(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0167782
(43) 공개일자 2023년12월12일

(51) 국제특허분류(Int. Cl.)
G06Q 50/26 (2012.01) *G06F 11/30* (2006.01)
G06F 3/14 (2006.01) *G06Q 50/10* (2012.01)
(52) CPC특허분류
G06Q 50/26 (2013.01)
G06F 11/3065 (2013.01)
(21) 출원번호 10-2022-0067546
(22) 출원일자 2022년06월02일
심사청구일자 2022년06월02일

(71) 출원인
인천대학교 산학협력단
인천광역시 연수구 갯벌로 27, 인천대학교 이노베이션센터(송도동)
(72) 발명자
김훈
인천광역시 연수구 아카데미로 119, 8호관 456호
반다리소빗
인천광역시 연수구 아카데미로 119, 8호관 457호 (송도동, 인천대학교)
(뒷면에 계속)
(74) 대리인
특허법인충정

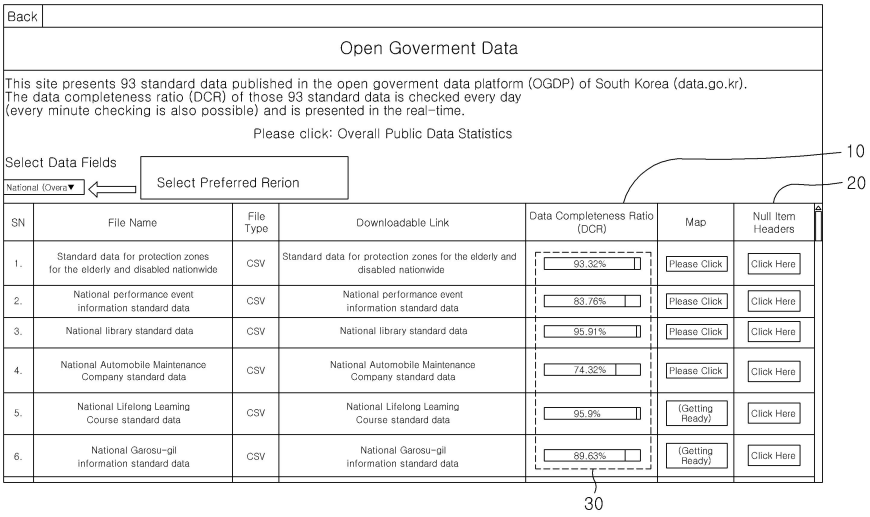
전체 청구항 수 : 총 9 항

(54) 발명의 명칭 **개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치**

(57) 요 약

개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치는 데이터 완전성 비율(Data Completeness Ration, DCR)과 관련하여 개방 공공 데이터 포털(Open Government Data Portal, ODGP)에서 데이터의 품질을 실시간으로 시각화하여 개방 공공 데이터의 DCR 결과를 모니터링하여 개방 공공 데이터의 품질을 개선할 수 있는 효과가 있다.

대 표 도



(52) CPC특허분류
G06F 3/14 (2020.08)
G06Q 50/10 (2015.01)
G06F 2201/80 (2013.01)
(72) 발명자
란잔나빈
인천광역시 연수구 아카데미로 119, 8호관 457호
(송도동, 인천대학교)
김영찬
인천광역시 연수구 아카데미로 119, 8호관 457호

명 세 서

청구범위

청구항 1

개방 공공 데이터를 제공하는 복수의 공공 데이터 제공 서버로부터 개방 공공 데이터 및 개방 공공 데이터의 등록 요청 신호를 수신하는 데이터 수신부;

상기 수신한 공공 데이터를 애플리케이션 프로그래밍 인터페이스를 통해 다운로드하는 표준 테이블 형식의 데이터 세트로 다운로드하는 애플리케이션 프로그래밍 인터페이스 저장부;

상기 데이터 세트를 표준 파일 포맷인 파일로 다운로드하는 자동 데이터 다운로드부; 및

상기 다운로드한 파일의 데이터 필드에 데이터가 차지하는 비율을 기초로 상기 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 계산하는 데이터 완전성 비율 계산부로 이루어진 개방 공공 데이터 인터페이스 서버를 포함하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 2

개방 공공 데이터를 제공하는 복수의 공공 데이터 제공 서버로부터 개방 공공 데이터 및 개방 공공 데이터의 등록 요청 신호를 수신하는 데이터 수신부;

상기 수신한 공공 데이터를 애플리케이션 프로그래밍 인터페이스를 통해 다운로드하는 표준 테이블 형식의 데이터 세트로 다운로드하는 애플리케이션 프로그래밍 인터페이스 저장부;

상기 데이터 세트를 표준 파일 포맷인 파일로 다운로드하는 자동 데이터 다운로드부;

