(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0067332
(43) 공개일자 2021년06월08일

(51) 국제특허분류(Int. Cl.)
G06Q 50/10 (2012.01) *G06F 16/951* (2019.01)
G06F 40/20 (2020.01) *G06N 3/08* (2006.01)
(52) CPC특허분류
G06Q 50/10 (2013.01)
G06F 16/951 (2019.01)
(21) 출원번호 10-2019-0156834
(22) 출원일자 2019년11월29일
심사청구일자 2019년11월29일

(71) 출원인
조선대학교산학협력단
광주광역시 동구 필문대로 309 (서석동)
(72) 발명자
김판구
광주광역시 북구 설죽로 600, 101동 904호 (일곡동, 삼호아파트)
유경호
광주광역시 풍암2로 62, 201-1802
(뒷면에 계속)
(74) 대리인
특허법인아이엠

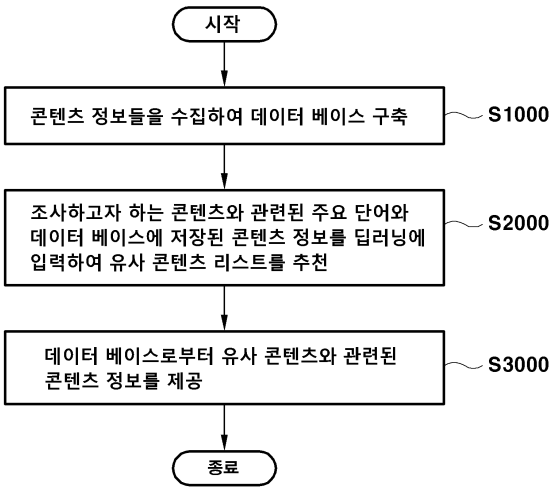
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 **스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.**

(57) 요 약

본 발명은 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법에 관한 것으로, 스토리텔링에 필요한 유효한 콘텐츠 정보들을 데이터베이스화하고, 데이터베이스화된 콘텐츠 정보 중 저작자가 필요한 콘텐츠의 정보를 딥러닝을 이용하여 제공할 수 있는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법이다.

대 표 도 - 도1



(52) CPC특허분류
G06F 40/205 (2020.01)
G06N 3/08 (2013.01)
(72) 발명자
김형주
광주광역시 동구 중앙로 160번길 16-33

노주현
광주광역시 북구 동문대로86번길 55, 202동 108호 (두암동, 두암동현대2차아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711081128
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	SW전문인력역량강화(R&D)
연구과제명	SW중심대학(조선대)
기 여 율	1/1
과제수행기관명	조선대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

명 세 서

청구범위

청구항 1

창작물, 전설, 설화 또는 민담과 관련된 콘텐츠 정보들을 수집하여 데이터베이스를 구축하는 단계;

저작자가 조사하고자 하는 콘텐츠와 관련된 주요 단어와 상기 데이터베이스에 구축된 콘텐츠 정보를 딥러닝에 입력하여 유사 콘텐츠 리스트를 추천하는 단계;

상기 데이터베이스에 저장된 유사 콘텐츠의 콘텐츠 정보를 제공하는 단계;를 포함하는 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.

청구항 2

제 1 항에 있어서,

상기 데이터베이스를 구축하는 단계:는

공공기관에 공개된 콘텐츠 정보를 수집하여 저장하는 제1 데이터베이스를 구축하는 단계; 및

상기 제1 데이터베이스에 구축된 콘텐츠 정보들을 기반으로 웹크롤링을 수행하여 웹상에서 콘텐츠 정보를 수집하여 저장하는 제2 데이터 베이스를 구축하는 단계;를 포함하는 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.

청구항 3

제 2 항에 있어서,

상기 제2 데이터 베이스를 구축하는 단계;는

상기 제1 데이터베이스에 수집된 콘텐츠 정보의 문장을 토큰화하는 단계;

토큰화된 문장의 불용어를 제거하는 단계;

불용어가 제거된 문장의 품사를 태깅하는 단계;

태깅된 품사 중 대명사를 추출하는 단계; 및

상기 대명사를 기반으로 웹크롤링을 수행하여 관련된 콘텐츠 정보를 수집하는 단계;를 포함하는 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.

