



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년03월14일
(11) 등록번호 10-1242369
(24) 등록일자 2013년03월05일

(51) 국제특허분류(Int. Cl.)
G06F 17/30 (2006.01)
(21) 출원번호 10-2007-7030613
(22) 출원일자(국제) 2006년06월27일
심사청구일자 2011년06월01일
(85) 번역문제출일자 2007년12월27일
(65) 공개번호 10-2008-0024157
(43) 공개일자 2008년03월17일
(86) 국제출원번호 PCT/US2006/024847
(87) 국제공개번호 WO 2007/005382
국제공개일자 2007년01월11일
(30) 우선권주장
11/172,121 2005년06월29일 미국(US)
(56) 선행기술조사문헌
US20050033777 A1
전체 청구항 수 : 총 17 항

(73) 특허권자
마이크로소프트 코퍼레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
더마이스, 수잔, 티.
미국 98052-6399 워싱턴주 레드몬드 원 마이크로
소프트 웨이
호비츠, 에릭, 제이.
미국 98052-6399 워싱턴주 레드몬드 원 마이크로
소프트 웨이
(74) 대리인
제일특허법인

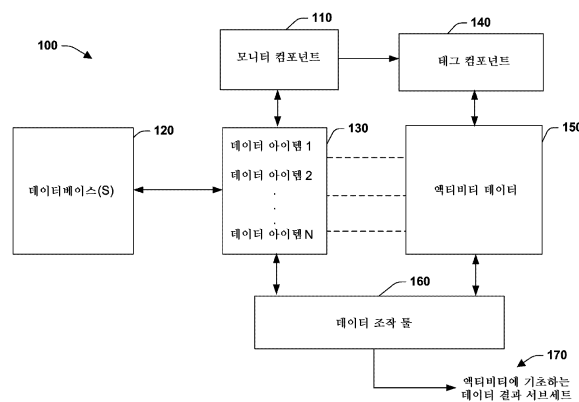
심사관 : 이명진

(54) 발명의 명칭 사용자 액티비티, 주의, 관심 측정을 활용하는 데이터감지, 저장, 인덱싱, 및 탐색

(57) 요약

다양한 컴포넌트들 및 프로세스들을 제공하여, 사용자 액티비티, 주의, 관심, 위치의 히스토리의 양태들, 또는 데이터와의 다른 상호작용이 결정되는 다수의 데이터 유형에 대하여 데이터 처리를 가능하게 하고, 정보 저장 및 액세스를 향상시킨다. 일 양태에서는, 데이터 조작 시스템을 제공한다. 이 시스템은, 하나 이상의 태그에 관련되고 사용자의 데이터와의 적어도 하나의 상호작용 또는 액티비티를 가리키는 하나 이상의 데이터 아이템을 포함한다. 조작 툴은, 사용자의 데이터 아이템과의 상호작용에 적어도 부분적으로 기초하여 데이터 아이템을 처리함으로써 데이터의 서브셋을 결정한다. 이러한 조작 툴을 이용하여, 인덱스의 용어들을 가중하고, 인덱스들을 압축하며, 검색에서 리턴된 아이템들의 랭킹에 영향을 끼치고, 데이터 아이템에 대한 추가 문의들을 자동적으로 또는 사용자 지시에 따라 발생시키고, 또는 데이터 아이템의 개선된 프리젠테이션을 위해 추가 문의들을 발생시키는 방법들을 설명한다.

대표도



특허청구의 범위

청구항 1

컴퓨터 구현된 데이터 조작 시스템으로서,

프로세서;

정보를 제시하는 디스플레이 디바이스; 및

상기 프로세서에 동작 가능하게 연결되어 컴퓨터 실행가능 명령어들을 저장하는 컴퓨터 판독가능 저장 매체 - 상기 컴퓨터 실행가능 명령어들은, 상기 프로세서에 의해 실행될 때에, 컴포넌트들을 구현함 -

를 포함하고,

상기 컴포넌트들은,

하나 이상의 데이터 아이템을 하나 이상의 태그와 자동으로 연관시키는 태깅(tagging) 컴포넌트 - 상기 연관된 태그는 적어도 하나의 사용자에게 의해 데이터 아이템과의 상호작용 또는 상기 데이터 아이템에 대한 주의(attention)의 양(quantity)이나 성질(nature)을 표시하도록 가중되어, 상기 데이터 아이템과 상기 적어도 하나의 사용자의 후속하는 상호작용에 응답하여 상기 디스플레이 디바이스를 통해 상기 적어도 하나의 사용자에게 데이터를 제시하는 것을 용이하게 함 - ; 및

상기 프로세서를 통해 상기 데이터 아이템을 처리하여 상기 데이터 아이템과 상기 적어도 하나의 사용자의 상호작용에 적어도 부분적으로 기초하여 상기 데이터 아이템의 서브셋을 결정하는 조작 틀 - 상기 조작 틀은 또한 데이터 아이템과의 상호작용 또는 주의를 나타내는 주석의 상태에 따른 검색을 위해 상기 데이터 아이템의 서브셋을 상이하게 가중하는 인덱싱 프로시저를 제공함 - ;

을 포함하고,

상기 컴퓨터 판독가능 저장 매체는, 상기 데이터, 상기 연관된 태그, 또는 가중값 중 적어도 하나를 보유하는 컴퓨터 구현된 데이터 조작 시스템.

청구항 2

제1항에 있어서, 상기 데이터 아이템은, 다큐먼트(documents), 파일, 이메일 메시지, 캘린더 약속(calendar appointment), 웹 페이지, 상기 데이터 아이템 내의 서브섹션, 또는 아이템 간 추상(cross-item abstractions)을 포함하는 컴퓨터 구현된 데이터 조작 시스템.

청구항 3

제1항에 있어서, 상기 하나 이상의 태그 중 제1 태그는 상기 적어도 하나의 사용자가 상기 연관된 데이터 아이템에 마지막으로 액세스한 위치를 나타내거나, 상기 적어도 하나의 사용자가 상기 연관된 데이터 아이템에 액세스하거나 상기 연관된 데이터 아이템과 상호작용한 횟수의 위치 이력(location history of times)을 나타내는 컴퓨터 구현된 데이터 조작 시스템.

청구항 4

제1항에 있어서, 상기 하나 이상의 태그 중 제1 태그는 상기 적어도 하나의 사용자가 상기 연관된 데이터 아이템에 마지막으로 액세스한 시간을 나타내고, 상기 연관된 데이터 아이템에 액세스한 총 횟수는, 과거로 연장되는 시간 기간 내에서 상기 연관된 데이터 아이템에 액세스한 빈도(frequency)를 나타내거나, 또는 하나 이상의 임의로 지정된 시간 기간 내에서 상기 연관된 데이터 아이템에 액세스한 빈도를 나타내는 컴퓨터 구현된 데이터 조작 시스템.

청구항 5

제1항에 있어서, 시간에 따른 액세스 빈도의 고차 통계값(higher-order statistics)을 인코딩하는 컴포넌트를 더 포함하는 컴퓨터 구현된 데이터 조작 시스템.

청구항 6

제1항에 있어서, 상기 적어도 하나의 사용자가 상기 하나 이상의 태그의 기능에 기초하여 상기 하나 이상의 데이터 아이템 중 적어도 하나의 데이터 아이템을 검색할 수 있게 하는 뷰어, 상기 적어도 하나의 사용자가 상기 하나 이상의 태그의 기능에 기초하여 검색된 데이터 아이템을 정렬하거나 필터링할 수 있게 하는 뷰어, 또는 상기 하나 이상의 태그의 기능에 기초하여 상기 하나 이상의 데이터 아이템의 검색된 데이터 아이템을 제시하는 뷰어를 더 포함하는 컴퓨터 구현된 데이터 조작 시스템.

청구항 7

데이터 아이템에 의해 사용자 활동(user activities)을 분석하는 컴퓨터 구현된 방법으로서,
 상기 데이터 아이템에 의해 사용자의 상호작용을 모니터링하는 단계;
 상기 데이터 아이템과 상기 사용자의 상호작용의 양 또는 성질을 나타내는 데이터 아이템을 자동으로 태깅하는 단계 - 태그는 상호작용의 정도(degree)를 나타내는 연관된 가중값을 가짐 - ;
 상기 태그를 이용하여, 데이터 또는 상기 데이터의 서브세트에 관련된 장래의 사용자 활동에 따라 상기 데이터 아이템을 데이터 아이템의 서브세트로 추가 처리하는 단계;
 데이터 아이템에 대한 주의 또는 상호작용을 나타내는 주석의 상태에 따른 검색을 위해 상기 데이터 아이템의 서브컴포넌트들을 상이하게 가중하는 인덱싱 프로시저를 제공하는 단계; 및
 상기 장래의 사용자 활동, 상기 태그, 또는 태그 가중값 중 적어도 하나에 기초하여 상기 데이터 또는 상기 데이터의 서브세트를 상기 사용자에게 제시하는 단계
 를 구현하도록 컴퓨터 관독가능 저장 매체 상에 저장된 컴퓨터 실행가능 명령어들을 실행하는 프로세서를 이용하는 단계를 포함하는 컴퓨터 구현된 방법.

청구항 8

제7항에 있어서, 상기 데이터 아이템과 연관된 주의 주석(attentional annotation) 내의 데이터를 별개의 데이터베이스 또는 상기 데이터 아이템에 임베디드된 데이터 구조 내에 저장하는 단계를 더 포함하는 컴퓨터 구현된 방법.

청구항 9

제8항에 있어서, 상기 데이터 아이템은 컴퓨터 관독가능 저장 매체를 구성(compose)하는 컴퓨터 구현된 방법.

청구항 10

제9항에 있어서, 상기 컴퓨터 관독가능 저장 매체는 텍스트, 그래픽, 및 관련 데이터 컴포넌트를 포함하는 데이터 아이템인 컴퓨터 구현된 방법.

청구항 11

제7항에 있어서, 상기 인덱싱 프로시저는 데이터 아이템에 대한 주의 또는 상호작용을 나타내는 주석의 상태에 따라 데이터 아이템의 정보를 간과(overlook)하거나 삭제(delete)하는 컴퓨터 구현된 방법.

청구항 12

제7항에 있어서, 상기 인덱싱 프로시저는 컴퓨터 사용자가 주의하지 않았거나 컴퓨터 사용자와 상호작용하지 않은 컴포넌트를 제거하여 압축되는 인덱스를 생성하는 컴퓨터 구현된 방법.

청구항 13

제12항에 있어서, 상기 인덱싱 프로시저는 데이터 아이템의 컴포넌트에 대한 주의가 적거나 데이터 아이템의 컴포넌트와의 상호작용이 적은 컴포넌트를 제거하여 압축되는 인덱스를 생성하는 컴퓨터 구현된 방법.

청구항 14

제12항에 있어서, 검색에 이용되는 랭킹 점수를 제공하여, 사용자가 주의했거나 상호작용한 데이터 아이템의 색선에 나타나는 용어(terms) 또는 개체(objects)에 사용자가 주의했거나 상호작용한 정도에 비례하는 가중값을 부여하는 단계를 더 포함하는 컴퓨터 구현된 방법.

청구항 15

제12항에 있어서, 주의 주석을 이용하여 과거 또는 현재 또는 이들 모두에서 주의되었거나 상호작용된 영역에 기초하여 문의(queries)를 자동으로 또는 반자동으로 생성하는 단계 또는 주의 주석을 이용하여 과거 또는 현재 또는 이들 모두에서 주의되었거나 상호작용된 아이템의 차별적인 디스플레이(differential display) 또는 아이템에 대한 차별적인 액세스(differential access)를 제공하는 단계를 더 포함하는 컴퓨터 구현된 방법.

청구항 16

데이터 아이템에 의해 사용자 활동을 분석하는 컴퓨터 구현된 시스템으로서,

프로세서;

사용자 인터페이스를 제시하는 디스플레이; 및

컴퓨터 실행가능 명령어들을 저장하는 컴퓨터 판독가능 저장 매체 - 상기 컴퓨터 실행가능 명령어들은, 상기 프로세서에 의해 실행될 때에, 컴포넌트들을 구현함 -

를 포함하고,

상기 컴포넌트들은,

상기 프로세서에 의해, 하나 이상의 데이터 아이템과 관련하여 사용자 활동을 결정하는 수단;

상기 사용자 활동에 적어도 부분적으로 기초하여 상기 데이터 아이템을 태깅하는 수단 - 태그는 상기 데이터 아이템과 사용자의 상호작용의 정도(degree)를 나타내는 연관된 가중값을 가짐 - ;

상기 태그에 부분적으로 기초하여 상기 컴퓨터 판독가능 저장 매체에 상기 데이터 아이템, 및 상기 데이터 아이템과 연관된 가중값의 레벨을 저장하는 수단;

상기 프로세서에 의해, 상기 데이터 아이템 및 연관된 가중된 태그를 처리하는 수단 - 상기 처리는, 상기 데이터 아이템에 대한 주의 또는 상호작용을 나타내는 주석의 상태에 따른 검색을 위해 상기 데이터 아이템 및 상기 데이터 아이템의 서브셋을 상이하게 가중하는 인덱싱 프로시저를 제공하는 것을 포함함 - ; 및

상기 데이터 아이템과 적어도 하나의 사용자의 후속하는 상호작용에 기초하여 상기 사용자 인터페이스를 통해 상기 적어도 하나의 사용자에게 상기 데이터 아이템 또는 관련 데이터 아이템을 제시하는 수단 - 상기 데이터 아이템은 상기 가중된 태그에 따라 제시됨 -

을 포함하는 컴퓨터 구현된 시스템.

