ASK 2024 논문집
Annual Symposium of KIPS 2024
www.kips.or.kr

신진학자 워크숍

개인정보 보호를 위한 동형암호화된 머신러닝 연구

이주희 교수
(성신여자대학교)

개인정보 보호를 위한 동형암호화된 머신러닝 연구

2024.5.24.

Sungshin Women's University

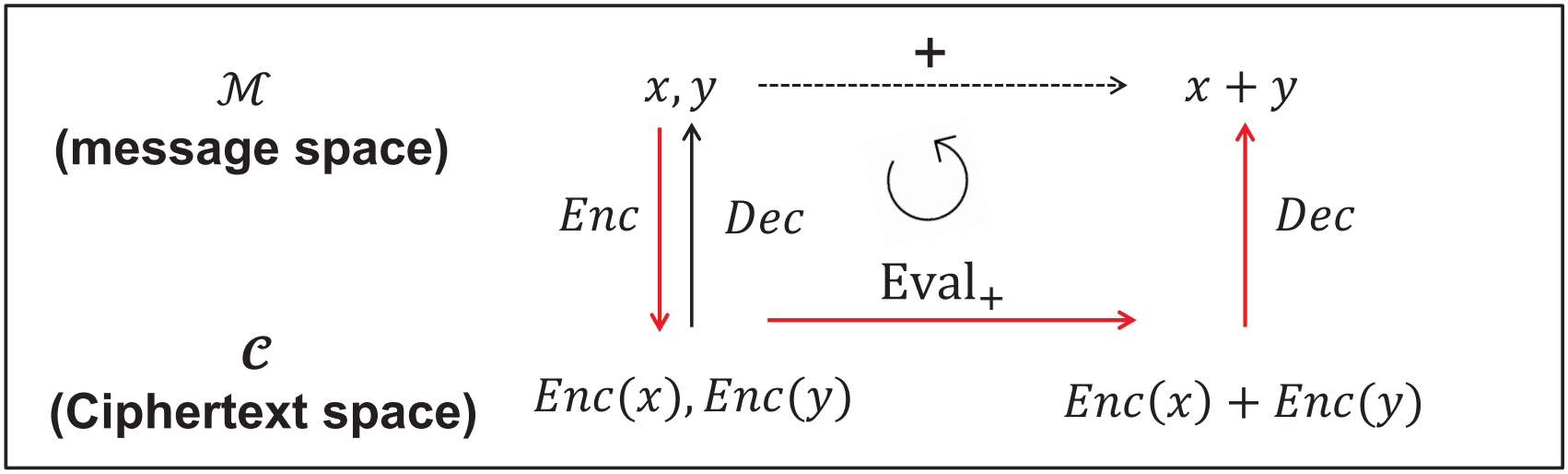
Joohee Lee

CONTENTS

- Fully Homomorphic Encryption
- Brief Intro for a study *"Privacy-Preserving Fair Learning of Support Vector Machine with Homomorphic Encryption (WWW'22)"*
- Brief Intro for an on-going study : *Privacy-Preserving Natural Language Processing with Homomorphic Encryption*