청구항 4

제 1 항에 있어서,

상기 딥러닝은 BERT-KNRM 알고리즘인 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.

청구항 5

제 1 항에 있어서,

상기 콘텐츠 정보를 제공하는 단계는 콘텐츠와 관련된 이미지, 영상, 문장 중 어느 하나를 포함하는 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법.

청구항 6

컴퓨터와 결합하여 제 1 항 내지 제 5 항 중 어느 한 항의 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법을 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법에 관한 것으로, 스토리텔링에 필요한 유효한 콘텐츠 정보들을 데이터베이스화하고, 데이터베이스화된 콘텐츠 정보 중 저작자가 필요한 콘텐츠의 정보를 딥러닝을 이용하여 제공할 수 있는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법이다.

배 경 기 술

[0002] 최근 5세대 이동통신 상용화를 필두로 사용자가 영화, 게임과 같은 콘텐츠와 이를 더욱 몰입도 높게 즐길 수 있는 가상현실/증강현실 기술이 주목받고 있다. 하지만 사용자가 지속적으로 즐길 수 있는 콘텐츠는 부족할 실정이다. 이러한 콘텐츠를 제작하기 위해서는 사용자가 빠져들 수 있는 스토리텔링이 콘텐츠에 포함되어야 한다.

[0003] 스토리텔링이란 스토리(Story)와 텔링(Telling) 의 합성어로서 말 그대로 "이야기하다 라는 의미를 지니며, 상대방에게 알리고자 하는 바를 재미있고 생생한 이야기로 설들력 있게 전달하는 행위이다.

[0004] 이러한, 스토리텔링은 문학, 영화, 교육학 등에 사용되고 있으며, 알리고자 하는 바를 단어, 이미지, 소리를 통해 사건, 이야기를 전달하는 것이다.

[0005] 한편, 스토리텔링과 같은 저작 작업은 유사 콘텐츠로부터 정보를 수집해야 하는데, 종래에는 많은 시간과 노력을 소모하여, 인터넷과 서적 등을 통해 유사 콘텐츠를 수작업으로 일일이 찾아야하는 문제가 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) KR10-1706167 B1 "콘텐츠 추천을 위한 장치 및 방법"

(특허문헌 0002) KR10-1559459 B1 "콘텐츠 추천 및 제공 방법"

발명의 내용

해결하려는 과제

[0007] 본 발명은 상술한 문제점을 해결하기 위해 안출된 것으로, 본 발명의 목적은 저작자가 스토리텔링의 저작 작업을 수행할 때 참고하고자 하는 콘텐츠와 유사한 콘텐츠 정보를 저작자에게 추천하여 이를 통해 창의적이고 자유로운 저작 활동에 도움을 줄 수 있는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법을 제공하는 데 있다.

과제의 해결 수단

[0008] 상술한 목적을 해결하기 위해 본 발명은 창작물, 전설, 설화 또는 민담과 관련된 콘텐츠 정보들을 수집하여 데이터베이스를 구축하는 단계; 저작자가 조사하고자 하는 콘텐츠와 관련된 주요 단어와 상기 데이터베이스에 구축된 콘텐츠 정보를 딥러닝에 입력하여 유사 콘텐츠 리스트를 추천하는 단계; 및 상기 데이터베이스에 저장된 유사 콘텐츠의 콘텐츠 정보를 제공하는 단계;를 포함하는 것을 특징으로 하는 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법을 제공한다.

[0009] 바람직한 실시예에 있어서, 상기 데이터베이스를 구축하는 단계:는 공공기관에 공개된 콘텐츠 정보를 수집하여 저장하는 제1 데이터베이스를 구축하는 단계; 및 상기 제1 데이터베이스에 구축된 콘텐츠 정보들을 기반으로 웹크롤링을 수행하여 웹상에서 콘텐츠 정보를 수집하여 저장하는 제2 데이터 베이스를 구축하는 단계;를 포함할 수 있다.

