



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년05월25일
 (11) 등록번호 10-1739539
 (24) 등록일자 2017년05월18일

(51) 국제특허분류(Int. Cl.)
 G06F 17/30 (2006.01)
 (52) CPC특허분류
 G06F 17/30303 (2013.01)
 G06F 17/30129 (2013.01)
 (21) 출원번호 10-2016-0009404
 (22) 출원일자 2016년01월26일
 심사청구일자 2016년01월26일
 (56) 선행기술조사문헌
 KR1020130140847 A*
 JP2013161472 A*
 *는 심사관에 의하여 인용된 문헌

(73) 특허권자
주식회사 솔트룩스
 서울특별시 강남구 언주로 538, 4층, 5층(역삼동, 대웅빌딩)
 (72) 발명자
이경일
 서울특별시 강남구 학동로77길 27, 102동 202호 (청담동, 대림아파트)
함영경
 서울특별시 강남구 언주로106길 23, 201호 (역삼동)
 (74) 대리인
리앤목특허법인

전체 청구항 수 : 총 8 항

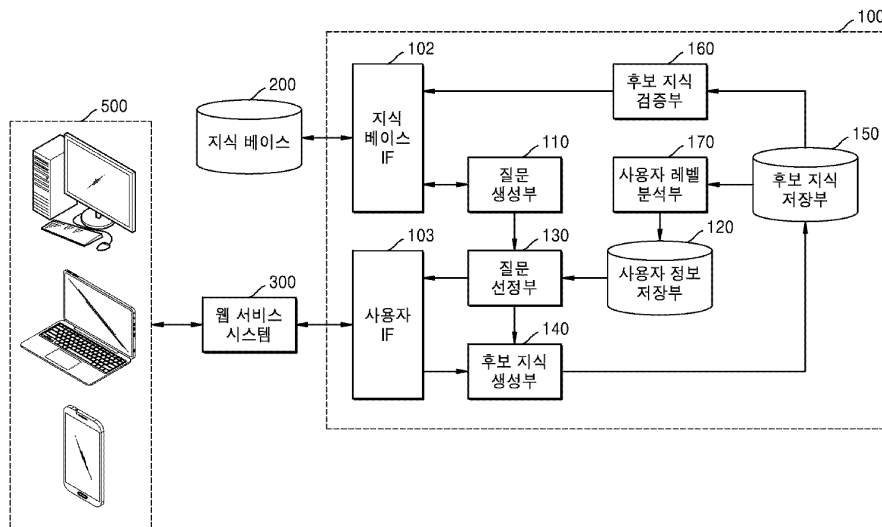
심사관 : 홍경아

(54) 발명의 명칭 지식 베이스의 검증 및 교정 시스템 및 방법

(57) 요약

지식 베이스의 검증 및 교정 시스템이 개시된다. 본 발명의 예시적 실시예에 따른 지식 베이스 교정 시스템은, 지식 베이스로부터 불완전한 지식 데이터를 검출하고 답변으로써 불완전한 지식 데이터를 교정하기 위한 질문을 생성하는 질문 생성부, 복수의 사용자들의 레벨 정보를 저장하는 사용자 정보 저장부, 및 복수의 사용자들의 레벨 정보에 기초하여 사용자들 각각에 배정될 질문의 개수 및 질문을 결정하는 질문 선정부를 포함할 수 있다.

대표도



(52) CPC특허분류

G06F 17/30371 (2013.01)

G06F 17/30539 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 R0101-15-0054

부처명 미래창조과학부

연구관리전문기관 정보통신기술연구진흥센터

연구사업명 SW컴퓨팅산업원천기술개발사업(SW)

연구과제명 WiseKB: 빅데이터 이해 기반 자가학습형 지식베이스 및 추론 기술 개발

기여율 1/1

주관기관 ㈜솔트룩스

연구기간 2015.03.01 ~ 2016.02.29

명세서

청구범위

청구항 1

지식 베이스로부터 불완전한 지식 데이터를 검출하고, 답변으로써 상기 불완전한 지식 데이터를 교정하기 위한 질문을 생성하는 질문 생성부;

복수의 사용자들의 레벨 정보를 저장하는 사용자 정보 저장부; 및

상기 복수의 사용자들의 레벨 정보에 기초하여 사용자들 각각에 배정될 질문의 개수 및 질문을 결정하는 질문 선정부를 포함하고,

상기 질문 생성부는, 복수의 값들이 기입된 속성을 가지는 제1 인스턴스를 포함하는 지식 데이터, 기입된 값이 속성의 형식과 일치하지 아니하는 제2 인스턴스를 포함하는 지식 데이터 또는 값이 누락된 속성을 가지는 제3 인스턴스를 포함하는 지식 데이터를 검출하는 오류 검출부를 포함하는 것을 특징으로 하는 지식 베이스 교정 시스템.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 질문 생성부는, 상기 제1 인스턴스, 상기 복수의 값들이 맵핑된 속성 및 상기 복수의 값들을 포함하는 객관식 질문, 상기 제2 인스턴스, 상기 기입된 값이 형식과 일치하지 아니하는 속성을 포함하는 주관식 질문 또는 상기 제3 인스턴스 및 상기 값이 누락된 속성을 포함하는 주관식 질문을 생성하는 질문 출력부를 더 포함하는 것을 특징으로 하는 지식 베이스 교정 시스템.

청구항 4

제1항에 있어서,

상기 질문 선정부에 의해서 선정된 질문 및 상기 질문에 대한 적어도 하나의 답변을 포함하는 후보 지식 데이터를 생성하는 후보 지식 생성부를 더 포함하는 지식 베이스 교정 시스템.

청구항 5

제4항에 있어서,

동일한 질문에 대응하는 적어도 하나의 후보 지식 데이터에 기초하여 후보 지식 데이터를 검증하고, 검증이 성공한 후보 지식 데이터에 기초하여 상기 지식 베이스에 저장된 지식 데이터를 교정하는 후보 지식 검증부를 더 포함하는 지식 베이스 교정 시스템.

청구항 6

제5항에 있어서,

상기 후보 지식 검증부는 검증이 실패한 후보 지식 데이터에 대응하는 질문을 상기 질문 선정부에 제공하고,

상기 질문 선정부는 상기 후보 지식 검증부에 의해 제공된 질문을 기배정된 사용자들과 상이한 복수의 사용자들에게 배정하도록 결정하는 것을 특징으로 하는 지식 베이스 교정 시스템.

청구항 7

제5항에 있어서,

상기 후보 지식 데이터 및 상기 후보 지식 검증부의 상기 후보 지식 데이터에 대한 검증 결과에 기초하여, 상기

사용자 정보 저장부에 저장된 사용자의 레벨 정보를 갱신하는 사용자 레벨 분석부를 더 포함하는 지식 베이스 교정 시스템.

청구항 8

제1항에 있어서,

상기 질문을 전송하고 상기 답변을 수신하는 사용자 인터페이스를 더 포함하는 지식 베이스 교정 시스템.

청구항 9

제1항에 있어서,

상기 사용자 정보 저장부에 저장된 복수의 사용자들의 레벨 정보를 리워드 시스템에 제공하는 리워드 인터페이스를 더 포함하는 지식 베이스 교정 시스템.

발명의 설명

기술 분야

[0001] 본 발명의 기술적 사상은 지식 베이스의 검증 및 교정 시스템 및 방법에 관한 것으로서, 자세하게는 불완전한 지식 데이터를 검출하고 교정하는 시스템 및 방법에 관한 것이다.

