



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2025년01월21일
(11) 등록번호 10-2757651
(24) 등록일자 2025년01월16일

- (51) 국제특허분류(Int. Cl.)
G06F 21/62 (2013.01) G06F 16/903 (2019.01)
G06F 16/907 (2019.01)
- (52) CPC특허분류
G06F 21/6254 (2013.01)
G06F 16/903 (2019.01)
- (21) 출원번호 10-2023-0133086
- (22) 출원일자 2023년10월06일
심사청구일자 2023년10월06일
- (56) 선행기술조사문헌
CN116346310 A*
KR1020230086536 A*
KR1020210118429 A*
*는 심사관에 의하여 인용된 문헌
- (73) 특허권자
이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
- (72) 발명자
배호
서울특별시 서초구 남부순환로 2311-12, 101동 801호 (방배동, 래미안방배아트힐)
- (74) 대리인
특허법인(유한)아이스스

전체 청구항 수 : 총 7 항

심사관 : 구대성

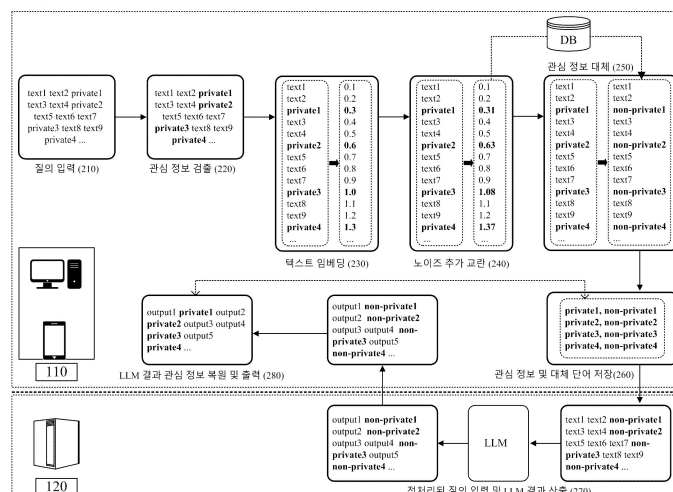
(54) 발명의 명칭 개인 정보 노출 없는 인공지능 기반 서비스 제공 방법 및 클라이언트 장치

(57) 요약

개인 정보 노출 없는 인공지능 기반 서비스 제공 방법은 클라이언트 장치가 사용자 질의를 입력받는 단계, 상기 클라이언트 장치가 상기 사용자 질의 중 개인 정보에 해당하는 타깃 단어를 대응하는 대체 단어로 변경하는 전처리를 수행하는 단계, 상기 클라이언트 장치가 상기 전처리한 사용자 질의를 인공지능 기반 서비스를 제공하는 서비스 서버에 전달하는 단계 및 상기 클라이언트 장치가 상기 서비스 서버로부터 수신한 상기 사용자 질의에 대응하는 답변에서 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 단계를 포함한다.

대표도

200



(52) CPC특허분류

G06F 16/907 (2019.01)

G06F 21/6263 (2013.01)

명세서

청구범위

청구항 1

클라이언트 장치가 사용자 질의를 입력받는 단계;

상기 클라이언트 장치가 상기 사용자 질의 중 개인 정보에 해당하는 타깃 단어를 대응하는 대체 단어로 변경하는 전처리를 수행하는 단계;

상기 클라이언트 장치가 상기 전처리한 사용자 질의를 인공지능 기반 서비스를 제공하는 서비스 서버에 전달하는 단계; 및

상기 클라이언트 장치가 상기 서비스 서버로부터 수신한 상기 사용자 질의에 대응하는 답변에서 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 단계를 포함하되,

상기 전처리를 수행하는 단계는

상기 클라이언트 장치가 상기 사용자 질의 중 상기 타깃 단어를 검출하는 단계;

상기 클라이언트 장치가 상기 타깃 단어를 임베딩하는 단계;