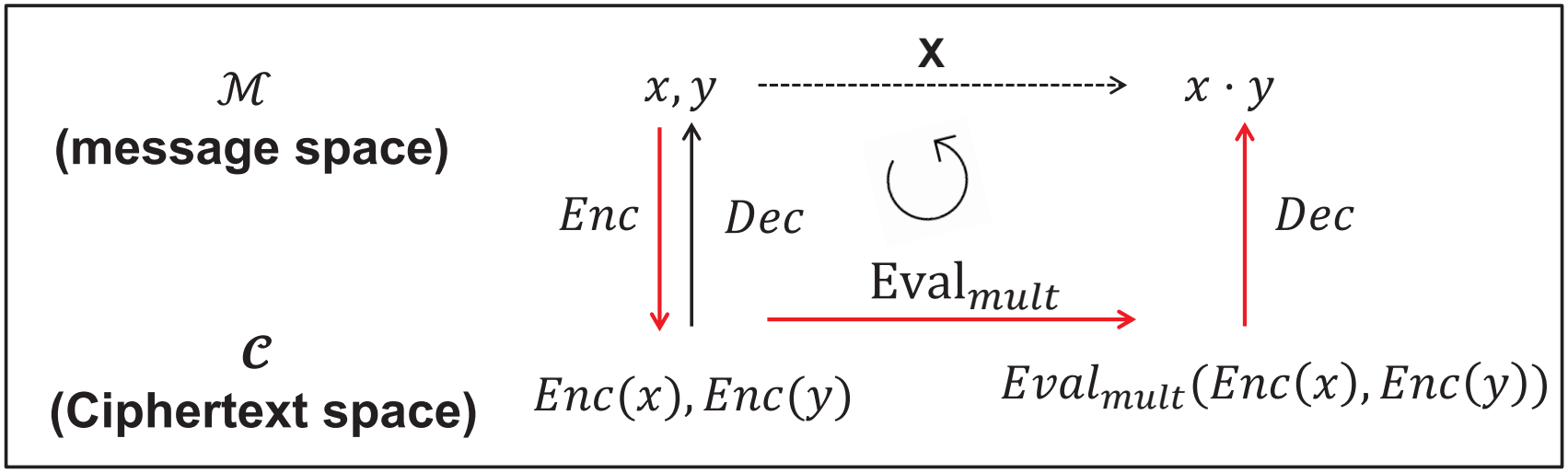
(FULLY) HOMOMORPHIC ENCRYPTION

- **Homomorphic Encryption (HE)** enables homomorphic evaluation over encrypted data without decryption (2009~)
  - E.g. from Enc(x) and Enc(y) compute Enc(x+y)
- Somewhat Homomorphic Encryption (SHE) : support a limited number of operations
- Fully Homomorphic Encryption (FHE) : Can evaluate any function on encrypted data



(FULLY) HOMOMORPHIC ENCRYPTION

- **Homomorphic Encryption (HE)** enables homomorphic evaluation over encrypted data without decryption (2010~)
  - E.g. from Enc(x) and Enc(y) compute Enc(x+y)
- Somewhat Homomorphic Encryption (SHE) : support a limited number of operations
- Fully Homomorphic Encryption (FHE) : Can evaluate any function on encrypted data

