



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0162284
(43) 공개일자 2024년11월15일

- | | |
|--|---|
| <p>(51) 국제특허분류(Int. Cl.)
 <i>G06Q 50/26</i> (2024.01) <i>G06F 16/951</i> (2019.01)
 <i>G06F 17/18</i> (2006.01) <i>G06F 17/40</i> (2006.01)
 <i>G06N 3/08</i> (2023.01) <i>G06Q 10/06</i> (2012.01)</p> <p>(52) CPC특허분류
 <i>G06Q 50/26</i> (2024.01)
 <i>G06F 16/951</i> (2019.01)</p> <p>(21) 출원번호 10-2023-0059141
 (22) 출원일자 2023년05월08일
 심사청구일자 2023년05월08일</p> | <p>(71) 출원인
 국립한밭대학교 산학협력단
 대전광역시 유성구 동서대로 125 (덕명동)</p> <p>(72) 발명자
 박성욱
 대전광역시 유성구 동서대로 725, 어울림하트@1212-1804</p> <p>(74) 대리인
 이은철, 이우영</p> |
|--|---|

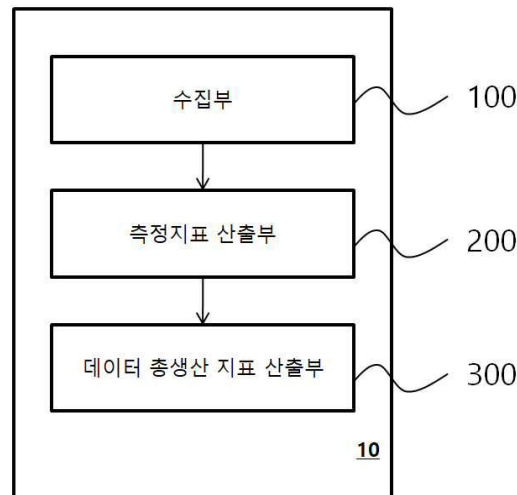
전체 청구항 수 : 총 4 항

(54) 발명의 명칭 **데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템**

(57) 요약

본 발명은 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템에 관한 것으로써, 데이터 경제 측면에서 새로운 경제 지표를 산출하기 위한 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템에 관한 것이다. 본 발명은, 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집하는 수집부, 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출하되, 각 측정지표마다 중요도에 따라 가중치를 부여하는 측정지표 산출부 및 가중치가 부여된 각 측정지표를 기반으로 새로운 데이터총생산(New GDP) 지표를 산출하여 다양한 측면에서 데이터 경제를 평가하는 지표를 고려하여 산출한 데이터 경제 측면의 새로운 경제 측정지표는 데이터 경제의 상황을 더욱 정확하게 파악할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06F 17/18 (2013.01)

G06F 17/40 (2020.08)

G06N 3/08 (2023.01)

G06Q 10/063 (2023.01)

명세서

청구범위

청구항 1

국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집하는 수집부,

상기 수집부를 통해 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출하되, 각 해당 측정지표마다 중요도에 따라 가중치를 부여하는 측정지표 산출부 및

상기 가중치가 부여된 각 측정지표를 기반으로 데이터총생산 지표를 산출하는 데이터총생산 지표 산출부를 포함하는 것을 특징으로 하는 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템.

청구항 2

제1항에 있어서,

상기 측정지표 산출부는

상기 수집부를 통해 수집한 데이터의 크기와 해당 데이터의 중요도에 따라 가중치를 부여하여 Volume 지표를 산출하는 데이터의 양 측정지표 산출모듈,

상기 수집한 데이터와 관련하여 사용자가 어떻게 데이터를 사용하는지 측정하기 위한 Usage 지표를 산출하는 사용자 수 측정지표 산출모듈,

상기 수집한 데이터와 관련하여 데이터가 얼마나 접근 가능한지를 평가하는 Accessibility 지표를 산출하는 데이터 접근성 측정지표 산출모듈,

상기 수집한 데이터와 관련하여 디지털 활동의 복잡성을 측정하는 요소로 1인당 광대역 소비량을 측정하여 Complexity 지표를 산출하기 위한 데이터 복잡성 측정지표 산출모듈 및

상기 수집한 데이터의 개방성을 평가하기 위해 데이터의 접근성 확인, 라이선싱 확인, 데이터의 재사용성 확인, 데이터의 품질 확인, 메타데이터 제공, 개인정보 보호 확인의 세부지표로 Openness 지표를 산출하는 데이터 개방성 측정지표 산출모듈을 포함하는 것을 특징으로 하는 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템.

청구항 3

제1항에 있어서,

상기 데이터 총생산 지표 산출부는 측정지표 산출부의 Volume, Usage, Accessibility, Complexity, Openness 지표들의 각 중요도에 따라 가중치를 부여하여 새로운 데이터총생산(New GDP)을 산출하는 것을 특징으로 하는 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템.

(수학적식)

$$\text{New GDP} = (w1 * \text{Volume}) + (w2 * \text{Usage}) + (w3 * \text{Accessibility}) + (w4 * \text{Complexity}) + (w5 * \text{Openness})$$

w1, w2, w3, w4, w5는 각각 Volume, Usage, Accessibility, Complexity, Openness 지표의 중요도에 따른 가중치

청구항 4

데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템을 이용한 방법에 있어서,

상기 시스템이 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집하는 단계,

상기 시스템이 수집한 데이터의 크기와 해당 데이터의 중요도에 따라 가중치를 부여하여 Volume 지표를 산출하

는 단계,

상기 시스템이 상기 수집한 데이터와 관련하여 사용자가 어떻게 데이터를 사용하는지 측정하기 위한 Usage 지표를 산출하는 단계,

상기 시스템이 상기 수집한 데이터와 관련하여 데이터가 얼마나 접근 가능한지를 평가하는 Accessibility 지표를 산출하는 단계,

상기 시스템이 상기 수집한 데이터와 관련하여 디지털 활동의 복잡성을 측정하는 요소로 1인당 광대역 소비량을 측정하여 Complexity 지표를 산출하는 단계,

상기 시스템이 상기 수집한 데이터의 개방성을 평가하기 위해 데이터의 접근성 확인, 라이선싱 확인, 데이터의 재사용성 확인, 데이터의 품질 확인, 메타데이터 제공, 개인정보 보호 확인지표로 중요도에 따라 가중치를 부여하여 Openness 지표를 산출하는 단계 및

상기 시스템이 Volume, Usage, Accessibility, Complexity, Openness 지표들의 각 중요도에 따라 가중치를 부여하여 새로운 데이터총생산(New GDP)을 산출하는 단계를 포함하는 것을 특징으로 하는 방법.

발명의 설명

기술 분야

[0001] 본 발명은 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템에 관한 것으로서, 보다 상세하게는 데이터 경제 측면에서 새로운 경제 지표를 산출하기 위한 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템에 관한 것이다.

배경 기술

[0002] 데이터 경제란, 데이터와 알고리즘을 이용하여 새로운 가치를 창출하고 이를 경제적 활동으로 이어나가는 것을 의미한다. 이를 위해서는 데이터의 수집, 가공, 분석, 활용 등의 과정이 필요하다. 이러한 과정을 위해서는 데이터 처리 기술인 빅데이터, 인공지능, 클라우드 컴퓨팅 등이 필요하다.

[0003] 빅데이터는 기존 데이터 처리 기술로는 처리할 수 없는 규모와 다양성을 갖는 대규모의 데이터를 다루는 기술이다. 인공지능은 기계 학습, 자연어 처리, 이미지 분석 등을 통해 데이터에서 지식을 추출하고 의사결정을 내리는 기술이다. 클라우드 컴퓨팅은 인터넷을 통해 서버, 스토리지, 데이터베이스, 네트워크, 소프트웨어 등의 컴퓨팅 자원을 제공하는 기술이다.

[0004] 이러한 기술들은 데이터 경제의 중요한 요소로 인식되어, 데이터 경제 관련 산업의 성장과 더불어 새로운 기술이 필요하다.