상기 클라이언트 장치가 상기 타깃 단어의 임베딩 값과의 유사도를 기준으로 상기 대체 단어를 결정하는 단계; 및

상기 클라이언트 장치가 상기 사용자 질의에서 상기 타깃 단어를 상기 대체 단어로 교체하는 단계를 포함하되,

상기 클라이언트 장치는 다수의 단어들 및 상기 다수의 단어들의 임베딩 값을 저장하는 테이블에서 상기 타깃 단어의 임베딩 값과 가장 유사한 임베딩 값을 갖는 단어를 상기 대체 단어로 결정하는 개인 정보 노출 없는 인공지능 기반 서비스 제공 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 클라이언트 장치는 사전에 구축한 학습모델을 이용하여 상기 사용자 질의 중 상기 타깃 단어를 검출하는 개인 정보 노출 없는 인공지능 기반 서비스 제공 방법.

청구항 4

삭제

청구항 5

제1항에 있어서,

상기 클라이언트 장치가 상기 타깃 단어와 상기 대체 단어를 테이블에 저장하는 단계를 더 포함하고,

상기 클라이언트 장치는 상기 답변에 상기 대체 단어가 존재하는 경우, 상기 테이블을 참조하여 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 개인 정보 노출 없는 인공지능 기반 서비스 제공 방법.

청구항 6

다수의 단어들 및 상기 다수의 단어들의 임베딩 값을 제1 테이블에 저장하는 저장장치;

사용자 질의를 입력받는 인터페이스 장치;

전처리된 사용자 질의를 송신하고, 상기 전처리된 사용자 질의에 대한 인공지능 기반 서비스 결과인 답변을 수

신하는 통신장치; 및

상기 입력된 사용자 질의 중 개인 정보에 해당하는 타깃 단어를 대체 단어로 변경하는 상기 전처리를 수행하고, 상기 답변에서 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 연산장치를 포함하되,

상기 연산장치는 상기 제1 테이블에서 상기 타깃 단어의 임베딩 값과 가장 유사한 임베딩 값을 갖는 단어를 상기 대체 단어로 결정하는 인공지능 기반 서비스 제공받는 클라이언트 장치.

청구항 7

제6항에 있어서,

상기 연산장치는 사전에 구축한 학습모델을 이용하여 상기 입력된 사용자 질의 중 상기 타깃 단어를 검출하는 인공지능 기반 서비스 제공받는 클라이언트 장치.

청구항 8

삭제

청구항 9

제6항에 있어서,

상기 타깃 단어의 임베딩 값은 노이즈를 포함한 값인, 인공지능 기반 서비스 제공받는 클라이언트 장치.

청구항 10

제6항에 있어서,

상기 저장장치는 상기 타깃 단어와 상기 대체 단어를 제2 테이블에 저장하고,

상기 연산장치는 상기 답변에 상기 대체 단어가 존재하는 경우 상기 제2 테이블을 참조하여 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 인공지능 기반 서비스 제공받는 클라이언트 장치.

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 인공지능 기반 서비스 제공 기법에 관한 것이다.

배경 기술

[0002] 최근 다양한 초거대 AI(artificial intelligence) 모델이 개발되고 있다. 예컨대, 초거대 AI 모델은 검색 엔진에 연동되어 개인 사용자의 질의 등을 입력받아 처리한다. 이때 사용자가 입력하는 내용은 개인 정보 또는 민감 정보가 포함될 수 있다. 이 경우 개인 정보 또는 민감 정보가 서버에 전달되기에 정보 노출의 위험성이 존재한다.

선행기술문헌

특허문헌

[0003] (특허문헌 0001) 한국공개특허 제10-2022-0050657호

발명의 내용

해결하려는 과제

[0004] 이하 설명하는 기술은 자연어를 처리하는 외부 서버에 개인 정보/민감 정보를 노출하지 않으면서 AI 기반 서비스를 제공하는 기법을 제공하고자 한다.