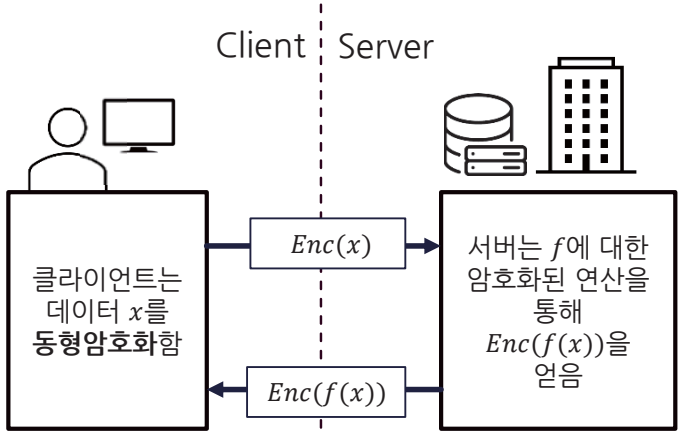


(FULLY) HOMOMORPHIC ENCRYPTION

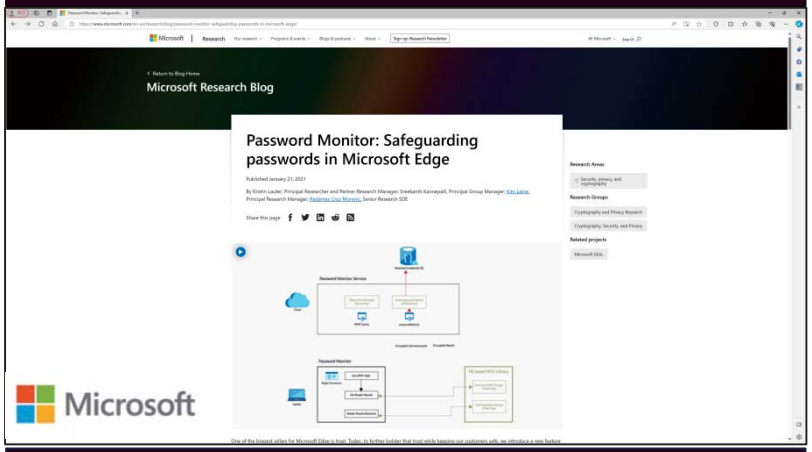
1. 완전동형암호: 암호화된 데이터의 연산이 가능

원본데이터 (평문)		암호화된 데이터 (암호문)
4	동형암호화 →	@##$d1
+ 9	동형암호화 →	+ %$)eo!#
13	← 동형복호화	&#J1*?%

2. 완전동형암호를 이용한 서비스 시나리오



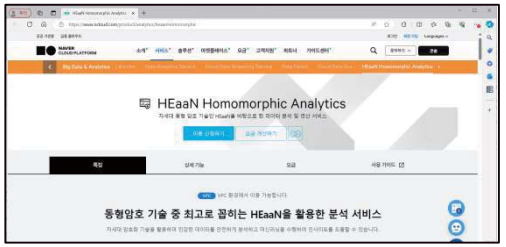
실제 산업 적용 사례 (해외): 美 Microsoft의 완전동형암호를 이용한 패스워드 모니터링 서비스(엣지 브라우저에 도입)



실제 산업 적용 사례 (국내):
① 네이버 클라우드의 완전동형암호를 이용한 암호화된 클라우드 서비스
② 크립토랩의 완전동형암호를 이용한 암호화된 동선 비교 서비스

NAVER CLOUD PLATFORM

HEAAN CRYPTO LAB



VARIANTS OF LATTICE-BASED HE SCHEMES

Scheme	Message Space	Operation	Packing	Bootstrapping
Brakerski-Gentry-Vaikuntanathan (**BGV**) Brakerski/Fan-Vercauteren (**B/FV**)	Finite field $\mathbb{Z}_p$	Modular (mod p) arithmetic	Yes	Optional
Ducas and Micciancio (FHEW) Chillotti-Gama-Georgieva-Izabachene (**TFHE**)	Single bit $\{0, 1\}$	Boolean (XOR, AND)	No	Default
Cheon-Kim-Kim-Song (**CKKS**)	Real/complex numbers $\mathbb{R}, \mathbb{C}$	Approximate (fixed-point) arithmetic	Yes	Optional

Brief Intro for a study

"Privacy-Preserving Fair Learning of Support Vector Machine with Homomorphic Encryption (WWW'22)"

Saerom Park, Junyoung Byun, Joohee Lee

SCENARIO : FAIR LEARNING FOR AI

RETAIL OCTOBER 11, 2018 / 8:04 AM / UPDATED 4 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

By Jeffrey Dastin 8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's AMZN.O machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

