(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0136982
(43) 공개일자 2019년12월10일

(51) 국제특허분류(Int. Cl.)
G16H 10/60 (2018.01)
(52) CPC특허분류
G16H 10/60 (2018.01)
(21) 출원번호 10-2019-0063326
(22) 출원일자 2019년05월29일
심사청구일자 2019년05월29일
(30) 우선권주장
1020180061518 2018년05월30일 대한민국(KR)

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이승룡
경기도 성남시 분당구 구미로144번길 25,102동 102호 (구미동, 삼환빌라)
수자트 후사인 쿠잘
경기도 용인시 기흥구 서천동 1 경희대학교국제캠퍼스
(74) 대리인
특허법인도담

전체 청구항 수 : 총 7 항

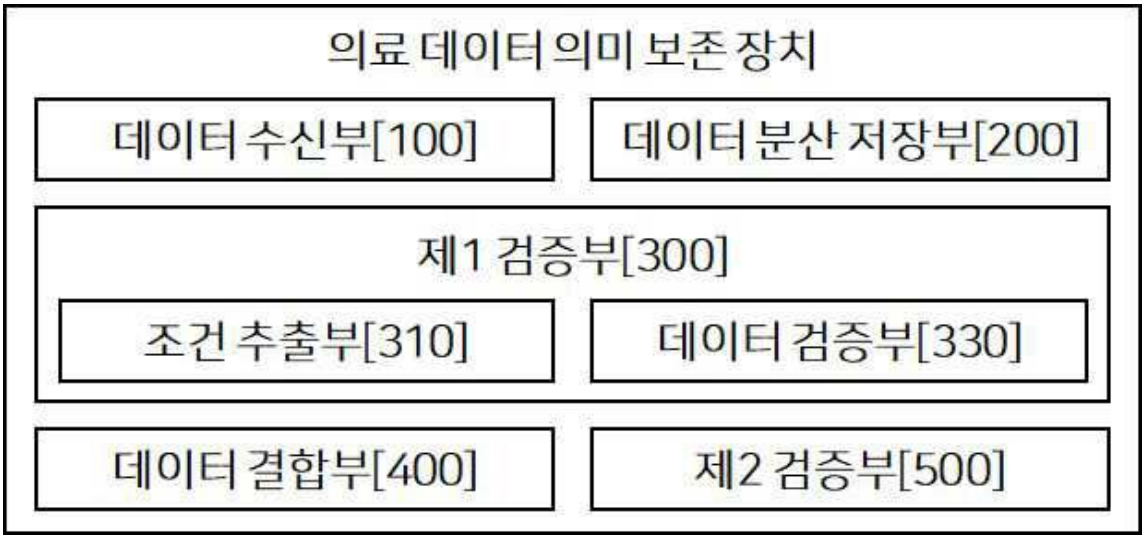
(54) 발명의 명칭 **표준화된 의료 데이터의 의미 보존 방법 및 그 장치**

(57) 요 약

본 발명은 의료 데이터를 분산함에 따라 발생하는 의미적 손실을 최소화하기 위해 의미적 손실이 발생한 의료 데이터를 수정하는 것을 목적으로 한다. 이러한 목적을 달성하기 위한 본 발명은 서버가 표준화된 의료 데이터의 의미를 보존하는 방법에 있어서, 표준화된 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를

(뒷면에 계속)

대 표 도 - 도1



생성하고, 상기 제2 의료 데이터를 분산 저장하는 단계, 기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 상기 검증 조건을 만족하는 지 판단하는 단계, 상기 제2 의료 데이터가 검증 조건을 만족하지 못하면 상기 제2 의료 데이터가 손상되었다고 판단하고, 상기 제2 의료 데이터를 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 구분하는 단계, 스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성하는 단계 및 상기 제3 의료 데이터를 검증하는 단계를 포함하는 것을 특징으로 한다.