과제의 해결 수단

- [0005] 개인 정보 노출 없는 인공지능 기반 서비스 제공 방법은 클라이언트 장치가 사용자 질의를 입력받는 단계, 상기 클라이언트 장치가 상기 사용자 질의 중 개인 정보에 해당하는 타깃 단어를 대응하는 대체 단어로 변경하는 전처리를 수행하는 단계, 상기 클라이언트 장치가 상기 전처리된 사용자 질의를 인공지능 기반 서비스를 제공하는 서비스 서버에 전달하는 단계 및 상기 클라이언트 장치가 상기 서비스 서버로부터 수신한 상기 사용자 질의에 대응하는 답변에서 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 단계를 포함한다.
- [0006] 인공지능 기반 서비스 제공받는 클라이언트 장치는 사용자 질의를 입력받는 인터페이스 장치, 전처리된 사용자 질의를 송신하고, 상기 전처리된 사용자 질의에 대한 인공지능 기반 서비스 결과인 답변을 수신하는 통신장치 및 상기 입력된 사용자 질의 중 개인 정보에 해당하는 타깃 단어를 대체 단어로 변경하는 상기 전처리를 수행하고, 상기 답변에서 상기 대체 단어를 대응하는 상기 타깃 단어로 대체하는 연산장치를 포함한다.

발명의 효과

- [0007] 이하 설명하는 기술은 초거대 AI 모델을 활용한 서비스에서 개인 정보/민감 정보가 제거하여 정보 노출 위험 없으면서도 원활한 서비스를 제공할 수 있다.

도면의 간단한 설명

- [0008] 도 1은 관심 정보 노출 없는 LLM 서비스 시스템에 대한 예이다.
- 도 2는 관심 정보 노출 없이 LLM 서비스를 제공하는 과정에 대한 예이다.
- 도 3은 클라이언트 장치에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

- [0009] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0010] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0011] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계 동작 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [0012] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.
- [0013] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0015] 이하 설명하는 기술은 초거대 AI 기반의 서비스에 활용할 수 있는 기법이다.
- [0016] 이하 설명하는 기법은 다양한 AI 모델에 적용할 수 있다. 예컨대, AI 모델은 LLM(Large Language Model)일 수

있다. LLM은 다양한 방법론으로 구현될 수 있다.