[0002] 본 발명은 미래창조과학부 SW컴퓨팅산업원천기술개발사업(SW)의 일환으로 (주)솔트룩스가 주관하고 연구하여 수행된 연구로부터 도출된 것이다. [연구기간: 2015.03.01~2016.02.29, 연구관리 전문기관: 정보통신기술연구진흥센터, 연구과제명: WiseKB: 빅데이터 이해 기반 자가학습형 지식베이스 및 추론 기술 개발, 과제 고유번호: R0101-15-0054]

배경 기술

[0003] 지식 데이터를 저장하고 저장된 지식 데이터를 제공하는 지식 베이스(knowledge base)는 다양한 방식으로 구축될 수 있다. 예를 들면, 특정한 도메인의 전문가 그룹에 의해서 그 도메인에 대한 지식 베이스가 구축될 수도 있고, 데이터로부터 지식을 추출함으로써 지식 베이스가 구축될 수도 있다. 전자의 경우, 높은 정확도의 지식 데이터를 제공할 수 있는 반면, 구축 가능한 지식 베이스의 규모가 제한적일 수 있다. 이에 따라, 데이터를 수집하고 수집된 데이터로부터 지식 베이스를 구축하는 방법이 고려되고 있다.

발명의 내용

해결하려는 과제

[0004] 본 발명의 기술적 사상은, 클라우드 소싱을 통해서 불완전한 지식 데이터를 정정함으로써 지식 베이스를 검증 및 교정하는 시스템 및 방법을 제공한다.

과제의 해결 수단

[0005] 상기와 같은 목적을 달성하기 위하여, 본 발명의 기술적 사상의 일측면에 따른 지식 베이스 교정 시스템은, 지식 베이스로부터 불완전한 지식 데이터를 검출하고 답변으로써 불완전한 지식 데이터를 교정하기 위한 질문을 생성하는 질문 생성부, 복수의 사용자들의 레벨 정보를 저장하는 사용자 정보 저장부, 및 복수의 사용자들의 레벨 정보에 기초하여 사용자들 각각에 배정될 질문의 개수 및 질문을 결정하는 질문 선정부를 포함할 수 있다.

[0006] 본 발명의 예시적 실시예에 따라, 질문 생성부는, 복수의 값들이 기입된 속성을 가지는 제1 인스턴스를 포함하는 지식 데이터, 기입된 값이 속성의 형식과 일치하지 아니하는 제2 인스턴스를 포함하는 지식 데이터 또는 값이 누락된 속성을 가지는 제2 인스턴스를 포함하는 지식 데이터를 검출하는 오류 검출부를 포함할 수 있다.

[0007] 본 발명의 예시적 실시예에 따라, 질문 생성부는, 제1 인스턴스, 복수의 값들이 맵핑된 속성 및 복수의 값들을 포함하는 객관식 질문, 제2 인스턴스, 기입된 값이 형식과 일치하지 아니하는 속성을 포함하는 주관식 질문 또는 제3 인스턴스 및 값이 누락된 속성을 포함하는 주관식 질문을 생성하는 질문 출력부를 더 포함할 수 있다.

[0008] 본 발명의 예시적 실시예에 따라, 질문 선정부에 의해서 선정된 질문 및 질문에 대한 적어도 하나의 답변을 포

합하는 후보 지식 데이터를 생성하는 후보 지식 생성부를 더 포함할 수 있다.

- [0009] 본 발명의 예시적 실시예에 따라, 동일한 질문에 대응하는 적어도 하나의 후보 지식 데이터에 기초하여 후보 지식 데이터를 검증하고, 검증이 성공한 후보 지식 데이터에 기초하여 지식 베이스에 저장된 지식 데이터를 교정하는 후보 지식 검증부를 더 포함할 수 있다.
- [0010] 본 발명의 예시적 실시예에 따라, 후보 지식 검증부는 검증이 실패한 후보 지식 데이터에 대응하는 질문을 질문 선정부에 제공할 수 있고, 질문 선정부는 후보 지식 검증부에 의해 제공된 질문을 기배정된 사용자들과 상이한 복수의 사용자들에게 배정하도록 결정할 수 있다.
- [0011] 본 발명의 예시적 실시예에 따라, 후보 지식 데이터 및 후보 지식 검증부의 후보 지식 데이터에 대한 검증 결과에 기초하여, 사용자 정보 저장부에 저장된 사용자의 레벨 정보를 갱신하는 사용자 레벨 분석부를 더 포함할 수 있다.
- [0012] 본 발명의 예시적 실시예에 따라, 질문을 전송하고 답변을 수신하는 사용자 인터페이스를 더 포함할 수 있다.
- [0013] 본 발명의 예시적 실시예에 따라, 사용자 정보 저장부에 저장된 복수의 사용자들의 레벨 정보를 리워드 시스템에 제공하는 리워드 인터페이스를 더 포함할 수 있다.

발명의 효과

- [0014] 본 발명의 기술적 사상에 따른 지식 베이스 검증 및 교정 시스템 및 방법은, 방대한 지식 베이스의 불완전한 지식 데이터를 교정하기 위한 수단을 제공함으로써 지식 베이스의 신뢰도 및 활용도를 향상시킬 수 있다.
- [0015] 또한, 본 발명의 기술적 사상에 따른 지식 베이스 검증 및 교정 시스템 및 방법은, 클라우드 소싱을 통해서 불완전한 지식 데이터에 대한 교정의 신뢰성을 높일 수 있다.
- [0016] 또한, 본 발명의 기술적 사상에 따른 지식 베이스 검증 및 교정 시스템 및 방법은, 불완전한 지식 데이터의 교정에 참여하는 사용자에게 이득을 주기 위한 수단을 제공함으로써, 지식 베이스가 지속적으로 고품질의 서비스를 제공하는 것을 가능하게 한다.

도면의 간단한 설명

- [0017] 도 1은 본 발명의 예시적 실시예에 따른 지식 베이스 교정 시스템의 블록도이다.
- 도 2는 본 발명의 예시적 실시예에 따라 도 1의 질문 생성부의 예시를 나타내는 블록도이다.
- 도 3은 불완전한 지식 데이터의 예시를 나타내는 도면이다.
- 도 4는 본 발명의 예시적 실시예에 따라 도 1의 후보 지식 검증부의 동작을 설명하기 위한 도면이다.
- 도 5는 본 발명의 예시적 실시예에 따른 지식 베이스 교정 시스템의 블록도이다.
- 도 6은 본 발명의 예시적 실시예에 따른 지식 베이스의 교정 방법을 나타내는 순서도이다.
- 도 7은 본 발명의 예시적 실시예에 따른 지식 베이스 교정 방법을 나타내는 순서도이다.