이 발명을 지원한 국가연구개발사업
과제고유번호 1711070401
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 연구개발고급인력지원
연구과제명 지능형 의료 플랫폼 개발
기 여 율 40/100
주관기관 경희대학교산학협력단
연구기간 2019.01.01 ~ 2019.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1711065120
부처명 과학기술정보통신부
연구관리전문기관 정보통신기술진흥센터
연구사업명 디지털콘텐츠원천기술개발
연구과제명 다양한 디지털 제품의 UX 평가를 위한 Lean UX 핵심기술 및 플랫폼
기 여 율 30/100
주관기관 경희대학교산학협력단
연구기간 2019.01.01 ~ 2019.12.31
이 발명을 지원한 국가연구개발사업
과제고유번호 1415146278
부처명 산업통상자원부
연구관리전문기관 한국산업기술평가관리원
연구사업명 지식서비스산업핵심기술개발
연구과제명 퍼스널 빅데이터를 활용한 마이닝마인즈 핵심기술 개발
기 여 율 30/100
주관기관 경희대학교산학협력단
연구기간 2017.06.01 ~ 2018.05.31

명 세 서

청구범위

청구항 1

서버가 표준화된 의료 데이터의 의미를 보존하는 방법에 있어서,

표준화된 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를 생성하고, 상기 제2 의료 데이터를 분산 저장하는 단계;

기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 상기 검증 조건을 만족하는 지 판단하는 단계;

상기 제2 의료 데이터가 검증 조건을 만족하지 못하면 상기 제2 의료 데이터가 손상되었다고 판단하고, 상기 제2 의료 데이터를 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 구분하는 단계;

스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성하는 단계; 및

상기 제3 의료 데이터를 검증하는 단계를 포함하는 표준화된 의료 데이터의 의미 보존 방법.

청구항 2

제1항에 있어서,

상기 제1 의료 데이터는 국제 의료 표준화 기구가 제정한 CDA를 통해 표준화된 것을 특징으로 하는 표준화된 의료 데이터의 의미 보존 방법.

청구항 3

제1항에 있어서,

상기 검증 모델은 상기 제2 의료 데이터의 유효성을 검증하기 위한 것으로, 표준 스키마 구조를 통해 모델링 된 것을 특징으로 하는 표준화된 의료 데이터의 의미 보존 방법.

청구항 4

제1항에 있어서,

상기 제1 손상 데이터는 상기 제2 의료 데이터의 상부가 손상되고,

상기 제2 손상 데이터는 상기 제2 의료 데이터의 하부가 손상된 것을 특징으로 하는 표준화된 의료 데이터의 의미 보존 방법.

청구항 5

제1항에 있어서,

상기 제3 의료 데이터를 검증하는 단계는,

상기 제3 의료 데이터의 상부 및 하부에서 컨텍스트를 각각 추출하는 단계;

상기 상단 및 하단 컨텍스트의 유사 여부를 판단하는 단계;

상기 제3 의료 데이터의 스키마 구조가 표준 스키마 구조를 따르는 지를 판단하는 단계;

상기 판단 결과에 따라 상기 제3 의료 데이터의 유효성을 판단하는 단계를 더 포함하는 표준화된 의료 데이터의 의미 보존 방법.

청구항 6

표준화된 의료 데이터의 의미를 보존하는 장치에 있어서,

데이터베이스로부터 표준화된 제1 의료 데이터를 로드하는 데이터 수신부;

상기 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를 생성하고, 상기 제2 의료 데이터를 분산 저장하는 데이터 분산 저장부;

기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 상기 검증 조건을 만족하는 지 판단하고, 상기 제2 의료 데이터가 검증 조건을 만족하지 못하면 상기 제2 의료 데이터가 손상되었다고 판단하고, 상기 제2 의료 데이터의 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 분리하는 제1 검증부;

스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성하는 데이터 결합부; 및

상기 제3 의료 데이터를 검증하는 제2 검증부를 포함하는 표준화된 의료 데이터의 의미 보존 장치.

