데이터보안과 동형암호 기술의 활용

2023. 5. 11

서울대 이석윤 교수

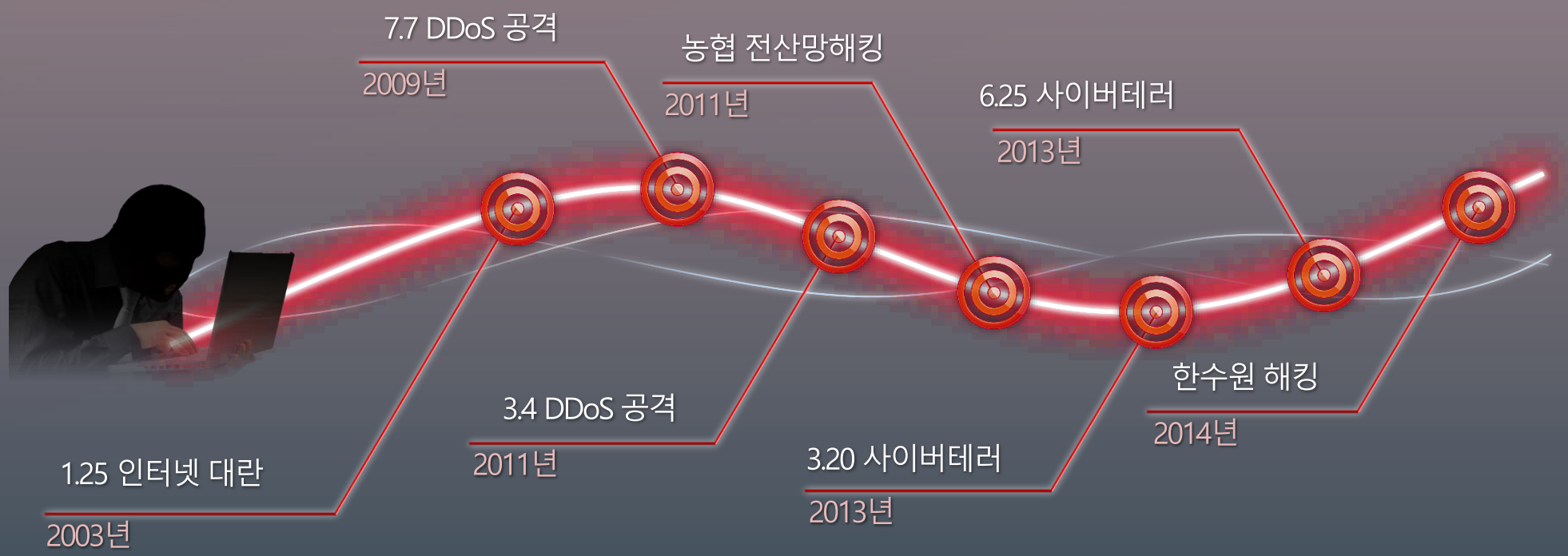
2023 HACKTHEON SEJONG

4차산업혁명,

데이터 활용과 보호간 딜레마

2023
HACKTHEON
SEJONG

대형 해킹사고!!


7.7 DDoS 공격
2009년
농협 전산망해킹
2011년
6.25 사이버테러
2013년
1.25 인터넷 대란
2003년
3.4 DDoS 공격
2011년
3.20 사이버테러
2013년
한수원 해킹
2014년

망분리 정책 불가피성!!

방화벽, 침입탐지/침입방지시스템, 안티 바이러스, 서버 접근 통제 등 많은 정보보호제품 설치 운용 중

부문별, 기관별 보안관제센터 운영/ 정기 취약점 분석평가(년 1회) 및 정보보호관리체계(ISMS) 인증 등

많은 기술적 관리적 보안 대책에 불구, 해킹사고 발생

➡ 업무망과 인터넷망의 물리적 분리 운영 고착화
(국가기관, 금융사)

초연결사회 딜레마 !!

4차 산업혁명시대, 사람과 사물이 인터넷을 매개로 연결돼 있어 정보의 생성과 수집, 공유, 활용이 이루어지는 초연결사회로 진화

물리적 망 분리 보안대책은 핀테크, 빅데이터, AI 등 4차산업 발전에 지장 초래

"인터넷을 매개로 사람과 사물에 대한 보호와 데이터의 공유와 활용이 이루어져야 한다"는 초연결사회의 기본원칙과 배치

그러나, IoT시대에서 대형 보안사고 발생시 지금까지 겪었던 보안사고와 비교할 수 없는 더 큰 사회적 경제적 혼란 초래 우려

데이터 활용과 보호 동시 필요

Big data 및 AI 시대의 도래로 개인정보 등 민감자료 분석 수요는 향후에도 폭발적으로 증가할 것으로 예측되나, 이와 더불어 데이터 유출사고 역시 급증

개인정보 분석 수요의 증대

Big data, AI의 발전으로 개인정보 분석 수요는 크게 확대됨

개인정보
(성별, 연령, 주소, 연락처, 소득, 진료기록, 소비패턴 등)

Big data, 인공지능 등 '지능정보 기술'에 기반한 개인정보 분석 수요 급증 ("4차 산업혁명")

Big data　AI　클라우드

AI 기반 데이터 분석 시장 개화 (AI-as-a-Service)

IBM　Microsoft　intel　amazon　Google

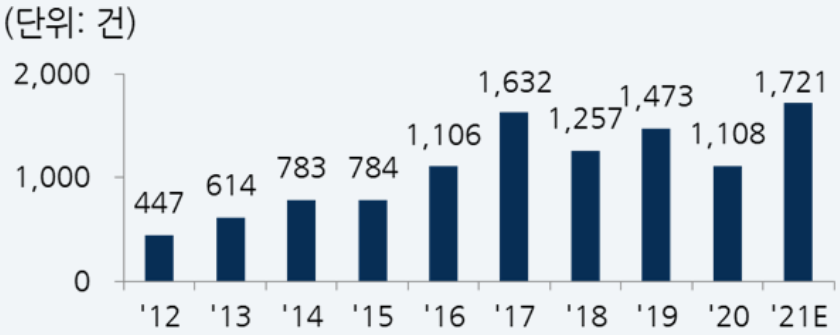
금융·의료·정부기관 등 개인정보 분석 활동 증가

금융기관　의료기관　정부기관

개인정보 유출 사고의 증가

단, 현 암호체계의 근원적 문제로 정보 유출 사고는 크게 증가 중

미국 Data 유출 사고 건 수



- 외부 해킹, 관리실수로 인한 유출, 내부자에 의한 고의 유출 등으로 인한 개인정보 유출 증가

업종별 Data 유출 비중 및 주요 사례

금융기관 15%　의료기관 18%　정부기관 15%

- Wells Fargo, Capital One의 대출 정보 2,400만건 유출
- Morgan Stanley는 '16, '19년 1,500만 고객 정보 유출

2023 HACKTHEON SEJONG

클라우드시스템 필요성

민간 클라우드를 활용한 경우 빠른 서비스 제공은 물론, 적은 비용으로 안정적인 서비스를 제공(매킨지 리포트, 2020년)

출처 : McKinsey & Company, 2020, https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-public-sector-tech-leaders-can-speed-up-the-journey-to-the-cloud

미국 공공분야 클라우드 적용사례

법무부(DOJ), 국토안보부(DHS) 등은 AWS, 국립보건원(NIH) 등은 구글

AWS

재향 군인회는 AWS GovCloud(미국)에 대해 FISMA High Authority to Operate(ATO)를 발행해서 리전을 사용하여 미국 재향 군인들에게 중요한 환자 데이터를 저장하고 보호합니다. 자세히 알아보기 »

GSA의 18F가 구축한 Cloud.gov는 다른 정부 기관이 기술 제품을 구축, 구매 및 공유하는 데 도움을 제공하며 스스로 실행하는 데 필요한 FedRAMP 규정 준수 작업을 최소화합니다. 자세히 알아보기 »

법무부는 미션 크리티컬 워크로드, 데브테스트 및 고급 기능 제공을 위해 AWS GovCloud(미국)를 활용합니다. 자세히 알아보기 »

Defense Digital Service는 미국을 관리합니다. AWS GovCloud(미국)에서 200개 이상의 전용 호스트와 1,000개의 개별 가상 머신을 안전하게 실행하는 공군의 차세대 GPS 위성 운영 제어 시스템입니다. AWS GovCloud(미국)에서 DoD IL5 워크로드에 대해 자세히 알아보고 확인하세요.

미국 재무부는 AWS GovCloud(미국)에서 디지털 트랜스포메이션을 지원하는 동시에 미션 어슈어런스를 제공합니다. 자세히 알아보기 »

미국 국토안보부의 HSIN 정보 공유 플랫폼은 AWS GovCloud(미국)에서 안전하게 활성화된 FedRAMP High 시스템입니다. 자세히 알아보기 »

캔자스 주와 AWS 파트너인 PayIt은 60일 이내에 온라인 및 모바일 라이선스 갱신 앱을 배포하여 정부 서비스에 대한 시민 경험을 향상했습니다. 자세히 알아보기 »

NASA 제트 추진 연구소는 AWS GovCloud(미국)로 거버넌스, 보안 및 규정 준수를 개선하면서 혁신합니다. 자세히 알아보기 »

Google

미국 공공분야 클라우드 적용사례

미 육군, 미시간주 등은 IBM, 상공회의소(Chamber of Commerce) 등은 MS

IBM

Microsoft

SEJONG

미국 정보기관도 클라우드 서비스 도입 운용?

- 어느 수준까지 민간 클라우드 활용?

[국방부]

지디넷코리아

美 국방부, 10조 규모 클라우드 프로젝트 연내 마무리

미국 국방부가 오는 12월 90억 달러(약 10조9천억원) 규모 클라우드 전환 프로젝트 계약을 체결할 계획이다. 논란 끝에 취소됐던 제다이(JEDI)를 대체...

Mar 30, 2022

[CIA]

Fox Business

CIA awards cloud computing contract worth billions to firms including Amazon, Microsoft, Google

CIA awards cloud computing contract worth billions to firms including Amazon, Microsoft, Google. Contract will go to Amazon Web Services,...

Nov 20, 2020

[NSA]

지디넷코리아

AWS, NSA 클라우드 계약 수주...MS 즉각 항의

미국 정보기관들이 클라우드 도입을 잇따라 추진하는 가운데, 아마존웹서비스(AWS)가 미국 국가안보국(NSA) 클라우드 사업을 수주했다.

Aug 11, 2021

미국 공공분야 클라우드 보안대책

- FedRAMP(Federal Risk and Authorization Management Program)

기밀성, 무결성, 가용성 측면에서 정보자산에 미치는 위험수준을 3등급(High, Moderate, Low)으로 구분하고 3가지 보안목표(Security Target)별 통제되는 항목 체계적으로 명시

	POTENTIAL IMPACT		
Security Objective	**LOW**	**MODERATE**	**HIGH**
Confidentiality Preserving authorized restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. [44 U.S.C., SEC. 3542]	The unauthorized disclosure of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized disclosure of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Integrity Guarding against improper information modification or destruction, and includes ensuring information non-repudiation and authenticity. [44 U.S.C., SEC. 3542]	The unauthorized modification or destruction of information could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The unauthorized modification or destruction of information could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.
Availability Ensuring timely and reliable access to and use of information. [44 U.S.C., SEC. 3542]	The disruption of access to or use of information or an information system could be expected to have a **limited** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **serious** adverse effect on organizational operations, organizational assets, or individuals.	The disruption of access to or use of information or an information system could be expected to have a **severe or catastrophic** adverse effect on organizational operations, organizational assets, or individuals.

Impact System Level	FedRAMP Controls
Low	125
Moderate	326
High	421

각 정부기관마다 별도 유지해 왔던 보안요건 및 인증제도를 FedRAMP 로 통합

모든 정부기관에 일관성있게 적용 가능

미국 FedRAMP와 한국 CSAP 비교

KISA 운영 클라우드서비스 보안인증제도(CSAP : Cloud Service Assurance Program)은 클라우드 서비스 유형별로 단순히 통제항목 명시

Control Type	Low	Moderate	High
1. Access Control	11	43	54
2. Awareness Training	4	5	7
3. Audit and Accountability	10	10	30
4. Security Assessment and Authorization	9	16	16
5. Configuration Management	11	26	36
6. Contingency Planning	6	23	35
7. Identification and Authentication	15	27	32
8. Incident Response	7	17	26
9. Maintenance	4	12	14
10. Media Protection	4	10	12
11. Physical and Environmental Protection	10	20	26
12. Planning	3	6	6
13. Personnel Security	8	9	10
14. Risk Assessment	4	10	12
15. System and Services Acquisition	6	22	26
16. System and Communications Protection	10	32	39
17. System and Information Integrity	7	28	38
sum	129	316	419

(a) FedRAMP 통제 타입 및 영향 수준별 통제항목 수 (2020년 4월기준)

통제분야	IaaS	SaaS 표준	SaaS 간편	DaaS
1. 정보보호 정책 및 조직	5	5	2	5
2. 인적보안	12	5	2	8
3. 자산관리	10	3	0	10
4. 서비스 공급망 관리	4	3	0	4
5. 침해사고 관리	7	7	2	7
6. 서비스 연속성 관리	7	6	2	7
7. 준거성	4	3	1	4
8. 물리적 보안	12	0	0	12
9. 가상화 보안	10	6	1	7
10. 접근통제	10	10	5	10
11. 네트워크 보안	6	5	2	6
12. 데이터 보호 및 암호화	10	8	4	10
13. 시스템 개발 및 도입 보안	12	10	2	12
14. 공공부문 추가 보안요구 사항	8	7	7	8
총계	117	78	30	110

(b) CSAP 통제 분야 및 클라우드 서비스 유형별 통제항목 수

출처 : 윤대훈, "공공기관의 민간클라우드 도입 활성화를 위한 전제 조건인 클라우드서비스 보안 인증 - FedRAMP *vs* CSAP," 2022

클라우드시스템을 활용하지 않았다면 ?

