합하여 빅데이터화 하는 문제도 많은 진전을 보았음



■ 개인정보보호 패러다임 이동

- 기업의 무분별한 타겟 마케팅 전개로 고객불만 야기
- 기업 내 보유정보를 내부인이 식별하여 프라이버시 침해 사례 발생
- 기업이 고객정보를 동의 없이 제3자에 제공하는 사례
- 해킹 등 사고로 인해 고객정보가 유출된 사례
- 기업보유정보를 연구 등 목적을 위해 익명화 후 공표,
개인식별사고 발생(넷플릭스, 미국 Group Insurance Commission 등)



국내의 경우, 상당량의 고객정보가 여러 경로로 유출되고 주민번호로 합쳐져 암시장
에서 거래 중 → 주민번호 키 사용 금지

■ 개인정보보호 패러다임 이동

- 경영분석: 경영활동을 위해 보유정보를 분석/사용
- 타겟 푸쉬 마케팅: 이메일, SMS, 전화 정보를 이용한 상품판매(보험, 통신가입 등)
- 기업집단: 그룹 내 고객정보를 합쳐 공유하고 이를 통해 거대 플랫폼화
- 제 3자 제공: 기업 채널을 이용해 제3자 제공 동의 획득
- 정보 판매: 통신사 상권정보 판매, 스타트업이 수익확보를 위해 앱 사용고객 정보 판매



- 개인화 광고: 개인정보에 기반한 인터넷, 휴대폰 앱 광고 사업
- 신용평가, 보험 사기 Detection 등 기업 활동에 필요한 지능형 모형 개발
- 병원, 건보원 등 보유정보를 합쳐 의료정보 빅데이터 플랫폼 구축
- 기업 간 보유정보 매시업 등
- CCTV 저장 정보에 의한 범인추적 등 공공목적 활용

■ 개인정보보호 패러다임 이동

데이터 후진국 전략, 네이버 카카오 헬스케어 핀테크 고사위기
개인정보 정보통신 신용정보법 3대 법안 개정 국회표류, 빅데이터 강국 요원

'카드사 트라우마'에 갇혀 빅데이터 경쟁서 뒤처진 한국

'개인정보 보호 vs. 데이터 활용' 두 마리 토끼 잡을 수 있을까?

정보인권이 존중 받아야 할 빅데이터 시대

저커버그가 7일(현지시각) 이같은 구상이 담긴
‘프라이버시에 초점을 맞춘 소셜 네트워킹의 비전’ 을 발표했다.

페이스북 대적하려던 구글플러스 '참담한 실패'

소프트웨어 버그(오류)로 사용자 5천200여만 명의 개인정보가 유출되는
대형 프라이버시 사고가 터졌다.

■ 개인정보보호 패러다임 이동

[데이터3법①] 지금 생각해도 오싹... '개인정보보호 강화' 10년의 역사

2020.03.01 09:52:53 / 이종현 bell@ddaily.co.kr

관련기사

- ▶ "써도 되나요?"... 데이터3법 통과
- ▶ 전문가들 "개인정보 비식별 조치하
- ▶ [취재수첩] 데이터3법, 정말로 개인
- ▶ 국회 행안위 2019년 결산... "개인
- ▶ 개인정보 보호 기본계획 발표... "데

[데이터3법②] "보호에서 활용으로"... 데이터경제 시대로 간다

2020.03.03 08:56:41 / 이종현 bell@ddaily.co.kr

관련기사

- ▶ [데이터3법①] 지금 생각해도 오싹
- ▶ "써도 되나요?"... 데이터3법 통과
- ▶ 전문가들 "개인정보 비식별 조치하
- ▶ [취재수첩] 데이터3법, 정말로 개인
- ▶ 국회 행안위 2019년 결산... "개인
- ▶ 개인정보 보호 기본계획 발표... "데

[데이터3법③] 결국 '데이터 활용'에 무게심는 이유... "더 나은 세상 만들 것"

2020.03.04 09:42:00 / 이종현 bell@ddaily.co.kr

관련기사

- ▶ [데이터3법②] "보호에서 활용으로"
- ▶ [데이터3법①] 지금 생각해도 오싹
- ▶ "써도 되나요?"... 데이터3법 통과
- ▶ 전문가들 "개인정보 비식별 조치하
- ▶ [취재수첩] 데이터3법, 정말로 개인