[0005] 한편, 제4차 산업혁명시대의 새로운 경제 패러다임에 맞는 새로운 경제지표가 필요하다. 데이터 경제가 2030년까지 13조달러 규모로 글로벌 경제활동의 새로운 패러다임을 가져올 것으로 맥킨지는 예측하고 있다. 이에 제4차 산업혁명시대에 맞는 새로운 경제 패러다임으로 기존의 국내총생산(GDP, Gross Domestic Product)을 새로운 데이터총생산(New GDP, New Gross Data Product)으로 지표를 구성할 필요가 있다.

선행기술문헌

특허문헌

[0006] (특허문헌 0001) 등록특허 10-2041915(2019.11.01.)

발명의 내용

해결하려는 과제

[0007] 본 발명은 데이터 경제 측면에서 새로운 경제 지표를 산출하기 위한 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템을 제공함에 목적이 있다.

과제의 해결 수단

[0008] 상기 목적을 달성하기 위하여 본 발명은, 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집하는 수집부, 상기 수집부를 통해 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출하되, 각 측정지표마다 중요도에 따라 가중치를 부여하는 측정지표 산출부 및 상기 가중치가 부여된 각 측정지표를 기반으로 새로운 데이터총생산 지표 산출하는 데이터총생산 지표 산출부를 포함한다.

발명의 효과

[0009] 본 발명에 따르면, 다양한 측면에서 데이터 경제를 평가하는 지표를 고려하여 산출한 데이터 경제 측면의 새로운 경제 지표는 데이터 경제의 상황을 더욱 정확하게 파악할 수 있다. 이를 통해 국가별 데이터 경제의 상황을 비교하고 분석함으로써, 데이터 경제 분야의 개선 방향을 찾을 수 있다. 또한, 이러한 경제 지표를 활용하여 적극적인 데이터 활용 방안을 모색하고, 데이터 경제 생태계를 발전시키는 데 기여할 수 있는 효과가 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템의 전체 구성도이다.

도 2는 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템의 측정지표 산출부의 세부구성을 설명하기 위한 도면이다.

도 3은 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템을 이용한 방법을 나타낸 개략적인 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하, 첨부된 도면들에 기재된 내용들을 참조하여 본 발명을 상세히 설명한다. 다만, 본 발명이 예시적 실시 예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일 참조부호는 실질적으로 동일한 기능을 수행하는 부재를 나타낸다.

[0012] 본 발명의 목적 및 효과는 하기의 설명에 의해서 자연스럽게 이해되거나 보다 분명해질 수 있으며, 하기의 기재만으로 본 발명의 목적 및 효과가 제한되는 것은 아니다. 또한, 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이, 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명을 생략하기로 한다.

[0013] 도 1은 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템의 전체 구성도이다.

[0014] 도 1에 도시된 바와 같이, 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산의 측정지표 개발을 위한 시스템(10)은 수집부(100), 측정지표 산출부(200), 데이터총생산 지표 산출부(300)를 포함한다.

[0015] 수집부(100)는 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집한다.

[0016] 측정지표 산출부(200)는 수집부를 통해 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출하되, 각 측정지표마다 중요도에 따라 가중치를 부여하는 구성이다.

[0017] 이러한 측정지표 산출부(200)는 수집부를 통해 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출을 위한 분석과정에서, 데이터 분석시 인공지능망을 이용하여 데이터를 학습하고 예측하는 방식으로 하여 각 측정 지표별 중요도에 따라 해당 측정 지표를 산출한다.

[0018] 새로운 데이터 총생산 지표 산출부(300)는 가중치가 부여된 각 측정지표를 기반으로 데이터총생산 지표를 산출하기 위한 구성이다. 여기서, 새로운 데이터 총생산 지표 산출부(300)라 표현하였지만, 데이터 총생산 지표 산출부를 나타낸다.

[0019] 각 구성에 대해 세부적으로 설명하면 다음과 같다.

[0020] 본 실시예에 따른 수집부(100)는 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터,

광대역 인터넷 사용량 데이터를 수집하며, 국가에서 생성된 모든 데이터를 추적하고 수집하는 것은 현재 기술적으로 불가능하므로 이는 비현실적인 목표이나, 일부 데이터는 수집이 가능하며 이를 기반으로 대략적인 추정이 가능하다.

- [0021] 예를 들어, 한국에서는 통계청이 국가통계포털(KOSIS)을 운영하고 있다. KOSIS는 국가통계정보의 종합적인 수집, 편람화, 분석, 제공 등을 담당하는 국가대표 통계포털로서, 국가에서 생성된 많은 데이터들을 수집하고 있다. KOSIS에서는 경제, 사회, 문화, 교육, 환경 등 다양한 분야에서 발생하는 데이터를 수집하고 있다. 예를 들어 경제 분야에서는 국내총생산, 생산량, 수출입 데이터 등을 수집하고, 사회 분야에서는 인구통계, 건강, 복지 등의 데이터를 수집하고 있다. 이렇게 수집된 데이터들은 Volume 지표를 산출하는 데 사용될 수 있다.
- [0022] 그러나, KOSIS에서 수집하는 데이터는 모든 데이터가 아니므로, 이러한 데이터를 수집하는 외부 기관들의 데이터도 함께 고려해야 한다. 예를 들어, 대한민국에서는 금융감독원, 고용노동부, 환경부 등 다양한 기관들이 각자의 분야에서 데이터를 수집하고 있다. 이러한 외부 데이터를 모두 수집하여 Volume 지표를 계산하는 것은 현재로서는 어렵지만, 이러한 외부 데이터들을 함께 고려함으로써 보다 정확한 추정이 가능하다.
- [0023] 도 2는 본 발명의 일 실시예에 따른 데이터 경제 시대 데이터총생산의 측정지표 개발을 위한 시스템의 측정지표 산출부의 세부구성을 설명하기 위한 도면이다.
- [0024] 도 2에 도시된 바와 같이, 측정지표 산출부(200)는 수집부를 통해 수집한 데이터로 데이터의 양, 사용자 수, 데이터 접근성, 데이터 복잡성, 데이터 개방성의 측정지표를 산출하되, 각 측정지표마다 중요도에 따라 가중치를 부여하는 구성으로서, 데이터의 양 측정지표 산출모듈(210), 사용자 수 측정지표 산출모듈(220), 데이터 접근성 측정지표 산출모듈(230), 데이터 복잡성 측정지표 산출모듈(240), 데이터 개방성 측정지표 산출모듈(250)을 포함한다.
- [0025] 본 실시예에 따른 다섯가지 측정지표는 Volume 지표, Usage 지표, Accessibility 지표, Complexity 지표, Openness 지표로 분류할 수 있다. 여기서, 각 측정 지표는 다음과 같다.
- [0026] Volume(양) : 데이터의 양
- [0027] Usage(사용자) : 데이터를 사용하는 사용자 수
- [0028] Accessibility(접근성) : 데이터에 접근할 수 있는 사용자 수
- [0029] Complexity(복잡성) : 데이터의 복잡성
- [0030] Openness(개방성) : 데이터에 대한 개방성
- [0031] 이렇게 가중치를 부여하여 각 측정 지표를 고려한 데이터 경제 측면의 새로운 경제 지표를 산출할 수 있다. 이러한 지표는 국가의 데이터 경제 발전 상황을 평가하고, 데이터 경제 산업을 발전시키는 데 도움을 줄 수 있다.
- [0032] 이러한 각 측정지표를 산출하기 위한 측정지표 산출부의 세부구성들에 대해 아래에서 설명하기로 한다.
- [0033] 데이터의 양 측정지표 산출모듈(210)은 데이터총생산 지표를 산출하기 위해 국가별 생성된 데이터의 양을 측정하기 위한 데이터 양 측정지표를 산출한다.
- [0034] 데이터의 양에 따른 Volume 지표를 계산하기 위해서는 수집부를 통해 수집한 데이터를 토대로 다음과 같은 과정이 필요하다:
- [0035] 데이터 수집: 먼저 국가에서 생성된 모든 데이터를 추적하고 수집해야 한다. 이를 위해 데이터 수집 도구나 시스템을 구축하거나 이미 구축된 데이터베이스를 활용할 수 있다. 예를 들어, 국가에서 발행된 모든 뉴스 기사, 공식 보고서, 인구통계 등을 수집할 수 있다.
- [0036] 데이터 정제: 수집된 데이터는 대부분 다양한 형태와 포맷으로 존재하므로, 일관된 형태로 정제해야 한다. 예를 들어, 뉴스 기사의 경우 HTML 태그를 제거하고 텍스트만 추출해야 한다. 또한, 데이터의 중복을 제거하거나 결측치를 처리해야 한다.
- [0037] 데이터 저장: 정제된 데이터를 데이터베이스나 스프레드시트 등의 저장 매체에 저장해야 한다. 이를 통해 데이터의 구조를 보다 효과적으로 관리하고 분석할 수 있다.
- [0038] 데이터 분석: 저장된 데이터를 바탕으로 Volume 지표를 계산할 수 있다. 이를 위해 데이터베이스나 스프레드시트에서 적절한 집계 함수를 사용하여 데이터를 분석한다. 예를 들어, 뉴스 기사의 경우 일별 또는 월별 기사 수

를 집계하여 Volume 지표를 계산할 수 있다.