- [0017] 이하 설명하는 기법은 개인 정보 및/또는 민감 정보 등의 노출 없이 LLM 기반의 서비스를 제공한다.
- [0018] 개인 정보와 민감 정보는 법률적으로는 서로 다른 개념이다.
- [0019] 개인 정보는 개인에 관한 성명, 주민등록번호 및 영상 등을 통하여 개인을 알아볼 수 있는 정보를 의미한다. 개인 정보는 주민등록번호, 여권번호, 운전면허번호, 외국인 등록 번호 등과 같은 고유식별정보를 포함할 수 있다.
- [0020] 민감 정보는 사상, 신념, 노동조합 또는 정당의 가입/탈퇴, 정치적 견해, 건강, 성생활 등에 관한 정보, 그 밖에 정보주체의 사생활을 현저히 침해할 우려가 있는 정보를 의미한다.
- [0021] 이하 AI 기반 서비스를 제공하는 서버 또는 AI 모델에 노출 우려가 있는 정보를 관심 정보라고 명명한다. 관심 정보는 개인 정보, 민간 정보 및 고유식별정보 중 적어도 하나를 포함할 수 있다.
- [0022] 다만, 개인 정보는 넓게는 민감 정보도 포함하는 의미로 사용될 수 있다. 즉, 개인 정보는 이하 설명하는 관심 정보와 동일한 의미도 사용될 수도 있다.
- [0024] 도 1은 관심 정보 노출 없는 LLM 서비스 시스템(100)에 대한 예이다.
- [0025] 클라이언트 장치(110)는 개인 사용자가 사용하는 단말을 의미한다. 클라이언트 장치(110)는 스마트 기기, PC, 노트북, 웨어러블 기기, 스마트 스피커 등 다양한 장치 중 어느 하나일 수 있다. 클라이언트 장치(110)는 텍스트 또는 음성으로 사용자 질의를 입력받을 수 있다(① 사용자 질의 입력). 음성으로 입력받는 경우 클라이언트 장치(110)는 음성을 텍스트로 변환한다.
- [0026] 클라이언트 장치(110)는 사용자가 입력한 질의를 일정하게 전처리한다(② 사용자 질의 전처리). 클라이언트 장치(110)는 사용자 질의 중 관심 정보를 검출한다. 클라이언트 장치(110)는 사용자 질의인 텍스트에서 검출한 관심 정보를 삭제하거나, 다른 단어로 변경할 수 있다. 이때 관심 정보를 대체하는 단어는 개인 정보 또는 민감 정보가 포함되지 않은 단어이다. 이하 관심 정보를 대체하는 단어를 대체 단어라고 명명한다.
- [0027] 클라이언트 장치(110)는 관심 정보가 삭제 내지 대체된 사용자 질의를 서비스 서버(120)에 전달한다(③ 전처리된 사용자 질의 전달). 이때 클라이언트 장치(110)는 관심 정보와 해당 관심 정보에 대응하는 대체 단어를 매칭하여 저장한다.
- [0028] 서비스 서버(120)는 전처리된 사용자 질의를 수신하고, LLM을 이용하여 질의에 대응하는 답변을 생성한다(④ 질의 대응하는 답변 생성). 일반적으로 서비스 서버(120)는 전처리된 질의를 임베딩 모델을 이용하여 일정하게 임베딩한다. 서비스 서버(120)는 임베딩된 벡터를 사전에 구축한 벡터 DB(database)를 이용하여 일정한 프롬프트(prompt)로 변환한다. 서비스 서버(120)는 프롬프트를 LLM에 입력하여 질의에 대응하는 답변을 생성한다.
- [0029] 클라이언트 장치(110)는 서비스 서버(120)로부터 질의에 대응하는 답변을 수신한다. 클라이언트 장치(110)는 수신한 답변 텍스트에서 대체 단어를 사전에 저장한 정보를 이용하여 관심 정보로 변경한다(⑤ 답변에서 관심 정보 복원). 클라이언트 장치(110)는 관심 정보가 복원된 답변을 출력할 수 있다.
- [0031] 도 2는 관심 정보 노출 없이 LLM 서비스를 제공하는 과정(200)에 대한 예이다.
- [0032] 클라이언트 장치(110)는 사용자의 질의를 입력받는다(210). 사용자 질의가 음성 데이터이면, 클라이언트 장치(110)는 음성 데이터를 텍스트 데이터로 변환한다. 클라이언트 장치(110)는 사용자 질의를 다수의 토큰들로 파싱(parsing)할 수 있다. 사용자의 질의는 {text1 text2 private1 text3 text4 private2 text5 text6 text7 private3 text8 text9 private4 ...}와 같은 다수의 토큰들로 구성된다고 가정한다. "text"는 관심 정보가 아닌 일반적인 토큰(단어)을 의미하고, "private"는 개인 정보 또는 민감 정보가 포함된 토큰을 의미한다.
- [0033] 클라이언트 장치(110)는 사용자 질의 중 관심 정보를 검출한다(220). 관심 정보를 검출하는 방법론은 다양할 수 있다.
- [0034] (1) 클라이언트 장치(110)는 사전에 구축된 관심 정보 테이블을 이용하여 사용자 질의 중 관심 정보를 검출할 수 있다. 관심 정보 테이블은 관심 정보에 해당하는 키워드들 내지 단어들로 구성될 수 있다. 관심 정보 테이블은 업데이트될 수 있다. 클라이언트 장치(110)는 관심 정보 테이블을 이용하여 사용자의 질의 중 일반 단어 및 관심 정보를 각각 라벨링할 수 있다.
- [0035] (2) 클라이언트 장치(110)는 사전에 구축된 별도의 자연어 처리 모델을 이용하여 사용자 질의 중 관심 정보를

검출할 수 있다. 즉, 이 경우 자연어 처리 모델은 입력되는 텍스트 중 관심 정보를 분류 내지 라벨링하는 모델에 해당하며 사전에 구축되어야 한다. 클라이언트 장치(110)는 사용자 질의 내지 사용자 질의에서 추출한 토큰을 자연어 처리 모델에 입력하여 사용자 질의 중 관심 정보를 분류 내지 검출할 수 있다.