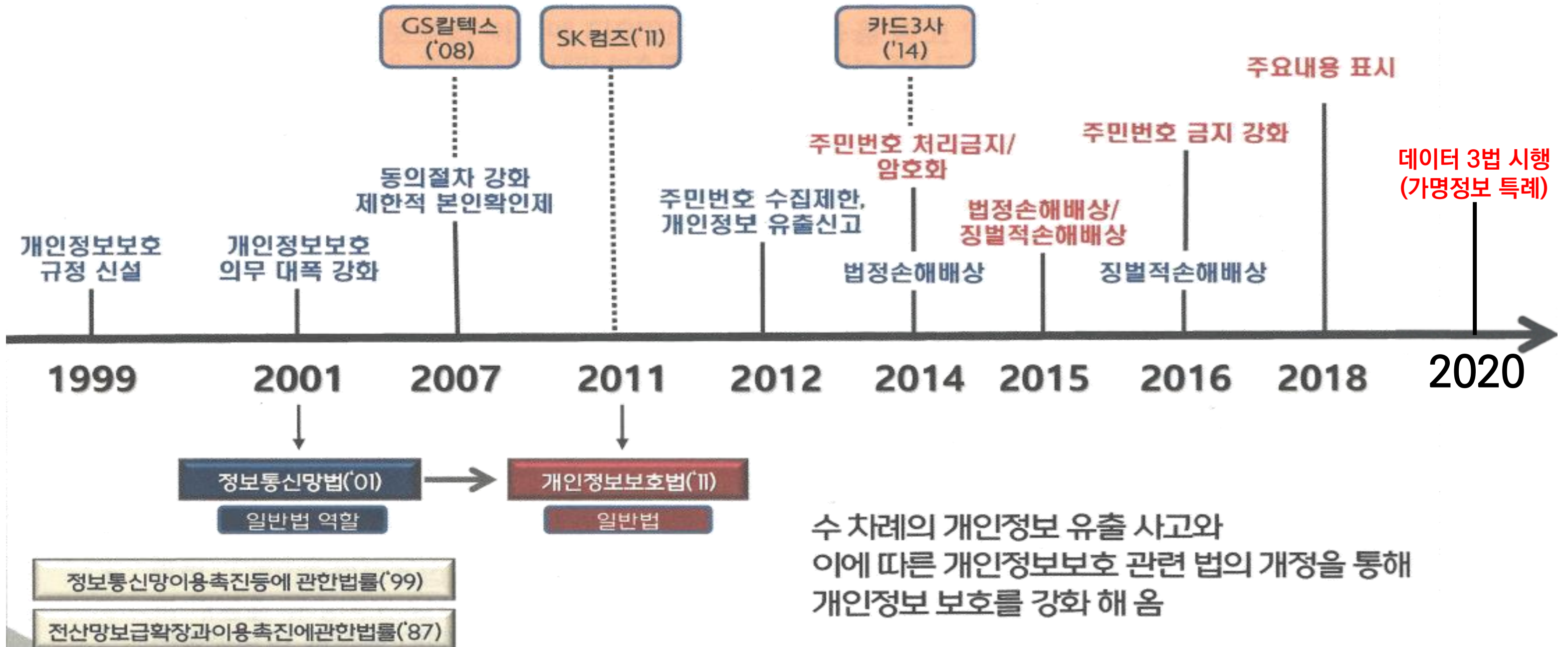
[데이터3법④] 걸음마는 댔지만... 데이터3법, 후속 과제는?

2020.03.05 09:13:38 / 이종현 bell@ddaily.co.kr

관련기사

- ▶ [데이터3법③] 결국 '데이터 활용'에 무게심는 이유... "더 나은 세상 만들 것"
- ▶ [데이터3법②] "보호에서 활용으로"... 데이터경제 시대로 간다
- ▶ [데이터3법①] 지금 생각해도 오싹... '개인정보보호 강화' 10년의 역사
- ▶ "써도 되나요?"... 데이터3법 통과에도 업계는 '뜨뜻미지근'
- ▶ 전문가들 "개인정보 비식별 조치하면 산업 활용 문제없어"...시민단체 우려에 반론
- ▶ [취재수첩] 데이터3법, 정말로 개인정보보호 약화되나?
- ▶ 개인정보 보호 기본계획 발표... "데이터 활용하려면 안전한 보호가 전제"
- ▶ 데이터 3법 통과로 '비식별 솔루션' 시장 청신호... "이미지·영상은 아직"
- ▶ 시민단체의 신중론 "데이터 경제 이전에 데이터 유출 처벌 강화해야"
- ▶ 속도내는 데이터 경제... "AI 시대로 한발 더"

■ 개인정보보호법 관계 법령 변천사



※ (출처) 권현영, 고려대학교 정보보호대학원, “개인정보보호법의 이해 및 동향”