[0010] 바람직한 실시예에 있어서, 상기 제2 데이터 베이스를 구축하는 단계;는 상기 제1 데이터베이스에 수집된 콘텐츠 정보의 문장을 토큰화하는 단계; 토큰화된 문장의 불용어를 제거하는 단계; 불용어가 제거된 문장의 품사를 태깅하는 단계; 태깅된 품사 중 대명사를 추출하는 단계; 및 상기 대명사를 기반으로 웹크롤링을 수행하여 관련된 콘텐츠 정보를 수집하는 단계;를 포함한다.

[0011] 바람직한 실시예에 있어서, 상기 딥러닝은 BERT-KNRM 알고리즘이다.

[0012] 바람직한 실시예에 있어서, 상기 콘텐츠 정보를 제공하는 단계는 콘텐츠와 관련된 이미지, 영상, 문장 중 어느 하나를 포함한다.

[0013] 또한, 본 발명은 컴퓨터와 결합하여 상기 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법을 실행시키기 위한 저장매체에 저장된 컴퓨터 프로그램을 제공한다.

발명의 효과

[0014] 본 발명은 다음과 같은 우수한 효과가 있다.

[0015] 본 발명의 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법에 따르면, 저작자가 참고하고자 하는 콘텐츠의 주요 단어를 입력하는 것만으로 해당되는 유사 콘텐츠를 추출하여 저작자에게 유사 콘텐츠의 콘텐츠 정보를 제공할 수 있으므로 콘텐츠 정보를 쉽고 편리하게 수집할 수 있어, 자료 수집에 투입되는 시간이 절약될 수 있다는 장점이 있다.

도면의 간단한 설명

[0016] 도 1은 본 발명의 일 실시예에 따른 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법의 순서도,

도 2는 본 발명의 일 실시예에 따른 데이터베이스 구축 방법의 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0017] 본 발명에서 사용되는 용어는 가능한 현재 널리 사용되는 일반적인 용어를 선택하였으나, 특정한 경우는 출원인이 임의로 선정한 용어도 있는데 이 경우에는 단순한 용어의 명칭이 아닌 발명의 상세한 설명 부분에 기재되거나 사용된 의미를 고려하여 그 의미가 파악되어야 할 것이다.

[0018] 이하, 첨부한 도면에 도시된 바람직한 실시예들을 참조하여 본 발명의 기술적 구성을 상세하게 설명한다.

[0019] 그러나 본 발명은 여기서 설명되는 실시예에 한정되지 않고 다른 형태로 구체화될 수도 있다. 명세서 전체에 걸쳐 동일한 참조번호는 동일한 구성요소를 나타낸다.

[0020] 도 1은 본 발명의 일 실시예에 따른 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법의 순서도로, 도 1을 참조하면, 본 발명의 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법은 저작자에게 스토리텔링 작업에 필요한 콘텐츠 정보들을 제공할 수 있는 방법이다.

[0021] 또한, 상기 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법은 실질적으로 컴퓨터에 의해 수행되며 상기 컴퓨터에는 상기 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법이 수행하기 위한 컴퓨터 프로그램이 저장된다.

[0022] 또한, 상기 컴퓨터는 일반적인 퍼스널 컴퓨터뿐만 아니라, 스마트폰이나 태블릿 PC와 같은 스마트 기기를 포함하는 광의의 컴퓨팅 장치를 의미한다.

[0023] 또한, 상기 컴퓨터 프로그램은 별도의 기록 매체에 저장되어 제공될 수 있으며, 상기 기록 매체는 본 발명을 위하여 특별히 설계되어 구성된 것들이거나 컴퓨터 소프트웨어 분야에서 통상의 지식을 가진 자에게 공지되어 사용 가능한 것일 수 있다.

[0024] 예를 들면, 상기 기록 매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD, DVD와 같은 광 기록 매체, 자기 및 광 기록을 겸할 수 있는 자기-광 기록 매체, 롬, 램, 플래시 메모리 등 단독 또는 조합에 의해 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치일 수 있다.