청구항 7

제6항에 있어서,

상기 제2 검증부는,

상기 제3 의료 데이터의 상단 및 하단에서 컨텍스트를 각각 추출하고, 상기 상단 및 하단 컨텍스트의 유사 여부를 판단하여 상기 제3 의료 데이터의 유효성을 판단하는 것을 특징으로 하는 의료 데이터의 의미 보존 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 표준화된 의료 데이터의 의미 보존 방법 및 그 장치에 관한 것으로, 보다 자세하게는 대규모의 의료 데이터를 분산 저장함에 따라 발생 가능한 의미적 손실을 보존하는 방법 및 그 장치에 관한 것이다.

배 경 기 술

[0002] 최근 의료 분야에 있어서, 의료업 종사자, 건강 관리 의사 결정권자, 그리고 보험사와 같은 이해 관계자는 다양한 의료 표준 데이터를 생성한다. 그러나 일반적으로 의료 데이터는 복잡한 스키마 구조를 갖기 때문에 데이터의 처리에 있어서 어려움이 있다. 특히 대규모의 의료 데이터는 데이터 처리 속도를 향상하기 위하여 분할하여 처리되는데, 이 때 일관성 없는 분할과 의료 데이터의 의미적 손실 등과 같은 문제점이 발생할 수 있다. 의료 데이터의 의미적 손실은 의료 데이터의 분할 시 중요한 부분이 절단되어 서로 다른 클러스터에 할당될 경우 치명적인 문제가 된다.

발명의 내용

해결하려는 과제

[0004] 본 발명은 전술한 문제점을 해결하기 위한 것으로서, 의료 데이터를 분산함에 따라 발생하는 의미적 손실을 최소화하는 것을 일 목적으로 한다.

[0005] 또한 본 발명은 의미적 손실이 발생한 의료 데이터를 수정하여 의료 데이터의 의미를 보존하는 것을 일 목적으로 한다.

과제의 해결 수단

[0006] 이러한 목적을 달성하기 위한 본 발명은 서버가 표준화된 의료 데이터의 의미를 보존하는 방법에 있어서, 표준화된 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를 생성하고, 상기 제2 의료 데이터를 분산 저장하는 단계, 기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 상기 검증 조건을 만족하는 지 판단하는 단계, 상기 제2 의료 데이터가 검증 조건을 만족하지 못하면 상기 제2 의료 데이터가 손상되었다고 판단하고, 상기 제2 의료 데이터를 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 구분하는 단계, 스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성하는 단계 및 상기 제3 의료 데이터를 검증하는 단계를 포함하는 것을 일 특징으로 한다.

[0007] 또한 상기 제1 의료 데이터는 국제 의료 표준화 기구(HL7)가 제정한 CDA를 통해 표준화된 것을 일 특징으로 한다.

[0008] 나아가 상기 검증 모델은 상기 제2 의료 데이터의 유효성을 검증하기 위한 것으로, 표준 스키마 구조를 통해 모델링 된 것을 일 특징으로 한다.

[0009] 또한 상기 제1 손상 데이터는 상기 제2 의료 데이터의 상부가 손상되고, 상기 제2 손상 데이터는 상기 제2 의료 데이터의 하부가 손상된 것을 일 특징으로 한다.

[0010] 나아가 상기 제3 의료 데이터를 검증하는 단계는, 상기 제3 의료 데이터의 상부 및 하부에서 컨텍스트를 각각 추출하는 단계, 상기 상단 및 하단 컨텍스트의 유사 여부를 판단하는 단계, 상기 제3 의료 데이터의 스키마 구조가 표준 스키마 구조를 따르는 지를 판단하는 단계, 상기 판단 결과에 따라 상기 제3 의료 데이터의 유효성을 판단하는 단계를 더 포함하는 것을 일 특징으로 한다.