발명을 실시하기 위한 구체적인 내용

- [0018] 이하, 첨부한 도면을 참조하여 본 발명의 실시 예에 대해 상세히 설명한다. 본 발명의 실시 예는 당 업계에서 평균적인 지식을 가진 자에게 본 발명을 보다 완전하게 설명하기 위하여 제공되는 것이다. 본 발명은 다양한 변경을 가할 수 있고 여러 가지 형태를 가질 수 있는 바, 특정 실시 예들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 개시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용한다. 첨부된 도면에 있어서, 구조물들의 치수는 본 발명의 명확성을 기하기 위하여 실제보다 확대하거나 축소하여 도시한 것이다.
- [0019] 본 출원에서 사용한 용어는 단지 특정한 실시 예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수개의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요

소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

- [0020] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 갖는다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0021] 이하 도면 및 설명에서, 하나의 블록으로 표시 또는 설명되는 구성요소는 하드웨어 블록 또는 소프트웨어 블록일 수 있다. 예를 들면, 구성요소들 각각은 서로 신호를 주고 받는 독립적인 하드웨어 블록일 수도 있고, 또는 하나의 프로세서에서 실행되는 소프트웨어 블록일 수도 있다.
- [0022] 도 1은 본 발명의 예시적 실시예에 따른 지식 베이스 교정 시스템(100)의 블록도이다.
- [0023] 데이터를 수집하고 수집된 데이터로부터 추출된 지식에 기반하여 구축된 지식 베이스는, 방대한 규모를 가지는 한편, 불완전한 지식 데이터, 예컨대 오류를 포함하거나 불충분한 지식 데이터를 포함할 수 있다. 예를 들면, 지식 베이스 구축의 기반이 되는 데이터는, 데이터베이스(예컨대, 위키피디아(Wikipedia)), CSV 파일 등과 같은 정형 데이터를 포함할 수도 있고, 뉴스, 블로그, SNS, 문서 파일 등과 같은 비정형 데이터를 포함할 수도 있다. 정형 데이터로부터 지식 베이스가 구축되는 경우, 지식 베이스의 스키마를 이해하고 있는 전문가에 의해서 작성된 매핑 규칙에 의해서 지식 베이스는 상대적으로 용이하게 구축될 수 있다. 다른 한편으로, 비정형 데이터로부터 지식 베이스가 구축되는 경우, 형태소 분석, 구문 분석 등을 포함하는 자연어 처리 기술을 사용하여 비정형 데이터로부터 지식을 추출하고 구조화하는 작업이 필요하고, 이에 따라 자연어 처리 기술의 한계 비정형 데이터의 신뢰성 등에 기인하여 지식 베이스는 오류를 포함할 수 있다.
- [0024] 인터넷 등을 통해서 수집된 방대한 양의 비정형 데이터로부터 구축된 지식 베이스에 포함된 불완전한 지식 데이터를 교정하기 위하여, 소수의 지정된 전문가가 지식 베이스에 포함된 불완전한 지식 데이터를 검출하고 정정하는 것은 현실적으로 불가능할 수 있다. 지식 베이스에 포함된 불완전한 지식 데이터의 교정은 지식 베이스의 신뢰도를 향상시키고 활용도를 높이기 위하여 필수적인 작업으로서, 본 발명은 지식 베이스에 포함된 불완전한 지식 데이터를 검출하고, 클라우드 소싱을 통해서 용이하게 교정할 수 있는 시스템 및 방법을 제공한다.
- [0025] 도 1을 참조하면, 지식 베이스 교정 시스템(100)은 지식 베이스(200) 및 웹 서비스 시스템(300)과 통신할 수 있고, 사용자 단말들(500)이 웹 서비스 시스템(300)과 통신할 수 있다. 지식 베이스 교정 시스템(100), 지식 베이스(200), 웹 서비스 시스템(300) 및 사용자 단말들(500)은 LAN(local area network) 및 WAN(wide area network)와 같은 네트워크에 접속됨으로써 서로 통신할 수도 있고, 전용 채널을 통해서 일대일 통신에 의해서 양자간 통신할 수도 있다. 도 1에 도시된 바와 같이, 지식 베이스 교정 시스템(100)은 지식 베이스 인터페이스(102), 사용자 인터페이스(103), 질문 생성부(110), 질문 선정부(120), 사용자 정보 저장부(130), 후보 지식 생성부(140), 후보 지식 저장부(150), 후보 지식 검증부(160), 사용자 레벨 분석부(170)를 포함할 수 있다.
- [0026] 지식 베이스 인터페이스(102)는 지식 베이스 교정 시스템(100)에 포함된 다른 구성요소들에 지식 베이스(200)를 액세스하기 위한 인터페이스를 제공할 수 있다. 예를 들면, 후술되는 바와 같이, 질문 생성부(110) 및 후보 지식 검증부(160)는 지식 베이스 인터페이스(102)를 통해서 지식 베이스(200)에 액세스함으로써, 지식 베이스(200)에 저장된 지식 데이터를 수신할 수도 있고, 교정된 지식 데이터를 지식 베이스(200)에 전송할 수도 있다.
- [0027] 유사하게, 사용자 인터페이스(103)는 지식 베이스 교정 시스템(100)에 포함된 다른 구성요소들에 웹 서비스 시스템(300)을 액세스하기 위한 인터페이스를 제공할 수 있다. 예를 들면, 후술되는 바와 같이, 질문 선정부(120) 및 후보 지식 생성부(140)는 사용자 인터페이스(103)를 통해서 웹 서비스 시스템(300)에 액세스함으로써, 질문을 전송할 수도 있고, 질문에 대한 답변을 수신할 수도 있다.
- [0028] 질문 생성부(110)는 지식 베이스(200)로부터 불완전한 지식 데이터, 오류를 포함하는 지식 데이터 또는 불충분한 지식 데이터를 검출할 수 있고, 답변으로써 불완전한 지식 데이터를 교정하기 위한 질문을 생성할 수 있다. 예를 들면, 질문 생성부(110)는 상충하는 내용의 지식 또는 누락된 지식을 검출할 수 있고, 이를 교정하기 위한 내용을 답변으로 하는 질문을 생성할 수 있다. 생성된 질문들 중 일부는 복수의 사용자들에게 제공될 수 있다. 질문 생성부(110)에 대한 상세한 내용은 도 2 및 도 3을 참조하여 후술될 것이다.
- [0029] 사용자 정보 저장부(120)는 복수의 사용자들의 레벨을 저장할 수 있다. 사용자의 레벨(또는 사용자 레벨)은 질문에 대한 사용자의 답변 태도(예컨대, 답변의 횟수, 주기 등) 및 답변의 신뢰도에 따른 사용자의 평가 점수에 대응할 수 있다. 후술되는 바와 같이, 사용자 레벨은 질문 생성부(110)에 의해서 생성된 질문들 각각을 어떤

사용자에 배정하는지를 결정하는데 기초로 사용될 수 있다. 또한, 사용자 정보 저장부(120)는 사용자의 식별 정보를 저장할 수 있고, 사용자의 식별 정보는 웹 서비스 시스템(300)과 같은 외부의 시스템(예컨대, 도 5의 리워드 시스템(400))과 사용자를 동기화하는데 사용될 수 있다.

[0030] 질문 선정부(130)는 사용자들의 레벨 정보, 즉 사용자 레벨 정보에 기초하여 사용자들 각각에 배정될 질문 및 질문의 개수를 결정할 수 있다. 예를 들면, 질문 선정부(130)는 높은 사용자 레벨을 가지고 있는 사용자에게 대하여 상대적으로 많은 수의 질문들을 배정할 수 있고, 난이도가 높은 질문(예컨대, 주관식 질문)을 배정할 수 있다. 이에 따라, 사용자의 수준을 고려하지 아니하고 사용자에게 질문을 제공함으로써 발생할 수 있는 문제점들, 예컨대 답변의 신뢰도 하락, 답변의 지체 등이 해소될 수 있다. 질문 선정부(130)는 사용자 인터페이스(103)를 통해서 선정된 적어도 하나의 질문을 지식 베이스 교정 시스템의 외부로 출력할 수 있고, 이 때 질문이 배정된 사용자 정보를 함께 출력할 수 있다.