[0025] 또한, 상기 컴퓨터 프로그램은 프로그램 명령, 로컬 데이터 파일, 로컬 데이터 구조 등이 단독 또는 조합으로 구성된 프로그램일 수 있고, 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라, 인터프리터 등을 사용하여 컴퓨터에 의해 실행될 수 있는 고급 언어 코드로 짜여진 프로그램일 수 있다.

[0026] 이하에서는 본 발명의 일 실시예에 따른 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법에 대해 상세히 설명한다.

[0027] 먼저, 본 발명의 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법은 다양한 콘텐츠 정보들을 수집하여 데이터베이스를 구축한다(S1000).

[0028] 도 2는 본 발명의 일 실시예에 따른 데이터베이스 구축 방법을 보여주는 순서도이다.

[0029] 도 2를 참조하면, 상기 데이터베이스는 저작자가 스토리텔링에 필요한 다양한 콘텐츠 정보들을 수집하여 저장되는 공간으로, 스토리텔링에 필요한 콘텐츠 정보를 수집하기 위해 먼저, 공공기관에 공개된 콘텐츠 정보를 수집하여 제1 데이터베이스를 구축한다(S1100).

[0030] 여기서, 상기 콘텐츠 정보는 소설, 영화, 드라마, 시나리오 등의 창작물, 전설, 설화 또는 민담과 관련된 데이터로, 문장, 이미지, 영상을 포함한다.

[0031] 다음, 수집된 콘텐츠 정보의 문장들을 단어 단위로 토큰화한다(S1200).

[0032] 다음, 토큰화된 문장들로부터 불용어를 제거한다(S1300).

[0033] 상기 불용어는 큰 의미가 없는 단어를 제거하는 것으로, 이를 통해 토큰화된 문장들에서 조사, 접미사 같은 단어들을 제거한다.

[0034] 다음, 불용어가 제거된 문장의 품사를 태깅한다(S1400).

[0035] 상기 태깅은 문장 내의 단어가 사용된 문맥에 따라서 품사를 할당하는 것으로, 통계 기반 접근 방법, 규칙 기반 접근 방법, 통합 접근 방법이 사용될 수 있다.

[0036] 다음, 태깅된 문장으로부터 대명사를 추출한다(S1500).

[0037] 상기 대명사는 사람, 사물, 장소를 대신해서 가리키는 단어로, 콘텐츠를 대표할 수 있는 단어일 수 있으며, 아래에서 설명할 제2 데이터베이스를 구축할 때 유효한 콘텐츠 정보를 찾을 수 있는 기준이 된다.

[0038] 다음, 대명사를 기반으로 웹크롤링을 통해 콘텐츠 정보들을 수집하여 제2 데이터베이스를 구축한다(S1600).

[0039] 상기 웹크롤링은 검색 엔진과 같은 여러 사이트에서 콘텐츠를 수집하는 작업으로, 공지된 다양한 라이브러리를 통해 수행될 수 있다.

[0040] 또한, 상기 웹크롤링을 수행할 때 검색하고자 하는 단어를 입력하게 되는데 추출된 대명사를 입력하여 상기 대명사와 관련된 유효한 콘텐츠 정보들을 수집한다.

[0041] 즉, 상기 웹크롤링을 할 때 대명사를 기반으로 콘텐츠 정보들을 수집함으로써, 연관성이 낮은 불필요한 콘텐츠 정보들이 수집되는 것을 최소화할 수 있다.

[0042] 또한, 상기 제1 데이터베이스와 상기 제2 데이터베이스에 저장된 콘텐츠 정보는 고유 인덱스 정보와 함께 정보된다.

[0043] 상기 고유 인덱스 정보는 각 콘텐츠를 대표하는 숫자로, 예를 들면 "0"은 "설씨녀와 가실"에 관한 콘텐츠, "1"은 "감생설화"에 관한 콘테츠로 표현될 수 있다.

[0044] 다음, 데이터 베이스를 구축한(S1000) 후에는, 딥러닝을 통해 저작자가 필요로 하는 콘텐츠와 유사한 콘텐츠를 추천한다(S2000).