[0036] 클라이언트 장치(110)는 사용자 질의인 텍스트를 일정하게 임베딩할 수 있다(230). 클라이언트 장치(110)는 텍스트에서 추출한 각 토큰을 일정하게 임베딩할 수 있다. 이때 텍스트 임베딩은 서비스 서버(120)에서 동작하는 LLM의 입력 전 처리하는 임베딩과 동일한 기법일 수 있다.

[0037] 클라이언트 장치(110)는 사용자 질의 중 관심 정보에 일정한 노이즈(noise)를 추가(노이즈 임베딩)하여 교란할 수 있다(240). 이때 노이즈는 일정한 규칙에 따라 결정되는 값일 수 있다. 또는 클라이언트 장치(110)는 DP(differential privacy) 기법에 따라 관심 정보(토큰)에 노이즈를 부가할 수도 있다.

[0038] 클라이언트 장치(110)는 노이즈를 임베딩한 값과 가장 가까운 토큰 내지 단어를 찾아 관심 정보를 대체할 수 있다(250). 도 2에서 DB는 특정 텍스트와 해당 텍스트를 임베딩한 값을 매칭한 테이블을 저장할 수 있다. 텍스트와 해당 텍스트의 임베딩 값을 매칭한 테이블을 텍스트 임베딩 테이블이라고 명명한다. 텍스트 임베딩 테이블은 대량의 단어들과 해당 단어들 각각의 임베딩 값을 매칭하여 저장한다.

[0039] 클라이언트 장치(110)는 해당 DB에서 노이즈를 임베딩한 값과 가장 가까운 토큰(단어)을 찾아, 사용자 질의에서 관심 정보를 대체한다. 도 2에서 관심 정보(private)들은 모두 대체 단어(non-private)로 대체되었다.

[0040] 클라이언트 장치(110)는 사용자의 질의 중 관심 정보와 해당 관심 정보를 대체한 대체 단어를 아래 표 1과 같이 테이블 형태로 저장한다(260). 표 1과 같은 테이블을 대체 단어 테이블이라고 명명한다.

표 1

[0041]

No.	관심 정보	대체 단어
1	privatel	non-privatel
2	private2	non-private2
3	private3	non-private3
4	private4	non-private4
...	...	...

[0042] 클라이언트 장치(110)는 관심 정보가 대체 단어로 변경된(전처리된) 사용자 질의를 서비스 서버(120)에 전달한다.

[0043] 서비스 서버(120)는 전처리된 사용자 질의를 일정하게 벡터화하여 LLM에 입력하고, LLM을 이용하여 사용자 질의에 대한 LLM 결과값(답변)을 산출한다(270). 사용자 질의에 대응하는 답변은 {output1 non-privatel output2 non-private2 output3 output4 non-private3 output5 non-private4 ...}으로 구성된다. 서비스 서버(120)는 사용자 질의에 대한 답변을 클라이언트 장치(110)에 전달한다.

[0044] 클라이언트 장치(110)는 수신한 답변에서 대체 단어를 검출한다. 클라이언트 장치(110)는 대체 단어 테이블을 이용하여 답변에서 대체 단어를 대응하는 관심 정보로 복원하고, 관심 정보를 복원한 답변을 출력할 수 있다(280).

[0046] 클라이언트 장치는 날짜, 나이, 식별정보 등을 포함하는 사용자 질의를 입력받을 수 있다. 예컨대, 클라이언트 장치는 "9월 4일 미팅 참석자는 AAA(실명) 포함 10명입니다."라는 질의를 입력받을 수 있다. 클라이언트 장치는 해당 질의에서 관심 정보를 검출하여 다른 대체 단어로 변경할 수 있다. 예컨대, 클라이언트 장치는 사용자 질의를 "[DATE_A]일 미팅 참석자는 [NAME_A] 포함 10명입니다."와 같이 전처리할 수 있다. 이 경우 대체 단어는 날짜, 이름 등을 일반화한 단어일 수도 있습니다. 서비스 서버는 전처리된 질의를 입력받아 일정한 답변을 생성할 수 있다. 예컨대, 클라이언트 장치가 질의와 함께 일정한 요청(내용 요약)을 서비스 서버에 전달할 수 있다. 이때 서비스 서버는 "(1)일시: [DATE_A], (2) 미팅 참석자: [NAME_A] 포함 10명"이라는 답변을 생성할 수 있다. 클라이언트 장치는 수신한 답변에서 대체 단어를 알고 있는 관심 정보로 대체하여 "(1)일시: 9월 4일, (2) 미팅 참석자: AAA 포함 10명"과 같이 출력할 수 있다.