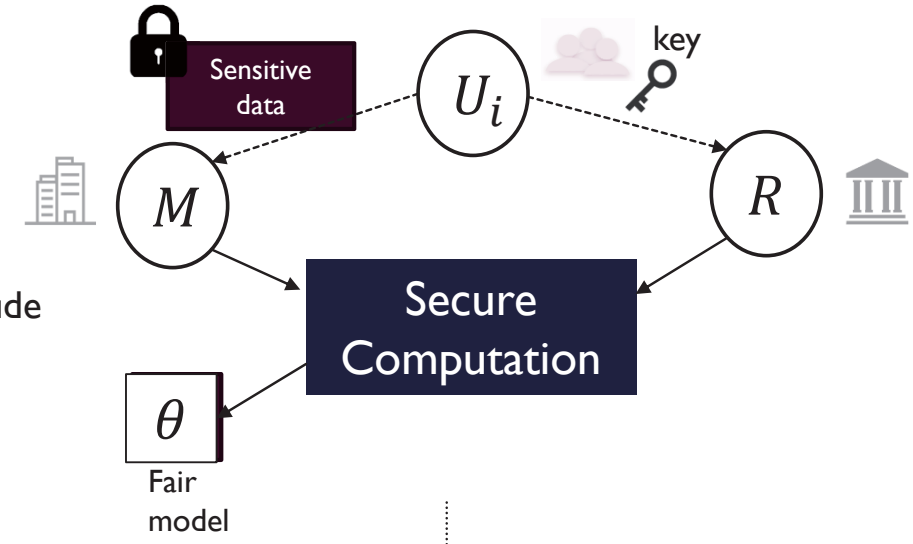
Amazon scraps secret AI recruiting tool that showed bias against women | Reuters

- Automatic decision from AI can lead to practical harms to a specific group.
  - Alert : the ML model can have bias without discriminatory intent (i.e., *disparate impact*).
- **Fair training** can help to correct algorithmic bias in the learning process!
- However, fair training also requires sensitive variables to identify groups.

[WWW'22] PRIVACY-PRESERVING FAIR LEARNING OF SUPPORT VECTOR MACHINE WITH HOMOMORPHIC ENCRYPTION

- U_i : i-th user
- M : modeler
- R : regulator

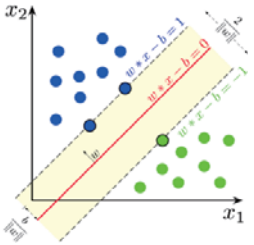
Assumption : M, R do not collude



Kilbertus et al.'s Solution

- General
- Secure computation with secure **Multi-Party Computation (MPC)**
- M, R interactively computes the output

Our Solution



- Focus on Support Vector Machine (SVM)
- Secure computation with approximate **Homomorphic Encryption (HE)**
  - HE enables computations over encrypted data
- M conducts homomorphic computations all by itself

Brief Intro for an on-going study : *Privacy-Preserving Natural Language Processing with Homomorphic Encryption*

NATURAL LANGUAGE PROCESSING

- Natural Language Processing (NLP)
  - One of the major fields of artificial intelligence that studies and implements human language phenomena so that they can be imitated using machines such as computers.
  - Applications: Sentiment Analysis, AI Assistant, Machine Translation, Chatbot, etc

Sentiment Analysis

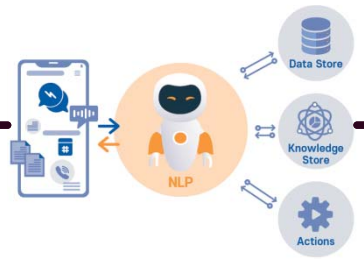
Machine Translation

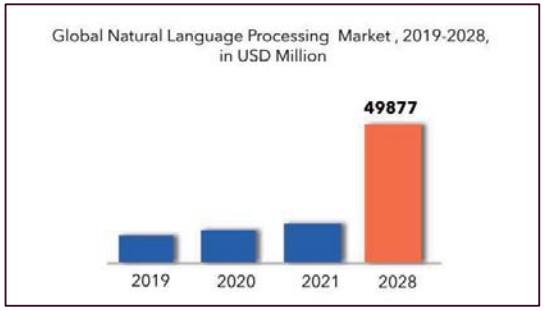
Google Translate

papago

DeepL

그림 출처: http://word.tta.or.kr/dictionary/dictionaryView.do?word_seq=049996-1





출처: 자연어 처리 시장 규모 및 성장 분석 (exactitudeconsultancy.com)