- 코로나19 팬데믹, ‘20년 3월 사상 초유의 온라인 개학
  - ✓ 기존 온라인 시스템의 수용능력 한계 (수천배 확장?)
  - ✓ 몇 주내 600만명이 동시 온라인 수업 교육에 필요한 컴퓨팅서비스 자원 제공

- 코로나19 백신 접종 사전예약시스템
  - ✓ 초당 수천명이 접속하는 사전 예약시스템의 트래픽 병목현상 해결
  - ✓ 시간당 200만명까지 예약 가능한 사전예약 시스템 구축

클라우드서비스의 사용은 증가하고 있으나 기밀성, 무결성, 가용성 등
보안문제 우려

공공분야 클라우드컴퓨팅 보안규정 개정

공공분야 클라우드 이용시 데이터의 가치 등 정보시스템의 중요도를 분류 (3등급), 차등화된 보안대책을 적용하는 방향으로 국가정보보안기본지침 개선(2023.1.31)

국가 정보보안기본지침 개정

제41조(클라우드컴퓨팅 보안) ② 각급기관의 장은 민간 클라우드컴퓨팅서비스를 이용하고자 할 경우 다음 각 호에 해당하는 사항을 준수하여야 한다.

2. 다음 각목의 요건에 따라 일반 이용자용 서비스와 영역이 분리되어 제공되는 서비스(이하 "공공 전용(專用) 민간 클라우드"라 한다)의 이용

가. 영역 분리는 일반 이용자용 서비스와 데이터 및 프로세스 등의 간섭없이 국가정보원 및 이용기관의 보안관제, 사고조사, 예방보안활동 유지를 위한 제반 환경을 만족해야 함

나. 영역 분리는 '시스템 중요도'에 따라 물리적 또는 논리적으로 구현

다. '시스템 중요도' 분류는 [별표4]의 기준 준용 <신설 2023.1.31.>

클라우드 도입시 시스템 중요도 분류 기준

국가기관은 클라우드서비스 도입시 시스템의 특성, 정보의 중요도 및 파급영향 등 시스템 보안 중요도를 판단하여 상/중/하 등급을 분류

등급	분류기준		영역분리
상	파급영향	해당 정보시스템에 대한 침해는 운영기관, 자산 및 개인에게 치명적 악영향을 미칠 수 있음	물리적
	분류기준	국가 중대 이익(안보, 국가안전, 국방, 통일, 외교 등), 수사·재판 등 민감정보를 포함하거나 행정 내부업무 등을 운영하는 시스템	
중	파급영향	해당 정보시스템에 대한 침해는 운영기관, 자산 및 개인에게 심각한 영향을 미칠 수 있음	물리적
	분류기준	비공개 업무자료를 포함 또는 운영하는 시스템	
하	파급영향	해당 정보시스템에 대한 침해는 운영기관, 자산 및 개인에게 제한적인 영향을 미칠 수 있음	물리적 또는 논리적
	분류기준	개인정보를 포함하지 않고 공개된 공공데이터를 포함 또는 운영하는 시스템	

2023 HACKTHEON SEJONG

국내 공공분야 클라우드 도입요건/절차

- 클라우드 서비스 도입요건
  - 사전 검증 : KISA의 보안인증제도(CSAP)에 의해 인증받은 클라우드 서비스 대상으로 국정원에 도입요건 확인 요청
  - 사후 검증 : 수요기관이 상급기관 또는 전문평가기관에게 안정성 확인후 상급기관 또는 전문평가기관이 국정원에 보안기준 적합여부 확인 요청

- 클라우드 서비스 도입 절차
  - 정보화사업 계획 수립: 각급기관은 국정원이 도입 요건을 확인한 민간 클라우드서비스(NCSC 홈피 게시) 대상 클라우드 환경을 통한 데이터 유출방지 방안 등 자체 보안대책 수립
  - 보안성 검토 : 국가 정보보안지침에 명시된 보안요구사항과 국가 클라우드 컴퓨팅 보안 가이드라인에 명시된 보안기준을 마련, 국정원에 의뢰

    * 시스템 중요도 분류 체크리스트 작성, 클라우드 영역별 보안 기본원칙 준수 등

행안부, 정보시스템 등급별 보안관리제도 사례(2019)

정보시스템 등급분류

① 국가. 사회적 중요도 평가 → ② 시스템 특성 반영 → ③ 기관 신뢰도 평가를 종합, 정보시스템 등급 분류

- 정보시스템이 취급하는 정보 또는 서비스의 중요도에 따라 5개 등급 부여
  - 1등급 : 국가 존립관련 국방, 외교, 통일분야 시스템
  - 2등급 : 지진, 항공, 자연재해 등 국민생명관련 시스템
  - 3등급 : 개인정보, 건강증진 등 국민건강관련 시스템
  - 4등급 : 업무 수행을 기관 단순 정보시스템
  - 5등급 : 공개서비스로 국민에게 미미한 영향을 주는 시스템

행안부 정보시스템 등급별 보안관리 제도사례(2109년)

- 보안관리 수준 3단계(H, M, L)로 차별, 등급별 차별화된 보안관리 수행 (예시)

기준	1~2등급 (H)	3등급 (M)	4~5등급 (L)
취약점 점검	정보통신기반시설 취약점 점검항목(**상, 중**)에 따라 **외부점검**(연1회), **대내외 웹서비스 모의침투시험**	정보통신기반시설 취약점 점검항목(**상**)에 따라 **자체점검**(연1회), **대외 웹서비스 모의침투시험**	정보통신기반시설취약점 점검항목(**상**)에 따라 **자체점검**(연1회), **대외 웹서비스** **취약점 진단도구**로 진단
취약점 조치	즉시조치 불가능한 취약점 모니터링(**분기 1회**)	즉시조치 불가능한 취약점 모니터링(**반기 1회**)	즉시조치 불가능한 취약점 모니터링(**연 1회**)
저장매체 불용처리	**물리적** 완전파쇄	**자기적 충격**에 의한 파쇄	**완전포맷 3회 이상** 수행
로그 기록 검토	로그검토 **(월 1회)**	로그검토 **(분기 1회)**	로그검토 **(반기 1회)**
로그 위변조 방지	로그 접근권한 최소화, 로그 별도백업**(3년 이상)**	로그 접근권한 최소화, 로그 별도백업**(6개월 이상)**	

2019.6.27, 부처 정보보호 및 시스템 운영담당 등에게 등급분류 및 등급별 보안관리기준 설명(80분)

정보시스템 등급별 분류에 따른 차별화된 보안관리 시행 가능?

2023 HACKTHEON SEJONG

공공분야 클라우드 서비스 보안제도 개선방향

국내 클라우드 인증프로그램 (CSAP)은 정보시스템의 기밀성, 무결성, 가용성 (3가지)에 미치는 위험수준을 모두 고려한 차등화된 보안통제를 제시토록 개선

기존 네트워크 보안대책 강화와 함께 동형암호, 기밀계산 (Confidential Computing) 등 새로운 보안기술 도입 활용

- 정보시스템의 중요도 분류와 별도의 데이터서버내 민감 데이터 공유 확대
  - 데이터의 중요도, 정보시스템 무단 접근 발생 위험도 등에 따라 시스템을 재구성하거나
  - 저등급 데이터부터 순차적으로 별도의 데이터서버에 저장하고 보안대책 마련후 클라우드 서비스 이용
  - 보안의 중요도가 매우 높은 외교, 국방 등 국가안보기관 및 원자력, 전력 등 제어시스템 을 운영하는 정보통신기반시설 운영기관의 망분리 정책은 유지

✓ 현 보안규정은 국가배후 사이버공격에 대비한 최소한의 대책이라는 사실 인식

동형암호 기술

2023
HACKTHEON
SEJONG

암호기술의 발전

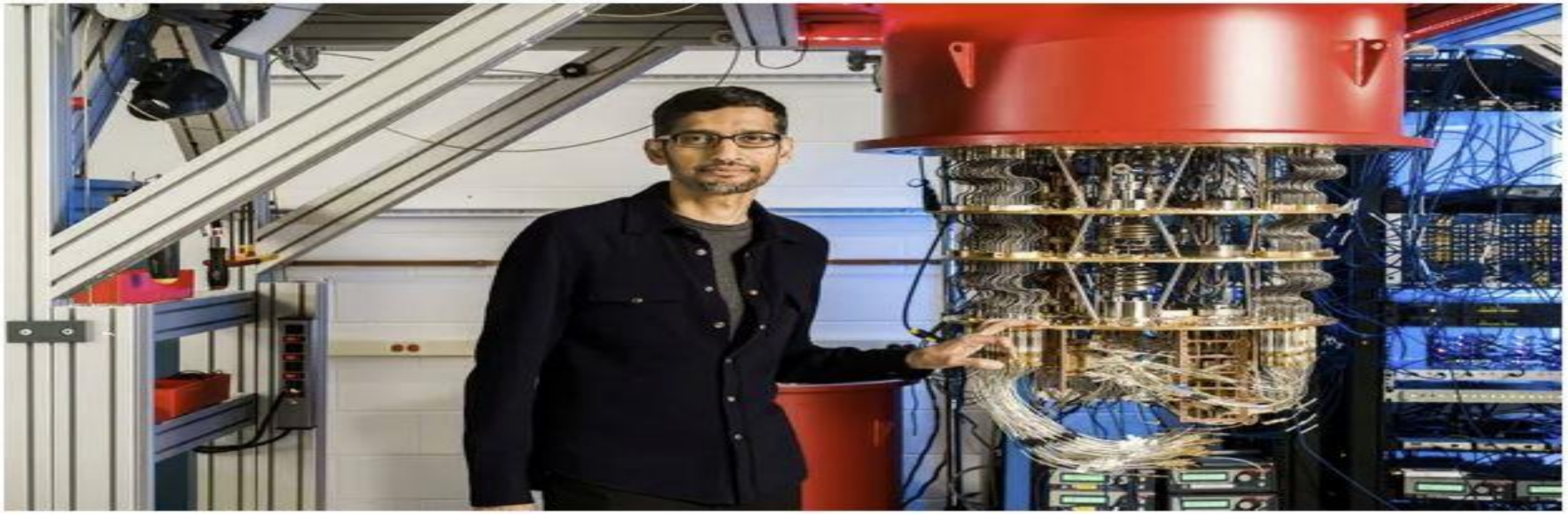
데이터 저장/전송 보호를 넘어 사용중에도 보호하는 것이 필요

1세대 암호	2세대 암호	3세대 암호	4세대 암호
			
PASSWORD(인증기술)	대칭키 암호(데이터암호화)	공개키 암호(키 암호화)	동형암호(키보호 암호)
	블록암호(AES)	RSA-암호화	암호화 상태 계산이 가능한 암호

양자컴퓨터 실용화 가능성 (2019년 구글 발표)

Google claims it has achieved 'quantum supremacy' – but IBM disagrees

Task that would take most powerful supercomputer 10,000 years 'completed by quantum machine in minutes'



▲ Sundar Pichai, pictured with the Sycamore Quantum processor, compared the feat to building the first rocket to reach space. Photograph: Reuters

출처: Guardian, 2019. 10.23. https://www.theguardian.com/technology/2019/oct/23.

국내 양자컴퓨터 개발 계획

SCIENCE Chosun

2030년 500큐비트 양자컴퓨터 만든다 신기술 로드맵 공개

신성장 4.0 전략 3대 분야 15대 프로젝트 발표
양자 기술개발 참여 기업에 세액공제·정책금융 지원
UAM 특별법 제정, 혁신형 SMR 기술개발도 본격화

이종현 기자

입력 2023.02.20 09:00

미래 산업 경쟁력의 핵심 기술로 꼽히는 양자(量子) 기술 육성에 정부가 본격적으로 나선다. 2030년까지 500큐비트 양자컴퓨터를 개발한다는 계획이다. 민간 도심항공 모빌리티(UAM)를 위한 특별법을 제정하고, 5000억원 규모의 K-바이오백신 펀드도 조성한다. 혁신형 소형모듈원자로(i-SMR) 기술개발에 본격 착수해 2028년까지 인가를 획득한다는 구상도 공개했다.