청구항 17

제16항에 있어서, 상기 데이터 아이템에 대한 주의 주석을 인코딩하는 수단 및 상기 데이터 아이템 내의 서브컴포넌트에 대한 주의 주석을 인코딩하는 수단을 더 포함하고, 상기 주석은 상기 서브컴포넌트 각각의 포인터 표시 및 상기 서브컴포넌트에 대해 수신된 결과로서 생성된 사용자 주의를 캡처하는 컴퓨터 구현된 시스템.

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

명세서

배경기술

[0001] 컴퓨터 플랫폼들은 많은 다양한 유형들의 데이터 세트들을 저장하고 처리하기 위한 많은 툴을 제공할 수 있다. 이러한 컴퓨터 플랫폼은, 워드 프로세싱 툴, 데이터 프리젠테이션 툴, 컴퓨터 보조 그래픽 툴, 전자 메일 핸들링 툴, 달력 및 스케줄링 툴, 및 수많은 데이터베이스 조작 툴을 포함한다. 플랫폼 상의 데이터에 대하여 다양한 사용을 제공함에 따라, 애플리케이션들은, 시간에 따라 다소 콘텐츠 중심으로 개발되었다. 다시 말하면, 데이터가 컴퓨터의 데이터베이스에 저장된 경우, 이에 후속하여, 이 데이터는 자신의 실제 콘텐츠에 기초하는 소정의 방식으로 검색되고 그리고/또는 조작된다. 하나의 특정 예에서, 이메일 인박스는, 검색 툴을 용어에 관련된 각 이메일에 링크하는 키워드에 기초하여 이전에 수신된 이메일들에 대하여 검색될 수 있으며, 여기서 용어는 저장된 이메일의 실제 콘텐츠에 링크된다. 따라서, 사용자가 "John"이라는 키워드를 검색하고자 하면, 이러한 키워드에 관련된 임의의 이메일을 탐색하여 사용자에게 제시하고, 이에 따라 사용자는 후속하여 "John"이라는 용어에 관련된 원하는 이메일을 위해 탐색된 리스트를 조사한다. 사용자가 검색하고 있는 특정한 이메일이 메일의 최종 리스트에서 탐색될 수 있지만, 원하는 이메일을 찾기 위해 많은 이메일들(예를 들어, John이라는 용어를 포함하는 30개의 이메일들)을 후속 검색해야 할 수 있다. 인식할 수 있듯이, 위의 예에서 설명한 이메일 처리는 많은 유형의 데이터 처리 및 파일 조작 액티비티를 포함하도록 확장될 수 있다. 예를 들어, 이러한 데이터 처리 유형은, 저장된 데이터의 인텍싱, 저장된 데이터의 프리젠테이션, 다양한 유형의 저장된 데이터 검색, 데이터 랭킹, 데이터 인텍싱 등을 포함한다.

[0002] 일반적으로 콘텐츠 중심 애플리케이션에 관련해서, 판독자에 의해 탐색되고, 뷰잉되며, 이용되는 완성된 다큐먼트의 하나의 공통적인 뷰(view)는, 일반적으로 지식 집약적 태스크를 적절히 지원하는 데 충분하지 않는다. 따라서, 사용자들 또는 사용자들의 그룹들은 자신들의 고유한 정보를 지식 소스에 추가할 수도 있어야 한다. 일 예에서, 역사가는 책의 한 챕터에 상세한 분석을 추가하길 원할 수 있다. 다른 사용자는, 책의 한 섹션에, 그러한 분석으로부터 모은 경험으로 주석달기를 원할 수 있다.

[0003] 거의 모든 다큐먼트들이 웹 상에서 또는 웹을 통해 이용가능하지만, 이 웹의 하이퍼텍스트 기능은, 현재, 기존의 정보(예를 들어, 책, 서류, 웹 페이지 등)를 직접적으로 수정하고 주석을 다는 데 광범위하게 사용되고 있지 않다. 오히려, 콘텐츠가 완성된 것으로 보일 때, 이 콘텐츠는 소정 유형의 아카이브(예를 들어, 디지털 라이브러리)에 저장되며, 결국에는 이러한 아카이브로부터 더 많은 콘텐츠의 생성을 위해 사용되는 모놀리식 엔티티로서 탐색된다. 게다가, 정보 탐색 태스크는 통상적으로 콘텐츠 개발 태스크와 통합되지 않는다. 따라서, 사용자는 태스크에 필요하다고 믿는 다큐먼트들을 탐색해야 하고 이후 콘텐츠 개발을 발견한 정보에 기초해야 한다. 새로운 다큐먼트 검색을 항상 수동으로 개시할 수 있지만, 콘텐츠 개발과 탐색이 통합되어야 하는 것이 훨씬 강력하게 전망된다. 사용자에게 의해 입력되는 새로운 텍스트를 연속적으로 스캐닝하고 분석하는 시스템은 추가 관련 정보를 검색하고 이를 사용자에게 제시할 수 있어야 하며, 이에 따라 사용자는 예를 들어 그 새로운 데이터를 검사하고, 통합하며, 상호 참조를 추가하고, 또는 제안된 소스들을 거절할 수 있다.

[0004] 또다른 양태는, 일반적으로 소스로부터의 지식이 다큐먼트의 생성자 및 다큐먼트의 판독자 둘 다의 컨텍스트의 설명 없이는 적용될 수 없다는 것이다. 2개의 컨텍스트 프레임들의 명시적인 표현만이 이 프레임들 간의 (반자동) 번역을 허용한다. 이 예에서, 구(old) 지식은 최근의 표준 및 어휘에 적용될 수 있지만, 현재 디지털 형태로 생성되고 저장되는 모든 다큐먼트들이 역사적인 지식 자체로 되는 경우, 유사한 문제점들이 중장기의 미래에 점점 더 발생할 수 있다.

[0005] 현재, 사용자들은 소정 유형의 인텍싱 및 랭킹 시스템들, 예를 들어, 웹 검색 엔진들 또는 디지털 라이브러리를 위한 소정 유형의 정보 탐색 시스템들(역사적으로, 이러한 시스템들은 서로 다른 루트들로부터 온 것이지만, 최근의 구현예에서는 이러한 기술들 간에 일부 중첩되는 현상을 보임)을 통해 다큐먼트들을 얻는다. 어느 경우에서든, 시스템들은, 일반적으로, 웹 페이지들, 서류들 또는 책 전체든 상관없이 완전한 다큐먼트들을 리턴한다. 이것은 처리할 정보의 소스가 실질적으로 끝이 없어 많은 사용자들이 공감하는 "정보 과부하"의 주요 이유들 중 하나이다.

발명의 상세한 설명

[0006] 다음에 따르는 설명은, 본 명세서에서 설명하는 일부 양태들의 기본적인 이해를 제공하고자 간략한 개요를 제시한다. 이 개요는, 명세서의 넓은 범위에 걸친 개요가 아니며, 본 발명의 주요한/중대한 요소들을 식별하지 않

으며 또는 본 명세서에서 설명하는 다양한 양태들의 범위를 제한하지 않는다. 이 개요의 유일한 목적은 일부 개념들을 후술하는 상세한 설명에 대한 도입부로서 간략한 형태로 제공하는 것이다.

[0007] 순수한 콘텐츠 중심 데이터 처리 애플리케이션과는 달리, 파일 또는 애플리케이션에 관련된 메타데이터 태그는, 정보의 효율적인 정보 저장 및/또는 이러한 정보에 대한 효율적인 액세스를 용이하게 하는 데 이용될 수 있다. 각 파일이나 애플리케이션에 관련된 것과 같은 데이터와의 상호작용 또는 사용자 액티비티는, 특별히 관심을 갖는 메타데이터의 효율적인 유형을 나타내며, 많은 애플리케이션들의 초점 대상이다. 사용자가 시간에 따라 데이터를 처리하게 되면, 사용자의 그 데이터에 대한 액티비티를 감시하고 액티비티의 강도 및 유형에 따라 가중시킬 수 있다. 예를 들어, 사용자가 특정 파일에 텍스트를 자주 추가 및 제거함으로써 그 특정 파일과 크게 상호작용하면, 그 파일에 점수 또는 가중값을 메타데이터 또는 기타 포맷으로 할당하여 이러한 액티비티를 가리킬 수 있다.

[0008] 다른 예에서, 파일이 거의 상호작용되지 않는 경우(예를 들어, 1년에 한번 오픈되는 경우), 그 파일에 대한 이러한 상대적 휴지(inaction)는 보다 낮은 가중값을 할당시킬 수 있어서 사용자에게 그 파일의 중요도가 낮음을 가리킬 수 있다. 인식할 수 있듯이, 점수 또는 가중값은, 파일 사용 액티비티의 넓은 범위로 그리고 생성, 오픈, 뷰잉, 스크롤링, 편집, 인쇄, 주석달기, 세이브, 포워딩 등과 같은 복수의 서로 다른 액티비티에 대하여 할당될 수 있다. 이에 따라, 액티비티 가중 또는 패턴은 데이터 아이템들, 아이템들의 서브섹션, 또는 아이템들의 그룹들에 관련될 수 있다(예를 들어, 데이터베이스의 컬럼에 태그(tagged)될 수 있다). 이후, 예를 들어 추후에 검색 유틸리티와 같은 데이터 조작 틀에 의해 액티비티 가중을 이용하여 보다 넓은 세트의 데이터 아이템들을 보다 좁거나 보다 관리가능한 세트의 아이템들로 세분할 수 있다. 예를 들어, 콘텐츠 중심 키워드에 대하여 단순히 데이터 아이템들의 세트를 검색하는 대신, 정보 검색은 액티비티 향상된 클루(clue)를 통해 증가되어 관심을 갖는 필요한 데이터를 보다 효율적으로 탐색할 수 있다(예를 들어, 특정 사용자에게 포워딩된 모든 파일들을 찾고, 기타 애플리케이션들을 위해 가장 많이 활용된 프리젠테이션들의 서브세트를 찾으며, 최종 편집된 단락을 결정하는 등을 행한다).

[0009] 상술한 목적 및 관련 목적을 달성하기 위해, 본 명세서에서는 다음에 따르는 상세한 설명 및 첨부 도면과 관련하여 소정의 예시적인 양태들을 설명한다. 이러한 양태들은 실시될 수 있는 다양한 방식들을 가리키며, 이들 모두는 본 명세서에서 포함된다. 다른 이점들 및 신규한 특징들은 도면과 함께 고려할 때 다음에 따르는 상세한 설명으로부터 명백할 것이다.

실시예

[0019] 사용자 액티비티 또는 데이터와의 상호작용이 결정되고 이용되는, 다수의 데이터 유형들에 대한 데이터 처리를 가능하게 하는 다양한 컴포넌트들 및 프로세스들을 제공함으로써, 액티비티에 따라 데이터를 더 처리할 수 있다. 예를 들어, 액티비티 또는 상호작용은, 감시될 수 있고 이에 후속하여 추후에 데이터베이스에 상주하는 다양한 데이터 아이템들(또는 아이템 서브세트들)의 검색, 인덱싱, 분류(cataloging), 랭킹, 또는 뷰잉에 이용될 데이터 아이템에 태그될 수 있다(예를 들어, 파일 상호작용들의 액티비티에 가중값이 할당되고 데이터베이스의 컬럼에 적용됨). 특정한 일 양태에서는, 데이터 조작 시스템을 제공한다. 이 시스템은, 하나 이상의 태그에 관련되며 사용자의 데이터 아이템에 대한 적어도 하나의 상호작용을 가리키는 하나 이상의 데이터 아이템을 포함한다. 조작 틀(예를 들어, 검색 틀)은, 데이터 아이템에 대한 사용자의 상호작용에 적어도 부분적으로 기초하여 데이터 아이템들을 처리하여 데이터 아이템들의 서브세트를 결정한다.