[0031] 후보 지식 생성부(140)는 질문 선정부(130)에 의해서 선정된 적어도 하나의 질문 및 그 질문에 대한 답변을 포함하는 후보 지식 데이터를 생성할 수 있다. 일 실시예에서, 후보 지식 생성부(140)는 질문 선정부(130)로부터 선정된 질문을 수신할 수 있고, 사용자 인터페이스(103)를 통해서 선정된 질문에 대한 답변을 수신할 수 있다. 다른 실시예에서, 후보 지식 생성부(140)는 사용자 인터페이스(103)를 통해서 질문 및 그에 대한 답변을 모두 수신할 수도 있다. 후보 지식 데이터는 지식 베이스(200)에 저장된 지식 데이터가 될 가능성이 있는 것으로서, 후술되는 바와 같이, 후보 지식 검증부(160)에 의해서 검증이 성공한 경우 지식 베이스(200)의 지식 데이터를 교정하는데 사용될 수 있다.

[0032] 후보 지식 저장부(150)는 후보 지식 생성부(140)에 의해서 생성된 후보 지식 데이터를 저장할 수 있다. 전술된 바와 같이, 후보 지식 생성부(140)는 질문 및 그에 대응하는 답변을 포함하는 후보 지식 데이터를 생성할 수 있고, 후보 지식 저장부(150)는 복수의 질문들 및 답변들에 대하여 후보 지식 생성부(140)에 의해서 생성된 후보 지식 데이터를 저장할 수 있다. 도 1에 도시된 바와 같이, 후보 지식 저장부(150)는 후보 지식 생성부(140)로부터 후보 지식 데이터를 수신할 수 있고, 후보 지식 검증부(160) 및 사용자 레벨 분석부(170)에 후보 지식 데이터를 제공할 수 있다.

[0033] 후보 지식 검증부(160)는 동일한 질문에 대응하는 적어도 하나의 후보 지식 데이터에 기초하여 후보 지식 데이터를 검증할 수 있다. 후보 지식 검증부(160)는 검증이 성공한 후보 지식 데이터에 기초하여 지식 베이스 인터페이스(102)를 통해서 지식 베이스에 저장된 지식 데이터를 교정할 수 있다. 도 4를 참조하여 후술되는 바와 같이, 후보 지식 검증부(160)는 동일한 질문에 대하여 복수의 사용자들로부터 제공된 답변들을 평가함으로써 후보 지식 데이터를 검증할 수 있다. 후보 지식 검증부(160)에 대한 상세한 내용은 도 4를 참조하여 후술될 것이다.

[0034] 사용자 레벨 분석부(170)는 후보 지식 데이터 및 후보 지식 검증부(160)의 후보 지식 데이터에 대한 검증 결과에 기초하여, 사용자 정보 저장부(120)에 저장된 사용자의 레벨 정보를 갱신할 수 있다. 예를 들면, 사용자 레벨 분석부(170)는 후보 지식 검증부(160)의 후보 지식 데이터에 대한 검증 결과에 기초하여, 검증이 성공한 답변을 제공한 사용자의 레벨이 상향 조정되도록 사용자 정보 저장부(120)에 저장된 사용자의 레벨 정보를 갱신할 수 있다. 또한, 사용자 레벨 분석부(170)는 일정한 주기(예컨대, 1주, 1개월 등)로 사용자의 레벨 정보를 갱신할 수 있다. 즉, 일정한 주기 내에 사용자가 제공한 답변의 개수에 기초하여, 답변의 개수가 미리 정해진 기준치 미만인 경우 사용자 레벨을 하향 조정할 수 있고, 기타의 경우 사용자 레벨을 유지하거나 상향 조정할 수 있다.

[0035] 도 1의 예시에서, 사용자 정보 저장부(120) 및 후보 지식 저장부(150)는 지식 베이스 교정 시스템(100)의 구성 요소들에 의해서 액세스되는 것으로 도시되었다. 일 실시예에서, 사용자 정보 저장부(120) 및 후보 지식 저장부(150)는 외부의 액세스를 처리하는 어댑터들을 각각 포함할 수 있고, 지식 베이스 교정 시스템(100)의 구성 요소들은 어댑터들을 통해서 사용자 정보 저장부(120) 및 후보 지식 저장부(150)에 데이터를 저장하거나 저장된 데이터를 읽을 수 있다. 다른 실시예에서, 사용자 정보 저장부(120) 및 후보 지식 저장부(150)에 액세스하는 구성 요소들은 사용자 정보 저장부(120) 및 후보 지식 저장부(150)를 액세스하기 위한 어댑터들을 각각 포함할 수도 있다.

[0036] 비록 도 1은 지식 베이스 교정 시스템(100)이 지식 베이스(200)에 액세스하는 예시를 도시하고 있으나, 본 발명의 예시적 실시예에 따라 지식 베이스 교정 시스템(100)은 지식 베이스(200)를 포함할 수도 있고, 이 경우 지식 베이스 교정 시스템(100)이 지식 베이스의 기능, 즉 지식 데이터의 제공 서비스를 수행할 수도 있다. 또한, 도 1은 지식 베이스 교정 시스템(100)이 웹 서비스 시스템(300)을 통해서 사용자 단말들(500)과 통신하는 예시를

도시하고 있으나, 본 발명의 예시적 실시예에 따라 지식 베이스 교정 시스템(100)은 사용자 인터페이스(103)를 통해서 사용자 단말들(500)과 직접 통신할 수도 있다. 또한, 지식 베이스 교정 시스템(100)은 복수의 지식 베이스들 및 복수의 웹 서비스 시스템들과 통신할 수도 있다.

[0037] 도 2는 본 발명의 예시적 실시예에 따라 도 1의 질문 생성부(110)의 예시(110')를 나타내는 블록도이고, 도 3은 불완전한 지식 데이터의 예시를 나타내는 도면이다. 도 1을 참조하여 전술된 바와 같이, 질문 생성부(110')는 도 1의 지식 베이스(200)로부터 불완전한 지식 데이터를 검출할 수 있고, 답변으로써 불완전한 지식 데이터를 정정하기 위한 질문을 생성할 수 있다.

[0038] 도 2를 참조하면, 질문 생성부(110')는 지식 데이터(10)를 수신할 수 있다. 예를 들면, 질문 생성부(110')는 지식 베이스(200)로부터 불완전한 지식 데이터를 검출하기 위하여 지식 베이스(200)에 저장된 지식 데이터를 스캔할 수 있고, 지식 데이터(10)를 수신할 수 있다. 도 1의 지식 베이스(200)에 저장된 지식 데이터는 온톨로지 데이터, 예컨대 RDF(Resource Description Framework) 등의 형식을 가지는 데이터일 수 있고, RDF 스키마에 의해서 생성된 인스턴스를 포함할 수 있다. 도 3을 참조하면, 분류가 사람인 인스턴스로서 '반기문'은 다수의 속성들을 가질 수 있다. 예를 들면, 도 3에 도시된 바와 같이, '반기문'이 속하는 사람은 '출생', '국적', '배우자', '성별', '소속' 등의 속성들을 가질 수 있고, 인스턴스인 '반기문'은 각 속성들에 대하여 고유한 값들(예컨대, 1994, 대한민국 등)을 가질 수 있다.