정부는 20일 비상경제장관회의를 열고 신성장 4.0 전략의 구체적인 추진 과제를 발표했다. 지난해 12월 21일 윤석열 대통령이 주재한 제12차 비상경제민생회의 때 발표된 신성장 4.0 전략은 미래기술 확보, 디지털 전환, 전략산업 초격차 확대를 중심으로 한 대규모 민관 협력 프로젝트다.

신성장 4.0 전략

- 2023년말 20 큐비트 양자컴퓨터 시연
- 2026년말 50 큐비트 양자컴퓨터 개발
- 2030년 500 큐비트 양자컴퓨터 개발

* IBM : 2021년 127 큐비트 양자컴퓨터 공개
2023년 1,121 큐비트 양자컴퓨터 개발 목표

출처 : 조선일보 (https://biz.chosun.com/science-chosun/science/2023/02/20/E5ZFMMIYP5FENNUOAKDVDQJY4U/)

양자컴퓨터의 암호해독 위협

양자컴퓨터, RSA 공개키 암호의 실시간 해독 가능성 증대

대칭키 암호는 암호키 길이 2배 증가해야 안전성 가능

RSA 2,048비트 해독예상시간	슈퍼컴퓨터: 수십억년 이상
	양자컴퓨터: 수백초 이내

알고리즘	암호	영향
Shor	공개키	더 이상 안전하지 않음
Grover	대칭키	키 사이즈 2배 이상 증가 필요
	해시	암호 알고리즘의 출력 길이 증가 필요

중국 암호전문가, RSA 암호해독 주장 논란?(2023년)

양자컴퓨터가 벌써 RSA 암호화 알고리즘을 깼다고?

중국의 암호화 전문가 수십 명이 논문을 하나 발표해 업계의 이목을 끌었다. 이미 개발된 양자컴퓨터를 가지고 RSA 암호를 쉽게 해독할 수 있다는 내용이다. 하지만 전문가들의 '독후감'은 그리 호의적이지 않다.

[보안뉴스 문정후 기자] 지난 주 양자컴퓨팅 기술 때문에 RSA-2048이라는 암호화 알고리즘이 무력해질 수 있다는 내용의 논문이 나와 IT 업계의 큰 관심을 끌었다. 논문 저자들은 획기적인 발견이라고 주장했지만, 해당 분야 전문가들의 관점은 사뭇 다른 것으로 보인다.

[이미지 = utoimage]

논문의 제목은 "초전도 양자 프로세서 상에서 선형 이하 자원으로 정수 인수분해 하기(Factoring integers with sublinear resources on a superconducting quantum processor)"이며, 일반적인 양자 알고리즘을 통해 아무 숫자의 소인수를 빠르게 찾아내는 방법을 제시하고 있다. 소인수를 발견해 내는 건 기존 공공키 기반 암호화에 있어서 중추적인 부분이라고 할 수 있으며, 따라서 이 과정의 속도를 높이면 사실상 기존 암호화 알고리즘을 무력화시킬 수 있게 된다.

(출처 : 보안뉴스 https://www.boannews.com/media/view.asp?idx=113431)

4세대 암호: 동형암호(Homomorphic Encryption)

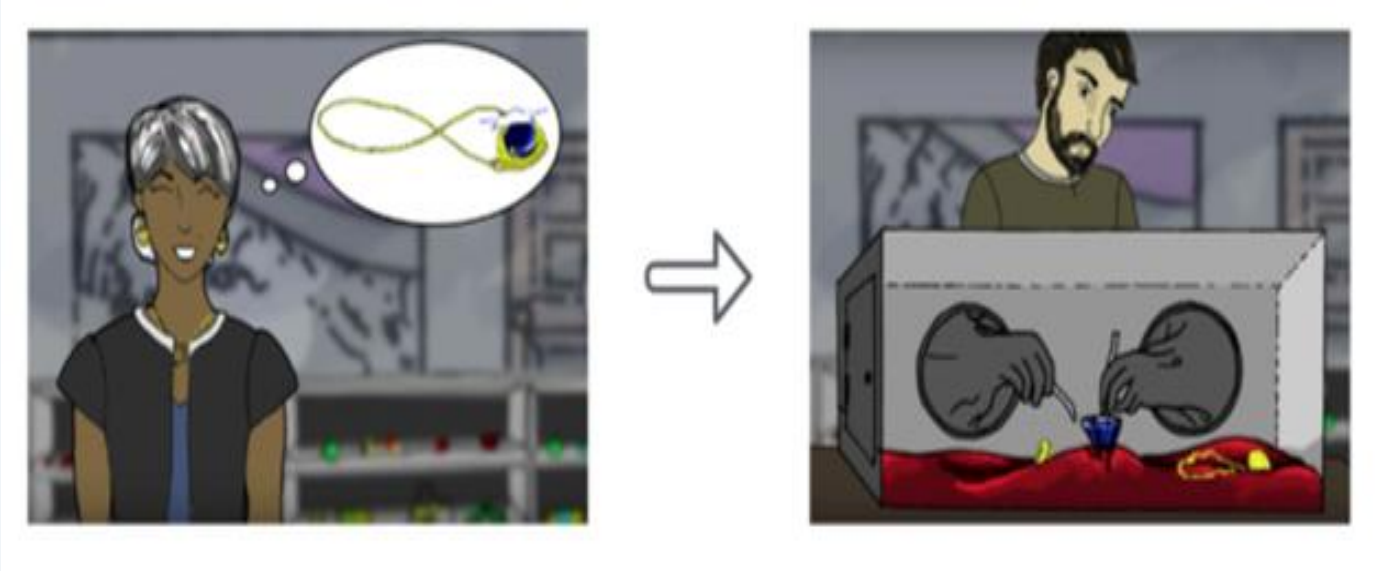
- 동형암호는 데이터를 **암호화후 복호화 없이 연산**하는 혁신기술
  - 기존의 암호는 복호화후 연산하고 재 암호화. 이 과정에서 비밀키나 평문 노출 위험
  - '완벽한 하인': 데이터 **보호**와 **활용**이 동시에 가능

- 안전성이 보장되는 **최고의 프라이버시보존 데이터 분석** 기술
  - 개인정보보호 규정(GDPR, CCPA, 데이터3법 등)을 준수하며 데이터 분석 수행가능

* 동형암호 : 양자안전암호

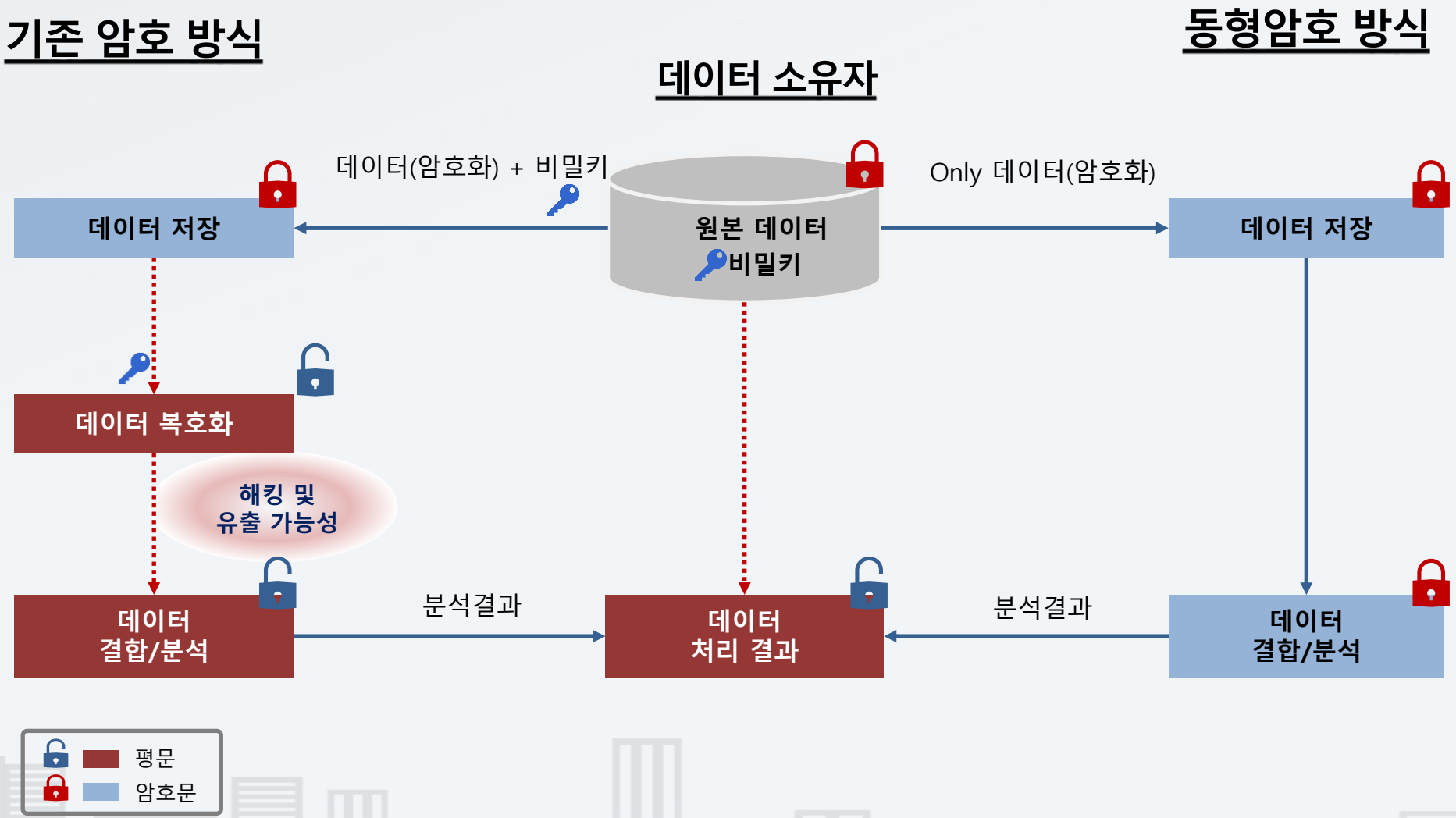
개념 설명:
- 보석상이 분실을 우려해 원석을 금고에 넣어 가공. 세공이 끝난 후 보석을 꺼냄

- 데이터를 프라이버시 보존을 위해 암호화하여 가공. 결과만을 복호화



2023 HACKTHEON SEJONG

기존 암호와 동형암호 사용시 데이터 처리 차이점



동형암호 개념

암호화된 상태에서 연산하려면?