[0011] 또한 본 발명은 표준화된 의료 데이터의 의미를 보존하는 장치에 있어서, 데이터베이스로부터 표준화된 제1 의료 데이터를 로드하는 데이터 수신부, 상기 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를 생성하고, 상기 제2 의료 데이터를 분산 저장하는 데이터 분산 저장부, 기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 상기 검증 조건을 만족하는 지 판단하고, 상기 제2 의료 데이터가 검증 조건을 만족하지 못하면 상기 제2 의료 데이터가 손상되었다고 판단하고, 상기 제2 의료 데이터의 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 분리하는 제1 검증부, 스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성하는 데이터 결합부 및 상기 제3 의료 데이터를 검증하는 제2 검증부를 포함하는 것을 일 특징으로 한다.

[0012] 나아가 상기 제2 검증부는, 상기 제3 의료 데이터의 상단 및 하단에서 컨텍스트를 각각 추출하고, 상기 상단 및 하단 컨텍스트의 유사 여부를 판단하여 상기 제3 의료 데이터의 유효성을 판단하는 것을 일 특징으로 한다.

발명의 효과

[0013] 전술한 바와 같은 본 발명에 의하면, 의료 데이터를 분산함에 따라 발생하는 의미적 손실을 최소화할 수 있다.

[0014] 또한 본 발명은 의미적 손실이 발생한 의료 데이터를 수정하여 의료 데이터의 의미를 보존할 수 있다.

도면의 간단한 설명

[0015] 도 1은 본 발명의 일 실시 예에 의한 표준화된 의료 데이터의 의미 보존 장치의 구성을 설명하기 위한 도면이다.

도 2는 본 발명의 일 실시 예에 의한 표준화된 의료 데이터의 의미 보존 방법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0016] 전술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다.

[0017] 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용되며, 명세서 및 특허청구의 범위에 기재된 모든 조합은 임의의 방식으로 조합될 수 있다. 그리고 다른 식으로 규정하지 않는 한, 단수에 대한 언급은 하나 이상을 포함할 수 있고, 단수 표현에 대한 언급은 또한 복수 표현을 포함할 수 있음이 이해되어야 한다.

[0018] 본 명세서에서 사용되는 용어는 단지 특정 예시적 실시 예들을 설명할 목적을 가지고 있으며 한정할 의도로 사용되는 것이 아니다. 본 명세서에서 사용된 바와 같은 단수적 표현들은 또한, 해당 문장에서 명확하게 달리 표시하지 않는 한, 복수의 의미를 포함하도록 의도될 수 있다. 용어 "및/또는," "그리고/또는"은 그 관련되어 나열되는 항목들의 모든 조합들 및 어느 하나를 포함한다. 용어 "포함한다", "포함하는", "포함하고 있는", "구비하는", "갖는", "가지고 있는" 등은 내포적 의미를 갖는 바, 이에 따라 이러한 용어들은 그 기재된 특징, 정수, 단계, 동작, 요소, 및/또는 컴포넌트를 특정하며, 하나 이상의 다른 특징, 정수, 단계, 동작, 요소, 컴포넌트, 및/또는 이들의 그룹의 존재 혹은 추가를 배제하지 않는다. 본 명세서에서 설명되는 방법의 단계들, 프로세스들, 동작들은, 구체적으로 그 수행 순서가 확정되는 경우가 아니라면, 이들의 수행을 논의된 혹은 예시된 그러한 특정 순서로 반드시 해야 하는 것으로 해석돼서는 안 된다. 추가적인 혹은 대안적인 단계들이 사용될 수 있음을 또한 이해해야 한다.

[0019] 또한, 각각의 구성요소는 각각 하드웨어 프로세서로 구현될 수 있고, 위 구성요소들이 통합되어 하나의 하드웨어 프로세서로 구현될 수 있으며, 또는 위 구성요소들이 서로 조합되어 복수 개의 하드웨어 프로세서로 구현될 수도 있다.

[0020] 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하기로 한다.