- [0039] 본 실시예에 따른 데이터의 양 측정지표 산출모듈(210)은 데이터 분석시 인공지능망을 이용하여 데이터를 학습하고 예측하는 방식으로 하여 수집 데이터를 학습하고 일별 또는 월별 등의 특정 선택기간별 데이터 패턴으로 Volume 지표를 분석한다.
- [0040] 결과 시각화: 계산된 Volume 지표를 그래프나 차트 등으로 시각화하여 보다 직관적으로 파악할 수 있다. 이를 통해 데이터의 변화 추세나 패턴 등을 쉽게 파악할 수 있다.
- [0041] 본 실시예에 따른 데이터의 양 측정지표 산출모듈(210)은 국가에서 발행된 모든 뉴스 기사, 공식 보고서, 인구통계 등을 수집하여 Volume 지표를 계산하는 경우, 다음과 같은 계산식을 사용할 수 있다.
- [0042]
$$\text{Volume} = \sum (\text{size} * \text{weight})$$
- [0043] 여기서, size는 수집된 데이터의 크기를 의미하고, weight는 해당 데이터의 중요도를 나타내는 가중치이다. 이 계산식에서 $\sum$ 는 모든 수집된 데이터에 대한 합을 의미한다.
- [0044] 뉴스 기사의 경우, 일별 또는 월별 기사 수를 size로 사용하고, 해당 기사의 중요도에 따라 가중치를 부여할 수 있다. 이 가중치는 다른 측정지표에 대해서도 적용될 수 있다.
- [0045] 인구통계 데이터의 경우, 인구수와 같은 수치적 데이터의 크기를 size로 사용하고, 해당 데이터의 중요도에 따라 가중치를 부여할 수 있다.
- [0046] 예를 들어, 한 국가에서 일별로 발행된 뉴스 기사의 수가 1000개, 1500개, 1200개, 1800개, 2000개였다면 이를 평균 내어 1500개 정도의 일별 기사 발행량으로 계산할 수 있다. 마찬가지로 월별로도 발행된 기사 수를 집계하여 월별 기사 발행량을 계산할 수 있다. 따라서, Volume 지표를 계산하는데 있어서는 일별 또는 월별 기사 수를 집계한 후, 일별 또는 월별 평균값을 계산하는 과정이 필요하다. 이를 수식으로 나타내면 다음과 같다.
- [0047] 일별 기사 발행량 = (각 날짜별 기사 수의 합) / (총 기간)
- [0048] 월별 기사 발행량 = (각 월별 기사 수의 합) / (총 기간)
- [0049] Volume 지표의 중요도에 대한 가중치 정하기에 대해 설명하면 다음과 같다.
- [0050] Volume 지표는 데이터의 양에 대한 지표로, 데이터의 양이 많을수록 분석의 정확성과 유의성을 높일 수 있다. 따라서 데이터 총생산을 계산하는 데 있어 Volume 지표 역시 중요한 역할을 한다. 일반적으로 데이터의 양은 매우 중요한 요소 중 하나로 간주된다. 데이터 양이 많을수록 해당 데이터로부터 유용한 정보를 추출할 가능성이 높아지기 때문이다. 그러나 데이터의 양이 많을수록 처리, 분석, 저장 및 관리 비용이 더 많이 들어간다.
- [0051] 따라서 Volume 지표에 대한 가중치는 다른 지표와 비교할 때 중간 정도로 설정될 수 있다. 즉, 다른 지표들처럼 높은 가중치는 부여하지 않아도 되며, 높은 가중치를 부여하기 위해서는 해당 국가의 데이터 총생산이 매우 큰 경우일 때 적용할 수 있다. 그러나 데이터의 양이 충분하지 않은 국가의 경우에도 Volume 지표는 중요한 역할을 한다. 따라서 일반적으로 중간 정도의 가중치를 부여하는 것이 합리적일 수 있다.
- [0052] Volume의 중요도에 대한 가중치 정하기는 데이터의 양이 국가의 경제적 성장과 밀접한 관련이 있으므로, Volume 지표는 경제 성장과 직결되어 있다. 따라서 Volume 지표에 가중치를 높게 부여하는 것이 적절하다. 예를 들어, Volume 지표의 중요도를 0.4로 설정할 수 있다.
- [0053] 이와 같이, 모든 수집된 데이터에 대해 크기와 중요도를 고려하여 가중 평균을 계산함으로써 Volume 지표를 산출할 수 있다.
- [0054] 사용자 수 측정지표 산출모듈(220)은 사용자가 어떻게 데이터를 사용하는지 측정하기 위한 Usage 지표를 산출하는 구성이다. 이를 위해 사용자 수 측정지표 산출모듈은 인터넷 사용자의 활동을 모니터링하고 분석하여 사용자의 요구와 콘텐츠에 대한 프록시 개수를 계산한다.
- [0055] Usage 지표는 데이터 분석에서 매우 중요한 지표 중 하나이다. 이 지표는 사용자가 데이터를 어떻게 사용하는지를 측정하기 위해 사용된다. 다음은 Usage 지표를 계산하기 위한 과정과 예시에 대해 설명하기로 한다.
- [0056] Usage 지표를 계산하기 위해서는 다음과 같은 과정이 필요하다.
- [0057] 데이터 수집: 데이터 수집은 인터넷 사용자의 활동을 모니터링하는 것으로 이루어진다. 이를 위해 웹사이트나 모바일 애플리케이션 등의 로그 데이터를 수집한다.