정수기반 동형암호 개념

- 비밀키= 큰 소수 p
- 암호화: $Enc(m) = m + pq$ (q: 임의의 자연수)
- 연산 및 복호화

• $\boldsymbol{Enc(m1) + Enc(m2)}$

$= (\boldsymbol{m1+pq1}) + (\boldsymbol{m2+pq2})$

$= (\boldsymbol{m1+m2}) + \boldsymbol{p}(\boldsymbol{q1+q2})$

$= \boldsymbol{Enc(m1+m2)}$

- 비밀키(p)=29, 임의의 자연수 q=15, 11
  - $Enc(3) = 3 + 29 \times 15 = 438$
  - $Enc(5) = 5 + 29 \times 11 = 324$
- 연산 및 복호화

• $\boldsymbol{Enc}(3) + \boldsymbol{Enc}(5) = 438 + 324$

$= 3+(29 \times 15) + 5+(29 \times 11) = 762$

$= (3+5) + 29(15+11) = 762 (\mathrm{mod}\ 29)$

$= \boldsymbol{Enc}(3+5) = 8$

...but **평문 쉽게 추측 가능**

2023 HACKTHEON SEJONG

동형암호의 종류

- 서울대 천정희교수 연구팀은 4세대 동형암호 알고리즘(스킴) CKKS 개발

	시작년도	특징	연산	ISO 표준
1세대	2009년	최초 완전동형암호	Bit	
2세대	2011년	최초의 사용 가능한 동형 암호	정수	BGV, BFV
3세대	2013년	작은 데이터 처리에 효과적인 동형암호	Bit	CGGI
4세대	2016년	최초의 실수 연산 지원하는 동형암호	실수, 정수	CKKS

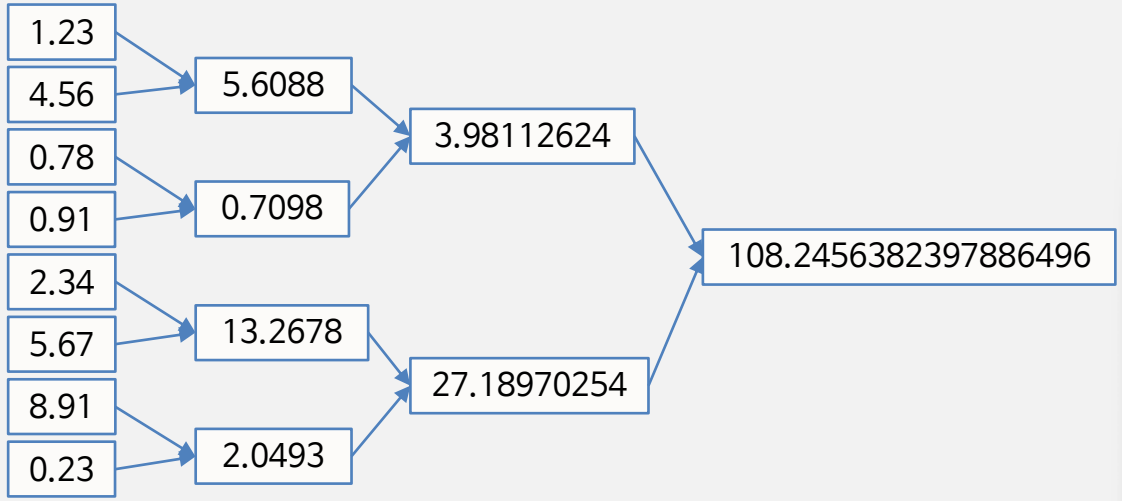
- 다수의 동형암호 오픈소스 라이브러리가 CKKS 지원 중

No	Library	Organization	Schemes used	URL
1	**Helib**	IBM	BGV / CKKS	https://github.com/homenc/HElib
2	**SEAL**	Microsoft	BFV / CKKS	https://github.com/microsoft/SEAL
3	**HEaaN**	CryptoLab	CKKS	https://github.com/snucrypto/HEAAN
4	**PALISADE**	Duality Technologies	BFV / BGV / CKKS / CGGI	https://palisade-crypto.org/software-library/
5	**Lattigo**	EPFL	BFV / CKKS	https://github.com/ldsec/lattigo

- CKKS: Cheon,Kim, Kim and Song
- BFV: Brakerski / Fan Vercauteren
- BGV: Brakerski Gentry Vaikuntanathan
- FHEW: Fastest Homomorphic Encryption in the West

실수연산 동형암호(CKKS)

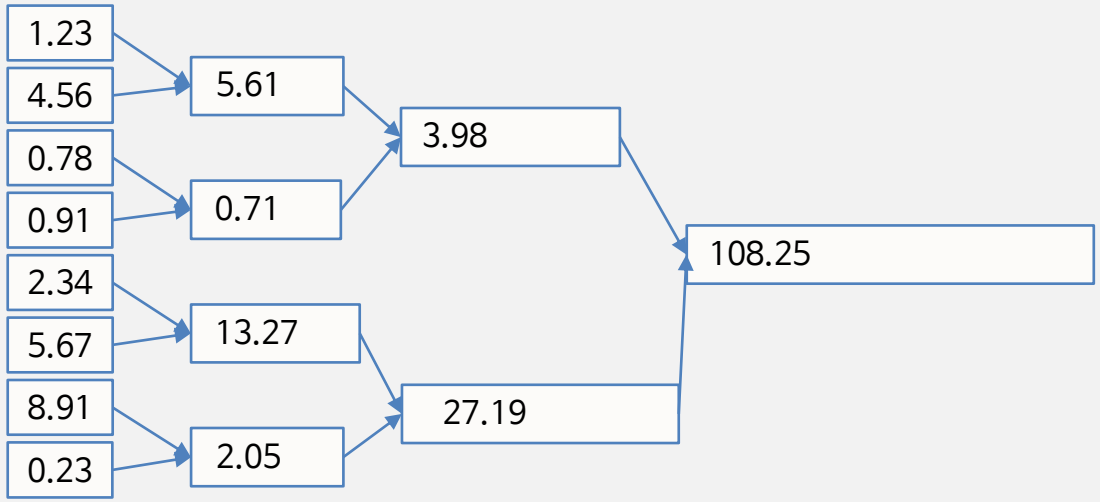
정수연산 동형암호



- 평문 크기가 연산마다 2배로 증가하면 20번 후에 백 만 비트

(쌀 한 톨의 비유: 한 톨로 시작하여 한달 후 $0.02g*2^{30}$ = 20톤)

실수연산 동형암호(CKKS)



- 근사계산을 지원
- 항상 같은 크기의 자릿수를 유지할 수 있음
- 대부분의 응용분야에서는 실수계산을 필요로 함
- 빅데이터 분석, 기계학습 등에 적합

출처 : 서울대학교 천정희 교수

2023 HACKTHEON SEJONG

iDASH 2017 : 서울대 우승

iDASH

2017년 국제게놈 보안경진대회(iDash)에서
美 마이크로소프트사, 프랑스 EPFL 공대 등이
자체 개발한 동형암호보다
월등한 계산속도와 정확도를 나타내어 **우승**.
세계최고의 **동형암호**임을 **입증**.

서울대학교

Microsoft

EPFL

* 국제 게놈 보안경진대회 : 정보를 유출하지 않고 클라우드 컴퓨팅으로 유전정보를 분석하는 대회

2017 TRACK 3: BEST-PERFORMING TEAMS

Evaluated on (three datasets of 1422 records for training/ 157 records for testing + 18 features)

Teams	AUC 0.7136	Encryption		Secure learning		Decryption		Overall time (mins)	Rank
		Size (MB)	Time (mins)	Time (mins)	Memory (MB)	Size (MB)	Time (mins)		
SNU	0.6934	537.667	0.060	10.250	2775.333	64.875	0.050	10.360	1
CEA LIST	0.6930	53.000	1.303	2206.057	238.255	0.350	0.003	2207.363	3
KU Leuven	0.6722	4904.000	4.304	155.695	7266.727	10.790	0.913	160.912	△
EPFL	0.6584	1011.750	1.633	15.089	1498.513	7.125	0.017	16.739	△
MSR	0.6574	1945.600	11.335	385.021	26299.344	76.000	0.033	396.390	2
Waseda*	0.7154	20.390	1.178	2.077	7635.600	20.390	2.077	5.332	X
Saarland	N/A	65536.000	1.633	48.356	29752.527	65536	7.355	57.344	X

* Interactive mechanism, no complete guarantee on 80-bit security at "analyst" side

** Program ends with errors

2023 HACKTHEON SEJONG

iDASH 2018: 라이브러리 "혜안" 우수성

2018 Track 2 : Secure Parallel Genome Wide Association Studies using Homomorphic Encryption

Team	Submission	Schemes	End to End Performance		Evaluation result (F1- Score) at different cutoffs							
			Running time (mins)	Peak Memory (M)	0.01		0.001		0.0001		0.00001	
					Gold	Semi	Gold	Semi	Gold	Semi	Gold	Semi
A*FHE	A*FHE -1 +	HEAAN	922.48	3,777	0.977	0.999	0.986	0.999	0.985	0.999	0.966	0.998
	A*FHE -2		1,632.97	4,093	0.882	0.905	0.863	0.877	0.827	0.843	0.792	0.826
Chimera	Version 1 +	TFHE & HEAAN (Chimera)	201.73	10,375	0.979	0.993	0.987	0.991	0.988	0.989	0.982	0.974
	Version 2		215.95	15,166	0.339	0.35	0.305	0.309	0.271	0.276	0.239	0.253
Delft Blue	Delft Blue	HEAAN	1,844.82	10,814	0.965	0.969	0.956	0.944	0.951	0.935	0.884	0.849
UC San Diego	Logistic Regr +	HEAAN	1.66	14,901	0.983	0.993	0.993	0.987	0.991	0.989	0.995	0.967
	Linear Regr		0.42	3,387	0.982	0.989	0.980	0.971	0.982	0.968	0.925	0.89
Duality Inc	Logistic Regr +	CKKS (Aka HEAAN), pkg: PALISADE	3.8	10,230	0.982	0.993	0.991	0.993	0.993	0.991	0.990	0.973
	Chi2 test		0.09	1,512	0.968	0.983	0.981	0.985	0.980	0.985	0.939	0.962
Seoul National University	SNU-1	HEAAN	52.49	15,204	0.975	0.984	0.976	0.973	0.975	0.969	0.932	0.905
	SNU-2		52.37	15,177	0.976	0.988	0.979	0.975	0.974	0.969	0.939	0.909
IBM	IBM-Complex	CKKS (Aka HEAAN), pkg: HElib	23.35	8,651	0.913	0.911	0.169	0.188	0.067	0.077	0.053	0.06
	IBM- Real		52.65	15,613	0.542	0.526	0.279	0.28	0.241	0.255	0.218	0.229

2023 HACKTHEON SEJONG

iDASH 2020 : 서울대 등 공동 우승

Final Rankings (Considering Time and AUC)

1st Place: SNU, Desilo, Chimera, SamsungSDS

2nd : Alibaba Gemini Lab, A*FHE

Team Name	Institution	Country	End-End Time (Sec.)	AUC
SNU	Seoul National University	South Korea	286.13	0.9857596265677173
Desilo	Desilo	South Korea	43.70	0.9774713444941853
Chimera	Inpher	Switzerland	0.74	0.9704708204593838
SamsungSDS	SamsungSDS	South Korea	3.39	0.9745466735892511
Alibaba Gemini Lab	Alibaba Group	China	3.82	0.9668795482408528
A*FHE	A*STAR	Singapore	186.01	0.9774805423336614

동형암호 실용화 걸림돌 – 속도

- Gartner Top Strategic Trends for 2021 – 동형암호의 문제는 속도 (`20.10)

People centricity

Privacy-enhancing computation

Privacy-enhancing computation comprises three types of technologies that protect data while it's being used to enable secure data processing and data analytics:

- The first provides a trusted environment in which sensitive data can be processed or analyzed. It includes trusted third parties and hardware-trusted execution environments (also called confidential computing).
- The second performs processing and analytics in a decentralized manner. It includes federated machine learning and privacy-aware machine learning.
- The third transforms data and algorithms before processing or analytics. It includes differential privacy, homomorphic encryption, secure multiparty computation, zero-knowledge proofs, private set intersection and private information retrieval.

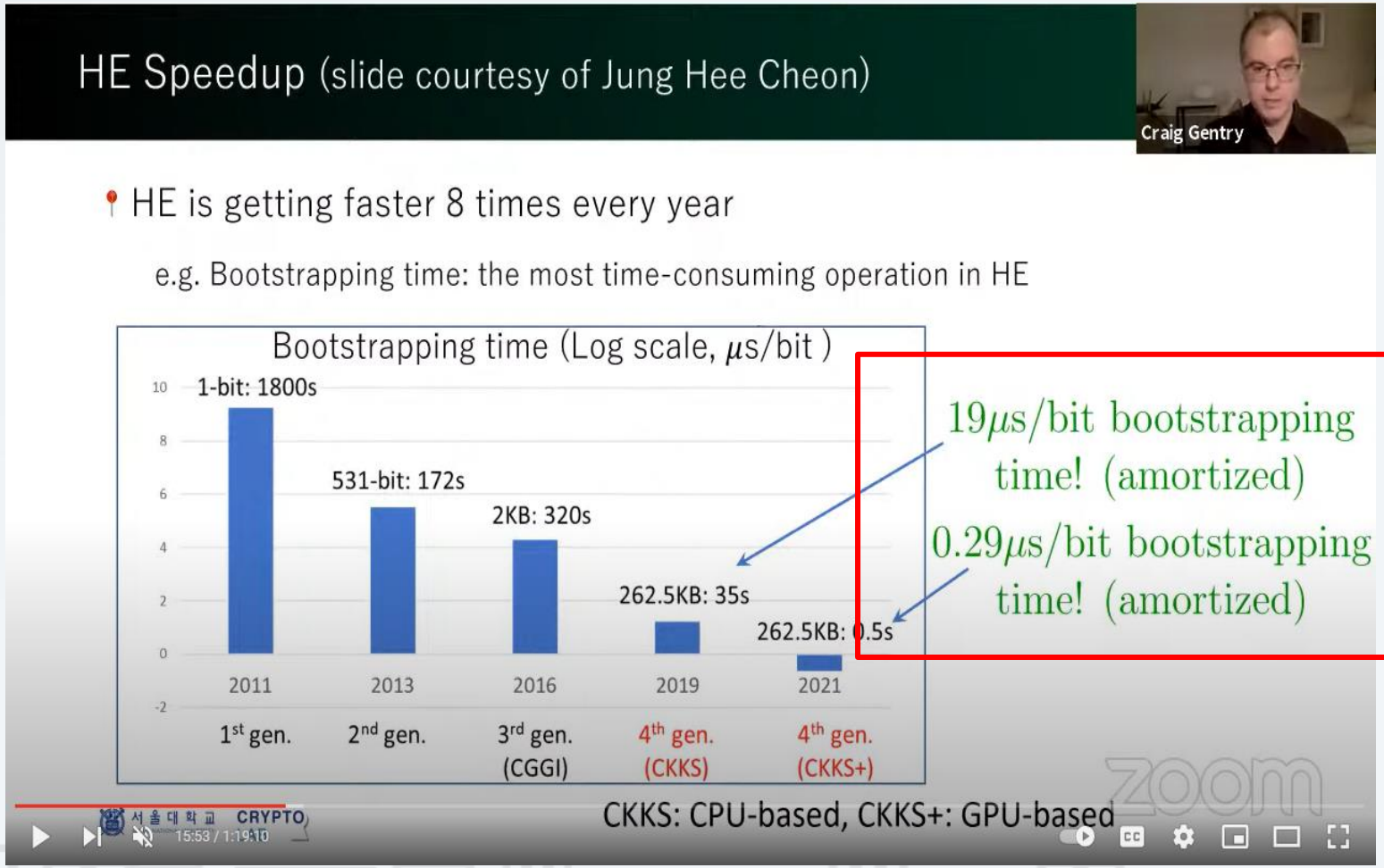
This enables organizations to safely share data in untrusted environments, an increasingly in-demand desire as the amount of data grows alongside the need to protect that data.