[0048] 도 3은 클라이언트 장치(300)에 대한 예이다. 도 3의 클라이언트 장치(300)는 도 1의 클라이언트 장치(110)에 해당한다. 클라이언트 장치(300)는 사용자의 질의를 입력받아 관심 정보를 제거하여 서비스 서버에 전달하고, 질의에 대한 답변에서 관심 정보를 복원하는 장치이다. 클라이언트 장치(300)는 스마트기기, PC, 웨어러블

기기, 스마트 스피커, 리모컨(Remote Controller) 등 다양한 형태 중 어느 하나로 구현될 수 있다.

- [0049] 클라이언트 장치(300)는 저장 장치(310), 메모리(320), 연산장치(330), 인터페이스 장치(340) 및 통신 장치(350)를 포함할 수 있다. 또는 클라이언트 장치(300)는 저장 장치(310), 메모리(320), 연산장치(330), 인터페이스 장치(340), 통신 장치(350) 및 출력 장치(360)를 포함할 수 있다.
- [0050] 저장 장치(310)는 사용자 질의를 저장할 수 있다.
- [0051] 저장 장치(310)는 대체 단어 테이블을 저장할 수 있다. 대체 단어 테이블은 관심 정보와 해당 정보에 매칭되는 대체 단어를 저장할 수 있다.
- [0052] 저장 장치(310)는 텍스트 임베딩 테이블을 저장할 수 있다. 텍스트 임베딩 테이블은 임의의 단어들과 해당 단어들 각각의 임베딩 값을 매칭하여 저장할 수 있다. 텍스트 임베딩 테이블의 임베딩 값은 해당 단어(토큰)에 노이즈가 포함된 값일 수 있다. 이때 노이즈는 DP 기법에 따른 값일 수 있다.
- [0053] 저장 장치(310)는 데이터 전처리, 관심 정보 검출, 관심 정보 대체, 관심 정보 복원 등을 제어하는 소스 코드 내지 프로그램을 저장할 수 있다.
- [0054] 메모리(320)는 클라이언트 장치(300)가 사용자 질의에서 관심 정보 제거 내지 대체 및 답변에서의 관심 정보 복원을 수행하는 과정에서 발생하는 데이터를 저장할 수 있다.
- [0055] 인터페이스 장치(340)는 일정한 정보 내지 데이터를 입력받는 장치이다.
- [0056] 인터페이스 장치(340)는 텍스트 데이터를 입력받는 장치 (키보드, 마우스, 터치 스크린 등) 또는 음성을 입력받는 장치(마이크)일 수 있다. 또는 인터페이스 장치(340)는 외부 객체 또는 연결된 장치로부터 내부로 정보를 전달받는 장치를 의미할 수도 있다.
- [0057] 인터페이스 장치(340)는 입력 장치 또는 외부 객체(외부 저장 장치 등)로부터 사용자 질의를 입력받을 수 있다. 이때 사용자 질의는 파일 형태일 수도 있다.
- [0058] 인터페이스 장치(340)는 관심 정보가 대체된 사용자 질의(전처리된 사용자 질의)를 외부 객체로 전달할 수도 있다.
- [0059] 인터페이스 장치(340)는 외부 객체로부터 사용자 질의에 대한 답변을 입력받을 수 있다.
- [0060] 통신 장치(350)는 유선 또는 무선 네트워크를 통해 일정한 정보를 수신하고 전송하는 구성을 의미한다.
- [0061] 통신 장치(350)는 외부 객체로부터 사용자 질의를 수신할 수 있다. 이때 사용자 질의는 파일 형태일 수도 있다.
- [0062] 통신 장치(350)는 관심 정보가 대체된 사용자 질의(전처리된 사용자 질의)를 외부 객체(서비스 서버)에 송신할 수도 있다.
- [0063] 통신 장치(350)는 LLM을 이용하여 결정된 사용자 질의에 대한 답변을 외부 객체(서비스 서버)로부터 수신할 수 있다.
- [0064] 통신 장치(350)는 관심 정보가 복원된 답변을 외부 객체에 송신할 수도 있다.
- [0065] 나아가, 인터페이스 장치(340)는 통신 장치(350)를 통해 수신하는 정보 내지 데이터를 클라이언트 장치(300) 내부로 전달하는 구성을 포함할 수 있다.
- [0066] 연산 장치(330)는 음성 데이터 형태의 사용자 질의를 텍스트 형태로 변환할 수 있다. 텍스트 변환은 전통적인 음성 처리 기법 또는 딥러닝 기반의 모델이 활용될 수 있다.
- [0067] 연산 장치(330)는 사용자 질의에 대한 토큰화를 수행할 수 있다.
- [0068] 연산 장치(330)는 사용자 질의(토큰)에서 관심 정보를 검출할 수 있다. 연산 장치(330)는 전술한 바와 같이 관심 정보 테이블을 이용하거나, 관심 정보 분류를 위해 사전에 구축된 별도의 학습모델을 이용하여 사용자 질의 중 관심 정보를 검출할 수 있다. 연산 장치(330)는 사용자 질의(토큰) 중 관심 정보를 라벨링할 수 있다.
- [0069] 연산 장치(330)는 도 2에서 설명한 과정을 통해 사용자 질의 중 관심 정보(토큰)를 대체 단어로 변경할 수 있다.
- [0070] 연산 장치(330)는 사용자 질의 중 관심 정보와 해당 관심 정보에 대응하는 대체 단어를 테이블(대체 단어 테이블)