- [0058] 데이터 전처리: 수집한 로그 데이터를 정리하고 필요한 정보를 추출하는 과정이다. 예를 들어, 사용자의 요청에 대한 URL, 시간대, 장치 종류, 브라우저 종류 등을 추출한다.
- [0059] 데이터 분석: 추출한 정보를 기반으로 데이터를 분석한다. 예를 들어, 어떤 콘텐츠가 가장 많은 조회수를 받았는지, 사용자가 어떤 검색어를 가장 많이 사용했는지 등을 분석한다.
- [0060] 지표 계산: 분석 결과를 바탕으로 지표를 계산한다. 예를 들어, 조회수가 많은 콘텐츠나 인기 검색어 등을 기준으로 사용자의 요구와 콘텐츠에 대한 프록시 개수를 계산한다.
- [0061] 시각화: 계산된 지표를 시각화하여 보고서 등으로 제공한다. 이를 통해 데이터를 이해하기 쉽게 만들어 사용자가 데이터를 활용할 수 있도록 돕는다.
- [0062] 본 실시예에 따른 사용자 수 측정지표 산출모듈(220)에서, Usage 지표를 계산하기 위한 계산식은 다음과 같이 표현할 수 있다.
- [0063]
$$\text{Usage} = (\text{데이터를 사용하는 사용자 수} / \text{총 인구 수}) \times 100$$
- [0064] 이 계산식은 데이터를 사용하는 사용자 수를 총 인구 수로 나눈 후 100을 곱하는 방식으로 계산된다. 이렇게 계산된 Usage 지표는 해당 데이터가 얼마나 널리 사용되고 있는지를 나타내는 지표이다. 예를 들어, 총 인구가 1,000,000명이고 해당 데이터를 사용하는 사용자가 10,000명이라면, Usage 지표는 다음과 같이 계산된다.
- [0065]
$$\text{Usage} = (10,000 / 1,000,000) \times 100 = 1\%$$
- [0066] 이 결과는 해당 데이터를 사용하는 인구의 1%에 해당하는 사용자가 있다는 것을 나타낸다. 이러한 Usage 지표를 이용하여 데이터의 사용성을 평가하고, 데이터 경제 활성화에 필요한 인사이트를 도출할 수 있다.
- [0067] Usage 지표 계산은 분석된 데이터를 사용하여 Usage 지표를 계산한다. 예를 들어, 사용자가 웹 페이지에서 데이터를 가져오는 데 걸린 시간을 측정하고, 이를 사용하여 웹 페이지의 성능 지표를 계산할 수 있다. 또는 사용자가 특정 기능을 사용하는 빈도를 측정하여 해당 기능의 인기도를 파악할 수 있다.
- [0068] 따라서, Usage 지표는 데이터 분석에서 매우 중요한 역할을 한다. 이를 통해 사용자의 요구에 맞는 콘텐츠를 제공하고, 적절한 인프라를 구축할 수 있다.
- [0069] Usage(사용자) 지표의 중요도에 대한 가중치 정하기에 대해 설명하면 다음과 같다.
- [0070] Usage(사용자) 지표는 데이터가 실제로 얼마나 많이 사용되는지를 나타내며, 이는 데이터가 중요한지 여부를 결정하는 데 중요한 역할을 한다. 따라서 Usage 지표의 중요도에 대한 가중치를 결정하는 것은 매우 중요하다. 가중치를 결정하는 방법에는 여러 가지가 있을 수 있지만, 다음과 같은 요인들이 고려된다.
- [0071] 데이터 사용량: 데이터가 많이 사용될수록 해당 데이터는 더욱 중요한 역할을 하게 된다. 따라서 데이터 사용량에 따라 가중치를 부여할 수 있다.
- [0072] 데이터 사용 빈도: 데이터가 자주 사용될수록 해당 데이터는 더욱 중요한 역할을 한다. 따라서 데이터 사용 빈도에 따라 가중치를 부여할 수 있다.
- [0073] 데이터 이용 분야: 데이터가 사용되는 분야에 따라 해당 데이터의 중요도가 달라질 수 있다. 예를 들어, 보건 데이터나 범죄 통계와 같은 분야에서는 데이터의 중요성이 더욱 높을 수 있다.
- [0074] 데이터 이용자: 데이터를 사용하는 사람들의 직종, 분야, 국가 등에 따라 해당 데이터의 중요도가 달라질 수 있다. 예를 들어, 정부나 연구 기관에서 데이터를 사용하는 경우 해당 데이터의 중요도가 더욱 높을 수 있다. 이러한 요인들을 고려하여 Usage 지표의 중요도에 대한 가중치를 결정할 수 있다. 예를 들어, 데이터 사용량, 사용 빈도, 이용 분야, 이용자 등을 종합하여 가중치를 부여하는 방법이 있다.
- [0075] 데이터 접근성 측정지표 산출모듈(230)은 국가에서 생성된 데이터가 얼마나 접근 가능한지를 평가하는 Accessibility 지표를 산출하기 위한 구성이다.
- [0076] Accessibility 지표는 국가에서 생성된 데이터가 얼마나 접근 가능한지를 평가하는 지표이다. 이를 평가하기 위해 여러 가지 요소를 고려해야 한다.
- [0077] 예를 들어, 한국에서는 공공 데이터 포털인 data.go.kr을 통해 다양한 공공 데이터를 제공한다. 이 데이터 포털에서는 다양한 데이터셋을 제공하며, 이 데이터를 다운로드하고 API를 통해 접근할 수 있다. 따라서 Accessibility 지표를 평가하기 위해서는 이러한 데이터 포털에서 제공되는 데이터셋의 수와 다운로드 및 API

접근성이 얼마나 용이한지 평가해야 한다.

- [0078] 또한, 데이터의 개인 정보보호와 관련된 문제도 고려해야 한다. 예를 들어, 국가에서 생성된 데이터 중에는 개인 정보를 포함하고 있는 데이터가 있을 수 있다. 이 경우, 해당 데이터를 외부에 공개할 때 개인 정보 보호 규정을 준수해야 한다. 이를 위해, 데이터셋에 대한 메타데이터를 작성하고, 개인 정보가 포함된 데이터에 대해서는 암호화 또는 익명화를 수행해야 한다.
- [0079] 또한, 데이터의 품질과 관련된 문제도 고려해야 한다. 데이터가 잘못된 정보를 포함하고 있다면, 이를 사용하는 연구자나 응용 프로그램에서 잘못된 결과를 도출할 수 있다. 따라서 데이터 포털에서는 데이터의 정확성과 일관성을 검증하기 위한 데이터 품질 평가 지표를 제공해야 한다.
- [0080] 이러한 요소들을 종합하여 Accessibility 지표를 계산할 수 있다. 이를 통해 국가에서 생성된 데이터가 얼마나 접근 가능한지를 평가하고, 데이터의 개인 정보보호와 품질 관리를 위한 개선 방안을 도출할 수 있다.
- [0081] 본 실시예에 따른 데이터 접근성 측정지표 산출모듈(230)에서, Accessibility 지표를 계산하기 위한 과정은 다음과 같은 단계로 이루어진다.
- [0082] 데이터 제공 형태 확인: 먼저, 평가 대상 데이터가 어떤 형태로 제공되는지 확인한다. 데이터가 공개적으로 제공되는 경우와 제한적으로 제공되는 경우가 있으며, 이에 따라 데이터 사용 가능성이 달라진다.
- [0083] 데이터 형식 확인: 데이터가 어떤 형식으로 제공되는지 확인한다. 예를 들어, 데이터가 텍스트, 이미지, 비디오, 오디오 등의 형식인지 확인한다. 데이터 형식에 따라 어떤 유형의 연구자나 응용 프로그램에서 활용할 수 있는지가 달라진다.
- [0084] 데이터 양 확인: 데이터가 얼마나 많이 제공되는지 확인한다. 데이터 양은 연구자나 응용 프로그램에서 데이터를 사용하여 모델을 구축하거나 예측을 수행할 때 모델의 정확도와 일반화 능력에 큰 영향을 미친다.
- [0085] 데이터 가용성 확인: 데이터가 얼마나 쉽게 사용할 수 있는지 확인한다. 예를 들어, 데이터가 API를 통해 제공되는 경우, API 사용 방법과 API의 한계를 확인해야 한다. 데이터가 다운로드 가능한 경우, 다운로드 링크가 유효한지, 다운로드 속도 등을 확인해야 한다.
- [0086] 개인 정보 보호 및 데이터 보호 확인: 데이터에는 개인 정보 보호와 관련된 문제와 데이터 보호와 관련된 문제가 있다. 개인 정보 보호 문제는 개인 식별 정보가 포함되어 있는지, 이를 제거할 수 있는지 등을 확인해야 하고, 데이터 보호 문제는 데이터가 암호화되어 있는지, 접근 권한이 제한되어 있는지 등을 확인해야 한다.
- [0087] 데이터 활용성 확인: 마지막으로, 데이터가 연구자나 응용 프로그램에서 어떻게 활용될 수 있는지, 어떤 분야에서 사용될 수 있는지 등을 고려하여 데이터의 활용성을 평가한다. 예를 들어, 특정 데이터셋이 자연어 처리 분야에서 사용될 수 있는지, 이미지 인식 분야에서 사용될 수 있는지 등을 확인한다.
- [0088] 이러한 과정을 통해 Accessibility 지표를 계산할 수 있다. 이를 통해 데이터가 얼마나 개방적이고, 다양한 분야에서 활용될 수 있는지를 평가할 수 있다.
- [0089] Accessibility 지표를 계산하기 위한 계산식을 예를 들어 설명하면 다음과 같다.
- [0090] Accessibility 지표는 데이터에 접근할 수 있는 사용자 수를 측정하는 지표이다. 이를 계산하기 위해서는 다음과 같은 계산식을 사용할 수 있다.
- [0091]
$$\text{Accessibility} = (\text{Accessible Data} / \text{Total Data}) \times 100$$
- [0092] 여기서 Accessible Data는 해당 데이터에 접근할 수 있는 사용자 수를 의미하며, Total Data는 전체 데이터 양을 의미한다. 계산 결과는 백분율로 나타내며, 접근성이 높을수록 더 높은 값이 나타난다.
- [0093] 예를 들어, 어떤 공공기관에서 100GB의 데이터를 보유하고 있으며, 이 중 70GB의 데이터가 공개되어 있고, 이를 인터넷 등의 전자적 수단을 사용하여 접근할 수 있는 사용자 수가 50명이라고 가정해보기로 한다. 이 경우 Accessibility 지표는 다음과 같이 계산된다.
- [0094]
$$\text{Accessibility} = (70\text{GB} / 100\text{GB}) \times 50 = 35$$
- [0095] 따라서, 이 공공기관의 데이터는 Accessibility 지표가 35이며, 데이터에 접근할 수 있는 사용자 수가 상대적으로 적다는 것을 알 수 있다.
- [0096] Accessibility(접근성) 지표의 중요도에 대한 가중치 정하기에 대해 설명하면 다음과 같다.