What is homomorphic encryption?

Homomorphic encryption (HE) is a cryptographic method that enables third parties to process encrypted data and return an encrypted result to the data owner, while providing no knowledge about the data or the results. HE enables algorithm providers to protect proprietary algorithms and data owners to keep data private. Homomorphic encryption is still maturing. In practice today, fully homomorphic encryption is not fast enough for most business implementations.

동형암호 실용화 걸림돌 – 빠른 속도로 해결 중

- [속도] `11년 1,000억 배 → `19년 1,000배 (매년 8 배씩 향상)



Craig Gentry
(the father of HE)

Bootstrapping is the slowest operation in HE.

The speed can be realized only in CryptoLab's HEaaN.lib

출처 : From a talk by Craig Gently, who first invented FHE, at EuroCrypt 2021 https://www.youtube.com/watch?v=487AjvFW1lk

동형암호의 필요성

- Gartner Top Strategic Trends for 2022 – 동형암호를 대표 사례로 선정 (`21.10)
- 2025년도 큰 기업의 60% 이상이 PET 기술 채택, 클라우드 컴퓨팅 분야 등에서 사용

TREND

Privacy-Enhancing Computation

The real value of data exists not in simply having it, but in how it's used for AI models, analytics, and insight.

Privacy-enhancing computation (PEC) approaches allow data to be shared across ecosystems, creating value but preserving privacy.

Approaches vary, but include encrypting, splitting or preprocessing sensitive data to allow it to be handled without compromising confidentiality.

How It's Used Today:

DeliverFund is a U.S.-based nonprofit with a mission to tackle human trafficking. Its platforms use homomorphic encryption so partners can conduct data searches against its extremely sensitive data, with both the search and the results being encrypted. In this way, partners can submit sensitive queries without having to expose personal or regulated data at any point.

By 2025, 60% of large organizations will use one or more privacy-enhancing computation techniques in analytics, business intelligence or cloud computing.

Source: Gartner

How to Get Started:

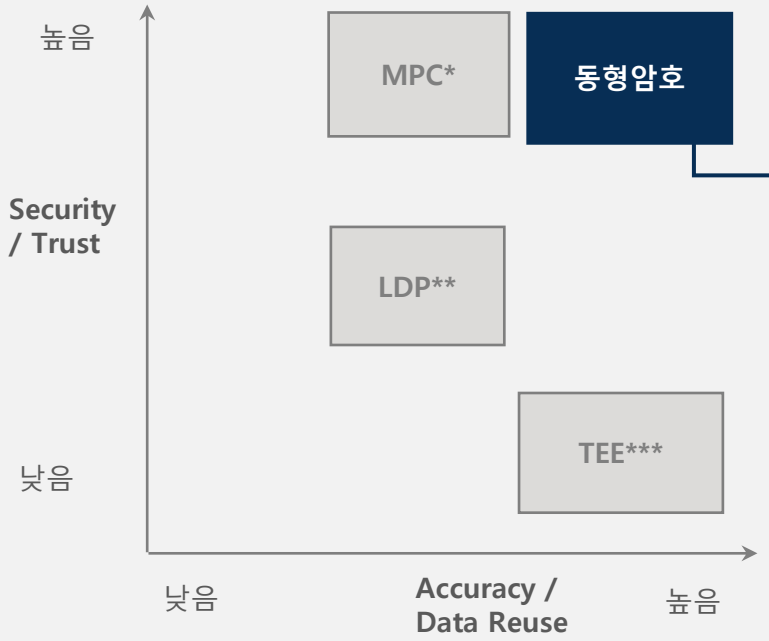
Investigate key use cases within the organization and the wider ecosystem where a desire exists to use personal data in untrusted environments or for analytics and business intelligence purposes, both internally and externally.

Prioritize investments in applicable PEC techniques to gain an early competitive advantage.

PET (Privacy Enhancing Technology) 기술의 비교

동형암호는 PET 기술 중 보안성 및 데이터 활용성 측면에서 가장 우수

PET (Privacy Enhancing Technology) 기술



Security

- 동형암호는 **외부 해킹 뿐만 아니라 내부자에 의한 데이터 유출도 원천적으로 방지**
  - 동형암호/MPC는 수학적으로 노출 차단
  - 반면, LDP는 암호 대비 보안성이 낮으며, TEE 는 H/W 안전성에 의존

Trust

- 동형암호는 **복호화 키만 안전하게 관리하면 됨.**
  - TEE는 Trust 측면에서 TEE용 Software 를 전적으로 신뢰해야 함

Accuracy

- 동형암호 기반 연산은 **원본 데이터를 활용한 연산과 거의 동일한 결과값을 산출함**
  - LDP는 단순 연산에만 적용 가능하며, 데이터 제공자가 적은 경우 (예: 100K 이하) 정확도 보장 어려움.

Data Reuse

- 동형암호 활용 시, **수집된 데이터를 다양한 용도로 활용 가능함**
  - MPC 및 LDP는 특정 용도로만 활용 가능

*Multi-Party Computation; **(Local) Differential Privacy; ***Trusted Execution Environment

2023 HACKTHEON SEJONG

동형암호 표준화

TTAK KO-12.0347 (2019. 12.11)
세계 최초 동형암호 국내 표준화

표준종류	정보통신단체표준(TTAS)		
표준번호	TTAK.KO-12.0347	구 표준번호	
제개정일	2019-12-11	총페이지	235
한글 표준명	근사연산 동형암호 알고리즘		
영문 표준명	Homomorphic Encryption for Arithmetic of Approximate Numbers		
한글 내용요약	이 표준은 근사연산 동형암호를 지원하는 동형암호를 정의하고 있다. 본문에서 근사연산 동형암호의 암호문 생성 방법과 암호문에 대한 근사연산방법을 제시하고 재부팅을 정의한다.		
영문 내용요약	The standard defines a homomorphic encryption algorithm that supports approximated operations. The standard defines a method for generating cipher-text of the algorithm and an approximation method for cipher-text, and the reboots.		

“CKKS” 동형암호 국제 표준화 추진중

- CKKS 을 포함 4개의 동형 암호스킴이 ´21.4월 열린 ISO/IEC JTC1 SC27 WG2 총회에서 만장일치로 신규 작업 표준안으로 통과(21년)
- 4세대 암호 CKKS 는 ISO 국제표준으로 제정될 전망

동형암호 응용분야

향후 동형암호 활용 비중 증가가 예상되는 주요 산업 및 업체

금융

- 다량의 고객금융정보를 수집 활용하는 **금융기관**
  - 은행, 보험사, 증권사, 신용평가사 등
- 개인 금융정보 및 행동데이터 등을 분석하는 **핀테크 업체**

AI / Big Data

- **AI-as-a-Service 제공 업체**
- 빅데이터 센터 및 클라우드 업체 등
  - 대량의 개인정보를수집▪보관▪활용하는 **대형 tech 업체**

자율주행

- 자율주행용 **AI 소프트웨어 개발사**
  - AI에 필수적인 데이터의 보호 및 Privacy 이슈 없는 데이터 수집/활용

동형암호

의료기관 및 헬스케어

- 진료기록을 수집▪보관하는 **대형병원**
- 진료기록 및 DNA 정보 등 의료 데이터를 수집▪보관▪활용하는 **헬스케어 업체**
- 임상데이터를 수집▪활용하는 **제약업체**

정부 및 공공기관

- 국민의 소득▪재산▪납세 정보 등을 수집▪보관하는 **정부기관**
- 빅데이터 기반 **통계분석을 수행하는 기관** : 통계청, 공공연구기관 등

스마트 팩토리

- 스마트 팩토리용 **장비 제조사**
  - 팩토리 업체는 회사 기밀 이슈로 데이터 제공 꺼림
- 공정 데이터 분석 **AI 분석 서비스 개발사**

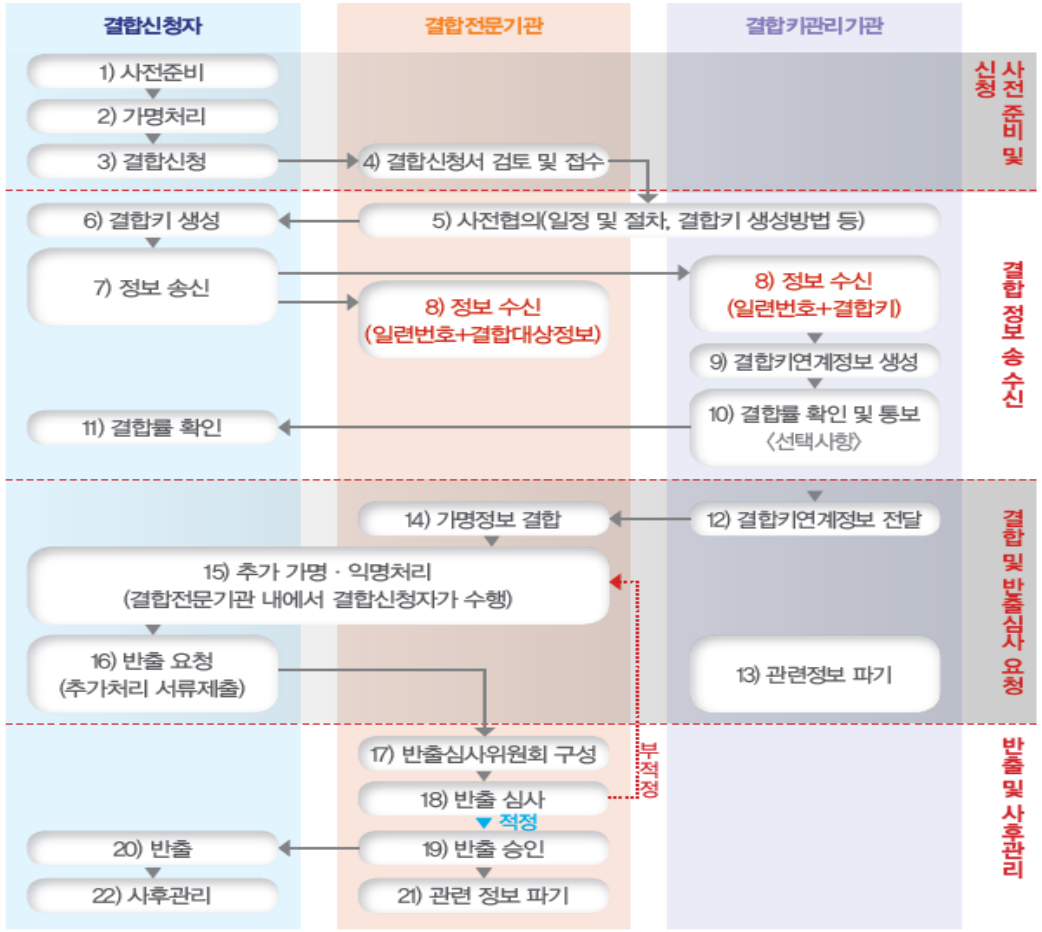
2023 HACKTHEON SEJONG

동형암호 활용방안 (Use Case)

2023
HACKTHEON
SEJONG

#.1 다기관 데이터 결합 후 데이터분석 모형 개발

여러 기관 민감 데이터 결합 시 안전하고 신속하게 데이터분석 모형 개발이 가능한가?