[0020] 본 명세서에서 사용되는 바와 같이, "컴포넌트", "시스템", "태그", "감시", "모델", "문의" 등의 용어들은 컴퓨터 관련 엔티티인, 하드웨어, 하드웨어 및 소프트웨어의 조합, 소프트웨어, 또는 실행중인 소프트웨어를 참조하도록 의도된 것이다. 예를 들어, 컴포넌트는, 프로세서 상에서 실행중인 프로세스, 프로세서, 개체, 실행가능 파일, 실행 스레드, 프로그램, 및/또는 컴퓨터일 수 있지만, 이러한 예로 한정되지는 않는다. 예를 들어, 서버 상에서 실행되는 애플리케이션 및 서버 둘 다 하나의 컴포넌트일 수 있다. 하나 이상의 컴포넌트는 프로세스 및/또는 실행 스레드 내에 상주할 수 있고, 하나의 컴포넌트는 컴퓨터 상에 위치할 수 있으며 그리고/또는 2개 이상의 컴퓨터 간에 분산될 수 있다. 또한, 이러한 컴포넌트들은, 다양한 데이터 구조가 저장된 다양한 컴퓨터 판독가능 매체로부터 실행될 수 있다. 이 컴포넌트들은, 하나 이상의 데이터 패킷을 갖는 신호에 따르는 것처럼 로컬 및/또는 원격 프로세스들을 통해 통신할 수 있다(예를 들어, 하나의 컴포넌트로부터의 데이터가, 로컬 시스템, 분산형 시스템, 및/또는 그 신호에 의해 다른 시스템들과 상호작용하는 인터넷과 같은 네트워크를 통해, 다른 컴포넌트와 상호작용함).

[0021] 본 명세서에서 사용되는 바와 같이, "추론"이라는 용어는 일반적으로 이벤트 및/또는 데이터를 통해 캡처된 바

와 같은 관측들의 세트로부터 시스템, 환경, 및/또는 사용자의 상태들을 논증하거나 이성적으로 판단하는 프로세스를 가리킨다. 예를 들어, 추론을 이용하여 특정한 컨텍스트나 액션을 식별할 수 있고, 또는 상태들에 대한 확률 분포를 발생시킬 수 있다. 추론은 확률적일 수 있으며, 즉, 관심 대상인 상태들에 대한 확률 분포의 컴퓨팅이 데이터 및 이벤트의 고려 사항에 기초한다. 또한, 추론은 이벤트 및/또는 데이터의 세트로부터 보다 높은 레벨의 이벤트들을 구성하는 데 이용되는 기술들을 참조할 수 있다. 이러한 추론에 의해, 이벤트들이 시간적으로 밀접하게 상관되는지 여부에 관계없이, 그리고, 이벤트 및 데이터가 하나 또는 여러 개의 이벤트 및 데이터 소스들로부터 온 것인지에 관계없이, 관측된 이벤트 및/또는 저장된 이벤트 데이터의 세트로부터 새로운 이벤트 또는 액션의 구성이 발생한다. 게다가, 추론은 논리적 모델 또는 규칙에 기초할 수 있고, 이에 따라 컴포넌트들 또는 데이터 간의 관계를, 데이터의 분석 및 이러한 분석에 의한 결론 도출에 의해 결정하게 된다. 예를 들어, 한 사용자가 네트워크를 통해 다른 사용자들의 서버세트와 상호작용하는 것을 관측함으로써, 이 사용자들의 서버세트가, 상호작용하지 않거나 거의 상호작용하지 않는 다수의 기타 사용자들과는 달리, 그 한 사용자를 위한 관심 대상인 필요한 사회적 네트워크에 속한다는 것을, 결정 또는 추론할 수 있다.

[0022] 먼저 도 1을 참조해 보면, 시스템(100)은, 사용자 액티비티 또는 상호작용 데이터를 이용하여 다양한 컴퓨터 관련 태스크들을 수행하는 데이터 처리 아키텍처를 예시한다. 감시 컴포넌트(110)는, 하나 이상의 데이터 아이템(130)을 저장하는 하나 이상의 데이터베이스(120)와의 데이터 상호작용 오버타임을 관측한다. 감시는, 백그라운드 및/또는 포그라운드 컴포넌트(도시하지 않음)를 통해 발생할 수 있고, 데이터 아이템(130)이 사용자에 의해 상호작용되는 때를 결정하는 데 이용될 수 있다. 예를 들어, 이것은, 파일이 오픈, 클로즈, 편집, 추가, 삭제, 판독, 기입, 컷, 페이스트, 최종 편집, 포워딩, 응답, 송신, 최종 뷰잉된 때, 뷰잉한 시간량, 시간 수평축을 초과하여 상호작용한 시간, 아이템(130) 또는 이 아이템의 서로 다른 서브컴포넌트들이 얼마나 오랫동안 주의 초점 대상으로 되었는지를 가리키는 주의 주석(attentional annotation) 등을 관측하는 것을 포함할 수 있다.

[0023] 파일의 애플리케이션이 오픈되고 특정한 데이터 아이템(130)이 애플리케이션 에 작용되고 있는 경우, 다양한 기술들을 이용하여 그 애플리케이션 내의 액티비티들을 결정할 수 있다. 이것은, 사용자가 데이터의 특별한 세트 또는 서브세트에 얼마나 오랫동안 배회하는지, 어떤 데이터가 수정 또는 관측되었는지, 얼마나 자주 그리고 얼마만큼의 시간 시간동안 데이터가 작용하였는지 등을 포함할 수 있다. 태그 컴포넌트(140)는 감시한 데이터 액티비티에 따라 가중값 또는 점수를 할당한다. 이러한 가중값이나 점수는, 필요시 (또는 다른 가중 분류에 따라) 확률적으로 할당될 수 있고, 데이터 또는 애플리케이션의 소정의 부분과의 사용자의 상호작용의 패턴이나 양을 반영할 수 있다. 예를 들어, 최소 파일 사용은 파일의 광범위한 편집보다 작은 가중값을 발생시킬 수 있다. 가중값 또는 점수가 결정된 경우, 액티비티를 위해 결정된 가중 정보는 하나 이상의 데이터 아이템에 관련되거나 태그되고 액티비티 데이터(150)로서 예시되어 있다. 이러한 관련은, 예를 들어, 데이터베이스(120)의 하나의 컬럼이나 컬럼들에서의 값들을 생성하거나 수정하여 관심 대상인 메타데이터를 위한 데이터베이스(120)의 행에서 식별된 특별한 아이템의 가중 또는 중요도를 가리키는 것처럼, 데이터베이스(120) 내에서 발생할 수 있다. 컬럼들은 논리적 엔티티이며 명시적으로 저장될 수 있고 또는 사용시 동적으로 컴퓨팅될 수 있음을 인식할 것이다. 다른 유형의 관련은, 하나 이상의 데이터 아이템(130)에 직접적으로 또는 간접적으로 할당되는 메타데이터 참조를 포함할 수 있다.

[0024] 태그 및 데이터 아이템(130)에는, 태그로부터 유도된 사용자 액티비티 정보를 이용하여 정보 저장(예를 들어, 효율적인 인덱스 생성), 정보 액세스(예를 들어, 아이템의, 검색, 필터링 또는 랭킹) 및 정보 프리젠테이션(예를 들어, 아이템들의 구성, 배열, 또는 제시)을 증가시킬 수 있는 하나 이상의 데이터 조작 툴이 적용될 수 있다. 결과(170)에서, 데이터 조작 툴(160)로부터의 결과들은 자동 발생하고, 보다 넓은 데이터 아이템들(130)의 세트로부터 저장된 데이터 아이템들의 서브세트를 포함할 수 있다. 본 명세서에서 적용될 때, 서브세트라는 용어는 데이터 아이템들(130)의 일부 또는 전부를 포함할 수 있음에 유의해야 한다. 또한, 결과(170)에 데이터 아이템들(130)을 보다 많이 또는 보다 적게 포함하도록, 액티비티 임계가 툴(160)내에서 설정될 수 있다.

[0025] 다른 양태들에서는, 다큐먼트를 위해 콘텐츠 기반 점수를 할당할 수 있다. 예를 들어, 다큐먼트의 콘텐츠에 대한 사용자 문의의 유사성에 기초하여 점수를 할당할 수 있다. 따라서, 랭킹을 행할 때, 예를 들어, 편집되었거나 사용자가 판독에 많은 시간을 소비한 하나의 다큐먼트의 섹션들 또는 다큐먼트에 보이는 용어들에 보다 많은 가중을 부여한다. 시스템(100)의 다른 양태에서는, 액티비티 데이터(150)를 이용하여 관심 대상인 용어들 또는 영역들을 식별할 수 있다. 따라서, 아래와 같은 경우들에 다큐먼트 영역들의 차분 가중을 활용할 수 있다.

[0026] 1) 관심 대상의 영역들의 용어들을 우선적으로 포함하도록 인덱스를 압축하는 경우

- [0027] 2) 랭킹을 위해 관심 대상인 영역들의 용어들에 서로 다른 가중들을 부여하는 경우
- [0028] 3) 상대성 피드백을 위해 관심 대상인 영역들에 서로 다른 가중들을 부여하는 경우
- [0029] 4) 현재 사용자 초점 대상의 영역들에 기초하여 문의들을 자동적으로 또는 반자동적으로 발생시키는 경우, 및/또는
- [0030] 5 관심 대상인 용어들의 영역들 또는 용어들을 (강조 또는 다른 기술에 의해) 서로 다르게 제시하는 경우
- [0031] 본 명세서에 설명하는 시스템 및 방법은 다수의 데이터 처리 애플리케이션을 지원한다. 이것은, 예를 들어, 다큐먼트, 파일, 이메일 메시지, 달력 약속, 웹 페이지, 데이터 아이템 내의 서브섹션 또는 상호 아이템 추상과 같은 데이터 아이템들의 처리를 포함한다. 태그를 데이터 아이템에 부가함으로써, 사용자가 마지막으로 아이템에 액세스한 위치를 나타낼 수 있고, 또는 사용자가 아이템에 액세스하였거나 상호작용한 시간들의 위치 히스토리를 나타낼 수 있다. 태그는, 사용자가 마지막으로 아이템에 액세스한 때, 아이템에 액세스한 전체 횟수를 나타내고, 과거로 연장되는 시간 주기 내에서 아이템에 액세스한 빈도를 나타내며, 또는 하나 이상의 임의의 특정한 시간 주기 내에서 아이템에 액세스한 빈도를 나타낸다. 시간 경과에 따른 액세스 빈도의 보다 높은 차수의 통계를 인코딩할 수 있는 다른 컴포넌트들을 제공할 수 있다. 하나의 경우에, 뷰어는, 하나 이상의 태그의 기능들에 기초하여 사용자가 아이템들을 탐색할 수 있게 하고, 하나 이상의 태그의 기능들에 기초하여 사용자가 탐색된 아이템들을 정렬하거나 필터링할 수 있게 하며, 또는 다른 방안으로 하나 이상의 태그의 기능들에 기초하여 탐색한 아이템들을 나타낸다. 다른 경우에, 태그는 액티비티 또는 관심의 확률적 표시일 수 있다.
- [0032] 다양한 프로세스는, 데이터 아이템들을 이용한 사용자 액티비티의 분석을 포함한다. 이것은, 데이터 아이템들이 컴퓨터 사용자들로부터 수신한 상호작용의 본질 또는 양에 자동적으로 태깅하고 이러한 태그들을 이용하여 미래의 데이터 액티비티에 따라 데이터 아이템들을 더 처리하는 것을 포함한다. 프로세스는, 별도의 데이터베이스에서 데이터 아이템들에 관련된 주의 주석 내에, 또는 데이터 아이템들에 임베딩된 데이터 구조 내에, 데이터를 저장하는 것을 포함할 수 있다. 또한, 주의 또는 데이터 아이템과의 상호작용을 가리키는 주석 상태에 따라 탐색을 위해 데이터 아이템들의 서브컴포넌트들을 서로 다르게 가중하는 인덱싱 프로시저를 제공할 수 있다. 이것은, 주의 또는 데이터 아이템과의 상호작용을 가리키는 주석 상태에 따라 데이터 아이템들의 정보를 간과하거나 삭제하는 인덱싱 프로시저를 포함한다. 인덱스는, 컴퓨터 사용자들에 의해 주의되지 않았거나 상호작용되지 않은 컴포넌트들 또는 데이터 아이템들의 컴포넌트들과 덜 상호작용하거나 덜 주의된 컴포넌트들을 제거함으로써 압축될 수 있다.
- [0033] 다른 양태에서, 데이터 탐색을 위해 랭킹 점수를 활용하여, 사용자가 주의한 또는 상호작용한 데이터 아이템의 섹션들에 보이는 개체들 또는 용어들에 보다 많은 가중값을 줄 수 있다. 이것은, 주의 주석을 이용하여, 과거 및/또는 현재 내에서 주의되었거나 상호작용된 영역들에 기초하여 문의를 자동적으로 또는 반자동적으로 발생시키는 것을 포함할 수 있다. 또한, 주의 주석을 이용함으로써, 과거 및/또는 현재 내에서 주의되었거나 상호작용된 아이템들을 차별적으로 표시할 수 있고 또는 아이템들에 대하여 차별적으로 액세스할 수 있다.
- [0034] 주의 주석이 데이터 아이템들 자체에 대한 주의로서 인코딩될 수 있을 뿐만 아니라 데이터 아이템들의 서브컴포넌트들에 대한 주의로서도 인코딩될 수 있으며, 여기서 주석은 서브컴포넌트들의 각각의 다른 표시 또는 포인터 및 수신된 주의를 캡처한다는 점에 주목한다. 예를 들어, 211 페이지의 다큐먼트와 같은 큰 다큐먼트를 고려해 본다. 이 다큐먼트는 23번 상호작용 및 주의된 상태로 오픈되었을 수 있으며, 예를 들어, 그 다큐먼트를 위한 주의 주석의 한 유형으로서 캡처된다. 그러나, 다른 주의 주석들은, 예를 들어, 사용자가 4-6 페이지, 89-93 페이지, 123-124 페이지, 198 페이지를 반복적으로 검사하였고 그 다큐먼트의 나머지 페이지들에 대해선 빨리 대충 보았음을 가리킨다. 따라서, 각 서브컴포넌트에게 열거될 수 있고, 텍스트의 각 부분에 대한 주의의 양이 주석에 인코딩될 수 있다.
- [0035] 도 2는 사용자 액티비티 또는 상호작용 데이터를 활용하는 예시적인 사용자 인터페이스(200)를 도시한다. 이 예에서, 조작 툴(210)(예를 들어, 데이터베이스에 적용되는 사용자 인터페이스)은 출력 또는 디스플레이(220)에 관련될 수 있다. 툴(210)은 하나 이상의 데이터베이스로부터 데이터를 처리하기 위한 많은 특징들을 포함할 수 있다. 예를 들어, 툴(210)은, 데이터 검색의 활성화, 데이터의 인덱싱 또는 분류, 데이터 랭킹 등을 위한 선택들을 포함할 수 있다. 이러한 데이터는 예를 들어 XML 데이터 또는 ASCII 데이터와 같은 텍스트 데이터를 포함할 수 있다. 다른 데이터는, 예를 들어, 이미지 데이터, 오디오 데이터, 비디오 데이터, 그래픽 데이터, 및/또는 일련의 슬라이드 내에 포함된 것과 같은 프리젠테이션 데이터를 포함한다. 스프레드 시트, 유니버설 리소스 로케이션(URL) 정보, 인터넷, 또는 웹 데이터 등을 포함하는, 실질적으로 어떠한 데이터 유형 또는 애플리케이션