■ 개인정보보호 관련 제도와 개인정보보호 조치




구분	개인정보보호 조치	개인정보보호 관련제도
인식단계 (~2010년)	<ul style="list-style-type: none"> ▪ 개인정보 필터링 ▪ 개인정보 노출진단 ▪ 데이터베이스 암호화 	<ul style="list-style-type: none"> ▪ [1995.01] 신용정보보호법 ▪ [1999.02] 정보통신망법 개정 (개인정보 관련 규정 신설)
관리단계 (~2015년)	<ul style="list-style-type: none"> ▪ 개인정보 생명주기 관리 ▪ 개인정보 검출/격리 	<ul style="list-style-type: none"> ▪ [2011.09] 개인정보보호법 <ul style="list-style-type: none"> - 표준 개인정보 보호지침 - 개인정보 영향평가에 관한 고시 - 개인정보의 안전성 확보조치 기준 - 정보보호 및 개인정보보호 관리체계 인증 등에 관한 고시
활용단계 (~현재)	<ul style="list-style-type: none"> ▪ 개인정보 접속기록 관리 ▪ 비정형 데이터 암호화 ▪ 개인정보 가명처리/익명처리 ▪ 가명정보 활용 및 결합 	<ul style="list-style-type: none"> ▪ [2016.06] 개인정보 비식별 조치 가이드라인 ▪ [2018.05] EU GDPR ▪ [2018.11] 데이터3법 개정안 발의 ▪ [2020.08] 데이터3법 개정안 시행

■ 개인정보 관련 개념과 활용범위

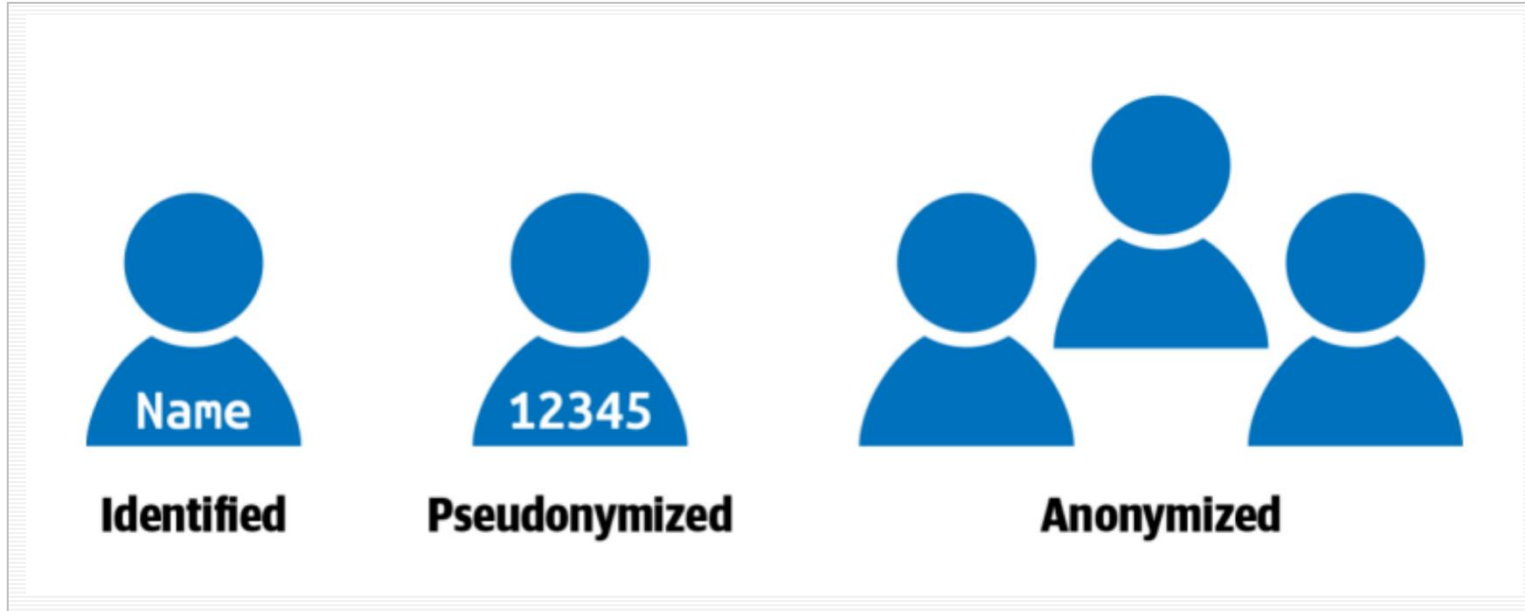
구분	개념	활용가능 범위
개인정보	<u>살아있는</u> 특정 개인에 관한 정보, 개인을 알아볼 수 있게 하는 정보	사전적이고 구체적인 동의를 받은 범위 내 활용 가능
가명정보	<u>추가정보</u> 사용 <u>없이</u> 는 특정 개인을 알아볼 수 없게 조치한 정보	다음 목적에 동의 없이 활용 가능(EU GDPR 반영) 1) 통계작성(상업적 목적 포함) 2) 연구(산업적 연구 포함) 3) 공익적 기록보존 목적 등
익명정보	<u>다른정보</u> 를 결합하여도 더 이상 개인을 알아볼 수 <u>없게</u> (복원 불가능할 정도로) 조치한 정보	개인정보가 아니기 때문에 제한없이 자유롭게 활용