[0021] 도 1은 본 발명의 일 실시 예에 의한 표준화된 의료 데이터의 의미 보존 장치(이하, 장치)의 구성을 도시한 도면이다. 도 1을 참조하면 장치는 데이터 수신부(100), 데이터 분산 저장부(200), 제1 검증부(300), 데이터 결합부(400), 그리고 제2 검증부(500)를 포함할 수 있다.

[0022] 데이터 수신부(100)는 데이터베이스로부터 표준화된 제1 의료 데이터를 불러올 수 있다. 이 때 제1 의료 데이터는 HL7(Health Level-7)이 제정한 통합 프레임워크의 표준에 따라 표준화될 수 있다. 이하에서는 설명의 편의성을 위하여 제1 의료 데이터가 HL7이 제정한 CDA(Clinical Document Architecture)에 따라 표준화된 것으로 가정하고 설명한다. HL7은 이기종 의료정보시스템 간 정보 호환을 가능하게 하는 표준을 제정하는 표준화 기구로, 글로벌 의료 데이터의 상호운용성을 향상시키기 위하여 의료 데이터의 정보 교환, 통합, 공유, 추출을 위한 통합 프레임워크와 관련된 표준을 제정한다. HL7이 제정한 표준에는 V2 Message, V3 RIM(Reference Information Model), CDA(Clinical Document Architecture) 등이 있다.

[0023] 데이터 분산 저장부(200)는 데이터 수신부(100)로부터 불러온 제1 의료 데이터를 하나 이상의 블록으로 분할하여 제2 의료 데이터를 생성하고, 생성된 제2 의료 데이터를 분산 저장할 수 있다. 이 때 데이터 분산 저장부(200)는 제2 의료 데이터를 HDFS에 저장할 수 있다. HDFS(Hadoop Distributed File System)는 대용량 데이터를 분산 처리할 수 있는 자바 기반 오픈 소스 프레임워크인 하둡(Hadoop)의 데이터 저장소이다. 하둡은 맵리듀스(MapReduce)를 통해 하나 이상의 노드에 제2 의료 데이터를 분배할 수 있다.

[0024] 제1 검증부(300)는 기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 상기 제2 의료 데이터가 검증 조건을 만족하는 지 판단하고, 제2 의료 데이터가 검증 조건을 만족하지 못하면 제2 의료 데이터가 손상되었다고 판단하고, 제2 의료 데이터의 손상 유형에 따라 제1 손상 데이터 또는 제2 손상 데이터로 분리할 수 있다.

[0025] 제1 검증부(300)는 제2 의료 데이터의 스키마를 검증하기 위한 검증 모델을 이용할 수 있다. 제1 검증부(300)는 검증 모델을 이용하여 제2 의료 데이터의 스키마가 표준 스키마를 따르는 지를 판단하여 제2 의료 데이터를 검증할 수 있다. 검증 모델은 표준 스키마 구조를 통해 모델링 된 것을 특징으로 하며, 서버에 이미 저장되어 있어 별도의 저장소에서 불러올 필요가 없다. 검증 모델은 표준 스키마에 기반하여 설정되어, 스키마의 구조에 대한 조건을 하나 이상 포함할 수 있다. 스키마는 시작 태그, 속성, 끝 태그, 요소 등을 포함할 수 있다. 이 때 복수 개의 검증 조건을 포함하는 검증 모델은 중첩된 속성과 중첩되지 않은 속성을 더 포함할 수 있다. 중첩된 속성은 그 자체로 복수 개의 속성을 갖는 것을 의미하고, 중첩되지 않은 속성은 내부에 상이한 속성을 포함하는 것을 의미한다. 이는 스키마에 포함되는 속성이 복수 개일 수 있기 때문이다.

[0026] 제1 검증부(300)는 검증 모델에 포함된 하나 이상의 검증 조건을 추출하는 조건 추출부(310)와 제2 의료 데이터와 검증 조건을 비교하여 제2 의료 데이터의 손상 여부를 검증하는 데이터 검증부(330)를 포함할 수 있다.