[0045] 본 발명에서는 상기 딥러닝은 "CEDR:Contextualized Embeddings for Document Ranking", San MacAvaney, Arman Cohan, Andrew Yates, Nazli Goharian 논문에 공지된 BERT-KNRM 알고리즘을 사용한다.

[0046] 상기 BERT-KNRM 알고리즘은 공지된 BERT 알고리즘과 KNRM 알고리즘을 융합한 것으로, 상기 BERT-KNRM 알고리즘을 통해 유사한 콘텐츠를 추천하기 위해서는 저작자가 필요로 하는 콘텐츠의 주요 단어와 상기 데이터베이스에 구축된 콘텐츠 정보를 필요로 한다.

[0047] 자세하게는, 상기 BERT-KNRM 알고리즘은 단어 또는 문장으로 이루어진 쿼리와 문서를 동시에 입력하여, 쿼리와 관련성이 높은 문서들의 리스트를 뽑을 수 있는 알고리즘으로, 본 발명에서는 상기 주요 단어를 쿼리로 입력하고, 상기 데이터베이스에 구축된 문장들과 고유 인덱스를 문서로 입력한다.

[0048] 즉, 상기 주요 단어와 상기 데이터베이스에 구축된 문장들을 상기 BERT-KNRM 알고리즘에 입력하게되면 학습을 통해 상기 주요 단어와 관련성이 높은 문장들을 보유한 고유 인덱스들이 추출되게 된다.

[0049] 또한, 상기 BERT-KNRM 알고리즘을 통해 추출되는 고유 인덱스의 개수는 저작자에 의해 다양하게 설정될 수 있다.

[0050] 다음, 추출된 고유 인덱스를 기반으로 상기 제1 데이터베이스와 상기 제2 데이터베이스로부터 콘텐츠 정보를 불러와 저작자에게 제공한다(S3000).

[0051] 상세하게는 상기 제1 데이터베이스와 상기 제2 데이터베이스에 저장된 콘텐츠 정보들 중 추출된 고유 인덱스와 동일한 고유 인덱스를 갖는 콘텐츠 정보들을 불러오며, 문장, 영상, 이미지가 제공될 수 있다.

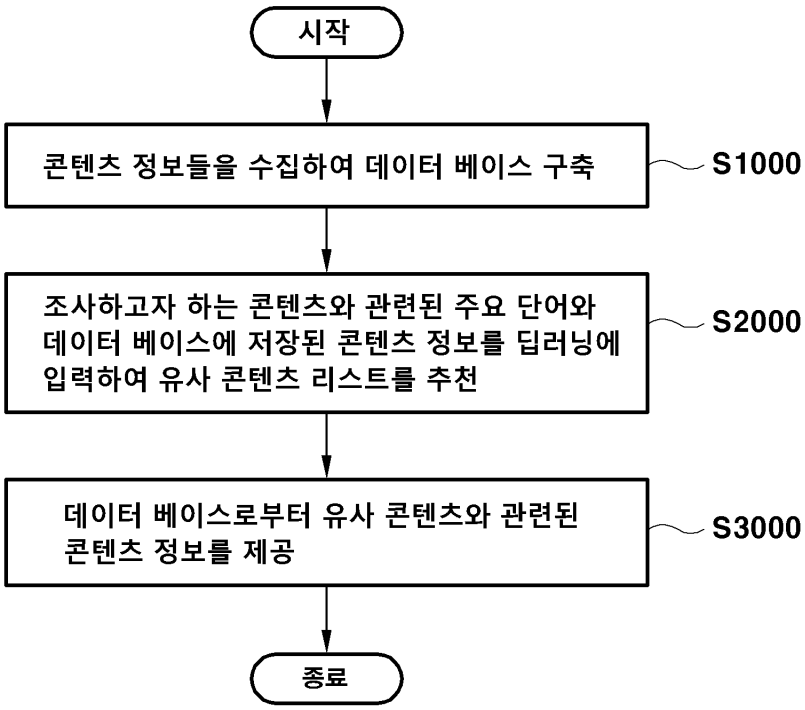
[0052] 따라서, 본 발명의 스토리텔링 저작을 위한 콘텐츠 추천 및 제공 방법은 저작자가 필요로 하는 콘텐츠의 주요 단어를 가지고 유사도가 높은 콘텐츠들을 리스트로 제공할 수 있기 때문에 자료 수집에 투입되는 시간이 절약될 수 있다는 장점이 있다.