3. 개인정보 가명·익명처리 국내 동향

■ 개인정보 비식별 처리 상태

 Identified	 Pseudonymized	 Anonymized
상태	설명	
식별 상태	데이터에 직접 식별자가 포함된 상태로, 이름이나 주소가 여기에 해당됨	
가명 상태	데이터 보호 메커니즘으로 직접 식별자를 토큰, 가짜 데이터, 삭제 기법 등 다양한 방법으로 제거된 상태로, <u>추가적인 정보가 있어야 재식별이 가능한 상태</u>	
익명화 상태	주어진 데이터 공개 모델에 대한 직접 및 간접 식별자를 제거하여 <u>데이터를 식별할 수 없다는 합리적인 확신을</u> 제공하는 상태	

■ 개인정보 / 가명정보 / 익명정보



구분	개념 및 활용가능 범위
개인 정보	특정 개인에 관한 정보, 개인을 알아볼 수 있게 하는 정보 → 사전적이고 구체적인 동의를 받은 범위 안에서 활용 가능
가명 정보	추가정보의 사용 없이는 특정 개인을 알아볼 수 없게 조치한 정보 → 통계작성(상업적 목적 포함), 연구(산업적 연구 포함), 공익적 기록보존
익명 정보	더 이상 개인을 알아볼 수 없게(복원 불가능할 정도로) 조치한 정보 → 개인정보가 아니기 때문에 제한없이 자유롭게 활용

개인정보 가명·익명처리 국내외 동향

3. 개인정보 가명·익명처리 국내 동향

개인(원본)정보										※ (참고) 금융보안원, 전자금융과 금융보안 (제20호, 2020-2Q)									
이름	생년월일	핸드폰	직장전화	자택전화	주소	직업	가족	예금 평균잔액	대출액										
홍길남	67년 1월 3일	010-9999-3333	055-747-0088	055-2345-6789	경남 창원시 마산구 합포동 123	국회의원	배우자 아들1 딸1	4,567,900원	45,896,789원										
(잘못된) 가명정보 예시										※ 이름, 생년월일 등을 삭제하고 핸드폰 번호를 암호화하였으나, '창원시 마산구에 주소를 둔 국회의원'이라는 극소수 사람에 관한 정보로 개인 식별 가능									
이름	생년월일	핸드폰	직장전화	자택전화	주소	직업	가족	예금 평균잔액	대출액										
- * 삭제	- * 삭제	ajeejkc93 * 암호화	- * 삭제	- * 삭제	경남 창원시 마산구 * 범주화	국회의원	배우자 아들1 딸1	4,567,900원	45,896,789원										
(제대로 된) 가명정보 예시										※ 사실상 해당 정보만으로는 개인식별 불가능, 다만, 핸드폰 번호 암호화 알고리즘을 알면 복원 가능하므로 가명정보에 해당									
이름	생년월일	핸드폰	직장전화	자택전화	주소	직업	가족	예금 평균잔액	대출액										
- * 삭제	- * 삭제	ajeejkc93 * 암호화	- * 삭제	- * 삭제	경남 창원시 마산구 * 범주화	- * 삭제	배우자 아들1 딸1	4,567,900원	45,896,789원										
익명정보 예시										※ 그 자체로, 또는 다른 정보와 결합해서도 개인을 식별할 수 없음									
이름	생년월일	핸드폰	직장전화	자택전화	주소	직업	가족	예금 평균잔액	대출액										
- * 삭제	- * 삭제	- * 삭제	- * 삭제	- * 삭제	경남 창원시 마산구 * 범주화	- * 삭제	가족 3명	400만 원 ~500만 원	4,000만 원 ~5,000만 원										