[0027] 조건 추출부(310)는 검증 모델에 포함된 스키마의 구조에 대한 검증 조건을 하나 이상 추출할 수 있다. 스키마의 구조에 대한 검증 조건은 예를 들어 제2 의료 데이터에 포함된 태그 모두가 시작 태그와 끝 태그를 모두 포함하고 있는 지의 여부일 수 있다.

[0028] 데이터 검증부(330)는 제2 의료 데이터가 조건 추출부(310)에서 추출된 하나 이상의 검증 조건을 만족하는 지 판단함으로써 제2 의료 데이터를 검증할 수 있다. 예를 들어, 제2 의료 데이터에 시작 태그인 <stream>이 포함되어 있으나 끝 태그인 </stream>가 포함되어 있지 않으면, 데이터 검증부(330)는 제2 의료 데이터가 검증 조건을 만족하지 못한다고 판단할 수 있다.

[0029] 데이터 검증부(330)는 검증 조건을 만족하지 못한 제2 의료 데이터가 손상되었다고 판단하고, 그 손상 유형을 식별할 수 있다. 데이터 검증부(330)는 제2 의료 데이터의 시작 태그가 누락되어 있으면 제2 의료 데이터의 상부 영역이 손상되었다고 판단하고, 끝 태그가 누락되어 있으면 제2 의료 데이터의 하부 영역이 손상되었다고 판단할 수 있다.

[0030] 데이터 검증부(330)는 손상된 영역에 따라 제2 의료 데이터를 제1 손상 데이터 또는 제2 손상 데이터로 구분할 수 있다. 데이터 검증부(330)는 상부 영역이 손상된 제2 의료 데이터를 제1 손상 데이터로, 하부 영역이 손상된 제2 의료 데이터를 제2 손상 데이터로 구분할 수 있다.

[0031] 데이터 검증부(330)는 제1 손상 데이터와 제2 손상 데이터에 손상 유형과 손상 영역의 수에 따른 태그를 더 추가할 수 있다. 예를 들어, 제2 의료 데이터가 </stream>과 </performer>의 두 끝 태그를 포함하지 않으면, 데이터 검증부(330)는 제2 의료 데이터를 제2 손상 데이터로 구분하고 2 개의 영역에 대한 손상이 발생하였음을 의미하는 태그를 추가할 수 있다.

[0032] 데이터 결합부(400)는 스키마 구조가 유사한 제1 손상 데이터 및 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성할 수 있다. 데이터 결합부(400)는 제3 의료 데이터를 생성하기 위하여 하나 이상의 제2 의료 데이터에서 식별된 제1 손상 데이터와 제2 손상 데이터의 스키마 구조를 판단할 수 있다. 의료 데이터에 대한 스키마 구조는 의료 데이터에 대응되는 질병의 종류에 따라 상이할 수 있기 때문에 제1 손상 데이터와 제2 손상 데이터의 결합을 위하여 먼저 스키마 구조를 판단할 필요가 있다. 데이터 결합부(400)는 제1 손상 데이터와 제2 손상 데이터의 스키마 구조가 동일하다고 판단하면, 제1 손상 데이터와 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성할 수 있다. 예를 들어, 제1 손상 데이터에 포함된 질병 코드가 알러지에 대응되는 것이면 데이터 결합부(400)는 알러지에 대응되는 질병 코드에 따른 스키마 구조를 갖는 제2 손상 데이터를 검색하여 제3 의료 데이터를 생성할 수 있다.