[0053] 이상에서 살펴본 바와 같이 본 발명은 바람직한 실시예를 들어 도시하고 설명하였으나, 상기한 실시예에 한정되지 아니하며 본 발명의 정신을 벗어나지 않는 범위 내에서 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변경과 수정이 가능할 것이다.

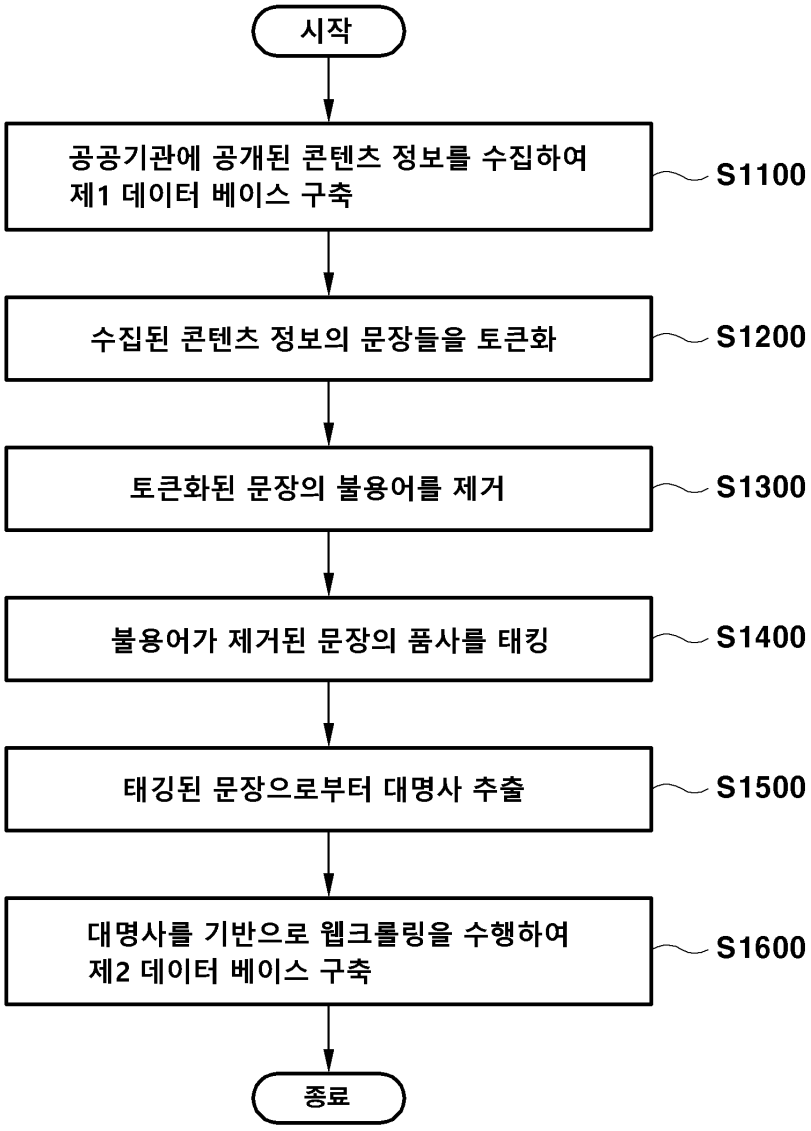
부호의 설명

도면

도면1



도면2


시작
공공기관에 공개된 콘텐츠 정보를 수집하여 제1 데이터 베이스 구축
S1100
수집된 콘텐츠 정보의 문장들을 토큰화
S1200
토큰화된 문장의 불용어를 제거
S1300
불용어가 제거된 문장의 품사를 태깅
S1400
태깅된 문장으로부터 대명사 추출
S1500
대명사를 기반으로 웹크롤링을 수행하여 제2 데이터 베이스 구축
S1600
종료