선도 이용할 수 있다. 도 1에 대하여 언급한 바와 같이, 이러한 데이터는 컬럼에 또는 파일 메타데이터로서 태깅되어 지난 사용 또는 상호작용을 나타내는 점수 또는 가중값을 가리킬 수 있다. 이에 따라, 조작 톨은 태깅된 데이터를 검색, 탐색, 또는 처리하여 사용자를 위한 보다 관리가능한 데이터의 서브세트들을 개선하거나 결정할 수 있다.

[0036] 톨(210)로부터의 출력(220)은 파일 또는 실제 사용자 인터페이스 디스플레이될 수 있다. 예를 들어, 톨이 데이터베이스 내의 검색 엔진으로서 이용되면, 그 출력은 리턴된 결과들의 디스플레이될 수 있다. 리턴된 정보는 블록(230)에서 예시된 바와 같이 그 본질상 보다 글로벌할 수 있다. 이것은, 하나의 파일 또는 파일들의 그룹이 사용자와의 액티비티가 증가되었기 때문에 선택되었음을 가리키도록 파일에 그래픽을 적용하거나 강조하는 것을 포함할 수 있다. 예를 들어 이메일 검색에 있어서, 톨(210)은, 키워드 컴퓨터를 갖고 지난 달에 파일에 관련된 적어도 하나의 그래픽 이미지를 갖고 있는 모든 파일에 대한 검색에 적용될 수 있다. 검색은 다수의 방식으로 세밀하게 행해질 수 있으며 콘텐츠 검색, 액티비티 기반 검색, 및/또는 이들의 조합을 포함할 수 있다. 예를 들어, 이 예에서, 10개의 이메일의 세트 중에서 3개의 이메일이, 다른 색으로 묘사되는 나머지 7개의 이메일보다 높은 액티비티 점수를 갖는 것처럼 하나의 색으로 강조될 수 있다. 블록(240)에서의 다른 양태에서는, 리턴된 파일 또는 데이터 세트 내의 정보(예를 들어, 다큐먼트 내의 사용 영역들을 강조하기 위해 다른 폰트 포맷으로 선택된 파일 내의 단락)가 강조되거나 주석이 부가되어 사용 액티비티를 가리킬 수 있다.

[0037] 도 3은 사용자 액티비티 또는 상호작용 데이터를 결정하고 적용하기 위한 프로세스(300)를 도시한다. 설명의 편의상, 일련의 액션들 또는 다수의 액션들로 방법들이 도시되어 설명되고 있지만, 주요 프로세스는 액션들의 순서에 의해 한정되지 않으며, 일부 액션들은 주요 프로세스에 따라 본 명세서에서 도시하고 설명하는 것으로부터의 다른 액션들과 서로 다른 순서들로 그리고/또는 동시에 발생할 수 있음을 이해 및 인식할 것이다. 예를 들어, 방법론이 다른 방안으로 상태도에서처럼 일련의 상호연관 상태들 또는 이벤트들로서 제시될 수 있다는 점을 당업자라면 이해 및 인식할 것이다. 게다가, 주요 프로세스에 따라 방법론을 구현하는 데 도시한 모든 액션들이 필요하지 않을 수 있다.

[0038] 블록(310)으로 진행하게 되면, 하나 이상의 로컬 또는 원격 데이터베이스와의 데이터 상호작용을 감시한다. 이러한 감시는, 백그라운드 및/또는 포그라운드 애플리케이션에서 발생할 수 있고, 데이터 또는 파일이 사용자에 의해 상호작용되는 때를 결정하는 데 이용된다. 예를 들어, 이것은, 파일들이 오픈, 클로즈, 편집, 추가, 삭제, 판독, 또는 기입되는 등의 때를 관측하는 것을 포함할 수 있다. 애플리케이션이 오픈되고 애플리케이션 내의 특정한 데이터가 동작하게 되면, 다양한 기술들을 이용하여 그 애플리케이션 내의 액티비티를 결정할 수 있다. 이러한 기술들은, 사용자가 데이터의 특정한 세트 또는 데이터의 서브세트에 얼마나 오랫동안 배회하는지, 어떤 데이터가 수정되었거나 관측되었는지 등을 감시하는 것을 포함할 수 있다. 또한, 외부 감시는, 데이터, 파일, 또는 애플리케이션에 대한 사용자의 상호작용을 결정하도록 애플리케이션에 관련될 수 있다. 예를 들어, 오디오 큐(cue), 자동화된 안면 인식 기술들, 또는 데이터 세트가 사용자에게 깊게 관련되어 있는 명시적 사용자 명령어를 이용할 수 있다.

[0039] 블록(320)에서, 감시한 액티비티에 대한 가중을 결정한다. 가중은, 확률적으로 할당될 수 있고, 주어진 데이터 또는 애플리케이션 부분과의 사용자의 상호작용의 양을 반영할 수 있다. 예를 들어, 다큐먼트를 매우 가볍게 속독하는 것은, 그 다큐먼트를 광범위하게 편집하는 것보다 낮은 가중값을 발생시킬 수 있다. 블록(330)에서, 액티비티를 위한 결정된 가중 정보는 하나 이상의 데이터 아이템에 관련되거나 태깅된다. 이러한 관련은, 예를 들어, 데이터베이스의 하나의 컬럼 또는 컬럼들을 태깅하여 그 데이터베이스의 행에서 식별된 특정 아이템의 가중이나 중요성을 가리키는 것처럼, 데이터베이스의 한정 내에서 발생할 수 있다. 다른 유형 관련은, 하나 이상의 데이터 아이템에 직접적으로 또는 간접적으로 할당되는 메타데이터 기준을 포함할 수 있다. 블록(340)에서, 태그 및 데이터 아이템은 데이터 처리 애플리케이션에 적용될 수 있다. 이것은, 태그들로부터 유도된 액티비티 정보를 활용하여, 아이템, 인덱스 아이템에 대한 검색을 증가시키고, 아이템들을 배열하며, 아이템들을 랭킹하고, 아이템들을 구성하는 등을 포함할 수 있다. 블록(350)에서, 데이터 처리 애플리케이션으로부터 결과들이 발생한다. 이것은, 보다 넓은 결과 세트를 보다 작은 서브세트로 필터링하는 것과 같은 명시적 액션들, 또는 디스플레이에 주석을 달아 사용자에게 의해 더 많이 상호작용된 그러한 아이템들을 가리키는 디스플레이 상의 파일들이나 데이터를 강조하는 것과 같은 보다 미세한 액션들을 포함할 수 있다.

[0040] 도 4를 참조해 보면, 시스템(400)은, 사용자 액티비티 데이터 처리와 함께 이용될 수 있는 정보 탐색 아키텍처를 예시하고 있다. 시스템(100)은 검색 결과들을 개인화하기 위한 일반적인 도면을 도시하고 있지만, 데이터 조작의 다른 형태를 전술한 바와 같이 수행할 수 있다. 개인화 컴포넌트(410)는, 사용자 액티비티에 기초하는 사용자 모델(420), 및 문의(430)를 수정함으로써 그리고/또는 검색으로부터 리턴된 결과(440)를 수정함으로써

검색 결과에 영향을 끼치는 데 그 사용자 모델을 이용하기 위한 처리 컴포넌트(예를 들어, 사용자 모델에 따라 수정된 탐색 알고리즘)를 포함한다. 사용자 인터페이스(450)는 문의(430)를 발생시키고, 개인화된 컴포넌트(410)에 의해 제공되는 문의 수정(470) 및 결과 수정(460)에 기초하여 수정되거나 개인화된 결과를 수신한다. 본 명세서에서 활용될 때, "문의 수정"이라는 용어는, 개인화된 결과(440)를 얻기 위해 문의(430)에서의 용어들에 대한 변경 및 문의(430)를 다큐먼트에 매칭시키는 알고리즘에서의 변경 둘다를 가리킨다. 수정된 문의 및/또는 결과(440)는 하나 이상의 로컬 검색 엔진 및/또는 원격 검색 엔진(480)으로부터 리턴된다. 사용자 통계의 글로벌 데이터베이스(490)는 사용자 모델(420)의 갱신이 용이하도록 유지될 수 있다. 인식할 수 있듯이, 사용자 모델(420) 및/또는 글로벌 통계(490)는, 데이터 조작이나 처리가 용이하도록 전문한 바와 같이 사용자 액티비티 또는 상호작용 데이터에 관련될 수 있다.

[0041] 일반적으로, 사용자 모델(420)에 기초하여 검색 결과에 적응하는 적어도 2개의 방식이 존재한다. 일 양태에서, 문의 수정은 초기 입력 문의를 처리하고, (사용자 모델을 통해) 그 문의를 수정하거나 재생하여 개인화된 결과를 출력한다. 관련성 피드백은 이 프로세스의 2-사이클 변동이며, 여기서 문의는 결과들을 발생시켜 (초기 결과들의 세트에 대한 명시적 또는 암시적 판정을 이용하여) 문의가 수정되게 하며, 이것은 문의 및 결과 세트에 기초하여 단(short) 용어 모델에 대하여 개인화된 결과를 출력하게 된다. 관련성 피드백의 컨텍스트에서 장(longer) 용어 사용자 모델을 이용할 수도 있다. 게다가, 문의 수정은, 문의를 다큐먼트에 매칭시키는 데 이용되는 알고리즘에 행해진 변경도 가리킨다. 다른 양태에서, 결과 수정은 사용자의 입력을 그대로 취하여 결과를 출력하는 문의를 발생시키고, 이후 이 결과는 (사용자 모델을 통해) 수정되어 개인화된 결과를 발생시키게 된다. 결과들의 수정은, 일반적으로 결정된 액티비티 데이터로부터의 가중 또는 고려 사항을 포함할 수 있는 보다 넓은 세트의 대체 방안들의 리랭킹(reranking) 및/또는 선택의 일부 형태를 포함한다는 것에 주목한다. 또한, 결과들의 수정은 결과들의 서브세트 또는 전부의 다양한 유형의 응집 및 요약을 포함한다.

[0042] 결과를 수정하기 위한 방법은, (사용자의 관심 및 콘텐츠가 벡터로 표현되고 아이템에 매칭되는) 통계적 유사성 매칭, 및 (보다 작은 세트의 설명어들을 이용하여 사용자의 관심 및 콘텐츠가 아이템으로 표현되고 아이템에 매칭되는) 카테고리 매칭을 포함한다. 문의 수정 또는 결과 수정의 이러한 프로세스들은, 독립적으로 조합될 수 있고, 또는 2개의 프로세스들 간에 의존성들이 도입되고 조절되는 집적된 프로세스에서 조합될 수 있다.

[0043] 도 5를 참조해 보면, 사용자 액티비티 또는 상호작용 데이터와 함께 이용될 수 있는 사용자 모델(500)이 예시되어 있다. 사용자 모델(500)은, 일반화된 탐색들로부터 개인화된 탐색들을 차별화하고 결정된 액티비티 데이터에 따라 리치(rich) 데이터 처리를 용이하게 하는 데 이용된다. 성공적인 개인화의 일 양태는, 사용자의 관심을 정밀하게 반영하고 유지가 쉽고 장 용어 및 단 용어 관심에 대한 변경에 적응하는 사용자의 모델을 만드는 것이다. 사용자 모델은, 아래와 같이 다양한 소스들로부터 얻을 수 있지만, 이러한 예로 한정되지는 않는다.