[0039] 도 2를 참조하면, 질문 생성부(110')는 속성 관리부(112), 오류 검출부(114) 및 질문 출력부(116)를 포함할 수 있다. 속성 관리부(112)는 지식 데이터를 구성하는 분류(예컨대, 사람, 조직, 지리, 이벤트 등)에 따라 속성들에 관한 정보를 저장할 수 있고, 오류 검출부(114)에 속성에 관한 정보를 제공할 수 있다. 예컨대 도 3을 참조하면, 속성 관리부(112)는 '출생', '국적', '배우자', '성별', '소속'을 사람에 대한 속성들로서 저장할 수 있고, 오류 검출부(114)에 제공할 수 있다. 또한, 속성 관리부(112)가 제공하는 속성들에 관한 정보는 속성들의 값이 가지는 형식, 예컨대 숫자, URL, 텍스트 등에 대한 정보를 포함할 수 있다.

[0040] 오류 검출부(114)는 불완전한 지식 데이터, 예컨대 오류를 포함하는 지식 데이터 또는 불충분한 지식 데이터를 검출할 수 있다. 즉, 속성 관리부(112)가 제공하는 속성 정보에 기초하여, 지식 데이터(10)가 불완전한 지식 데이터인지 여부를 판단할 수 있다. 예를 들면, 오류 검출부(114)는 복수의 값들이 기입된 속성을 가지는 인스턴스를 포함하는 지식 데이터를 검출할 수 있다. 도 3에 도시된 바와 같이, 도 1의 지식 베이스(200)에 저장된 지식으로서 인스턴스인 '반기문'은 속성 '소속'의 값으로써 2개 이상의 값들('UN', '외교통상부')이 기입될 수 있고, 오류 검출부(114)는 '반기문 - 소속 - UN, 외교통상부'를 포함하는 지식 데이터를 검출할 수 있다. 다른 예시로서, 오류 검출부(114)는 지식 데이터에 포함된 속성(예컨대, '출생')이 속성의 값이 가지는 형식(예컨대, 4자리 숫자)과 상이한 형식의 값(예컨대, 영문자)을 가지는 경우, 오류를 포함하는 것으로 결정할 수도 있다.

[0041] 또한, 오류 검출부(114)는 불충분한 지식 데이터, 즉 값이 누락된 속성을 가지는 인스턴스를 포함하는 지식 데이터를 검출할 수 있다. 도 3에 도시된 바와 같이, 도 1의 지식 베이스(200)에 저장된 지식으로서 인스턴스인 '반기문'은 속성 '배우자'의 값이 누락될 수 있고, 오류 검출부(114)는 '반기문 - 배우자 - empty'를 포함하는 불충분한 지식 데이터를 검출할 수 있다.

[0042] 질문 출력부(116)는 오류 검출부(114)에 의해서 검출된 지식 데이터에 기초하여 질문(20)을 생성할 수 있다. 예를 들면, 질문 출력부(116)는 복수의 값들이 기입된 속성을 가지는 인스턴스를 포함하는 지식 데이터의 경우, 인스턴스, 복수의 값들이 기입된 속성, 복수의 값들을 포함하는 객관식 질문을 생성할 수 있다. 또한, 질문 출력부(116)는 값이 누락된 속성을 가지는 인스턴스를 포함하는 지식 데이터의 경우, 인스턴스, 값이 누락된 속성을 포함하는 주관식 질문을 생성할 수도 있다. 도 3을 참조하면, 질문 출력부(116)는 '반기문 - 소속 - UN, 외교통상부'를 포함하는 지식 데이터에 대하여, 도 4에 도시된 바와 같이, "반기문의 소속은? ① UN ② 외교통상부 ③ 기타"를 포함하는 질문을 생성할 수 있다. 또한, 질문 출력부(116)는 '반기문 - 배우자 - empty'를 포함하는 지식 데이터에 대하여, "반기문의 배우자는?"을 포함하는 질문을 생성할 수 있다. 도 1을 참조하여 전술된 바와 같이, 질문 출력부(116)에 의해서 생성된 질문은, 도 1의 질문 선정부(120)에 의해서 사용자에게 배정될 수 있다.

[0043] 도 4는 본 발명의 예시적 실시예에 따라 도 1의 후보 지식 검증부(160)의 동작을 설명하기 위한 도면이다. 구체적으로, 도 4는 후보 지식 검증부(160)가 후보 지식 저장부(150)로부터 수신하는 후보 지식 데이터의 예시를 나타낸다. 도 1을 참조하여 전술된 바와 같이, 후보 지식 검증부(160)는 질문 및 답변을 포함하는 후보 지식 데이터를 검출할 수 있고, 검증에 성공한 후보 지식 데이터에 기초하여 지식 베이스(200)에 저장된 지식 데이터

를 교정할 수 있다.