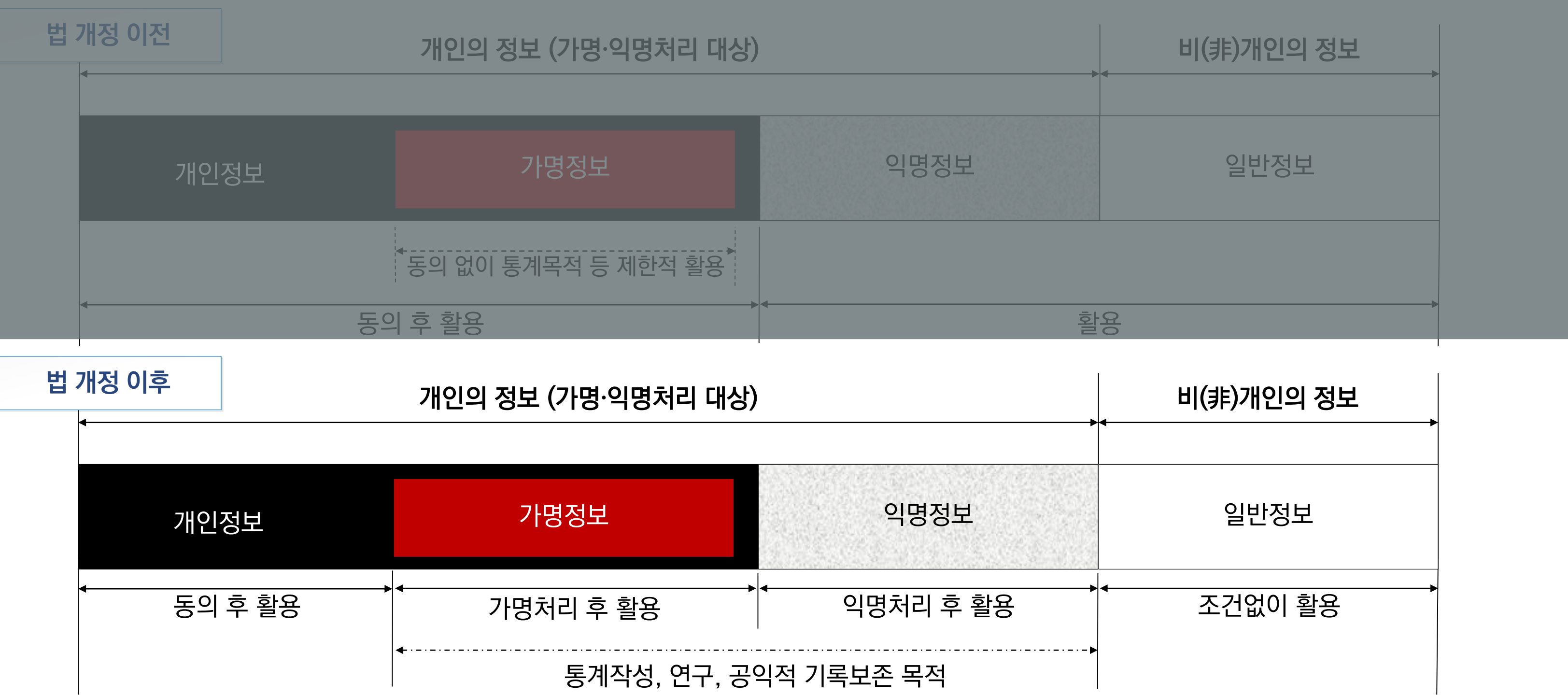
3. 개인정보 가명·익명처리 국내 동향

■ 개인정보 비식별 처리 대상

대상	데이터 유형	데이터 특징
정형데이터	<ul style="list-style-type: none"> ■ 텍스트 데이터 (CSV) ■ 엑셀 데이터 ■ 데이터베이스 유형 	<ul style="list-style-type: none"> ■ 개인정보 비식별 조치 가이드 라인의 직접적인 적용 대상 ■ 프라이버시 보호 모델 적용
반정형데이터	<ul style="list-style-type: none"> ■ XML 포맷 데이터 ■ JSON 포맷 데이터 ■ NoSQL 계열(Mongo DB, Hadoop 등) 	<ul style="list-style-type: none"> ■ 개인정보 검출 및 필터링에 집중
비정형데이터	<ul style="list-style-type: none"> ■ 텍스트 데이터 (Free Text Form) ■ 음성 데이터 ■ 영상/이미지 데이터 	<ul style="list-style-type: none"> ■ 한글, 워드, PDF 등 오피스 문서 ■ 음성 파일, CCTV 등 ■ 이미지 필터링(OCR), 인페인팅 등 별도 가명/익명처리 조치 필요

3. 개인정보 가명·익명처리 국내 동향

■ 개인정보보호 기준에 의한 데이터의 분류와 활용



개인정보 가명·익명 처리 국내외 동향

End of Document

감사합니다.