- [0097] 접근성(Accessibility) 지표는 공공데이터를 이용하고자 하는 사람들이 해당 데이터를 쉽게 이용할 수 있는지를 나타내는 지표이다. 예를 들어, 공공데이터를 제공하는 웹사이트가 잘 구성되어 있고, 검색 기능이 제공되어 있으며, API나 SDK 등 개발도구를 제공한다면 이는 높은 접근성을 가진 데이터라고 할 수 있다.
- [0098] 접근성은 데이터 활용에 있어서 매우 중요한 요소 중 하나이다. 따라서 국가별 데이터총생산(GDP)을 계산하는데 있어서도 이에 대한 중요성을 고려하여 가중치를 부여할 수 있다. 예를 들어, 한국에서는 공공기관에서 운영하는 공공데이터 포털에 데이터가 집중적으로 제공되기 때문에, 이러한 포털의 구성, 검색 기능, API 제공 등 접근성에 대한 요소가 상대적으로 높은 중요도를 가질 수 있다. 따라서, 국가별 데이터총생산(GDP)을 계산하는데 있어서 접근성의 중요도에 따른 가중치를 부여할 수 있다. 예를 들어, 접근성 지표의 중요도를 0.2로 설정하고, 해당 국가의 접근성 점수를 100점 만점 중 80점으로 산정했다면, 이 국가의 데이터총생산(GDP)은 $0.2 \times 80 = 16$ 이 된다. 이러한 방식으로 국가별 데이터총생산(GDP)을 계산할 수 있다.
- [0099] 데이터 복잡성 측정지표 산출모듈(240)은 디지털 활동의 복잡성을 측정하는 요소로 1인당 광대역 소비량을 측정하여 Complexity(복잡성) 지표를 산출하기 위한 구성이다.
- [0100] 데이터의 복잡성을 측정하기 위해 Complexity 지표를 산출하기 위해, 1인당 광대역 소비량을 측정하여 디지털 활동의 복잡성을 계산한다.
- [0101] 데이터 복잡성을 평가하기 위해 사용되는 Complexity 지표는 데이터 속성과 사용자 경험 측면에서 평가된다. 이 중에서, 디지털 활동의 복잡성은 사용자 경험 측면에서의 지표로서 1인당 광대역 소비량으로 측정될 수 있다.
- [0102] 예를 들어, 국가에서 제공하는 인터넷 서비스 사용량 데이터를 수집하고, 이를 1인당 광대역 소비량으로 계산한다. 이때 광대역 소비량이 높을수록, 해당 국가의 인터넷 사용자들이 더 많은 디지털 활동을 수행한다는 것을 의미한다. 디지털 활동의 종류는 다양하지만, 예를 들어 온라인 쇼핑, 비디오 스트리밍, 소셜 미디어 사용 등이 해당된다.
- [0103] 또한, 데이터 복잡성을 평가하는 데에는 다른 측면도 고려된다. 예를 들어, 데이터의 복잡성은 데이터의 다양성, 분포, 밀도 등의 속성으로 평가할 수도 있다. 이러한 측면은 데이터를 처리하고 분석하는 데에 영향을 미치기 때문에, 데이터의 복잡성을 평가하고 이를 고려하는 것이 중요하다.
- [0104] 본 실시예에 따른 데이터 복잡성 측정지표 산출모듈(240)에서, Complexity 지표를 계산하는 방법은 데이터와 측정 목적에 따라 다르지만, 다음과 같은 과정으로 계산된다.
- [0105] 데이터 수집: 복잡성 지표를 계산하기 위해서는 먼저 필요한 데이터를 수집해야 한다. 복잡성 지표를 계산하기 위해선 데이터가 어떤 디지털 활동과 관련되어 있는지 파악하는 것이 중요하다. 예를 들어, 디지털 활동의 복잡성을 측정하기 위해 1인당 광대역 소비량을 측정하고자 한다면, 인터넷 서비스 제공업체(ISP)로부터 광대역 인터넷 사용량 데이터를 수집해야 한다.
- [0106] 데이터 전처리: 수집된 데이터는 종종 노이즈(noise)나 결측치(missing data) 등의 문제가 있을 수 있다. 이를 해결하기 위해 데이터 전처리 과정이 필요하다. 예를 들어, 광대역 인터넷 사용량 데이터에 결측치가 있는 경우, 이를 대체하는 방법이나 제거하는 방법을 고려할 수 있다.
- [0107] 데이터 분석: 데이터 전처리가 완료되면, 복잡성 지표를 계산하기 위해 데이터를 분석해야 한다. 이를 위해 적절한 분석 방법을 선택하고, 데이터를 시각화하여 분석 결과를 쉽게 파악할 수 있도록 한다.
- [0108] 복잡성 지표 계산: 데이터 분석 결과를 바탕으로, 복잡성 지표를 계산한다. 예를 들어, 1인당 광대역 소비량을 측정하여 디지털 활동의 복잡성을 계산하고자 한다면, 광대역 인터넷 사용량을 인구수로 나누어 계산할 수 있다.
- [0109] 결과 해석: 복잡성 지표를 계산하면, 그 결과를 해석하여 디지털 활동의 복잡성을 파악할 수 있다. 이를 통해, 디지털 환경에서 발생하는 문제를 해결하기 위한 대응책을 마련하거나, 보다 효과적인 디지털 전략을 수립할 수 있다.
- [0110] Complexity 지표를 계산하기 위한 계산식을 예를 들어 설명하면 다음과 같다.
- [0111] Complexity 지표는 데이터의 복잡성을 나타내는 지표이다. 데이터가 복잡할수록 처리하기 어려워지고 분석하는데 많은 시간과 비용이 소요된다. 따라서 데이터 경제에서 Complexity 지표를 계산하여 데이터의 가치를 파악하는 것이 중요하다.