가명정보 처리 가이드라인

가명처리	정보 일부를 삭제하거나 전부를 대체 Ex)대체, 범주화, 상하단 코딩, 부분삭제
가명정보 처리 목적	① 통계작성 ② 과학적연구 ③ 공익적 기록 보전
가명결합 절차	① 결합 전문기관 통해서만 가능 ② 절차 복잡 ③ 소요시간 오래 걸림

2023 HACKTHEON SEJONG

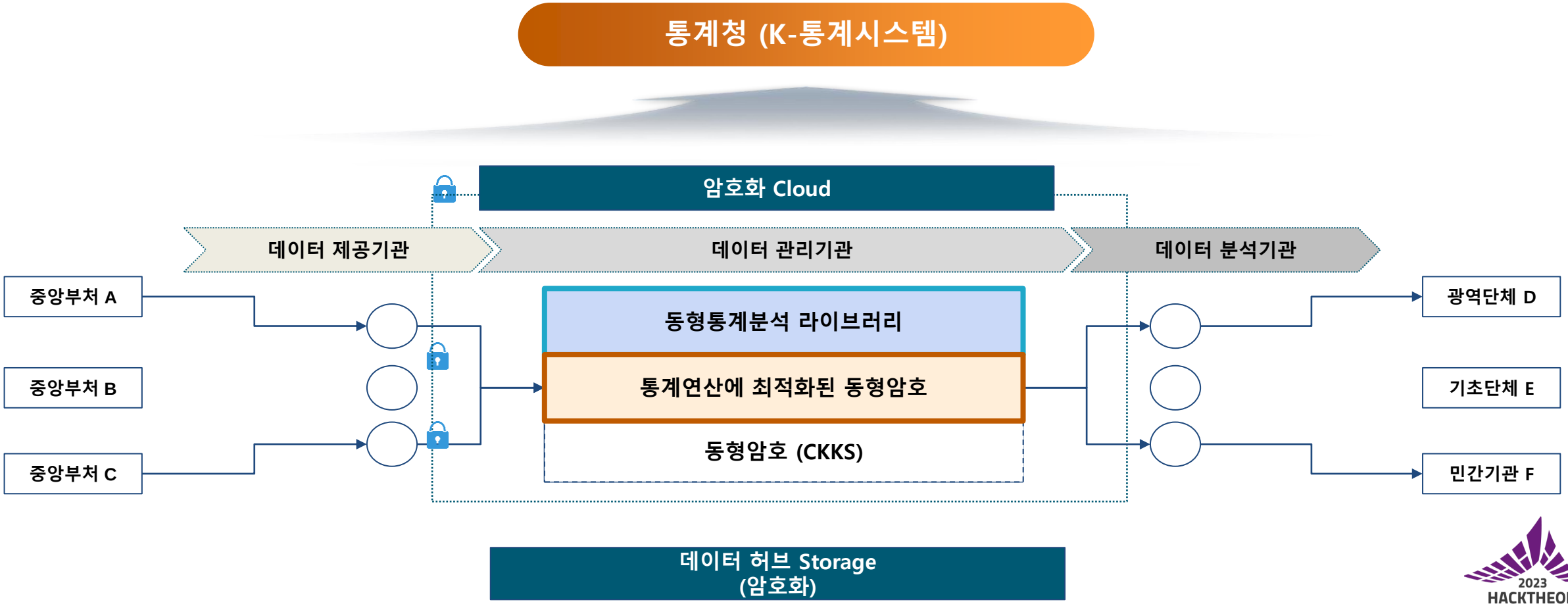
가명정보 처리기술 비교

	K-익명성 (K-anonymity)	차분 프라이버시 (DP: Differential Privacy)	연합학습 (FL: Federated Learning)	동형암호 (HE: Homomorphic Encryption)
장점	o 변환과 활용이 효율적 o 직관적이고 단순한 과정	o 단순 통계 계산 용이 o 안전성 정량적 분석 가능 o 결과값에서 유출되는 정보 최소화	o 효율적인 기계학습 가능 (사용자별 충분한 데이터 확보 시 효과적) o 데이터 직접적 유출 없음 o 계산 병렬화 통한 연산 속도 향상	**o 모든 계산 가능 (튜링 완전성)** **o 증명 가능한 안전성 (양자 내성 암호)**
단점	o 범주화로 인한 데이터 손실 o 범주화 불가능 데이터 존재 o 추가 데이터 활용 시 재식별화 가능성 (Netflix 사례)	o 복잡한 분석(기계학습 등) 수행 어려움 o 분석 결과값의 정확도 손실	o 이종 데이터 결합하여 분석 불가능 o 모델 학습에만 사용 가능 (예측 및 상관관계 분석 불가)	o 암호화 후 데이터 크기 증가 o 평문 대비 데이터 처리 속도 느림
안전성 평가 방법	o 시행 후 전문가 평가 필요 o 추가 데이터 결합 시 재평가	o 시행 후 전문가 평가 필요 o 추가 데이터 결합 후 재평가 불필요	o 모델 수립 시 평가 필요	o 표준 동형암호 사용

#.1 다기관 데이터 결합 후 데이터분석모형 개발

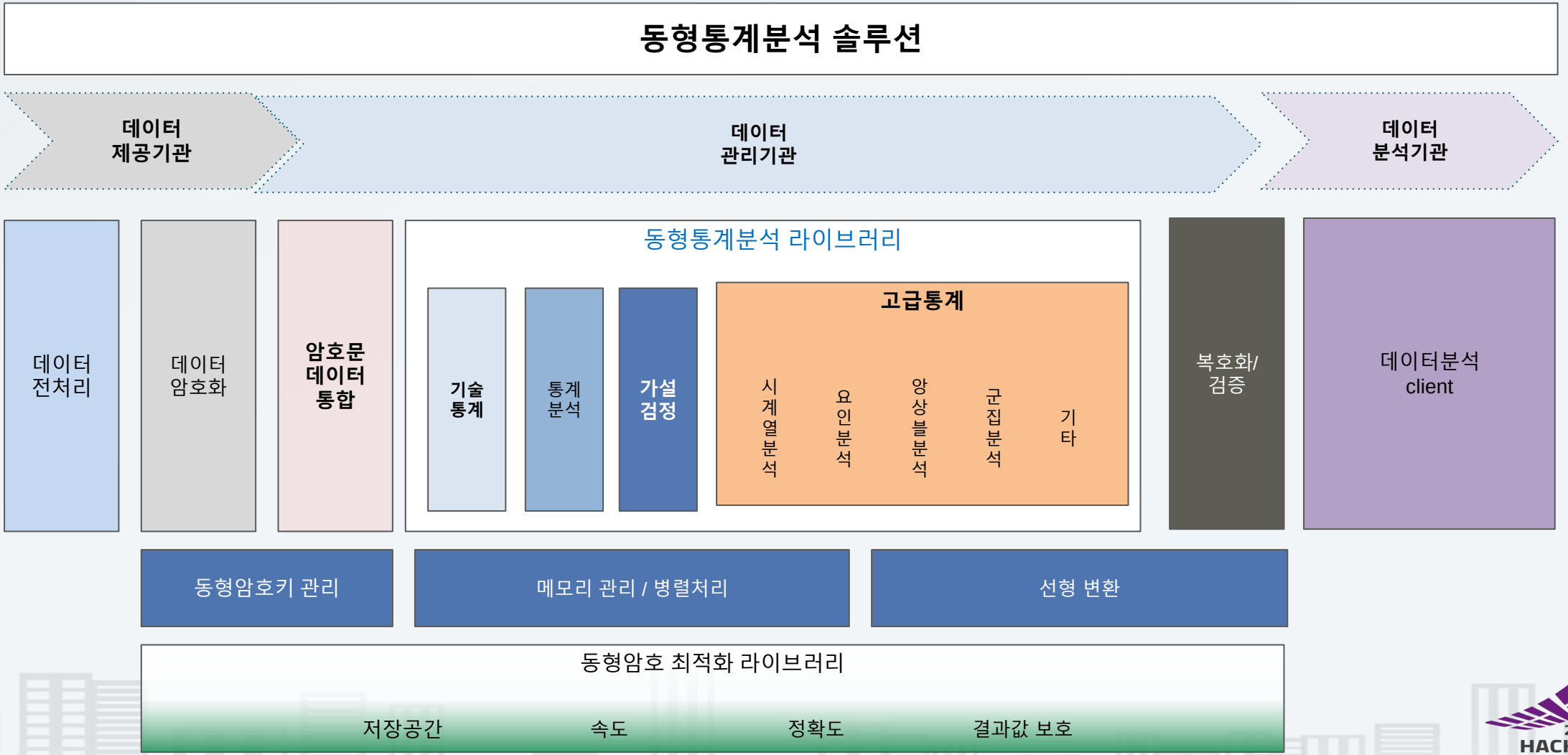
➔ *암호화 상태로 여러기관의 데이터를 결합한 상태에서 데이터분석(기계학습) 모형 개발*

➔ **정보노출(X), 정보손실(X), 결합절차 단순화 및 기간단축(O)**



2023 HACKTHEON SEJONG

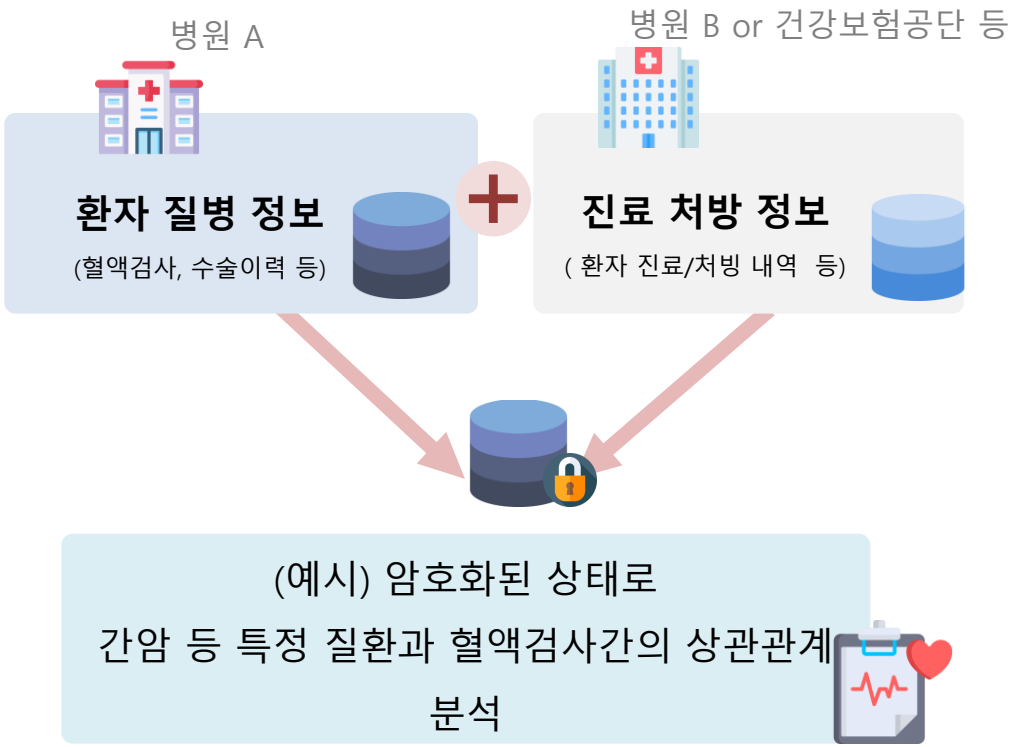
K-통계시스템의 동형통계 분석



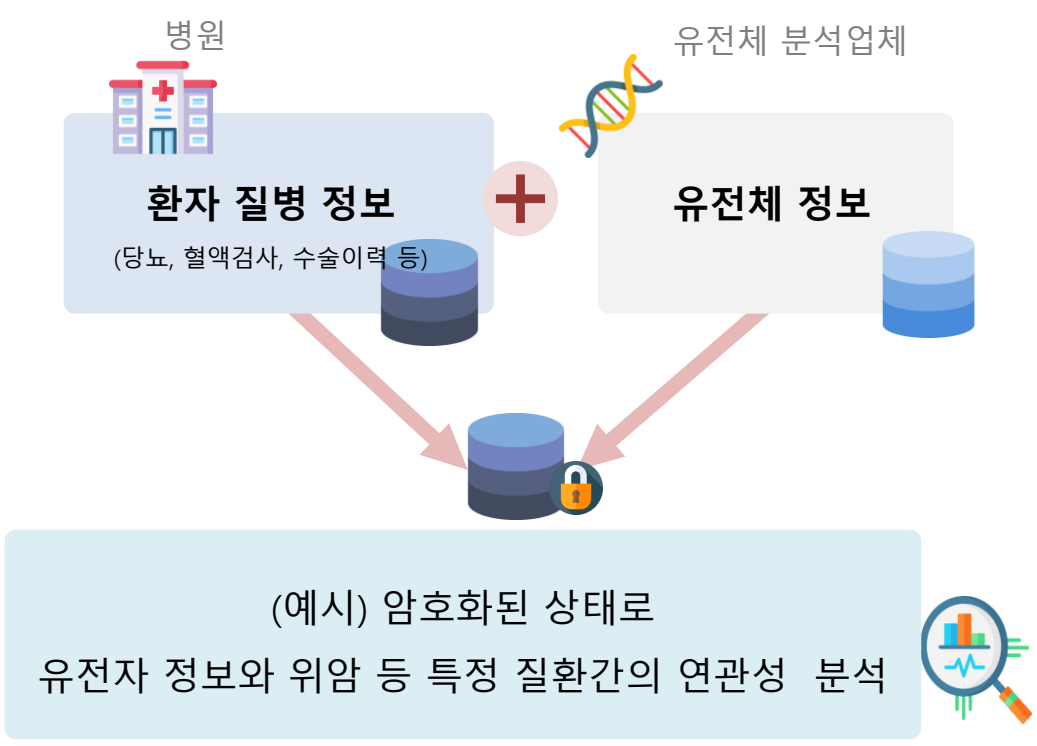
1. 민감한 의료데이터 결합분석 정밀 진단

병원 또는 보건기관, 유전자 분석 업체 등에서 분산 보관된 환자 질병정보, 처방내역, 유전자 정보 등을 안전하게 결합 분석하여 특정질병 개인 맞춤형 정밀진단 분석모형 개발