AI Assistant

Bixby

GiGA Genie

Google Assistant

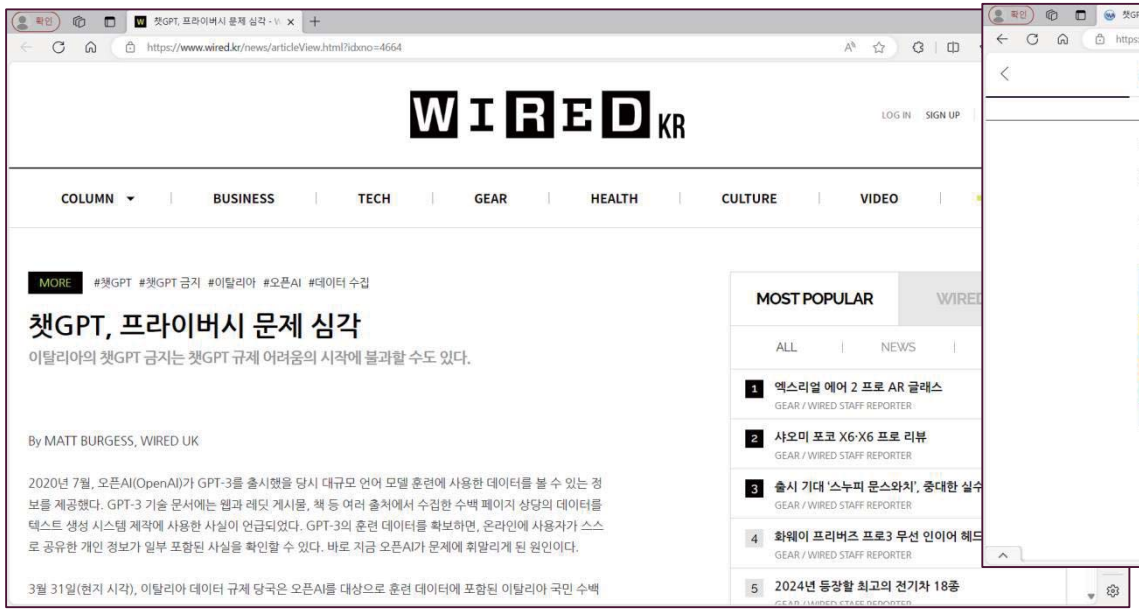
Siri

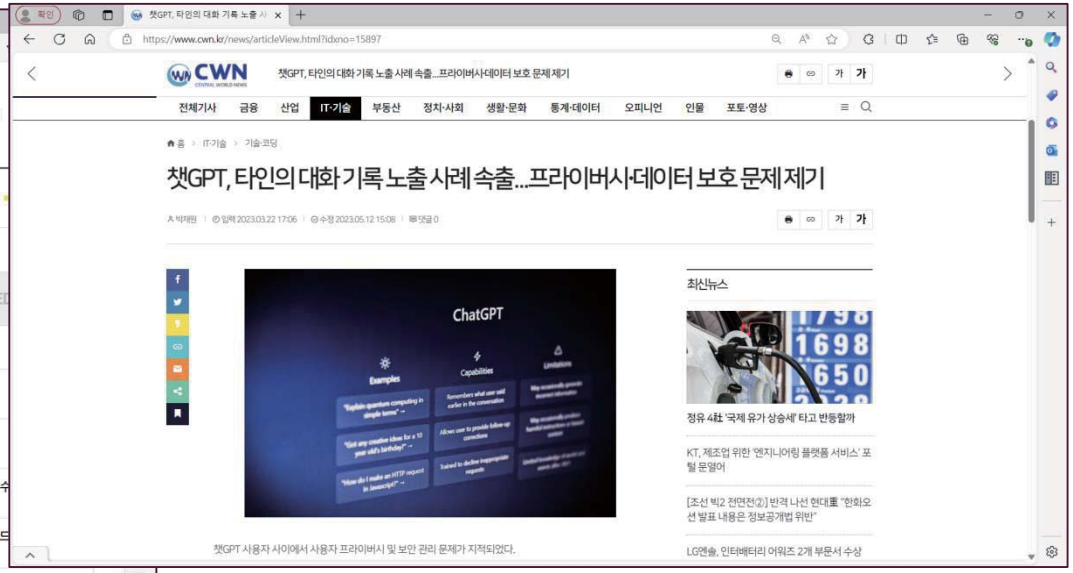
Chatbot

ChatGPT

PRIVACY ISSUES IN NLP

- Privacy Issues in NLP
  - Many applications of natural language processing use user-generated data.
  - **Sensitive user data** such as gender and age can be extracted from text data [CNC18]
  - Recent research has shown that 50-70% of input text data can be recovered from text embeddings [SR20]



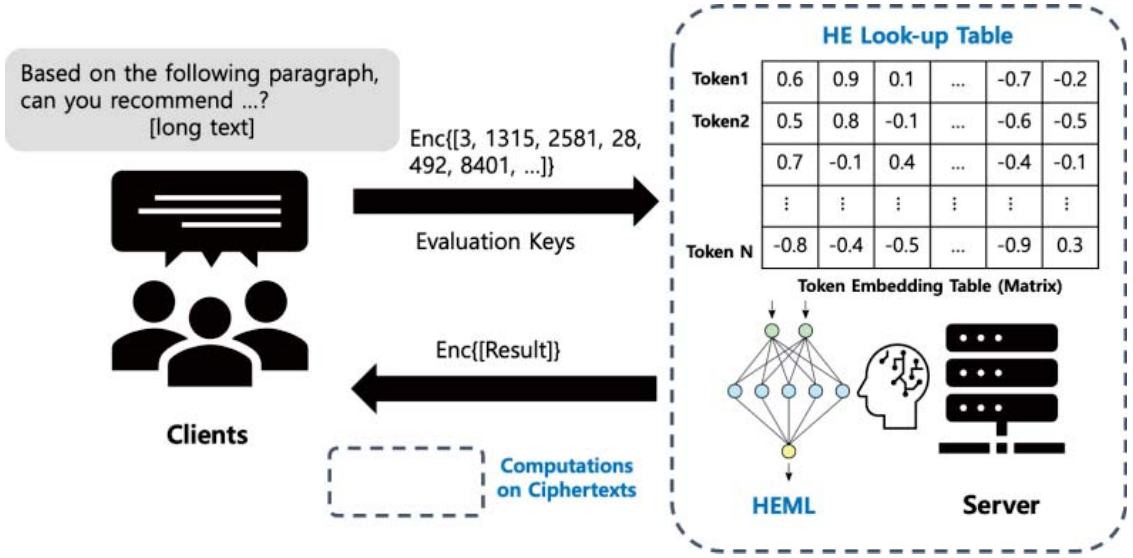


https://www.wired.kr/news/articleView.html?idxno=4664

챗GPT, 타인의 대화 기록 노출 사례 속출...프라이버시·데이터 보호 문제 제기 < 기술·코딩 < IT·기술 < 기사본문 - CWN

PRIVACY-PRESERVING NATURAL LANGUAGE PROCESSING WITH HOMOMOR PHIC ENCRYPTION

- **Privacy-Preserving NLP with HE**
  - Goal: Proposal and Implementation of an encrypted NLP operation algorithm that protects end-to-end privacy
    - To suggest an optimized HE computation for encrypted token embedding + RNN computation
    - Here, encrypted token embedding incurs table look-up evaluation which is a bottleneck for HE computation



SUMMARY

- Homomorphic Encryption enables computations over encrypted data
- Recent work [PBL'22] suggests an algorithm for fair ML training over encrypted data using HE
- We are also working on the problem to address the privacy issues in NLP with HE

Thank you!

Any Question?

Contact: jooheelee@sungshin.ac.kr