[0044] 1) 로컬 소스, 모바일 소스, 또는 원격 소스로부터 얻을 수 있는, 블록(510)에서 컴퓨팅 컨텍스트의 리치 히스토리 (예를 들어, 오픈된 애플리케이션, 이러한 애플리케이션의 콘텐츠, 위치를 포함하는 이러한 상호작용의 상세한 히스토리)

[0045] 2) 블록(520)에서 이전에 발생한 콘텐츠의 리치 인덱스 (예를 들어, 다큐먼트, 웹 페이지, 이메일, 인스턴트 메시지, 노트, 달력 약속 등)

[0046] 3) 블록(530)에서 키워드, 구성 차트에서의 관계, 약속 등으로부터 유도된 관심 대상인 최근의 또는 빈번한 컨택트, 토픽을 비롯한 클라이언트 상호작용의 감시

[0047] 4) 블록(540)에서 이전 검색 문의들의 히스토리를 포함하여 방문한 로컬/원격 데이터 사이트들 또는 이전 웹 페이지들의 로그 또는 히스토리

[0048] 5) 블록(550)에서 백그라운드 감시를 통해 유도되어 명시적으로 또는 암시적으로 특정될 수 있는 사용자 관심 프로파일

[0049] 6) 블록(560)에서 인구통계적 정보 (예를 들어, 위치, 장르, 연령, 백그라운드, 직업 카테고리 등)

[0050] 전문한 예들로부터, 사용자 모델(500)이 서로 다른 많은 정보 소스들에 기초할 수 있다는 것을 인식할 수 있다. 예를 들어, 사용자 모델(500)은, GPS와 같은 장치에 의해 감시될 때, 시간 경과에 따라 사용자가 방문한 위치들의 로그 또는 히스토리로부터 소스를 제공받을 수 있다. GPS로 감시할 때, 미가공 공간적 정보를, 텍스트로 된 도시명, 및 우편 번호로 변환할 수 있다. 미가공 공간적 정보는, 예를 들어, 사용자가 일시 정지하였거나 배회하였거나 GPS 신호 손실을 발생시킨 위치들에 대한 우편 번호 및 텍스트로 된 도시명으로 변환될 수 있다. 사용자가 일시 정지하였거나 배회하였거나 GPS 신호 손실을 발생시킨 위치들은, 식별될 수 있고, 관심 대상 포인

트들 및 비즈니스 데이터베이스를 통해 텍스트 라벨들로 변환될 수 있다. 다른 요인들은, 위치 및 관심 대상 포인트들을 결정하기 위해 그 날의 시간 또는 요일을 로깅하는 것을 포함한다.

[0051] 다른 양태들에서, 컴포넌트들은, 사용자의, 정보, 약속, 다큐먼트 또는 파일의 뷰잉, 액티비티, 또는 위치의 코퍼스(corpus)가 서브세트들로 그룹화될 수 있거나 유형, 연령, 또는 기타 조합에 기초하여 개인화를 위한 매칭 프로시저에서 서로 다르게 가중될 수 있는 방식을 제어하기 위한 파라미터들을 조작하는 데 제공될 수 있다. 예를 들어, 탐색 알고리즘은, 문의에 속하는 사용자 전부의 그러한 양태들로 제한될 수 있다(예를 들어, 문의 용어 또는 데이터와의 지난 상호작용을 포함하는 다큐먼트). 마찬가지로, 이메일은 이전 1개월로부터 분석될 수 있는 반면, 웹 액세스는 이전 3일로부터 분석될 수 있고, 사용자의 콘텐츠는 작년 내에 생성될 수 있다. 오늘날 또는 다른 시간 기간에서 GPS 위치 정보를 이용하는 것이 바람직할 수 있다. 파라미터들은 (예를 들어, 파라미터들을 변경하고 사용자나 시스템으로부터의 응답을 테스트하는 최적화 프로세스를 통해) 서브세트들을 생성하도록 자동 조작될 수 있고, 또는 사용자는 사용자 인터페이스를 통해 이러한 파라미터들 중 하나 이상의 변경할 수 있고, 여기서 이러한 설정은, 문의 본질의 기능, 그 날의 시간, 요일, 또는 기타 컨텍스트나 액티비티 기반 관측일 수 있다.

[0052] 모델은, 블록(570)에서 개별인들 또는 개별인들의 그룹들의 유사성 분석에 의해 프로파일을 전개하는 협력적 필터링 기술에 의해, 개별인이나 개별인들의 그룹으로부터 유도될 수 있다. 유사성 컴퓨팅은 콘텐츠 및/또는 아이템의 사용에 기초할 수 있다. 인프라스트럭처의 모델링 및 관련 처리는, 클라이언트, 다수의 클라이언트, 하나 이상의 서버, 또는 서버와 클라이언트의 조합에 의존할 수 있음에 주목한다.

[0053] 블록(580)에서, 머신 학습 기술은, 데이터가 사용자에 의해 상호작용되는 방식 및 시기 뿐만 아니라 시간 경과에 따른 학습 사용자 특성들 및 관심들에 적용될 수 있다. 학습 모델은, 예를 들어, 통계적/수학적 모델과 같은 실질적으로 임의의 유형의 시스템, 및 사용자를 모델링하고 베이스(Bayesian) 학습의 이용을 포함한 관심과 선호를 결정하기 위한 프로세스를 포함할 수 있고, 이 베이스 학습은, 베이스 네트워크와 같은 베이스 의존성 모델, 순수한(naive) 베이스 분류자, 및/또는 지원 벡터 머신(SVM)을 비롯한 기타 통계적 분류 방법론을 발생시킬 수 있다. 다른 유형의 모델 또는 시스템은 예를 들어 신경 네트워크 및 히든 마르코프 모델(Hidden Markov Model)을 포함할 수 있다. 복잡한 이성적 모델을 이용할 수 있지만, 다른 방식들도 활용할 수 있음을 인식할 것이다. 예를 들어, 보다 철저한 확률적 방식 보다는, 결정적 가정을 이용할 수도 있다(예를 들어, 특정 웹 사이트의 시간량 Z에 대하여 최근의 검색이 없다는 것은 규칙에 의하면 사용자가 각 정보에 더 이상 관심을 갖지 않는다는 것을 의미할 수 있다. 따라서, 불확실성에 의한 추론에 더하여, 사용자의, 상태, 위치, 컨텍스트, 관심, 집중 등에 관하여 논리적 결정을 내릴 수도 있다.

[0054] 학습 모델은, 서로 다른 데이터 소스들로부터 데이터를 수집하거나 응집하는 사용자 이벤트 데이터 저장소(도시하지 않음)로부터 훈련받을 수 있다. 이러한 소스들의 예로는, 사용자 이벤트 데이터를 기록하거나 로깅하는 다양한 데이터 획득 컴포넌트들(예를 들어, 핸드폰, 마이크에 의해 기록되는 음향 액티비티, GPS, 전자 달력, 비전 감시 장비, 데스크탑 액티비티, 웹 사이트 상호작용 등)을 포함할 수 있다. 시스템들은 개인 문의 및 결과 처리를 지원하는 실질적으로 임의의 방식으로 구현될 수 있음에 주목한다. 예를 들어, 시스템은, 클라이언트 애플리케이션 내에서, 서버, 서버 팜(farm)으로서 구현될 수 있고, 또는 사용자 인터페이스 및 검색 엔진처럼 검색 기능과 상호작용하는 웹 서비스 또는 기타 자동화 애플리케이션을 포함하도록 보다 일반화될 수 있다.

[0055] 진행하기 전에, 블록(570)에서 적용된 사용자 모델(500)의 협력적 필터 기술들을 보다 상세히 설명한다. 이 기술들은, 데이터를 분석하고 사용자를 위한 프로파일을 결정하는 데 협력적 필터의 이용을 포함한다. 협력적 필터링 시스템은 일반적으로 사용자 선호에 대하여 집중화된 데이터베이스를 이용하여 사용자가 원할 수 있는 추가 토픽들을 예측한다. 협력적 필터링을 사용자 모델(500)에 적용함으로써, 시스템의 새로운 사용자들을 위해 발생할 가능성이 있는 프로파일들을 예측하는 소정의 사용자에 대한 선호를 가리킬 수 있는, 사용자들의 그룹으로부터의 이전의 사용자 액티비티를 처리한다. 상관 계수, 벡터 기반 유사성 계산, 통계적 베이스 방법에 기초하는 기술들을 포함하는 여러 알고리즘들을 이용할 수 있다.

[0056] 도 6을 참조해 보면, 시스템(600)은 사용자 액티비티 또는 상호작용 데이터에 따라 액세스 기반 정보 탐색을 예시하고 있다. 시스템(600)은 사용자에 의해 액세스되거나 고려되는 하나 이상의 정보 소스(610)를 포함한다. 이러한 소스들(610)은 유사할 수 있으며 또는 다른 정보 콘텐츠를 갖는 본질에 있어서 개별적일 수 있으며, 이에 따라, 이 정보 소스들 중 일부는, 파일, 폴더, 애플리케이션, 이미지, 오디오 파일, 약속, 이메일 등과 같은 로컬 데이터 위치들을 나타낼 수 있고, 다른 소스들(610)은 예를 들어 웹 정보와 같은 원격 소스들을 나타낼 수 있다. 시간 경과에 따라 사용자가 서로 다른 유형들의 정보에 액세스하게 되면, 사용 분석기(614)는, 이 정보

를, 클라이언트 머신 상에서 로컬로 그리고/또는 서버에 따라 원격으로 백그라운드 태스크로서 동작할 수 있는 콘텐츠 분석기(620)(또는 모니터)에 전달하고, 데이터로부터 콘텐츠를 파싱하기 위한 필터(624) 및 액세스된 데이터 아이템들의 콘텐츠 인덱스(640)(또는 액티비티 태그)를 생성하는 자동 인덱서(630)를 통해 액세스된 데이터를 처리한다.

[0057] 일반적으로, 분석기(620)는 인덱스(640)에서 액세스된 데이터의 리프리젠테이션을 생성한다. 예를 들어, 사용자가 웹 페이지에 액세스한 경우, 콘텐츠 분석기(620)는 그 웹 페이지의 썸네일 리프리젠테이션을 생성할 수 있고 페이지 및 썸네일에 대한 하이퍼링크 참조를 메타데이터 파일의 일부로서 관련지을 수 있다. 리프리젠테이션은 페이지와의 사용자 상호작용의 상세한 패턴들을 반영하도록 더 구성될 수 있다. 다른 경우에, 사용자가 내부에 이미지들을 갖는 텍스트 다큐먼트에 액세스한 경우, 분석기(620)는 텍스트 또는 이 텍스트의 부분들을 추출할 수 있고, 파일 경로와 같은 데이터베이스 링크를 메타데이터의 일부로서 관련지을 수 있다. 이후, 인덱서(630)는 콘텐츠 인덱스(640)에서 2개의 아이템을 갖는 인덱스를 자동 생성(또는 기존의 인덱스에 추가)할 수 있으며, 썸네일 리프리젠테이션 및 텍스트 다큐먼트 리프리젠테이션은 메타데이터를 포함한다. 일반적으로, 필터는 아이템에 관련된 메타데이터 및 아이템의 콘텐츠를 분석한다. 따라서, 워드(Word) 다큐먼트에 있어서, 예를 들어, 필터(624)는, 다큐먼트에서의 단어들과 함께 파일명, 타이틀, 저자, 키워드, 생성일 등과 같은 메타데이터를 추출한다. 이것은 인덱스(640)를 구축하는 데 사용되는 것이다. 필요하다면, 썸네일의 생성 및 이미지의 분석도 필터(624)내에 포함될 수 있다. 콘텐츠 인덱스(640)에 저장된 아이템들을 설명하는 암시적 태그들 및/또는 사용자와 같은 기타 아이템들을 포함할 수 있는 액티비티 또는 상호작용 메타데이터를 이용할 수 있다. 인덱서(630)가 필터(624) 기능을 수행할 수도 있음을 인식할 것이다(예를 들어, 인덱서가 메타데이터를 필터링된 콘텐츠와 관련짓는다).

[0058] 콘텐츠 인덱스(640)에 포함된 정보 아이템에 대한 사용자 문의(654)를 수신하는 검색 컴포넌트(650)를 제공한다. 검색 컴포넌트(650)는, 링크 및/또는 리프리젠테이션(660)에서 문의(654)에 응답하여 액세스된 아이템들의 링크 및/또는 리프리젠테이션을 사용자에게 리턴하는 사용자 인터페이스의 일부로서 제공될 수 있다. 예를 들어, 사용자는, "작년의 성능 리뷰에 관한 아이템들"을 문의할 수 있고, 여기서 검색 컴포넌트(650)는, 인덱스(640)로부터, 문의(654)의 컨텍스트에 관련된 이메일, 동료 평가, 지난 해 발행된 다큐먼트, 웹 페이지 이미지, 오디오 리코딩 등과 같은 아이템들을 추출한다. 다른 예에서, 암시적 문의는 문의(654)로부터 유도될 수 있다(예를 들어, 사용자로부터 전화를 받을 때마다, 이 사용자로부터 최종 5개의 이메일을 추출한다).