상기 다운로드한 파일의 데이터 필드에 데이터가 차지하는 비율을 기초로 상기 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 계산하는 데이터 완전성 비율 계산부; 및

상기 완전성 비율 계산부에서 계산한 DCR 값을 기준으로 DCR ≥ 90%인 경우, 파란색으로, 70% ≤ DCR < 90%인 경우, 녹색으로, 50% ≤ DCR < 70%인 경우, 노란색으로, DCR < 50%인 경우, 빨간색을 DCR 값의 배경색으로 하고, 100%를 기준으로 상기 DCR 값만큼 슬라이드바 형태로 표시하는 웹사이트 프레임워크 처리부를 포함하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 3

청구항 1 또는 청구항 2에 있어서,

상기 데이터 완전성 비율 계산부는 상기 파일과, 상기 파일에서 특정 지역 $\mathcal{L} = \{l_1, l_2, \ldots, l_L\}$ 의 리스트와, 전체 데이터 필드와 필수 데이터 필드의 리스트를 입력받고, 지역 기반 DCR을 포함시킨 데이터 프레임 $\mathcal{R} = \{r_1, r_2, \ldots, r_M\}$ 을 출력하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 4

청구항 1 또는 청구항 2에 있어서,

상기 데이터 완전성 비율 계산부는 상기 파일에 있는 전체 필드의 데이터 포인트의 개수(O_{f_i})를 카운트하고,

상기 파일에 있는 전체 필드의 널 데이터 포인트(Null Data Point)의 개수($N_{f_i}^{l_j,overall}$)를 카운트하고, 하

기의 수학식 1을 이용하여 전체 데이터 필드의 제1 데이터 완전성 비율($DCR_{f_i}^{l_j,overall}$)을 계산하며, 상기 제1 데이터 완전성 비율을 상기 각 파일의 데이터 프레임에 포함시키는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

[수학식 1]

$$DCR_{f_i}^{l_j,overall} = \left(1 - \frac{N_{f_i}^{l_j,overall}}{O_{f_i}}\right) \times 100\ \%$$

청구항 5

청구항 1 또는 청구항 2에 있어서,

상기 데이터 완전성 비율 계산부는 상기 파일에 있는 필수 필드의 데이터 포인트의 개수(M_{f_i})를 카운트하고, 상기 파일에 있는 필수 필드의 널 데이터 포인트의 개수($N_{f_i}^{l_j,mandatory}$)를 카운트하고, 하기의 수학식 2를 이용하여 필수 데이터 필드의 제2 데이터 완전성 비율($DCR_{f_i}^{l_j,mandatory}$)을 계산하며, 상기 제2 데이터 완전성 비율을 상기 각 파일의 데이터 프레임에 포함시키는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

[수학식 2]

$$DCR_{f_i}^{l_j,mandatory} = \left(1 - \frac{N_{f_i}^{l_j,mandatory}}{M_{f_i}}\right) \times 100\ \%$$

청구항 6

청구항 1에 있어서,

상기 개방 공공 데이터 인터페이스 서버는,

상기 개방 공공 데이터는 상기 데이터 완전성 비율, 행이 비어 있는 파일의 헤더를 개수와 함께 표시하는 널 아이템 헤더를 행별로 이루어진 개방 공공 데이터 포털(Open Government Data Portal, ODGP)을 생성하여 대시보드 홈페이지 형태로 시각화하여 디스플레이부를 통해 출력하는 웹사이트 프레임워크 처리부를 더 포함하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 7

청구항 6에 있어서,

상기 웹사이트 프레임워크 처리부는 상기 완전성 비율 계산부에서 계산한 DCR 값을 기준으로 DCR ≥ 90%인 경우, 파란색으로, 70% ≤ DCR < 90%인 경우, 녹색으로, 50% ≤ DCR < 70%인 경우, 노란색으로, DCR < 50%인 경우, 빨간색을 DCR 값의 배경색으로 하고, 100%를 기준으로 상기 DCR 값만큼 슬라이드바 형태로 표시하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 8

청구항 6에 있어서,

상기 웹사이트 프레임워크 처리부는 상기 ODGP의 '널 아이템 헤더' 버튼 신호를 수신하면, 행이 비어 있는 파일

의 헤더를 개수(Null Value)와 함께 테이블 형태의 페이지로 생성하여 디스플레이부를 통해 출력하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

청구항 9

청구항 6에 있어서,

상기 웹사이트 프레임워크 처리부는 상기 ODGP에 저장된 파일의 DCR 값을 기준으로 상기 DCR 값을 오름차순으로 정렬하고, 상기 정렬한 DCR 값에 대응하는 개방 공공 데이터를 상기 ODGP에 업로드하여 디스플레이부를 통해 출력하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치.