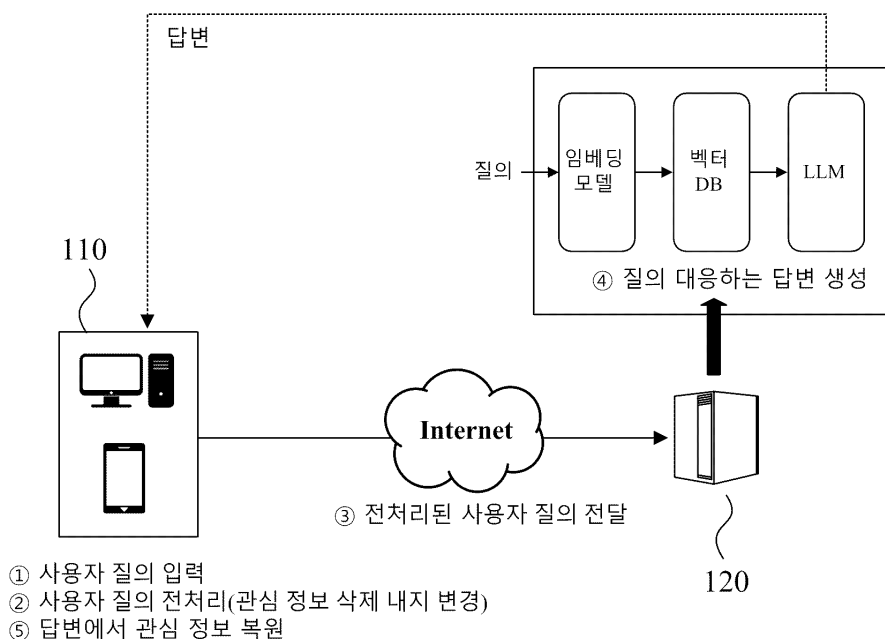
블) 형태로 저장할 수 있다.