[0033] 제2 검증부(500)는 데이터 결합부(400)에서 생성된 제3 의료 데이터를 검증할 수 있다. 제2 검증부(500)는 제3 의료 데이터의 스키마 구조가 표준 스키마 구조를 만족하는 지, 제3 의료 데이터의 상부 및 하부의 컨텍스트가 유사한 지 판단하여 제3 의료 데이터를 검증할 수 있다. 예를 들어, 제2 검증부(500)는 제3 의료 데이터에 대응되는 질병 코드에 대응되는 하나 이상의 태그가 작성되어 있는 지, 제3 의료 데이터에 포함된 태그 모두가 시작 태그와 끝 태그를 포함하고 있는지, 각 태그에 속성이 포함되어 있는지 등을 이용하여 제3 의료 데이터를 검증할 수 있다. 또한 제2 검증부(500)는 제3 의료 데이터의 상부 및 하부에 대한 컨텍스트를 추출하여 상부의 컨텍스트와 하부의 컨텍스트를 비교하여 제3 의료 데이터를 검증할 수 있다.

[0034] 제2 검증부(500)는 제3 의료 데이터가 의미적으로 유효하다고 판단되면 제3 의료 데이터를 데이터베이스에 저장하고, 그렇지 않다면 데이터 결합부(400)에 제3 의료 데이터의 재생성을 요청할 수 있다.

[0035] 이하에서는 도 2를 참조하여 본 발명의 일 실시 예에 의한 표준화된 의료 데이터의 의미 보존 방법을 설명한다. 표준화된 의료 데이터의 의미 보존 방법에 관한 설명에 있어서 전술한 표준화된 의료 데이터의 의미 보존 장치와 중복되는 세부 실시 예는 생략될 수 있다. 또한 표준화된 의료 데이터의 의미 보존 방법의 주체인 의미 보존 장치는 서버로 구현될 수 있는 바, 이하에서는 설명의 편의를 위하여 서버로 명명한다.

[0036] 서버는 데이터베이스로부터 표준화된 제1 의료 데이터를 불러올 수 있다(S100).

[0037] 서버는 제1 의료 데이터를 복수 개의 블록으로 분할하여 제2 의료 데이터를 생성하고, 제2 의료 데이터를 HDFS에 분산 저장할 수 있다(S200).

[0038] 서버는 기 저장된 검증 모델에서 하나 이상의 검증 조건을 추출하여 제2 의료 데이터가 검증 조건을 만족하는 지 판단할 수 있다(S300). 서버는 제2 의료 데이터의 스키마 구조가 표준 스키마를 따르는 지를 판단하여 제2 의료 데이터를 검증할 수 있다. 예를 들어 서버는 제2 의료 데이터에 포함된 태그 모두가 시작 태그와 끝 태그

를 포함하고 있는 지에 대한 여부로 제2 의료 데이터를 검증할 수 있다.

[0039] 서버는 제2 의료 데이터가 검증 조건을 만족하지 못하면 제2 의료 데이터가 손상되었다고 판단(S400)하여 제2 의료 데이터를 제1 손상 데이터 또는 제2 손상 데이터로 구분할 수 있다(S500). 만약 제2 의료 데이터가 검증 조건을 만족하면 서버는 제2 의료 데이터를 데이터베이스에 저장할 수 있다. 서버는 제2 의료 데이터에 포함된 태그 중 하나 이상이 시작 태그를 포함하지 않으면 제2 의료 데이터의 상부 영역이 손상되었다고 판단하고 제2 의료 데이터를 제1 손상 데이터로 구분할 수 있다. 나아가 서버는 제2 의료 데이터에 포함된 태그 중 하나 이상이 끝 태그를 포함하지 않으면 제2 의료 데이터의 하부 영역이 손상되었다고 판단하고 제2 의료 데이터를 제2 손상 데이터로 구분할 수 있다.