- [0044] 도 3 및 도 4를 참조하면, ‘반기문 - 소속 - UN, 외교통상부’ 를 포함하는 지식 데이터를 교정하기 위하여, “반기문의 소속은?” 이라는 질문이 도 1의 질문 생성부(110)에 의해서 생성되고, 질문 선정부(120)에 의해서 5명의 사용자들(갑, 을, 병, 정, 무)에게 배정되고, 5명의 사용자들(갑, 을, 병, 정, 무)로부터 답변을 수신함으로써 후보 지식 생성부(140)에 의해서 도 4에 도시된 바와 같은 후보 지식 데이터가 생성될 수 있다. 도 4에 도시된 예시에서, 후보 지식 데이터는 사용자 및 사용자 레벨 정보를 포함할 수 있다.
- [0045] 후보 지식 검증부(160)는 동일한 질문에 대하여 동일한 답변을 제공한 사용자의 수에 기초하여 후보 지식 데이터를 검증할 수 있다. 즉, 후보 지식 검증부(160)는 동일한 답변을 제공한 사용자들의 수가 미리 정해진 비율 이상인 경우, 그러한 답변을 포함하는 후보 지식 데이터에 대하여 검증이 성공한 것으로 결정할 수 있다. 예를 들면, 도 4에 도시된 예시에서, 후보 지식 검증부(160)는 ‘①’ 을 답변한 사용자들(갑, 을, 병, 정)의 수 및 답변한 전체 사용자의 수의 비율(4/5)이 미리 정해진 비율(2/3)을 초과하였으므로, ‘반기문 - 속성 - UN, 외교통상부’ 를 포함하는 지식 데이터를 ‘반기문 - 속성 - UN’ 을 포함하도록 정정할 수 있다.
- [0046] 또한, 후보 지식 검증부(160)는 동일한 답변을 제공한 사용자들의 수뿐만 아니라 사용자들의 레벨 정보에 더 기초하여 후보 지식 데이터를 검증할 수도 있다. 즉, 높은 레벨을 가지는 사용자의 답변에 대하여 가중치를 부가함으로써 후보 지식 데이터 검증의 신뢰성을 향상시킬 수 있다. 예를 들면, 도 4의 예시에서, 후보 지식 검증부(160)는 ‘①’ 을 답변한 사용자들(갑, 을, 병, 정)의 레벨 합 및 ‘②’ 를 답변한 사용자(무)의 레벨 합을 계산할 수 있고, 가장 높은 레벨 합 of 비율이 미리 정해진 비율을 초과하는 경우, 그러한 레벨 합에 대응하는 답변을 포함하는 후보 지식 데이터에 대하여 검증이 성공한 것으로 결정할 수 있다.
- [0047] 후보 지식 검증부(160)는 후보 지식 데이터의 검증이 실패한 경우, 예컨대 동일한 답변을 제공한 사용자들의 수가 미리 정해진 비율 미만인 경우 또는 가장 높은 레벨 합 of 비율이 미리 정해진 비율 미만인 경우, 후보 지식 데이터의 검증이 실패한 것으로 결정할 수 있다. 후보 지식 검증부(160)에 의해서 검증이 실패한 후보 지식 데이터에 대한 정보는 질문 선정부(130)에 제공될 수 있다. 질문 선정부(130)는 검증이 실패한 후보 지식 데이터에 대응하는 질문을 기배정된 사용자들과 상이한 복수의 사용자들에게 배정하도록 결정할 수 있다. 이에 따라, 검증에 실패한 후보 지식 데이터는 기배정된 사용자들과 상이한 사용자들로부터 제공된 답변들에 기초하여 후보 지식 검증부(160)에 의해서 추후 검증될 수 있다.
- [0048] 도 5는 본 발명의 예시적 실시예에 따른 지식 베이스 교정 시스템(100a)의 블록도이다. 도 1의 지식 베이스 교정 시스템(100)과 비교할 때, 도 5의 지식 베이스 교정 시스템(100a)은 리워드 시스템(400)과 통신하기 위한 리워드 인터페이스(104a)를 더 포함할 수 있다. 이하 도 5의 설명에서, 도 1을 참조하여 설명한 내용과 중복되는 내용은 생략될 것이다.
- [0049] 리워드 인터페이스(104a)는 지식 베이스 검증 시스템(100a) 외부의 리워드 시스템(400)이 사용자 정보 저장부(120a)에 액세스하기 위한 인터페이스를 제공할 수 있다. 예를 들면, 후술되는 바와 같이, 사용자 정보 저장부(120a)에 저장된 사용자 레벨 정보는 리워드 시스템(400)에 의해서 활용될 수 있고, 리워드 인터페이스(104a)는 리워드 시스템(400)을 포함하는 복수의 리워드 시스템들에 대하여 사용자 레벨 정보를 제공하는 인터페이스를 제공할 수 있다.
- [0050] 리워드 시스템(400)은 지식 베이스 검증 시스템(100a)(또는 사용자 정보 저장부(120))의 사용자 레벨 정보에 기초하여 사용자에게 리워드를 제공할 수 있다. 예를 들면, 리워드 시스템(400)은 답변을 제공한 사용자에 대하여 사용자의 레벨에 대응하는 리워드를 제공할 수도 있고, 주기적으로 사용자 레벨을 체크함으로써 사용자에게 주기적으로 사용자 레벨에 대응하는 리워드를 제공할 수 있다. 리워드 시스템(400)이 제공하는 리워드는 사용자가 유익하게 사용할 수 있는 것으로서, 비제한 적인 예시로서, 쿠폰, 금전, 가맹점들에 사용할 수 있는 포인트 등을 포함할 수 있다.
- [0051] 지식 베이스 검증 시스템(100a)으로부터 제공된 질문에 답변을 제공한 사용자는 리워드를 획득할 수 있고, 이에 따라 사용자의 답변에 대한 의욕은 고취될 수 있다. 도 1을 참조하여 기술된 바와 같이, 사용자 정보 저장부(120a)에 저장된 사용자 레벨 정보는 사용자 레벨 분석부(170a)에 의해서 갱신될 수 있고, 사용자 레벨 정보는 사용자의 답변 태도 및 답변의 신뢰도에 기초하여 산정되므로, 높은 효율성 및 신뢰성을 제공하는 지식 베이스(200)의 검증이 실현될 수 있다.
- [0052] 도 6은 본 발명의 예시적 실시예에 따른 지식 베이스의 교정 방법을 나타내는 순서도이다. 구체적으로, 도 6은 도 1의 지식 베이스(200)를 교정하기 위해 사용자에게 제공되는 질문을 생성하는 방법을 나타내는 순서도이다.

이하에서 도 6은 도 1 및 도 2를 참조하여 설명될 것이다. 도 6에 도시된 바와 같이, 지식 베이스(200)의 교정 방법은 복수의 단계들(S12, S14, S16, S18)을 포함할 수 있다.