[0059] 이하에서 보다 상세히 설명하듯이, 액세스된 아이템들은, 이전에 액세스된 정보 아이템들의 효율적이고 시의적절한 탐색이 용이하도록 설계된 복수의 서로 다른 포맷들로 제시될 수 있다. 또한, 링크 및/또는 리프리젠테이션(660)은, 이전에 액세스된 아이템들이 아니라 사용자가 보길 원할 수 있는 정보 아이템들을 제공하는 것처럼 사용자에게 대하여 관심 대상인 다른 아이템들을 포함할 수 있다(예를 들어, 시스템은, 현재의 문의에 기초하는 또는 현재의 문의로부터 추론되는 관심 대상인 다른 콘텐츠에 대한 링크를 제공하며, 예를 들어, 성능 리뷰 아이템들을 보여주는 것에 더하여, 이러한 아이템들이 사용자에게 의해 이전에 액세스되었을 수 있거나 액세스되지 않았을 수 있더라도, 콘텐츠의 다른 인덱스에 기초하여 리뷰 정책을 설명하는 인력 자원들에 대한 링크를 선택 사항으로서 제공한다).

[0060] 일 양태에서는, (예를 들어, 사용 분석기(614)에 관련된 사용자 액티비티를 감시하는 백그라운드 태스크인) 이벤트 컴포넌트(도시하지 않음)를 제공할 수 있다. 이벤트 컴포넌트는, 예를 들어, 세이브, 관독, 편집, 복사, 정보 배회, 정보 선택, 정보 조작 및/또는 파일 삭제와 같은 사용자 액티비티를 감시하고, 사용자 액션에 대하여 결정을 내린다. 이것은, 마이크, 카메라, 및 사용자 액션이나 목적을 결정하는 감시 데스크탑 액티비티를 갖는 기타 장치와 같은 센서를 포함할 수 있다. 일 예에서, 확률적 모델 및/또는 논리적 결정을 적용하여, 사용자가 정보를 관측한 때 또는 고려한 때와 같은 이벤트를 결정할 수 있다. 논리적 및/또는 통계적 모델(예를 들어, 베이스 의존 모델(Bayesian dependency model), 결정 트리, 지원 벡터 머신)은, 사용자 액티비티의 패턴들에 관련된 다음에 따르는 증거 클래스들의 예를 고려하도록 구성될 수 있다.

[0061] - 주의 초점 : 다큐먼트 전체에 걸친 스크롤링 후 다큐먼트의 부분들 또는 특정 서브텍스트에서 배회하는 아이템 배회 및/또는 아이템의 선택

[0062] - 내성(Introspection) : 액티비티의 주기 후 일시 정지 또는 상호작용의 비율의 상당한 저감

[0063] - 불필요한 정보 : 간략히 본 후 다큐먼트를 즉각 클로즈함으로써, 정보 액세스 액션 후 종래 상태로 리턴하려 한다. 이러한 관측은 최근 액션의 영향을 취소하는 것, 커맨드를 생성하고 취소하는 것, 아이템을 삭제하는 것을 포함한다.

- [0064] - 도메인 특정 구문 및 시맨틱 콘텐츠 : 다큐먼트의 콘텐츠 또는 구조에서의 특별한 구별 및 사용자가 이러한 특징이나 아이템과 상호작용하는 방식을 고려한다. 이것은, 태스크에 관련된 도메인 특정 특징들을 포함한다 (예를 들어, 사용자의 주의 초점에서 볼 때, 메시지의 저자로부터, 제목 주제(subject heading)의 메시지들의 개수 또는 시간의 시기, 이메일 메시지들의 비율 및 빈도를 고려함). 인식할 수 있듯이, 이벤트 컴포넌트는, 사용자 액티비티에 기초하여 다양한 유형들의 정보를 인덱싱하는 데 이용될 수 있다. 정보 개체에 대한 사용자의 액티비티도 활용하여 정보 프리젠테이션을 개선할 수 있다.
- [0065] 도 7은 다양한 탐색 서비스 애플리케이션(700)을 예시한다. 일 양태에서는, 명시적 문의(710) 및/또는 암시적 문의(714)를 지원할 수 있다. 명시적 문의(710)는, 사용자에 의해 관심 대상인 정보를 찾으라고 지시받는다(예를 들어, 회의 또는 날짜에 관한 모든 데이터 참조를 도시한다). 암시적 문의(714)는 일부 경우에 명시적 문의(710)로부터 유도될 수 있다. 예를 들어, 사용자는 자신의 데스크탑 폰 메시지를 자신의 이메일 시스템 또는 기타 메시지 시스템에 링크할 수 있다. 선택된 개별인으로부터 전화가 오는 경우, 이메일 시스템은 암시적 문의(714)를 통해 그 개별인에 대한 이메일을 자동 탐색할 수 있다. 다른 예에서, 다가올 회의에 앞서 소정의 간격으로, 사용자의 달력 시스템이 문의를 트리거하여 과거 회의로부터의 데이터 또는 다가올 회의에 참석하는 개별인들에 관한 정보를 리콜할 수 있다. 또한, 암시적 문의(714)는 사용자의 현재 컨텍스트 또는 문의(예를 들어, 최근에 관독된 단락들에서의 중요한 단어들로 구성된 문의)에 관련된 이성적 프로세스에 기초하여 발생할 수 있다.
- [0066] 블록(716)으로 진행하게 되면, 다른 유형의 문의는 컨텍스트 민감 문의를 지원한다. 이러한 유형의 문의는, 검색을 편집하거나 세분하는 추가 선택 옵션을 제공하는 것을 포함한다. 예를 들어, 문의는 특정 유형의 애플리케이션 또는 위치에 대하여 지시된 것일 수 있다(예를 들어, 이 문의를 메일 폴더에만 적용한다). 블록(720)에서, 문의를 수행할 때 애플리케이션의 컨텍스트를 고려할 수 있다. 예를 들어, 포토 애플리케이션이 사용중이라면, 문의는 이미지들만을 검색하도록 세분될 수 있다. 블록(724)에서, 아이템 집중식 통합을 수행할 수 있다. 이것은, 마우스 클릭 기능, 아이템 태깅, 메타데이터 파일 갱신, 아이템 삭제, 아이템이나 콘텐츠 편집 등과 같은 인터페이스 액션들을 지원하는 운영 체제 액션들을 포함한다.
- [0067] 블록(730)에서는, 파일 공유를 실행할 수 있다. 예를 들어, 사용자는 하나 이상의 다른 사용자들이 자신들의 문의/인덱스 데이터베이스의 서브세트 또는 모두를 검사하거나 액세스를 가질 수 있다는 점을 특정할 수 있다 (예를 들어, 그 사용자의 프로젝트 팀의 모든 사용자들이 그 사용자의 프로젝트 노트들에 액세스할 수 있다). 블록(734)에서는, 인덱스 스크러빙이 발생할 수 있다. 시간 경과에 따라, 사용자들은 자신들의 인덱스로부터 하나 이상의 아이템을 제거하길 원할 수 있다. 이러한 액티비티에 따라, 사용자들은 특정항 아이템들을 특정하여, 시스템에 의해 자동으로 스크러빙될 수 있는 일반적인 토픽 영역들을 제거하거나 특정할 수 있다(예를 들어, 2년 전 사용자의 생일에 관련된 썸네일을 제거한다). 소정 주기의 소정 횟수보다 적게 아이템이 액세스되고 이후 필요시 자동 제거될 수 있는 것처럼, 논리적 또는 이성적 프로세스들에 기초하여 다른 액션들이 발생할 수 있다.
- [0068] 블록(740)에서는, 효율적인 시간 컴퓨팅을 고려한다. 일 예로서, (사용자에 대한 데이터 제시 동안) 파일에 관련된 또는 파일에 대하여 유용한 날짜는 변경된 날짜이고, 메일을 제시하기 위한 날짜는 일반적으로 전달된 (이에 따라 대략 사용자가 그 메일을 본 때의) 날짜이며, 약속을 위한 유용한 날짜는 약속이 발생하는 날짜이다. 모든 시간 정보가 기록되고 인덱싱되며 정보 제시를 위해 유용한 날짜 정보가 활용될 수 있다는 점에 주목한다. 따라서, 약속에 있어서는, 예를 들어, 메일이 송신된 시간, (송신된 경우) 메일이 갱신된 시간, 사용자가 허용/거부한 시간, 회의가 발생한 시간을 인덱싱하는 것처럼 다양한 태스크가 발생할 수 있다. 그러나, 통상적으로 하나보다 많은 시간을 제공할 수 있지만 표시를 위해 하나의 시간이 선택된다.
- [0069] 진술한 바와 같이, 소정의 데이터는 파일 유형에 관련된 파일 요소들을 분석함으로써 이전에 관측된 것으로서 표시될 수 있다. 예를 들어, 텍스트 다큐먼트는 파일이 오픈된 때 또는 마지막으로 편집된 때를 가리키는 필드를 포함할 수 있다. 그러나, 달력 약속에 대하여, 달력이 생성된 때로부터 인덱스를 생성하는 것은 사람들에게 덜 유의할 수 있으며, 그 이유는 만남이 때때로 실제 만남일에 훨씬 앞서 생성되기 때문이다. 따라서, 달력 약속을 인덱싱하는 경우, 생성 시기에 대립되는 실제 만남일을 추적할 수 있다. 사용자는 이러한 유형의 효과적인 시간 고려 사항에 의해 메모리 리콜에 보다 적합한 방식으로 정보를 탐색할 수 있다. 단계(744)에서, 데이터의 휘발성을 고려하고 처리한다. 이러한 유형의 처리에는, 간헐적인 동작들 동안 영구적인 형태로 데이터를 인덱싱하는 것이 포함된다. 인식할 수 있듯이, 다양한 자동화 백그라운드 동작들이 가능하다.
- [0070] 도 8을 참조해 보면, 본 명세서에 설명하는 다양한 양태들을 구현하기 위한 예시적인 환경(810)이 컴퓨터(812)

를 포함하고 있다. 컴퓨터(812)는 처리 장치(814), 시스템 메모리(816), 및 시스템 버스(818)를 포함한다. 시스템 버스(818)는 시스템 메모리(816)를 비롯한 시스템 컴포넌트들을 처리 장치(814)에 결합하지만, 이러한 예로 한정되지는 않는다. 처리 장치(814)는 다양한 이용가능 프로세서들 중 임의의 것일 수 있다. 듀얼 마이크로프로세서 및 기타 멀티프로세서 아키텍처도 처리 장치(814)로서 이용할 수 있다.

[0071] 시스템 버스(818)는, 메모리 버스 또는 메모리 컨트롤러, 주변 버스나 외부 버스, 및/또는 이용가능한 버스 아키텍처들 중 임의의 것을 이용하는 로컬 버스를 포함하는 버스 구조의 몇몇 유형들 중 임의의 것일 수 있고, 예로서, 이러한 아키텍처는 11-비트 버스, ISA(industry standard architecture) 버스, MCA(micro channel architecture) 버스, EISA(Enhanced ISA) 버스, IDE(Intelligent Drive Electronics), VESA(video electronics standard association) 로컬 버스, PCI(Peripheral Component Interconnect), USB(Universal Serial Bus), AGP(Advanced Graphics Port), PCMCIA(Personal Computer Memory Card International Association bus), SCSI(Small Computer Systems Interface)를 포함하지만, 이러한 예들로 한정되지는 않는다.

[0072] 시스템 메모리(816)는, 휘발성 메모리(820) 및 비휘발성 메모리(822)를 포함한다. 기동(start-up) 동안처럼, 컴퓨터(812) 내의 요소들 간에 정보를 전달하는 기본 루틴을 포함하는 기본 입력/출력 시스템(BIOS)은, 비휘발성 메모리(822)에 저장되어 있다. 예를 들어, 비휘발성 메모리(822)는, ROM, PROM, EPROM, EEPROM, 또는 플래시 메모리를 포함할 수 있지만, 이러한 예로 한정되지는 않는다. 휘발성 메모리(820)는 외부 캐시 메모리로서 기능하는 RAM을 포함한다. 예를 들어, RAM은, SRAM, DRAM, SDRAM, DDR SDRAM, ESDRAM, SLDRAM, DRRAM과 같은 많은 형태로 이용가능하지만, 이러한 예로 한정되지는 않는다.