- [0071] 연산 장치(330)는 서비스 서버로부터 수신한 질의에 대한 답변에서 관심 정보를 복원할 수 있다. 연산 장치(330)는 대체 단어 테이블을 이용하여 답변에서 대체 단어를 대응하는 관심 정보로 변경할 수 있다.
- [0072] 연산 장치(330)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드된 칩과 같은 장치일 수 있다.
- [0073] 출력 장치(360)는 사용자 질의를 입력받는 인터페이스 화면 및/또는 관심 정보가 복원된 답변을 출력할 수 있다.
- [0075] 또한, 상술한 바와 같은 사용자 질의 처리 방법, LLM의 답변에서 관심 정보를 복원하는 방법 내지 관심 정보 노출 없는 LLM 서비스 제공 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [0076] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(read-only memory), PROM(programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [0077] 일시적 판독 가능 매체는 스테틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(Synclink DRAM, SDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.
- [0078] 본 실시예 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

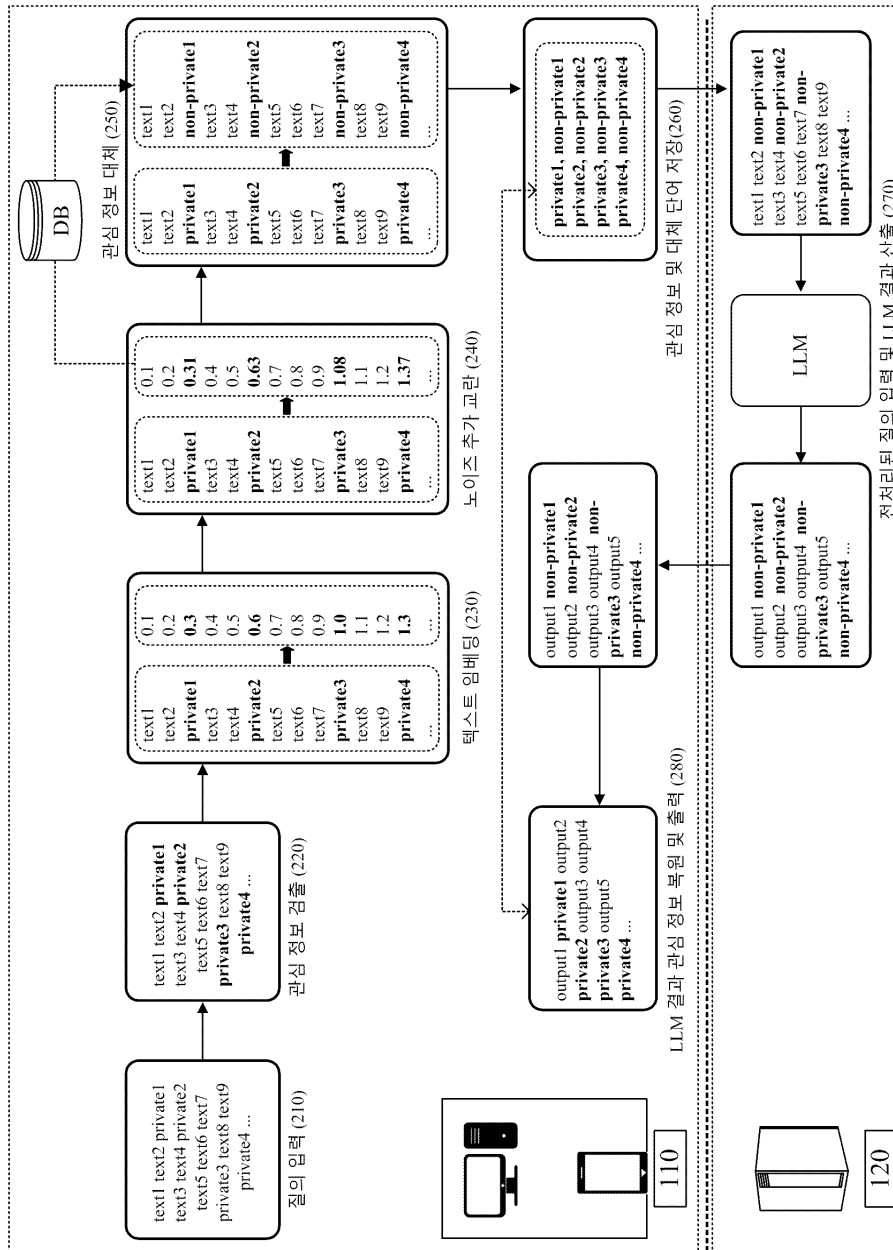
도면1

100



도면2

200



도면3