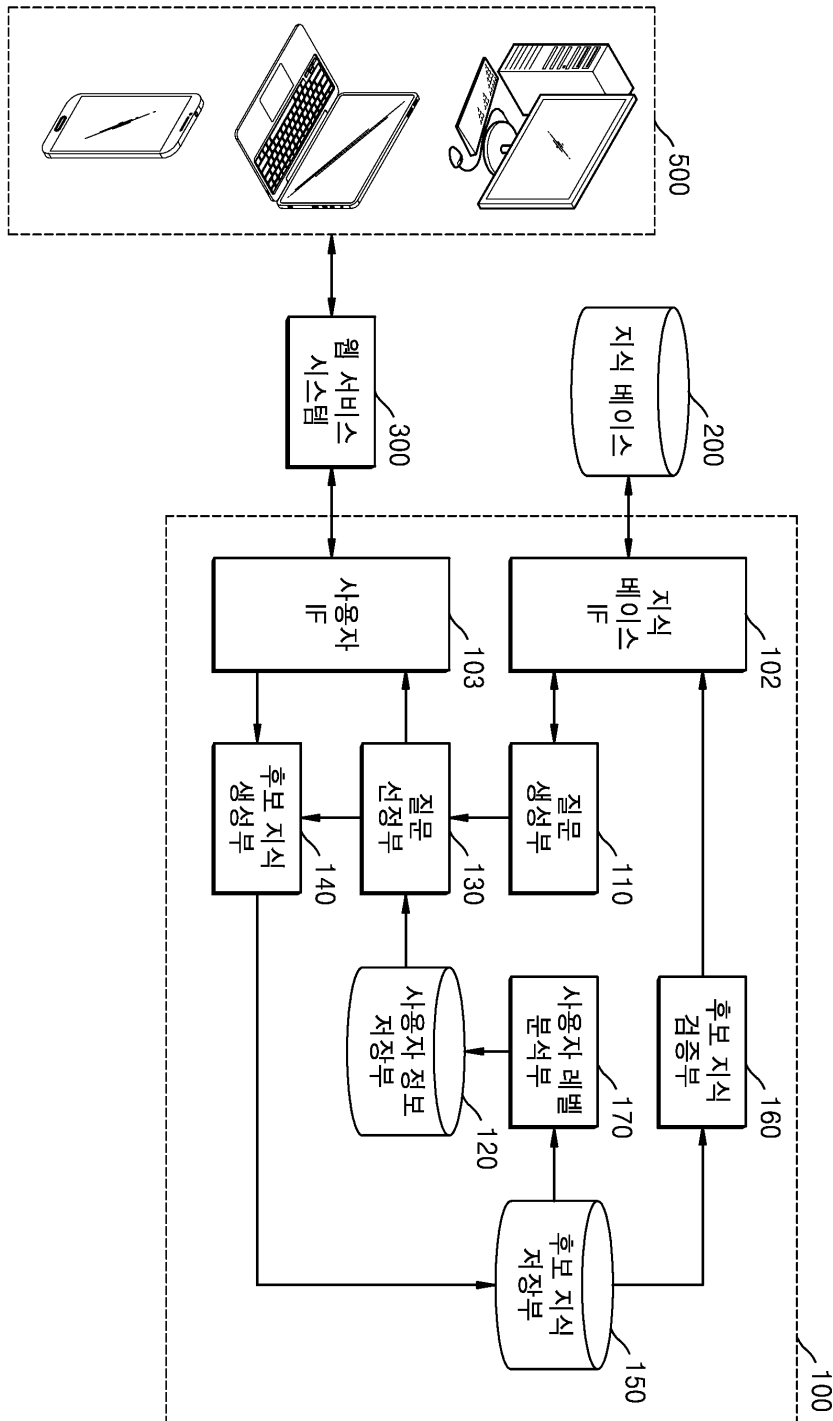
- [0053] 단계 S12에서, 지식 베이스(200)로부터 불완전한 지식 데이터를 검출하는 동작이 수행될 수 있다. 예를 들면, 도 2의 질문 생성부(110')에 포함된 오류 검출부(114)는 속성 관리부(112)가 제공하는 속성 정보에 기초하여 지식 데이터(10)가 불완전한 지식 데이터, 예컨대 오류를 포함하는 지식 데이터 또는 불충분한 지식 데이터인지 여부를 판단할 수 있고, 불완전한 지식 데이터를 검출할 수 있다.
- [0054] 단계 S14에서, 불완전한 지식 데이터의 교정을 위한 질문을 생성하는 동작이 수행될 수 있다. 예를 들면, 도 2의 질문 생성부(110')에 포함된 질문 출력부(116)는 오류 검출부(114)에 의해서 검출된 불완전한 지식 데이터로부터 질문을 생성할 수 있다. 전술된 바와 같이, 생성된 질문은 복수의 선택가능한 답변들을 포함하는 객관식 질문일 수도 있고, 주관식 질문일 수도 있다.
- [0055] 단계 S16에서, 사용자들의 레벨에 기초하여 질문의 개수 및 질문을 선정하는 동작이 수행될 수 있다. 예를 들면, 도 1의 질문 선정부(120)는 질문 생성부(110)에 의해서 생성된 복수의 질문들에 대하여, 사용자 정보 저장부(130)에 저장된 사용자 레벨 정보에 기초하여, 사용자에게 배정될 질문의 개수 및 사용자에게 배정될 질문을 결정할 수 있다. 유사하게는, 질문 선정부(120)는 사용자 정보 저장부(130)에 저장된 사용자 레벨 정보에 기초하여, 질문 생성부(110)에 의해서 생성된 질문을 배정할 복수의 사용자들을 결정할 수도 있다.
- [0056] 단계 S18에서, 선정된 질문을 전송하는 동작이 수행될 수 있다. 예를 들면, 도 1의 질문 선정부(120)에 의해서 선정된 질문은 사용자 정보와 함께 사용자 인터페이스(103)를 통해서 웹 서비스 시스템(300)에 전송될 수 있고, 웹 서비스 시스템(300)은 사용자 정보에 따라 선택된 사용자 단말들(500) 중 하나 이상에 수신된 질문을 전송할 수 있다.
- [0057] 도 7은 본 발명의 예시적 실시예에 따른 지식 베이스 교정 방법을 나타내는 순서도이다. 구체적으로, 도 7은 사용자들로부터 수신된 답변들을 평가함으로써 지식 베이스(200)를 교정하는 방법을 나타내는 순서도로서, 도 6의 순서도에 대응하는 방법에 이어서 수행될 수 있다. 이하에서, 도 7은 도 1을 참조하여 설명될 것이다. 도 7에 도시된 바와 같이, 지식 베이스(200)의 교정 방법은 복수의 단계들(S21, S23, S25, S27, S29)을 포함할 수 있다.
- [0058] 단계 S21에서, 질문에 대한 답변을 수신하는 동작이 수행될 수 있다. 도 1 및 도 6을 참조하여 전술된 바와 같이, 불완전한 지식 데이터를 교정하기 위하여 생성된 질문이 (예컨대, 웹 서비스 시스템(300)을 통해서) 사용자에게 제공될 수 있고, 사용자로부터 질문에 대한 답변이 수신될 수 있다. 예를 들면, 도 1의 웹 서비스 시스템(300)은 사용자 단말들(500)로부터 답변들을 수신할 수 있고, 지식 베이스 교정 시스템(100)에 수신된 답변들을 전송할 수 있다. 지식 베이스 교정 시스템(100)의 후보 지식 생성부(140)는 사용자 인터페이스(103)를 통해서 답변들을 수신할 수 있다.
- [0059] 단계 S23에서, 후보 지식 데이터를 생성하는 동작이 수행될 수 있다. 예를 들면, 도 1의 후보 지식 생성부(140)는 질문, 그에 대한 사용자의 답변 및 사용자 정보를 포함하는 후보 지식 데이터를 생성할 수 있다. 또한, 일실시예에서, 후보 지식 생성부(140)는 추가적으로 답변한 사용자의 레벨 정보를 더 포함하는 후보 지식 데이터를 생성할 수도 있다. 후보 지식 생성부(140)에 의해서 생성된 후보 지식 데이터는 후보 지식 저장부(150)에 저장될 수 있다.
- [0060] 단계 S25에서, 후보 지식 데이터를 검증하는 동작이 수행될 수 있다. 예를 들면, 도 1의 후보 지식 검증부(160)는 후보 지식 데이터를 저장하는 후보 지식 저장부(150)를 액세스함으로써 동일한 질문에 대한 복수의 답변들을 포함하는 후보 지식 데이터를 수신할 수 있고, 후보 지식 데이터를 검증할 수 있다. 도 4를 참조하여 전술된 바와 같이, 후보 지식 검증부(160)는 동일한 답변을 제공한 사용자들의 수에 기초하여 후보 지식 데이터를 검증할 수도 있고, 동일한 답변을 제공한 사용자들의 레벨들에 기초하여 후보 지식 데이터를 검증할 수도 있다.
- [0061] 후보 지식 데이터의 검증이 성공한 경우, 단계 S27에서 지식 데이터를 정정하는 동작이 수행될 수 있다. 예를 들면, 도 1의 후보 지식 검증부(160)는 2개 이상의 값들이 하나의 속성에 기입되어 오류를 포함하는 지식 데이터에 대하여, 검증된 답변에 대응하는 값만을 포함하도록 지식 데이터를 정정할 수 있다. 또한, 누락된 값을 가지는 속성에 의해서 불충분한 지식 데이터에 대하여, 해당 속성이 검증된 답변에 대응하는 값을 가지도록 지식 데이터를 정정할 수도 있다.
- [0062] 후보 지식 데이터의 검증이 실패한 경우, 단계 S29에서 기배정된 사용자와 상이한 사용자에게 질문을 배정하는

동작이 수행될 수 있다. 예를 들면, 도 1의 질문 선정부(120)는 후보 지식 검증부(160)로부터 검증이 실패한 후보 지식 데이터(또는 그에 대응하는 질문)를 수신할 수 있고, 질문 선정부(120)는 이전에 질문이 배정되지 아니한 복수의 사용자들에 질문을 배정할 수 있다. 배정된 질문은 새로운 사용자에게 제공될 수 있고, 그들로부터 수신된 답변들이 후보 지식 데이터의 재검증에 사용될 수 있다.