[0073] 컴퓨터(812)는 분리식/비분리식, 휘발성/비휘발성 컴퓨터 저장 매체도 포함한다. 도 8은, 예를 들어, 디스크 저장소(824)를 예시하고 있다. 디스크 저장소(824)는, 자기 디스크 드라이브, 플로피 디스크 드라이브, 테이프 드라이브, Jaz 드라이브, Zip 드라이브, LS-100 드라이브, 플래시 메모리 카드, 또는 메모리 스틱과 같은 장치들을 포함하지만, 이러한 예로 한정되지는 않는다. 또한, 디스크 저장소(824)는, 기타 저장 매체와는 별도로, 또는 콤팩트 디스크 ROM 장치(CD-ROM), CD 기록가능 드라이브(CD-R 드라이브), CD 재기입가능 드라이브(CD-RW 드라이브), 또는 디지털 버서타일 디스크 ROM 드라이브(DVD-ROM)와 같은 광 디스크 드라이브를 비롯한 기타 저장 매체와 조합하여 저장 매체를 포함할 수 있지만, 이러한 예로 한정되지는 않는다. 디스크 저장 장치들(824)을 시스템 버스(818)에 쉽게 접속시키고자, 분리식 또는 비휘발성 인터페이스가 통상적으로 인터페이스(826)로서 사용된다.

[0074] 도 8은 적합한 운영 환경(810)에서 설명되는 사용자와 기본 컴퓨터 리소스 간의 매개체로서 기능하는 소프트웨어를 설명하고 있음을 인식할 것이다. 이러한 소프트웨어는 운영 체제(828)를 포함한다. 운영 체제(828)는, 디스크 저장소(824)에 저장될 수 있으며, 컴퓨터 시스템(812)의 리소스들을 제어하고 할당하도록 기능한다. 시스템 애플리케이션(830)은, 시스템 메모리(816) 또는 디스크 저장소(824)에 저장된 프로그램 모듈(832) 및 프로그램 데이터(834)를 통해 운영 체제(828)에 의한 리소스들의 관리를 유익하게 한다. 본 명세서에서 설명한 다양한 컴포넌트들은 다양한 운영 체제 또는 운영 체제들의 조합으로 실현될 수 있음을 인식할 것이다.

[0075] 사용자는, 입력 장치(836)를 통해 커맨드 또는 정보를 컴퓨터(812)에 입력한다. 입력 장치(836)는, 마우스, 트랙볼, 스타일러스, 터치패드와 같은 포인팅 장치, 키보드, 마이크, 조이스틱, 게임 패드, 위성 안테나, TV 튜너 카드, 디지털 카메라, 디지털 비디오 카메라, 웹 카메라 등을 포함하지만, 이러한 예로 한정되지는 않는다. 이러한 입력 장치 및 다른 입력 장치는 인터페이스 포트(838)를 통해 시스템 버스(818)에 의해 처리 장치(814)에 접속된다. 인터넷 포트(838)는, 예를 들어, 직렬 포트, 병렬 포트, 게임 포트, 유니버설 직렬 버스(USB)를 포함한다. 출력 장치(840)는 입력 장치(836)와 동일한 유형의 포트들 중 일부를 이용한다. 따라서, 예를 들어, USB 포트는 컴퓨터(812)에 입력을 제공하고 컴퓨터(812)로부터 출력 장치(840)로 정보를 출력하는 데 사용될 수 있다. 출력 아답터(842)는, 기타 출력 장치들(840) 중에서 특별한 아답터들을 필요로 하는 모니터, 스피커, 프린터와 같은 일부 출력 장치들(840)이 존재하고 있음을 예시하고자 제공된 것이다. 출력 아답터(842)는, 예시를 위해, 출력 장치(840)와 시스템 버스(818) 간의 접속 수단을 제공하는 비디오 카드 및 사운드 카드를 포함하지만, 이러한 예로 한정되지는 않는다. 기타 장치들 및/또는 장치들의 시스템들이 원격 컴퓨터(844)처럼 입력 기능 및 출력 기능 둘 다를 제공한다는 점에 주목하길 바란다.

[0076] 컴퓨터(812)는 원격 컴퓨터(844)처럼 하나 이상의 원격 컴퓨터에 대한 논리적 접속부를 이용하여 네트워크화된 환경에서 동작할 수 있다. 원격 컴퓨터(844)는 퍼스널 컴퓨터, 서버, 라우터, 네트워크 PC, 워크스테이션, 마이크로프로세서 기반 제품, 피어 장치, 또는 기타 공통 네트워크 노드 등일 수 있으며, 통상적으로 컴퓨터(812)에 대하여 설명되는 요소들의 상당수 또는 전부를 포함한다. 간략히 설명하고자, 원격 컴퓨터(844)에 메모리

저장 장치(846)만을 예시하고 있다. 원격 컴퓨터(844)는 네트워크 인터페이스(848)를 통해 컴퓨터(812)에 논리적으로 접속된 후 통신 접속부(850)를 통해 물리적으로 접속된다. 네트워크 인터페이스(848)는 LAN 및 WAN과 같은 통신 네트워크들을 포함한다. LAN 기술들은, 광섬유 분산 데이터 인터페이스(FDDI), 구리 분산 데이터 인터페이스(CDDI), 이더넷/IEEE 802.3, 토큰 링/IEEE 802.5 등을 포함한다. WAN 기술들은, 점대점 링크, 종합 정보 통신망(ISDN)과 같은 회로 스위칭 네트워크 및 그 변형, 패킷 스위칭 네트워크, 디지털 가입자 회선(DSL)을 포함하지만, 이러한 예들로 한정되지는 않는다.

[0077] 통신 접속부(850)는 네트워크 인터페이스(848)를 버스(818)에 접속하는 데 이용되는 하드웨어/소프트웨어를 참조한다. 통신 접속부(850)가 예시를 위해 컴퓨터(812) 내에 있는 것으로 도시되어 있지만, 컴퓨터(812) 외부에 있을 수도 있다. 네트워크 인터페이스(848)로의 접속에 필요한 하드웨어/소프트웨어는, 단지 예시를 위해, 전화선급 모뎀(regular telephone grade modem), 케이블 모뎀, DSL 모뎀을 비롯한 모뎀, ISDN 아답터, 및 이더넷 카드와 같이 내부 및 외부 기술들을 포함한다.

[0078] 도 9는 이용할 수 있는 샘플 컴퓨팅 환경(900)의 개략적인 블록도이다. 이 시스템(900)은 하나 이상의 클라이언트(910)를 포함한다. 클라이언트(910)는 하드웨어 및/또는 소프트웨어일 수 있다. 또한, 시스템(900)은, 하나 이상의 서버(930)를 포함한다. 서버(930)도 하드웨어 및/또는 소프트웨어(예를 들어, 스레드, 프로세스, 컴퓨팅 장치)일 수 있다. 서버(930)는 스레드를 수용하여, 예를 들어 본 명세서에서 설명하는 컴포넌트들을 이용함으로써 변환을 수행할 수 있다. 클라이언트(910)와 서버(930) 간의 한 가지 가능한 통신은 2개 이상의 컴퓨터 프로세스 간에 송신되도록 구성된 데이터 패킷의 형태를 취할 수 있다. 시스템(900)은, 클라이언트(910)와 서버(930) 간의 통신을 용이하게 하는 데 이용될 수 있는 통신 프레임워크(950)를 포함한다. 클라이언트(910)는, 클라이언트(910)에게 로컬인 정보를 저장하는 데 이용될 수 있는 하나 이상의 클라이언트 데이터 저장소(960)에 동작가능하게 접속된다. 마찬가지로, 서버(930)는, 서버(930)에게 로컬인 정보를 저장하는 데 이용될 수 있는 하나 이상의 서버 데이터 저장소(940)에 동작가능하게 접속된다.

[0079] 전술한 바에서는 다양한 예시적인 양태들을 포함하고 있다. 물론, 이러한 양태들을 설명하기 위해 생각할 수 있는 모든 조합들 또는 방법들을 설명하는 것은 가능하지 않지만, 당업자라면 더 많은 조합들 및 순열들이 가능하다는 점을 인식할 수 있다. 이에 따라, 본 명세서에서 설명하는 양태들은 청구범위의 사상 및 범위 내에 속하는 이러한 모든 변경, 수정, 변형을 포함하고자 한다. 게다가, "포함한다(includes)"라는 용어가 상세한 설명 또는 청구범위에서 사용된다는 점에서, 이러한 용어는, "포함한다(comprising)"가 청구항에서 과도적인 단어로 사용될 때 해석되기 때문에 "포함한다(comprising)"는 용어와 유사한 방식으로 포괄적인 의미로 사용된다.

도면의 간단한 설명

[0010] 도 1은 사용자 액티비티 또는 상호작용 데이터를 이용하는 데이터 처리 시스템을 예시하는 개략적인 블록도이다.

[0011] 도 2는 사용자 액티비티 또는 상호작용 데이터를 활용하는 예시적인 사용자 인터페이스를 도시한다.

[0012] 도 3은 사용자 액티비티 결정 및 프로세스를 예시하는 흐름도이다.

[0013] 도 4는 사용자 액티비티 데이터 처리와 함께 이용될 수 있는 정보 탐색 아키텍처 시스템의 일 예를 도시한다.

[0014] 도 5는 사용자 액티비티 또는 상호작용 데이터와 함께 이용될 수 있는 사용자 모델의 일 예를 도시한다.

[0015] 도 6은 사용자 액티비티 또는 상호작용 데이터에 따라 액세스 기반 정보 탐색을 행하는 시스템을 도시한다.

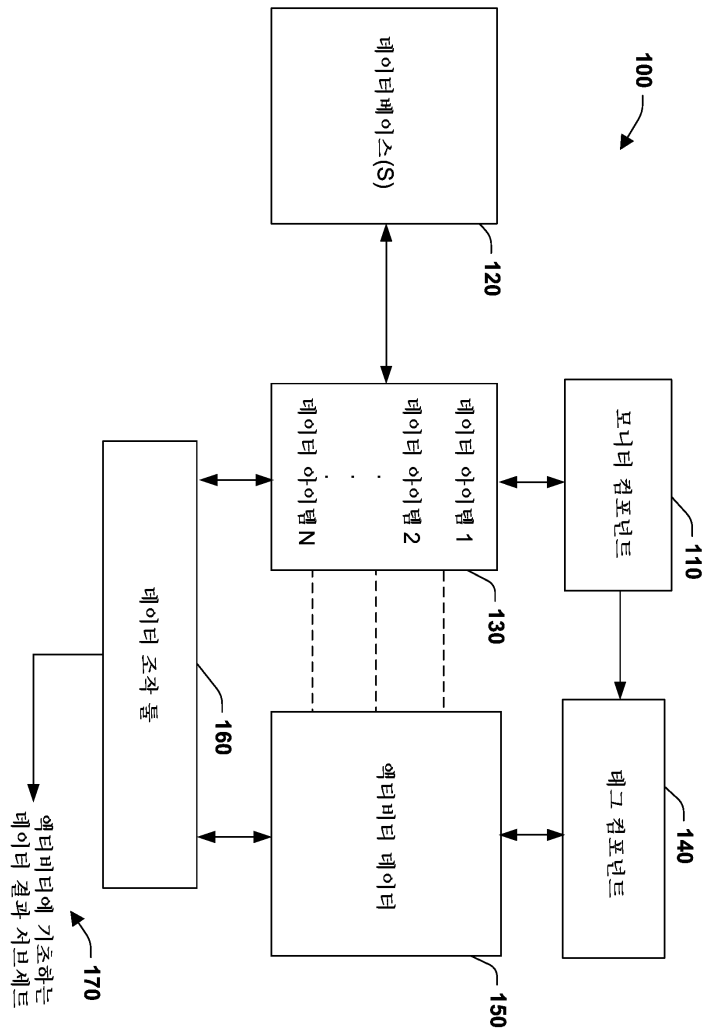
[0016] 도 7은 사용자 액티비티 또는 상호작용 데이터와 함께 이용될 수 있는 탐색 서비스 애플리케이션을 도시한다.

[0017] 도 8은 적합한 운영 환경을 예시하는 개략적인 블록도이다.

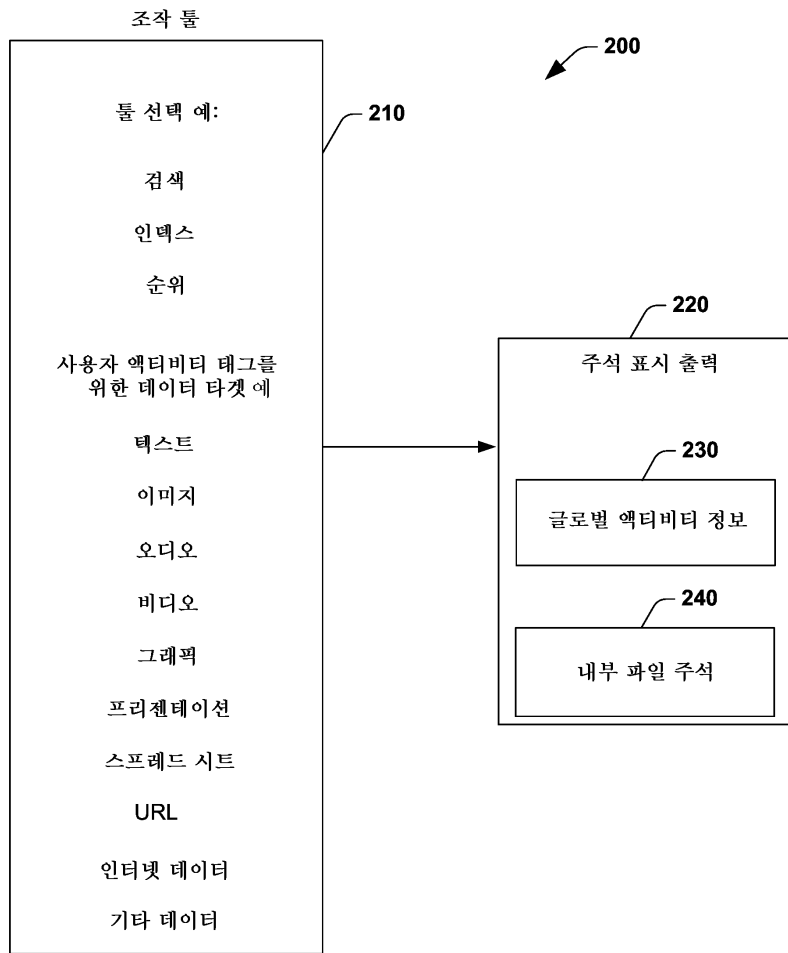
[0018] 도 9는 샘플 컴퓨팅 환경의 개략적인 블록도이다.