- [0112] Complexity 지표를 계산하기 위해서는 다음과 같은 계산식을 사용할 수 있다.
- [0113] $Complexity = (F + A + T) * S$
- [0114] 여기서, F는 데이터의 형식을 나타내는 Factor(계수) 값이며, A는 데이터의 다양성을 나타내는 Attribute(속성) 값이며, T는 데이터의 정확성을 나타내는 Truth(진실) 값이며, S는 데이터의 크기를 나타내는 Size(크기) 값이다.
- [0115] 각 계수 값은 1 내지 5 사이의 점수를 부여하여 Complexity 값을 계산한다. 예를 들어, F 값이 3, A 값이 4, T 값이 5, S 값이 2일 경우 Complexity 지표를 계산하면 다음과 같다.
- [0116] $Complexity = (3 + 4 + 5) * 2 = 24$
- [0117] 따라서 이 데이터의 Complexity 지표는 24이다. 이렇게 계산된 Complexity 지표를 통해 데이터의 복잡성 수준을 파악할 수 있다.
- [0118] 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템에서, 측정지표 산출부의 Complexity(복잡성) 지표 중요도에 대한 제1 가중치 정하기에 대해 설명하면 다음과 같다.
- [0119] Complexity(복잡성) 지표의 중요도에 대한 가중치는 국가별 새로운 데이터총생산(New GDP)을 계산하는데 있어서 중요한 역할을 한다. 복잡성이 높은 데이터는 보다 정교한 분석과 가치 있는 정보를 도출해낼 가능성이 높기 때문이다. 따라서, 복잡성 지표의 중요도에 대한 가중치를 정할 때에는 다음과 같은 요소들을 고려할 수 있다.
- [0120] 데이터의 다양성, 분석방법의 복잡성, 데이터 처리의 어려움의 요소를 고려할 수 있다. 이에 대해 설명하면 다음과 같다.
- [0121] 데이터의 다양성: 다양한 형식과 속성을 가진 데이터는 보다 복잡한 구조를 가질 가능성이 높다. 이에 따라서, 데이터의 다양성이 높을수록 복잡성 지표의 중요도는 높아질 수 있다.
- [0122] 분석 방법의 복잡성: 분석 방법이 보다 복잡하고 정교할수록, 데이터의 복잡성 지표는 높은 가중치를 받아야 한다. 예를 들어, 머신러닝 알고리즘 등을 이용한 복잡한 분석 방법을 적용해야 하는 경우, 데이터의 복잡성은 높아질 수밖에 없다.
- [0123] 데이터 처리의 어려움: 데이터의 처리가 보다 어렵고 복잡한 경우에는, 데이터의 복잡성 지표에 높은 가중치를 부여할 필요가 있다. 예를 들어, 대용량 데이터를 다루는 경우에는 복잡한 데이터 처리 기술이 요구될 수 있다.
- [0124] 이러한 요소들을 종합적으로 고려하여, 각 국가별로 데이터의 복잡성 지표에 대한 중요도를 결정할 수 있다. 이를 토대로 국가별 새로운 데이터총생산(New GDP)을 계산하는데 사용되는 복잡성 지표의 가중치를 부여할 수 있다.
- [0125] 데이터 개방성 측정지표 산출모듈(250)은 수집한 데이터의 개방성을 평가하기 위해 데이터의 접근성 확인, 라이선스 확인, 데이터의 재사용성 확인, 데이터의 품질 확인, 메타데이터 제공, 개인정보 보호 확인지표로 중요도에 따라 가중치를 부여하여 Openness 지표를 산출하기 위한 구성이다.
- [0126] 본 실시예에 따른 데이터 개방성 측정지표 산출모듈(250)은 Openness 지표를 산출하기 위해 다음의 과정이 필요하다.
- [0127] 데이터 수집 및 분류: 공공기관이 보유하고 있는 데이터베이스 및 전자화된 파일을 수집하고 분류한다. 이때, 데이터의 종류와 규모를 파악하여 적절한 형식으로 변환한다.
- [0128] 데이터 정제 및 가공: 수집된 데이터는 정제 과정을 거쳐 오류를 제거하고 불필요한 정보를 제거한다. 이후, 데이터를 가공하여 다양한 형태로 제공할 수 있도록 한다.
- [0129] 데이터 배포: 데이터를 이용자에게 제공할 수 있도록 배포한다. 이때, 데이터의 형식과 사용 방법 등에 대한 정보를 제공하여 이용자가 쉽게 접근할 수 있도록 한다.
- [0130] 라이선스 제공: 데이터를 재사용하고 이용할 수 있는 권한을 명시한 라이선스를 제공한다. 이때, 상업적 또는 비영리적인 용도에 따라 라이선스 조건을 다르게 설정할 수 있다.
- [0131] 데이터 이용 모니터링: 데이터를 이용하는 모든 사용자의 이용 내역을 모니터링하여 데이터의 재사용 횟수와 범

위를 파악한다. 이를 통해 데이터 제공의 효과와 문제점을 파악하고 개선할 수 있다.

- [0132] 보안 관리: 데이터 개방성을 제공하는 경우, 개인정보 보호 및 데이터 보안에 대한 문제가 발생할 수 있다. 이를 방지하기 위해 데이터 보안 관리 시스템을 운영하고, 데이터 이용자에게도 보안 관리에 대한 책임을 부여한다.
- [0133] 위와 같은 과정을 거쳐 공공기관은 데이터 개방성을 제공하며, 이용자는 다양한 목적에 따라 데이터를 재활용할 수 있다. 이는 데이터의 활용도를 높이고, 사회적 가치를 창출하는데 큰 역할을 한다.
- [0134] 또한 데이터 개방성 측정지표 산출모듈(250)에서 Openness 지표를 계산하기 위해서는 다음과 같은 과정이 필요하다.
- [0135] (01)데이터의 접근성 확인: 공공기관이 보유하고 있는 데이터베이스 및 전자화된 파일에 대한 접근성을 확인해야 한다. 데이터에 접근하기 위해서는 인터넷 또는 기타 전자적 수단을 사용하여 이용이 가능해야 한다.
- [0136] (02)라이선싱 확인: 데이터를 이용하기 위해서는 해당 데이터에 대한 라이선스를 확인해야 한다. 공공기관이 제공하는 데이터에 대해서는 라이선스가 없거나 자유로운 이용이 가능한 CC0 등의 라이선스가 사용되어야 한다.
- [0137] (03)데이터의 재사용성 확인: 데이터를 이용하기 위해서는 해당 데이터가 어떤 형식으로 제공되는지, 어떤 분석 도구를 사용해야 하는지 등을 확인해야 한다. 이러한 정보를 제공함으로써 이용자는 해당 데이터를 쉽게 재사용할 수 있게 된다.
- [0138] (04)데이터의 품질 확인: 데이터의 품질에 대한 정보를 제공해야 한다. 예를 들어, 데이터의 정확성, 일관성, 완전성 등을 평가할 수 있는 정보가 제공되어야 한다.
- [0139] (05)메타데이터 제공: 데이터를 이용하기 위해서는 해당 데이터에 대한 메타데이터가 제공되어야 한다. 메타데이터는 데이터의 소스, 작성자, 생성일자, 파일 크기 등의 정보를 포함하며, 이러한 정보를 제공함으로써 이용자는 해당 데이터를 쉽게 이해하고 활용할 수 있게 된다.
- [0140] (06)개인정보 보호 확인: 데이터를 이용함에 있어서 개인정보 보호 문제가 발생할 수 있다. 이를 방지하기 위해서는 개인정보 보호에 대한 정보가 제공되어야 하며, 데이터 이용에 있어서 반드시 개인정보 보호 규정을 준수해야 한다.
- [0141] 이러한 과정을 통해 Openness 지표를 계산할 수 있으며, 이를 통해 공공기관이 제공하는 데이터의 개방성을 평가할 수 있다.
- [0142] 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템에서, 측정지표 산출부의 데이터 개방성 측정지표 가중치 부여에 대해 설명하면 다음과 같다.
- [0143] Openness 지표의 가중치 부여를 위한 데이터총생산 지표 산출부는 Openness의 각 세부 지표의 중요도에 따라 가중치를 계산하기 위한 계산식은 다음과 같다.
- [0144]
$$\text{Openness} = (O1 * W1) + (O2 * W2) + (O3 * W3) + (O4 * W4) + (O5 * W5) + (O6 * W6)$$
- [0145] 여기서 O1, O2, O3, O4, O5는 각각 데이터의 접근성 확인, 라이선싱 확인, 데이터의 재사용성 확인, 데이터의 품질 확인, 메타데이터 제공, 개인정보 보호 확인 등의 지표를 나타내며, W1, W2, W3, W4, W5, W6은 각각 해당 지표의 중요도를 나타내는 가중치이다.
- [0146] 가중치는 사용자의 요구사항에 따라 다르게 부여될 수 있다. 예를 들어, 데이터의 개방성이 중요하다고 판단되면 Openness 지표의 가중치를 높일 수 있다. 가중치는 일반적으로 0에서 1 사이의 값으로 설정되며, 모든 가중치의 합은 1이 되어야 한다.
- [0147] 따라서, 각각의 지표별 중요도를 고려하여 적절한 가중치를 부여하고 이를 곱하여 계산하면 데이터총생산 지표 산출부의 데이터 개방성 측정지표 가중치 부여Openness 지표를 계산할 수 있다. 이를 통해 데이터의 개방성을 평가하고, 개선할 수 있는 방안을 모색할 수 있다.
- [0148] 데이터 개방성을 제공하기 위해 Openness 지표를 사용한다. 이를 위해, 공공기관이 보유하고 있는 데이터베이스 및 전자화된 파일을 이용자에게 자유롭게 재활용할 수 있도록 제공하고, 상업적 또는 비영리적으로 이용할 권한을 부여한다.
- [0149] Openness(개방성) 지표는 데이터가 얼마나 개방적으로 공유되고 사용 가능한지를 나타내는 지표이다. 이러한 지