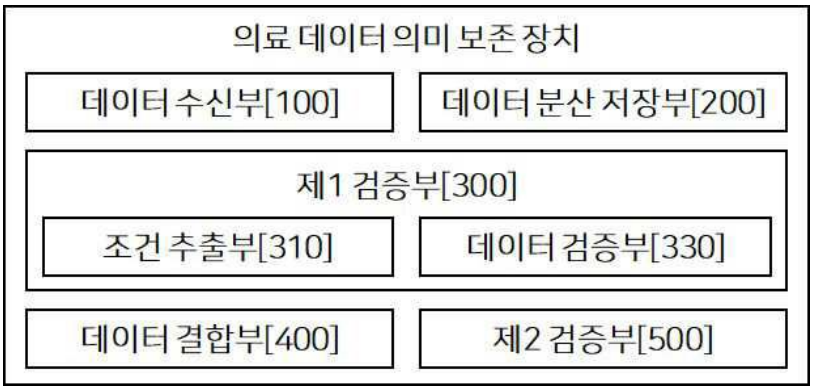
[0040] 서버는 하나 이상의 제2 의료 데이터에서 구분된 제1 손상 데이터와 제2 손상 데이터 중 스키마 구조가 유사한 제1 손상 데이터와 제2 손상 데이터를 결합하여 제3 의료 데이터를 생성할 수 있다(S600).

[0041] 서버는 제3 의료 데이터가 표준 스키마 구조를 만족하는 지를 판단하여 제3 의료 데이터를 검증할 수 있다(S700). 제3 의료 데이터가 표준 스키마 구조를 만족하면 서버는 제3 의료 데이터를 데이터베이스에 저장(S800)하고, 만족하지 않으면 다른 손상 데이터를 이용하여 제3 의료 데이터를 생성할 수 있다.

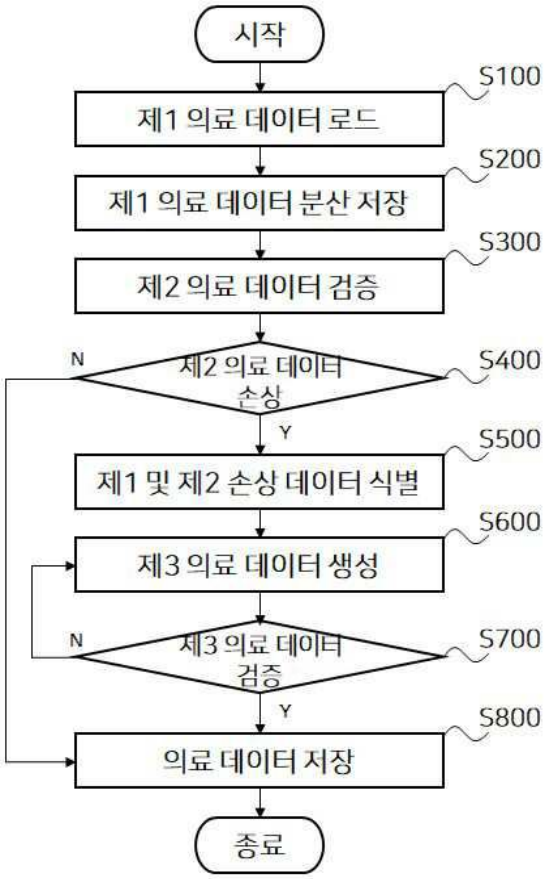
[0042] 본 명세서와 도면에 개시된 본 발명의 실시 예들은 본 발명의 기술 내용을 쉽게 설명하고 본 발명의 이해를 돕기 위해 특정 예를 제시한 것뿐이며, 본 발명의 범위를 한정하고자 하는 것은 아니다. 여기에 개시된 실시 예들 이외에도 본 발명의 기술적 사상에 바탕을 둔 다른 변형 예들이 실시 가능하다는 것은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 자명한 것이다.

도면

도면1



도면2


시작
S100
제1 의료 데이터 로드
S200
제1 의료 데이터 분산 저장
S300
제2 의료 데이터 검증
N
제2 의료 데이터 손상
S400
Y
S500
제1 및 제2 손상 데이터 식별
S600
제3 의료 데이터 생성
N
제3 의료 데이터 검증
S700
Y
S800
의료 데이터 저장
종료