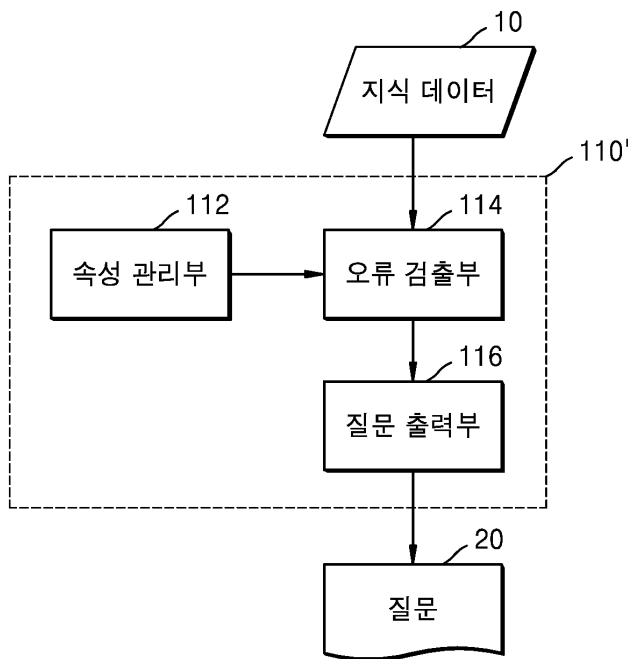
[0063] 이상에서와 같이 도면과 명세서에서 예시적인 실시예들이 개시되었다. 본 명세서에서 특정한 용어를 사용하여 실시예들을 설명되었으나, 이는 단지 본 발명의 기술적 사상을 설명하기 위한 목적에서 사용된 것이지 의미 한정이나 특허청구범위에 기재된 본 발명의 범위를 제한하기 위하여 사용된 것은 아니다. 그러므로 본 기술분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 따라서, 본 발명의 진정한 기술적 보호범위는 첨부된 특허청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

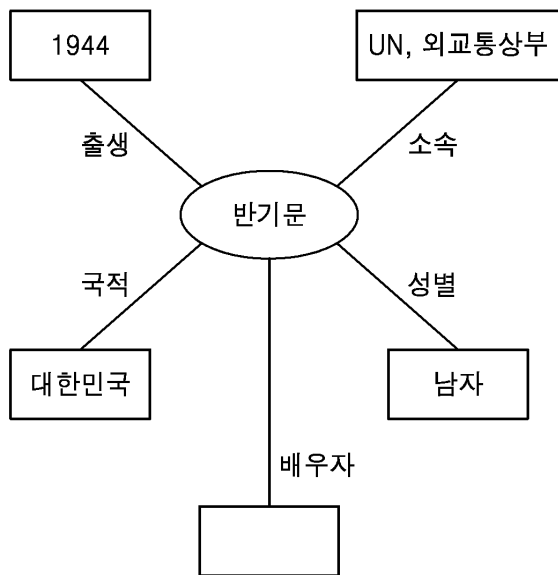
도면1



도면2



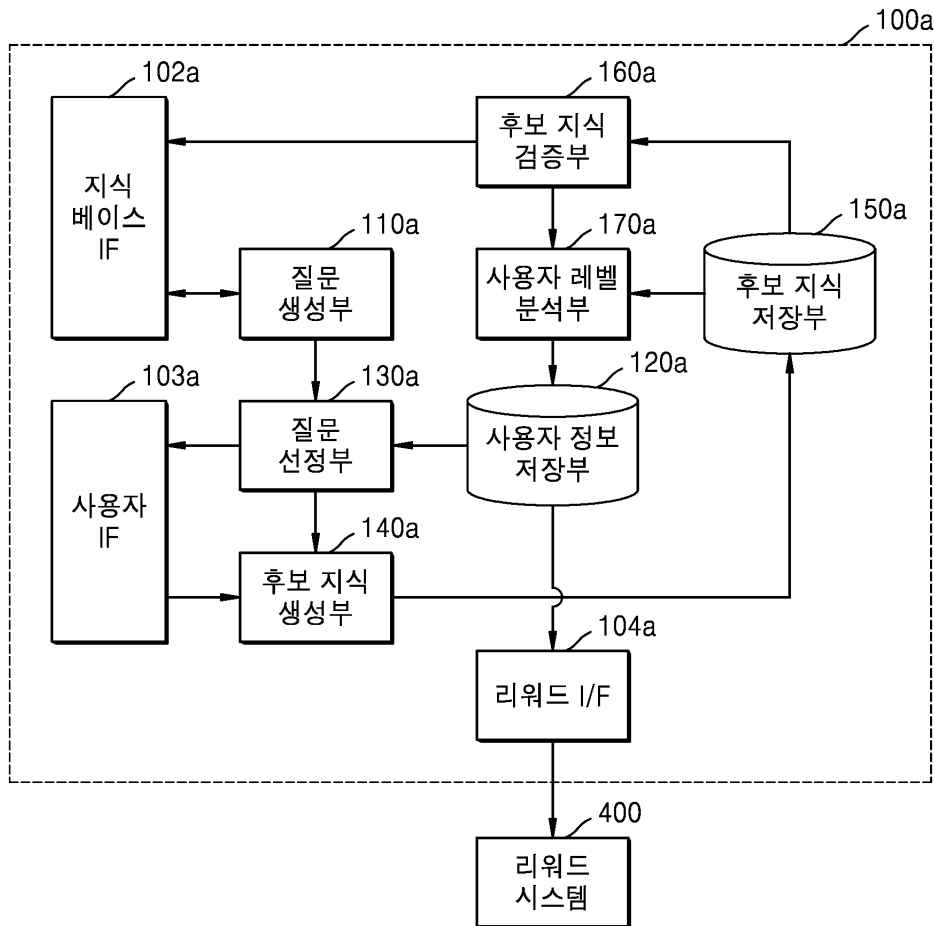
도면3



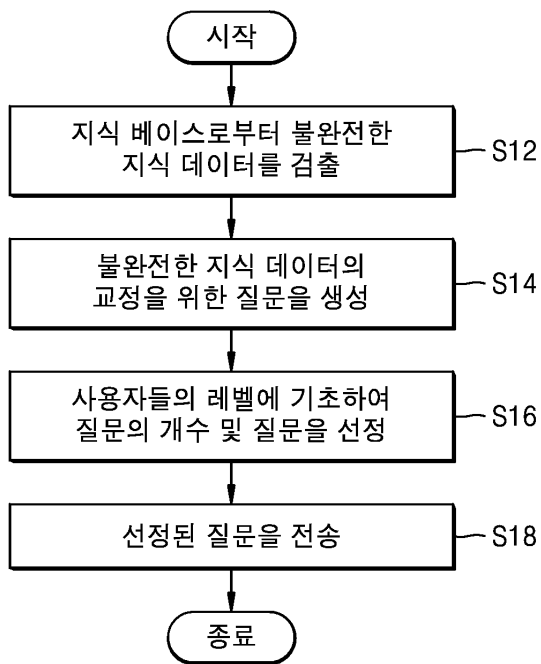
도면4

질문	사용자	레벨	답변
'반기문'의 소속은? ① UN ② 외교통상부 ③ 기타	갑	10	①
	을	8	①
	병	9	①
	정	5	①
	무	6	②

도면5



도면6



도면7