표는 데이터에 대한 액세스, 공개, 공유 및 재사용에 대한 정책 및 법률적 요건을 포함한다.

- [0150] 개방성은 데이터의 가치를 극대화하기 위해 매우 중요하다. 개방성 지표의 중요도에 대한 가중치를 결정하는데 있어서는 각 국가의 개방성 정책 및 개방성 지수, 즉 Open Data Barometer와 같은 지수를 고려해야 한다. 이러한 지수는 데이터가 얼마나 개방적으로 공유되고 사용 가능한지를 측정하기 위한 정량적 측정 지표이다.
- [0151] 따라서, Openness 지표의 중요도에 대한 가중치는 국가별 개방성 지수, 즉 Open Data Barometer 등의 지수를 참조하여 결정될 수 있다.
- [0152] 예를 들어, Open Data Barometer에서 높은 순위를 차지하는 국가는 개방성 지표의 중요도에 높은 가중치를 부여할 수 있다. Openness(개방성) 지표는 데이터에 대한 개방성을 나타내는 지표이다. 이는 다른 지표에 비해 더 높은 중요도를 가지는 경우가 있다.
- [0153] 예를 들어, 공공기관에서 제공하는 데이터가 높은 수준의 개방성을 가지고 있다면, 이 데이터를 활용하는 다양한 사용자들이 쉽게 이용할 수 있으며, 이를 통해 새로운 가치를 창출할 수 있다. 따라서, Openness 지표에는 높은 가중치를 부여하여 국가별 데이터 총생산을 계산하는 공식에서 이를 반영할 수 있다.
- [0154] 본 실시예에 따른 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템은 각 5가지 지표(Volume, Usage, Accessibility, Complexity, Openness)를 산출하는 과정에서 각 측정지표마다 중요도에 따라 가중치를 부여할 수도 있고, 새로운 데이터총생산 지표(New GDP)를 산출하는 과정에서, 각각 Volume, Usage, Accessibility, Complexity, Openness 지표의 중요도에 따른 가중치를 부여할 수 있다.
- [0155] 본 실시예에 따른 새로운 데이터총생산 지표 산출부(300)는 데이터 경제 시대에 새로운 데이터총생산 지표(New GDP)를 산출하기 위해 5가지 지표(Volume, Usage, Accessibility, Complexity, Openness)를 사용한다. 이때, 각각의 해당 지표에서 중요도에 따라 가중치를 부여한다. 이 과정에서의 가중치는 해당 지표에 대한 제1 가중치이다. 이후에는 제1 가중치가 부여된 5가지 지표를 사용하여 새로운 데이터총생산 지표(New GDP)를 산출할 때에도, 5가지 지표의 중요도에 따라 제2 가중치를 부여한다. 이때의 제2 가중치는 5가지 지표별로 중요도에 따라 다른 값을 가진다. 여기서, 새로운 데이터총생산 지표 산출부로 표현하였지만, 데이터총생산 지표 산출부이다.
- [0156] 데이터 경제 시대에서 기업이나 정부 등 다양한 조직에서 다양한 종류의 데이터를 수집, 분석하여 활용함으로써 새로운 가치를 창출할 수 있다. 이를 위해서는 데이터 총생산 지표를 산출할 수 있는 시스템이 필요하다. 데이터 총생산 지표는 기존의 경제지표와는 다르게 데이터 활용에 대한 가치를 측정하는 지표이다. 이를 위해서는 데이터의 생산, 처리, 분석, 활용 등과 관련된 활동들을 종합적으로 고려하여 지표를 산출해야 한다.
- [0157] 데이터총생산 지표 산출부(300)는 가중치가 부여된 각 측정지표를 기반으로 데이터총생산 지표를 산출하기 위한 구성이다.
- [0158] 앞서 언급한 5가지 지표(Volume, Usage, Accessibility, Complexity, Openness)는 국가의 새로운 데이터총생산(New GDP)을 산출하는데 사용될 수 있는 측정 지표이다. 이 측정 지표들은 각각의 중요도에 따라 제2 가중치를 부여하고, 국가별 새로운 데이터총생산(New GDP)을 계산하는 공식은 다음과 같다.
- [0159]
$$\text{New GDP} = (w1 * \text{Volume}) + (w2 * \text{Usage}) + (w3 * \text{Accessibility}) + (w4 * \text{Complexity}) + (w5 * \text{Openness})$$
- [0160] 여기서 w1, w2, w3, w4, w5는 각각 Volume, Usage, Accessibility, Complexity, Openness 지표의 중요도에 따른 제2 가중치를 나타낸다. 이 가중치는 0부터 1사이의 값을 가지며, 총합은 1이 되어야 한다.
- [0161] 앞에서 언급했지만, 제2 가중치를 정하는 방법을 예를 들어 정리하면, 다음과 같다.
- [0162] Volume의 중요도에 대한 가중치 정하기는 데이터의 양이 국가의 경제적 성장과 밀접한 관련이 있으므로, Volume 지표는 경제 성장과 직결되어 있다. 따라서 Volume 지표에 가중치를 높게 부여하는 것이 적절하다. 예를 들어, Volume 지표의 중요도를 0.3으로 설정할 수 있다.
- [0163] Usage의 중요도에 대한 가중치 정하기는 데이터 사용자 수는 데이터 경제에 중요한 역할을 한다. 따라서 Usage 지표에도 높은 가중치를 부여하는 것이 적절하다. 예를 들어, Usage 지표의 중요도를 0.25로 설정할 수 있다.
- [0164] Accessibility의 중요도에 대한 가중치 정하기: 데이터 접근성은 데이터 경제를 촉진하기 위해 매우 중요하다. 따라서 Accessibility 지표에도 높은 가중치를 부여하는 것이 적절하다. 예를 들어, Accessibility 지표의 중요도를 0.2로 설정할 수 있다.
- [0165] Complexity의 중요도에 대한 가중치 정하기: 데이터의 복잡성은 데이터 경제에서 중요한 역할을 한다. 따라서

Complexity 지표에도 일정한 가중치를 부여하는 것이 적절하다. 예를 들어, Complexity 지표의 중요도를 0.1로 설정할 수 있다.

[0166] Openness의 중요도에 대한 가중치 정하기: 데이터의 개방성은 데이터 경제에서 매우 중요한 역할을 한다. 따라서 Openness 지표에도 높은 가중치를 부여하는 것이 적절하다. 예를 들어, Openness 지표의 중요도를 0.15로 설정할 수 있다.

[0167] 예를 들어, 가중치를 다음과 같이 설정해볼 수 있다.

[0168] Volume(양) : 0.3

[0169] Usage(사용자) : 0.25

[0170] Accessibility(접근성) : 0.2

[0171] Complexity(복잡성) : 0.1

[0172] Openness(개방성) : 0.15

[0173] 이제 이 가중치를 이용하여 각 측정 지표를 계산하고, 이를 새로운 데이터총생산(New GDP)을 계산하는 공식에 대입할 수 있다.