병원과 병원 or 보건기관 데이터 결합 분석



유전자와 환자 질병 데이터 결합 분석



#.1 동형암호기반 신용평가서비스 제공(2020년)

동형 데이터 분석도구 활용(HEaaN. Stat) 국민연금공단과 KCB의 데이터 결합/분석 후 국민연금 성실납부자의 신용 평가개선 (2020, 금융규제 샌드박스)

- 신용 취약계층 및 소외계층, 금융데이터 부족군에 대한 합리적 신용평가 기준 마련

뉴스홈 | 최신기사

'국민연금 성실납부' 신용평가에 반영...55만명 신용점수 오를 듯

송고시간 | 2020-07-14 12:00

신재우 기자

기자 페이지

KCB 10월부터 적용...성실납부 36개월이면 최대 41점 가점

정부 "금융 이력 많지 않은 사회초년생 다수 혜택 볼 듯"

동형암호기술 세계 첫 적용..."개인정보 유출 없이 데이터 분석"

△[사진 제공 = KCB]

금융분야 본인신용정보관리업, 이른바 `마이데이터` 사업이 오는 8월 시행 예정인 가운데 개인정보 유출을 원천 차단하고 이를 활용한 빅데이터 분석까지 동시에 할 수는 없을까.

이런 물음에 국내 연구진이 세계 최초로 동형암호 상용화로 화답했다. 동형암호(同形暗號, Homomorphic Encryption)는 평문과 암호문에서 같은 성질이 유지된다는 의미로, 정보를 암호화한 상태에서 각종 연산을 했을 때 그 결과가 암호화 하지 않은 상태의 연산 결과와 동일하게 나오는 4세대 암호체계다.

출처 : 연합뉴스(https://www.yna.co.kr/view/AKR20200714068600530)
매일경제 (https://www.mk.co.kr/news/economy/view/2020/06/616446/

2023 HACKTHEON SEJONG

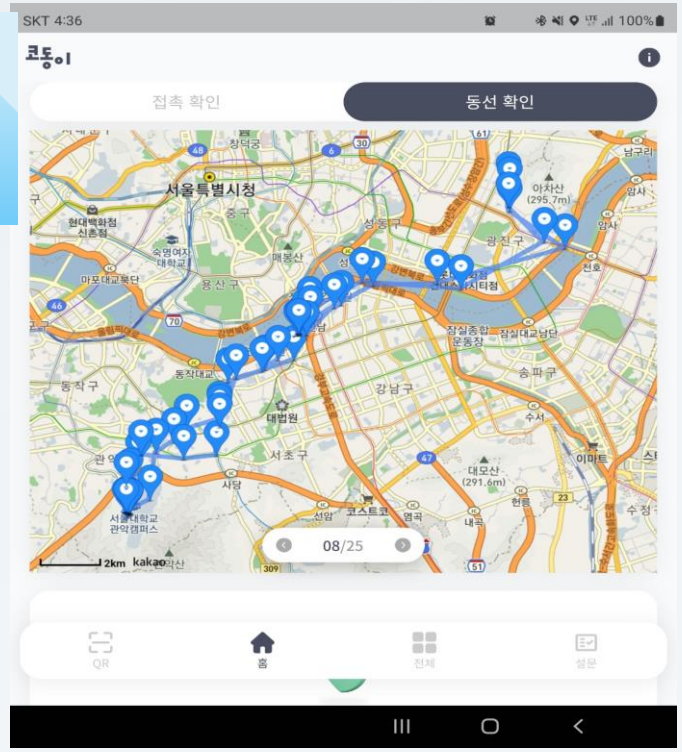
#.2 위치정보 보호 개인화 서비스

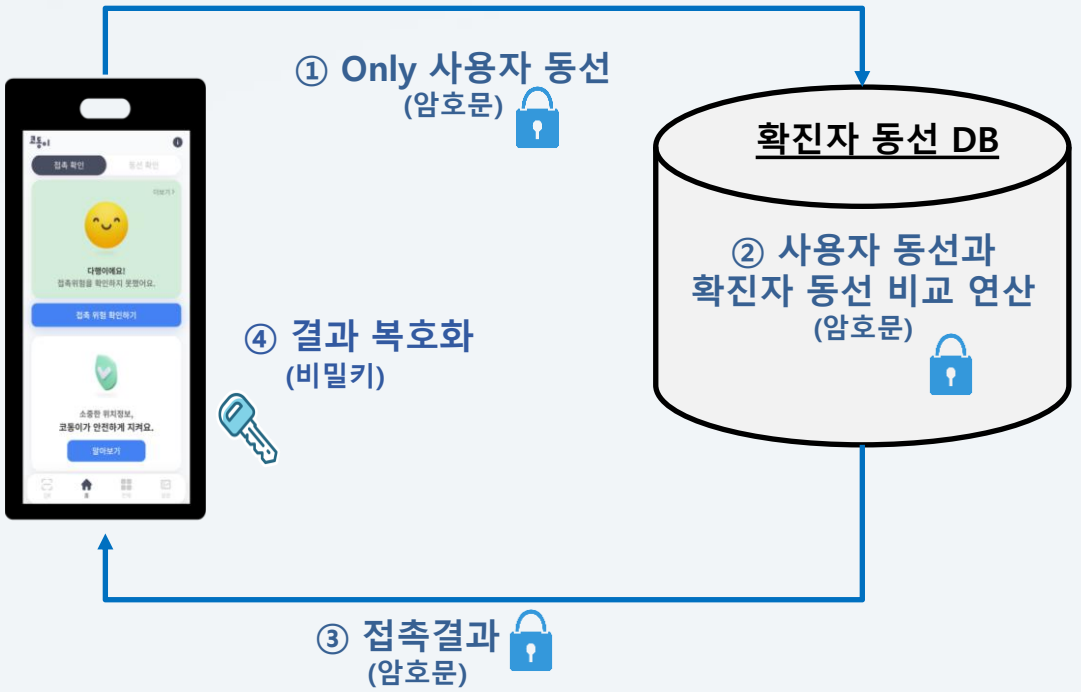
무심코 제공하는 내 위치정보는 안전한가?

2023 HACKTHEON SEJONG

#.2 위치정보 보호 개인화 서비스

코동이(코로나 동선 안심이) ➔ *사용자 위치(프라이버시)를 보호하면서 위치기반 서비스 이용이 가능*





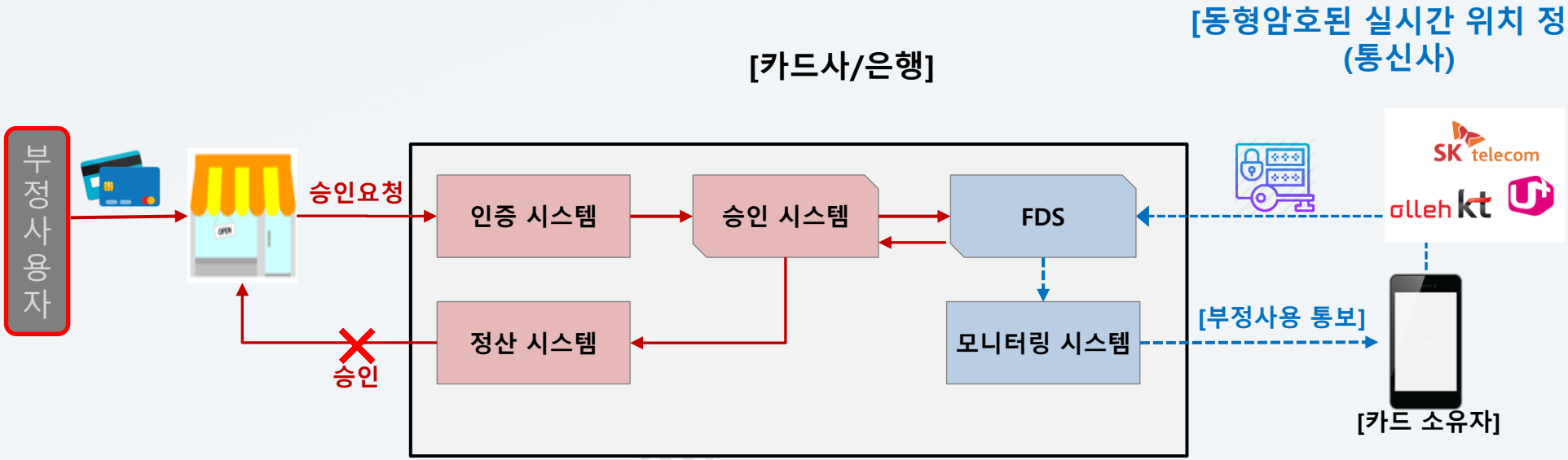
활용 방안

1. *[접근 알림] 범죄자 접근금지, 위험물차량 상수원 보호구역 진입 등*
2. *[이탈 알림] 치매노인 방지, 미아방지 서비스*

#.2 위치정보 보호 개인화 서비스

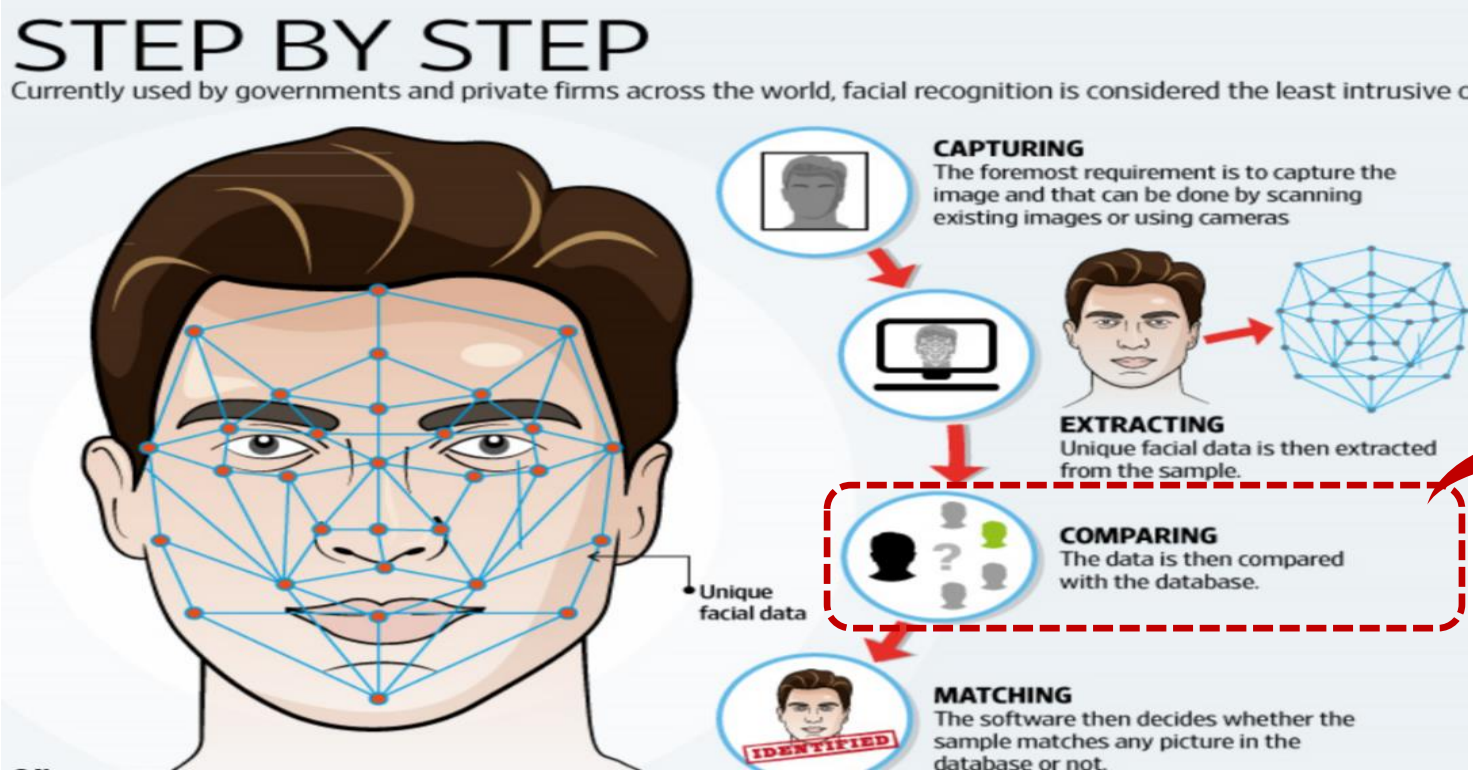
위치기반 Fraud Detection System 고도화

카드사 FDS에 통신사를 통해 고객의 정확한 실시간 위치 정보를 결합하여 **오프라인 상의 부정사용 차단**/
통신사가 제공하는 위치정보는 **동형 암호화되어** 전송되어 유출 위험 전무