도면

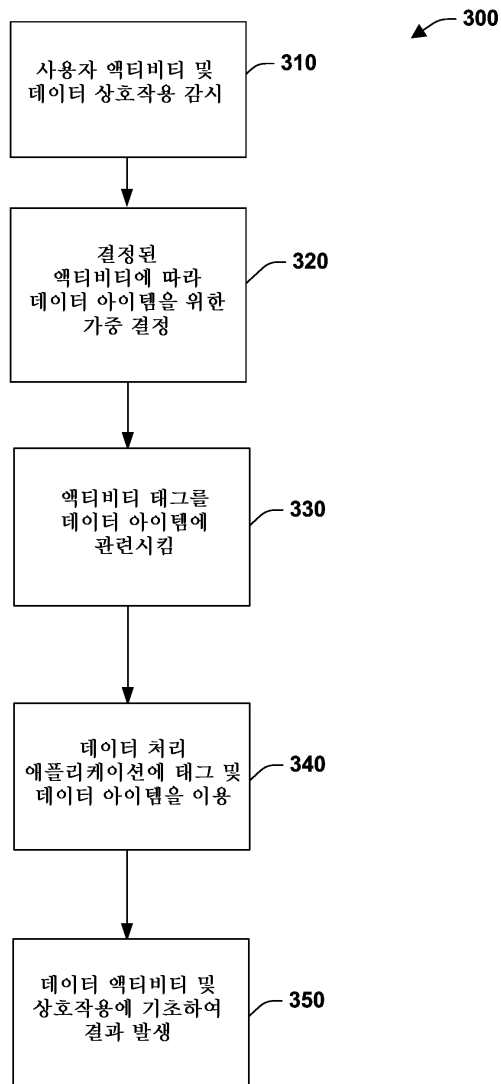
도면1



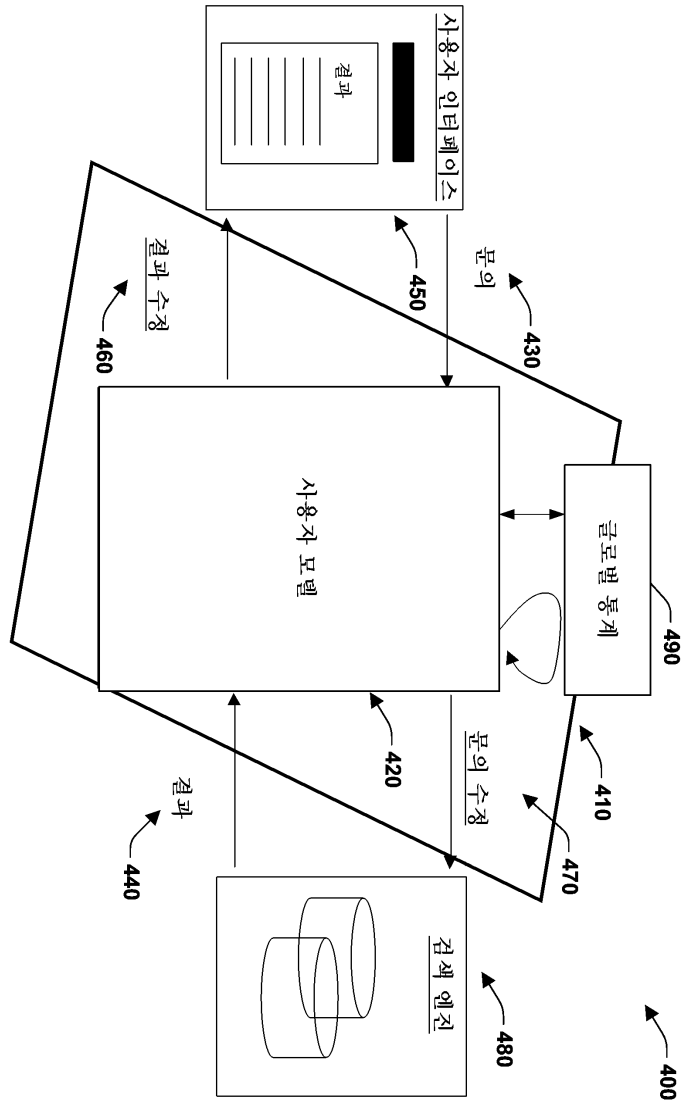
도면2



도면3



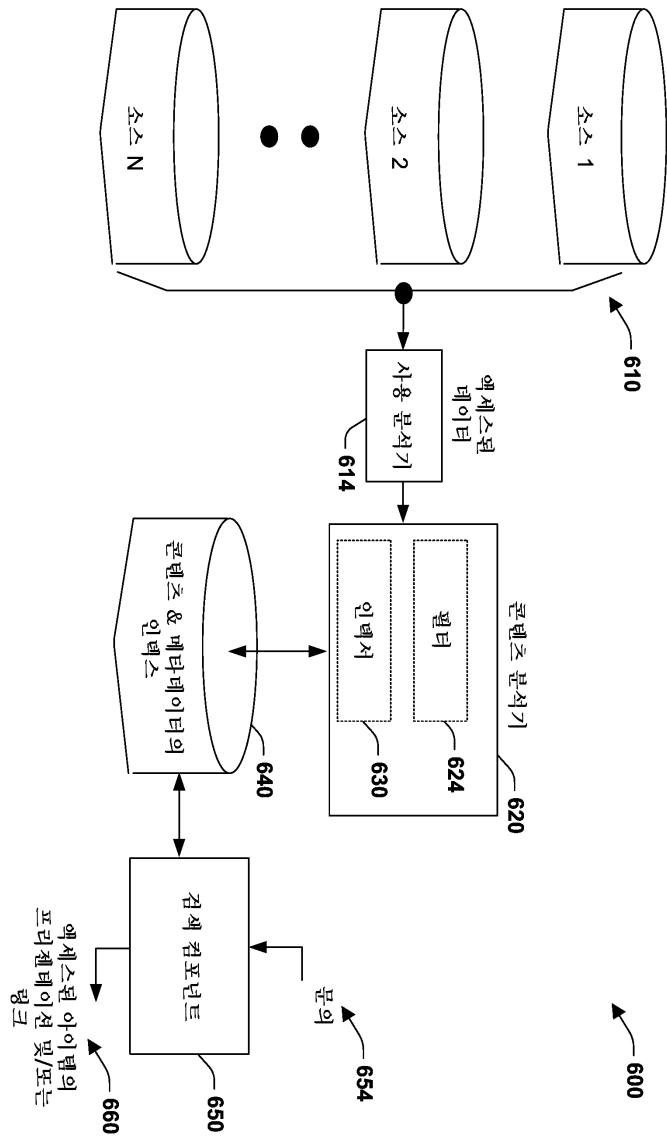
도면4



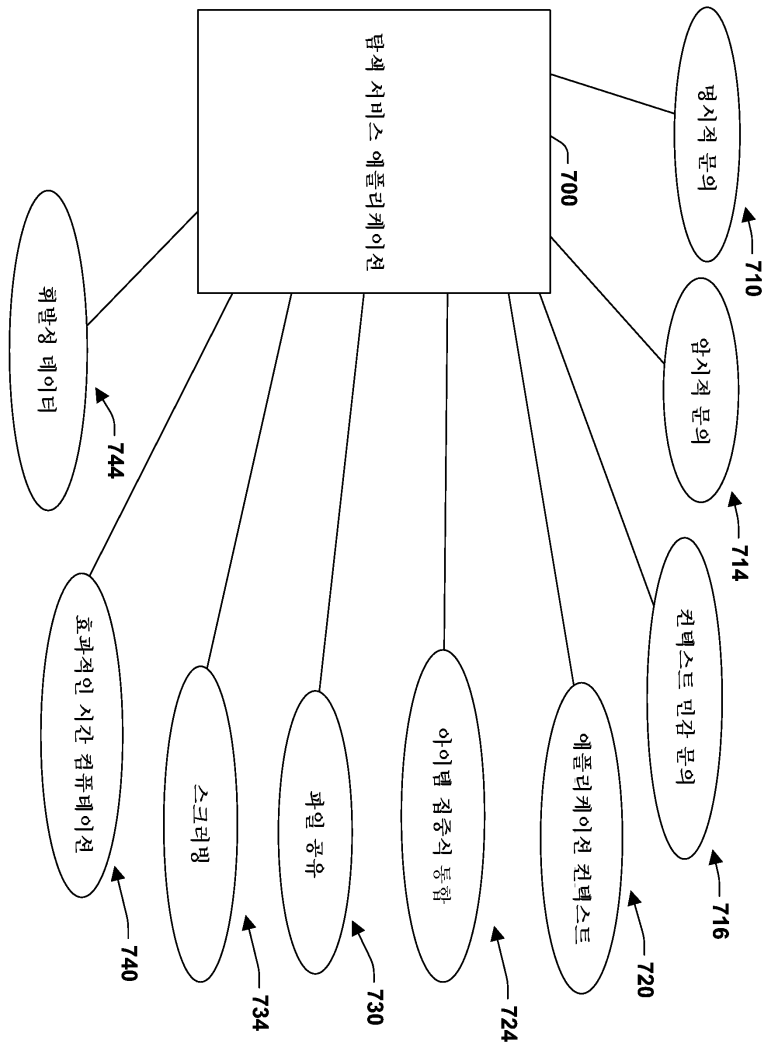
도면5

사용자 모델 & 데이터		500
컨텍스트	510	
리치 인덱스	520	
리치 클라이언트 상호작용	530	
검색 히스토리/로그	540	
관심 대상인 프로파일	550	
인구통계	560	
기타 사용자	570	
기계 학습	580	

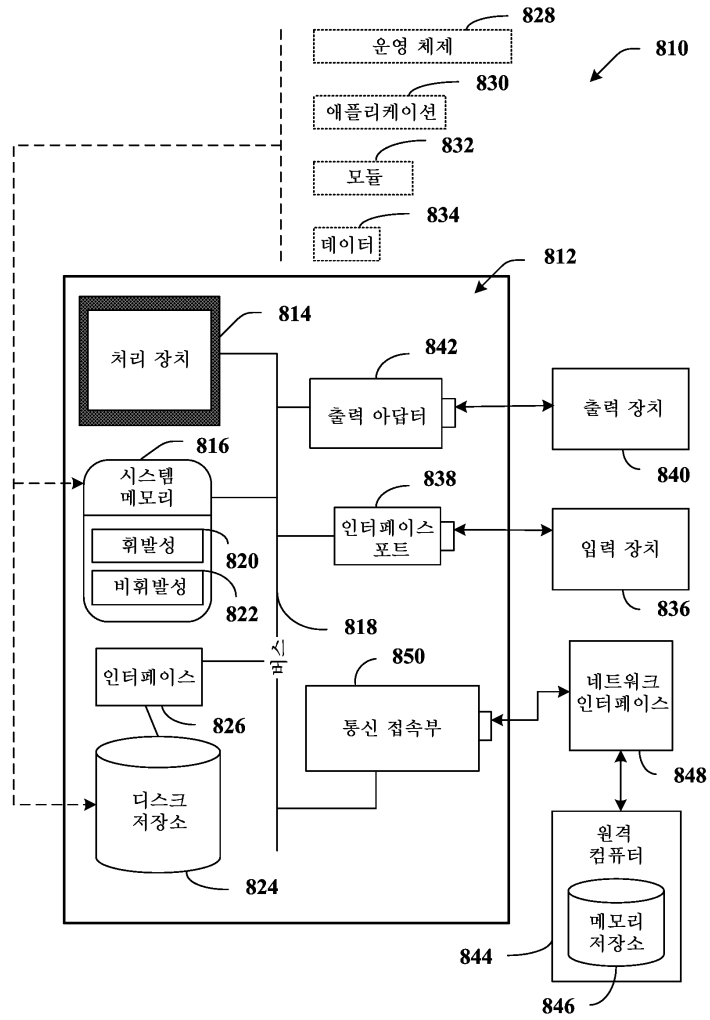
도면6



도면7



도면8



도면9