[0174] 새로운 데이터총생산(New GDP) = (Volume x 0.3) + (Usage x 0.25) + (Accessibility x 0.2) + (Complexity x 0.1) + (Openness x 0.15)

[0175] 여기서, 각 측정 지표는 앞서 설명한 대로 다음과 같다.

[0176] Volume(양) : 데이터의 양

[0177] Usage(사용자) : 데이터를 사용하는 사용자 수

[0178] Accessibility(접근성) : 데이터에 접근할 수 있는 사용자 수

[0179] Complexity(복잡성) : 데이터의 복잡성

[0180] Openness(개방성) : 데이터에 대한 개방성

[0181] 이렇게 가중치를 부여하여 각 측정 지표를 고려한 데이터 경제 측면의 새로운 경제 지표를 산출할 수 있다. 이러한 지표는 국가의 데이터 경제 발전 상황을 평가하고, 데이터 경제 산업을 발전시키는 데 도움을 줄 수 있다.

[0182] 이렇게 구성된 데이터 경제 측정기반 지표구축 시스템은 새로운 데이터총생산(New GDP)을 산출할 수 있다. 이 새로운 지표는 기존의 국내총생산(GDP, Gross Domestic Product) 측정 방식에 데이터 경제 측면을 더하여 새로운 경제 지표를 제공하며, 국가의 데이터 경제 발전 상황을 평가할 수 있는 중요한 지표로 활용될 수 있다.

[0183] 도 3은 본 발명의 일 실시예에 따른 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템을 이용한 방법을 나타낸 개략적인 흐름도이다.

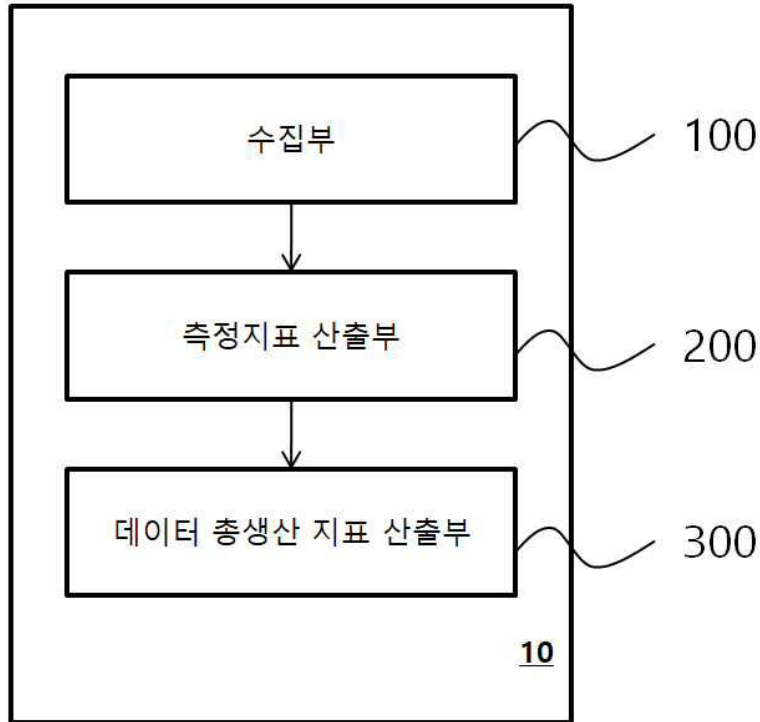
[0184] 도 3에 도시된 바와 같이, 데이터경제시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템을 이용한 방법은 데이터 경제 시대 새로운 데이터총생산(New GDP)의 측정지표 개발을 위한 시스템(이하, 시스템이라 함)이 국가별로 발행되어 생성되는 데이터, 인터넷 사용자 활동에 의한 로그 데이터, 광대역 인터넷 사용량 데이터를 수집하는 단계(a), 상기 시스템이 수집한 데이터의 크기와 해당 데이터의 중요도에 따라 가중치를 부여하여 Volume 지표를 산출하는 단계(b), 상기 시스템이 상기 수집한 데이터와 관련하여 사용자가 어떻게 데이터를 사용하는지 측정하기 위한 Usage 지표를 산출하는 단계(c), 상기 시스템이 상기 수집한 데이터와 관련하여 데이터가 얼마나 접근 가능한지를 평가하는 Accessibility 지표를 산출하는 단계(d), 상기 시스템이 상기 수집한 데이터와 관련하여 디지털 활동의 복잡성을 측정하는 요소로 1인당 광대역 소비량을 측정하여 Complexity 지표를 산출하는 단계(e), 상기 시스템이 상기 수집한 데이터의 개방성을 평가하기 위해 데이터의 접근성 확인, 라이선싱 확인, 데이터의 재사용성 확인, 데이터의 품질 확인, 메타데이터 제공, 개인정보 보호 확인지표로 중요도에 따라 가중치를 부여하여 Openness 지표를 산출하는 단계(f) 및 상기 시스템이 Volume, Usage, Accessibility, Complexity, Openness 지표들의 각 중요도에 따라 가중치를 부여하여 새로운 데이터총생산(New GDP)을 산출하는 단계(g)를 포함한다.

[0185] 본 발명에 따르면, 다양한 측면에서 데이터 경제를 평가하는 지표를 고려하여 산출한 데이터 경제 측면의 새로

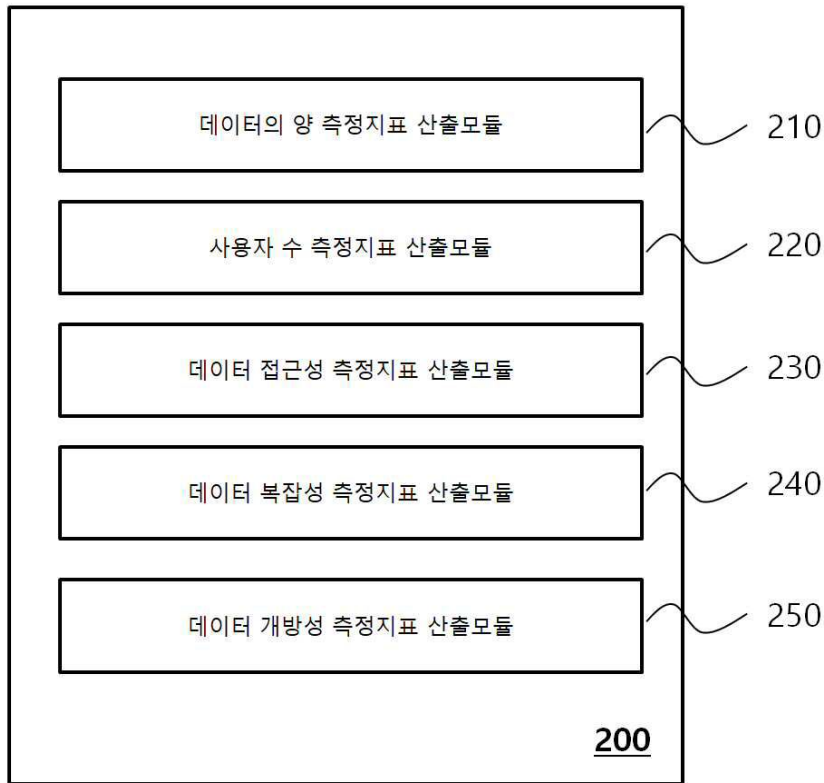
은 경제 지표는 데이터 경제의 상황을 더욱 정확하게 파악할 수 있다. 이를 통해 국가별 데이터 경제의 상황을 비교하고 분석함으로써, 데이터 경제 분야의 개선 방향을 찾을 수 있다. 또한, 이러한 경제 지표를 활용하여 적극적인 데이터 활용 방안을 모색하고, 데이터 경제 생태계를 발전시키는 데 기여할 수 있는 효과가 있다.

도면

도면1



도면2



도면3