#.3 Privacy보호 AI서비스 활용(이미지)

동형암호 기반 이미지 Feature 비교연산으로 원본 이미지를 서버에 보관할 필요 없음



동형암호 비교 연산

활용 방안

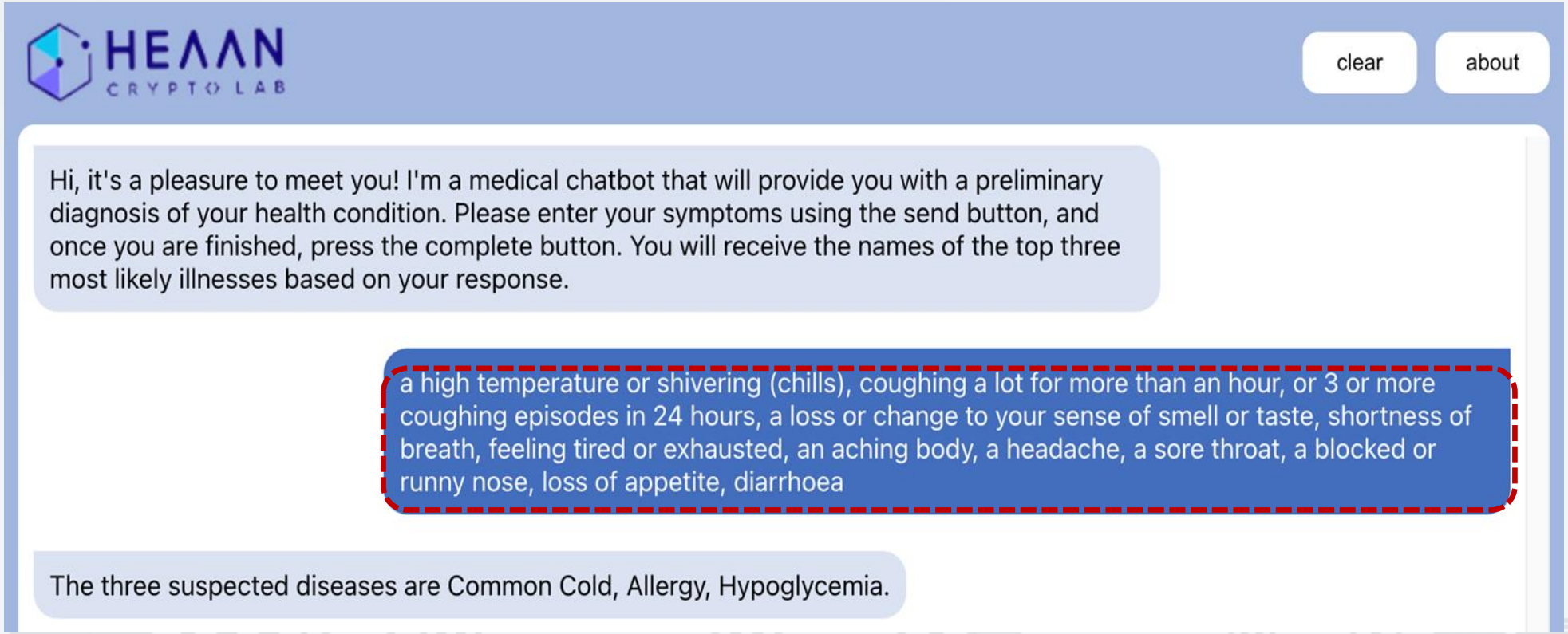
1. [FaceID] 지문, 홍채, 안면인식 등 본인 생체 인증 보호
2. [응급알람] 병실내 낙상, 병원내 화장실/목욕탕 쓰러진 알람

2023 HACKTHEON SEJONG

#.3 Privacy보호 AI서비스(text) :Medical Chabot

- **정신과 치료, 성병 등 민감한** 환자의 증상은 챗봇 과정에서 보호가 필요
- 암호화 된 증상 정보를 서버에 전달하면 암호화 된 상태로 데이터를 처리하여 환자에게 필요한 의료정보 제공

https://heaan.hemedicalchat.shop/



#.4 민감정보 보유기관의 AI 모형 개발

군사용 등 안보기관 AI모형 개발 시 내부 리소스만으로 가능?

(중략)

- (민관 협력을 제약하는 요소들) 인공지능 기술의 원천과 폭넓은 활용 측면에서 민간 주체가 우위에 있으므로 민간과의 협력은 군사적 활용에 있어 필수적이나, 이를 현실적으로 제약하는 요소가 존재

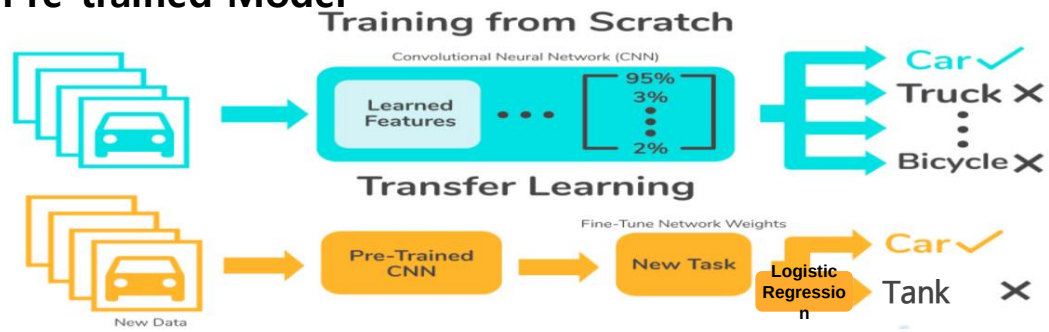
* 군사규격, 보안문제, 협력 대상 기업에 대한 과도한 자격요건 충족 요구, 군에 대한 문민통제, 기타 비용문제와 인공지능 무기화, 도입 당위성에 사회적 공감대 이슈 등

2023 HACKTHEON SEJONG

#.4 민감정보 보유기관의 AI 모형 개발

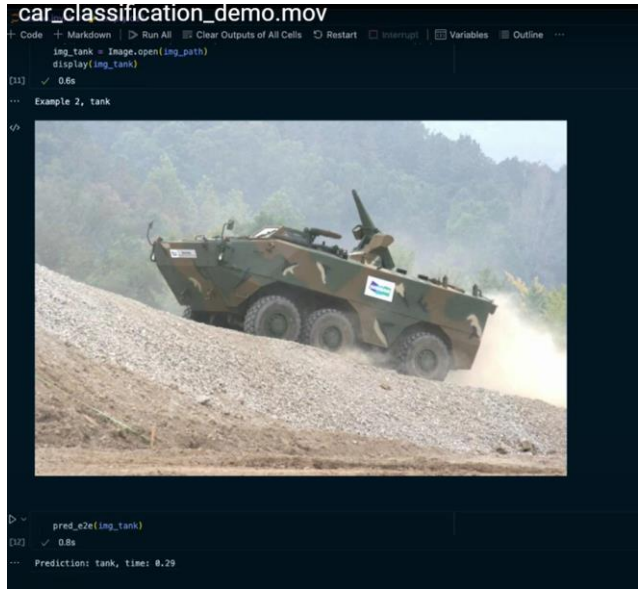
Pre-trained Model기반 민감정보만 동형암호화하여 추가 학습하여 AI모형 개발이 가능

1) Pre-trained Model



2) Encrypted Image Data 추가 학습

: 동형암호화된 image 1천건 추가 학습(20초 소요)



- Input [Enc(Tank image)] → Result 0.29 sec

활용 방안

1. [의료] 유전체(DNA) 데이터 이용 질병예측모형 개발
2. [HR] 입사·인적성 점수와 인사평가 분석(CDP :Career Development Plan)

2023 HACKTHEON SEJONG

동형암호 데이터의 전이학습(Transfer Learning) 성능

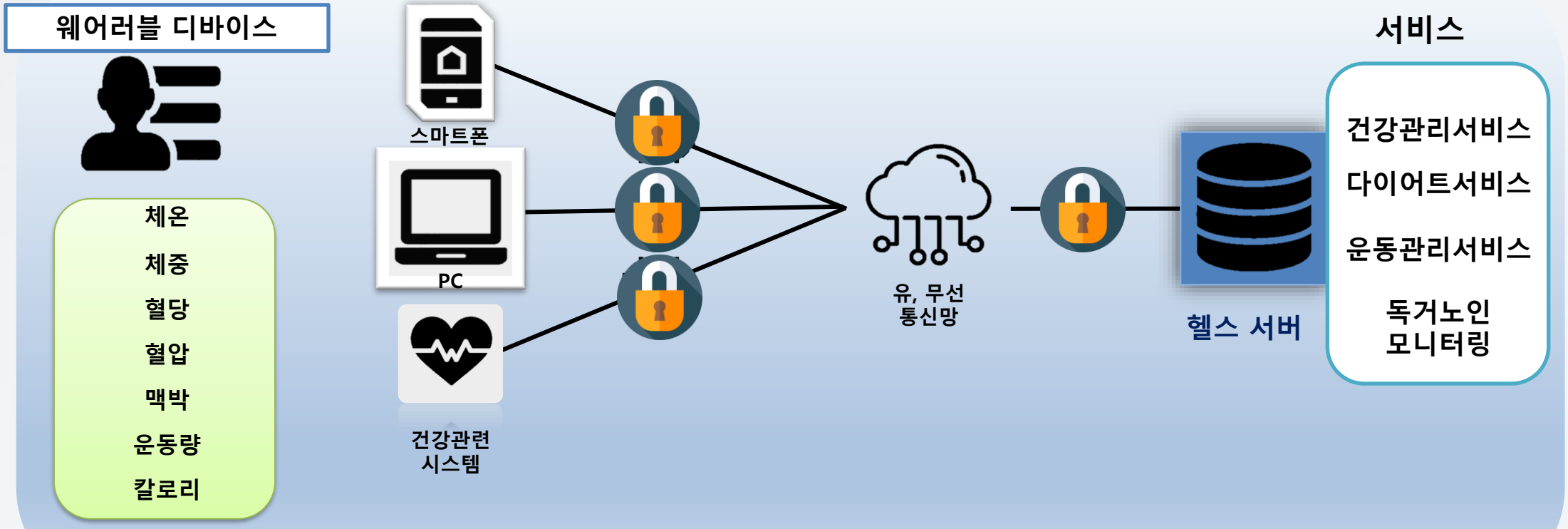
- 전이 학습 소요 시간 : 1시간 미만
- 원본 데이터와 정확도 차이 : 0.5% 미만

dataset	encrypted			not encrypted	
	Running time		ACC (a)	ACC (b)	ACC loss ((b) - (a))
	Total (s)	Time / Iter (s)			
MNIST	2477.66	8.26	93.05%	93.10%	0.05%
CIFAR-10	1032.75	8.26	91.90%	91.65%	-0.25%
Face Mask Detection	499.78	1.78	95.46%	95.46%	0.00%
DermaMNIST	1221.05	3.05	75.06%	75.06%	0.00%
SST-5	614.68	3.07	53.67%	53.44%	-0.23%
SNIPS	421.05	3.01	94.71%	94.71%	0.00%

* GPU: NVIDIA Ampere A40 기준

5. 클라우드시스템상 보건의료분석 서비스

클라우드 컴퓨팅와 웨어러블 디바이스을 통해 헬스서버에서 암호화된 상태로 보건의료 데이터를 연산하여 맞춤형 개인 건강데이터 분석서비스 제공 가능



감사합니다

seokylee@snu.ac.kr