발명의 설명

기 술 분 야

[0001] 본 발명은 개방 공공 데이터 관리 장치에 관한 것으로서, 더욱 상세하게는 데이터 완전성 비율(Data Completeness Ration, DCR)과 관련하여 개방 공공 데이터 포털(Open Government Data Portal, ODGP)에서 데이터의 품질을 실시간으로 시각화하여 개방 공공 데이터의 DCR 결과를 모니터링하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치에 관한 것이다.

배 경 기 술

[0002] 최근 세계 주요국들은 데이터를 또 하나의 자원으로 인식하고 있으며, 정부의 투명성과 효율성을 제고하고, 경제 활성화 효과를 기대하기 위하여 공공 데이터를 적극 개방하는 정책을 경쟁적으로 시행하고 있다.

[0003] 최근의 몇 년 동안 많은 국가의 정부는 경제를 부양하는데 데이터의 중요성을 인식했다. 이러한 결과는 공공 데이터를 표준화된 형식으로 누구나 쉽고 자유롭게 사용할 수 있도록 개방 공공 데이터(Open Government Data, OGD)를 제공한다.

[0004] 좋은 품질의 OGD는 한 국가의 경제를 부양할 수 있는 반면, 낮은 품질의 OGD는 효율적인 사용과 재사용을 위태롭게 할 수 있다. 따라서, 개방형 공공 데이터 포털(Open Government Data Portals, OGDP)에 저장된 데이터의 품질을 유지하는 것이 매우 중요하다.

[0005] 전 세계 정부 기관에서 공개하는 공개 데이터의 양과 다양성이 눈에 띄게 증가함에 따라 OGDP에 공개되는 데이터의 품질이 미래 잠재력을 결정한다.

[0006] OGDP의 품질이 높은 데이터는 21세기의 가장 중요한 자원이기 때문에 사용자가 필요한 데이터를 더 쉽게 찾을 수 있도록 도와준다

[0007] OGDP의 품질이 낮은 데이터는 개방형 데이터의 효율적인 사용과 재사용을 위태롭게 한다. 또한, 낮은 데이터 품질은 데이터 액세스 및 해석 비용을 증가시키고, 잘못된 정보를 제공하여 결과적으로 OGDP 사용을 줄이는 문제점이 발생한다.

선행기술문헌

특허문헌

[0008] (특허문헌 0001) 한국 등록특허번호 제10-1458820호

발명의 내용

해결하려는 과제

[0009] 이와 같은 문제점을 해결하기 위하여, 본 발명은 데이터 완전성 비율(Data Completeness Ration, DCR)과 관련하여 개방 공공 데이터 포털(Open Government Data Portal, ODGP)에서 데이터의 품질을 실시간으로 시각화하여 개방 공공 데이터의 DCR 결과를 모니터링하는 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치를 제공하는데 그 목적이 있다.

과제의 해결 수단

[0010] 상기 목적을 달성하기 위한 본 발명의 특징에 따른 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치는,

[0011] 개방 공공 데이터를 제공하는 복수의 공공 데이터 제공 서버로부터 개방 공공 데이터 및 개방 공공 데이터의 등록 요청 신호를 수신하는 데이터 수신부;

[0012] 상기 수신한 공공 데이터를 애플리케이션 프로그래밍 인터페이스를 통해 다운로드하는 표준 테이블 형식의 데이터 세트로 다운로드하는 애플리케이션 프로그래밍 인터페이스 저장부;

[0013] 상기 데이터 세트를 표준 파일 포맷인 파일로 다운로드하는 자동 데이터 다운로드부; 및

[0014] 상기 다운로드한 파일의 데이터 필드에 데이터가 차지하는 비율을 기초로 상기 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 계산하는 데이터 완전성 비율 계산부로 이루어진 개방 공공 데이터 인터페이스 서버를 포함한다.

[0016] 본 발명의 특징에 따른 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치는,

[0017] 개방 공공 데이터를 제공하는 복수의 공공 데이터 제공 서버로부터 개방 공공 데이터 및 개방 공공 데이터의 등록 요청 신호를 수신하는 데이터 수신부;

[0018] 상기 수신한 공공 데이터를 애플리케이션 프로그래밍 인터페이스를 통해 다운로드하는 표준 테이블 형식의 데이터 세트로 다운로드하는 애플리케이션 프로그래밍 인터페이스 저장부;

[0019] 상기 데이터 세트를 표준 파일 포맷인 파일로 다운로드하는 자동 데이터 다운로드부;

[0020] 상기 다운로드한 파일의 데이터 필드에 데이터가 차지하는 비율을 기초로 상기 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 계산하는 데이터 완전성 비율 계산부; 및

[0021] 상기 완전성 비율 계산부에서 계산한 DCR 값을 기준으로 DCR ≥ 90%인 경우, 파란색으로, 70% ≤ DCR < 90%인 경우, 녹색으로, 50% ≤ DCR < 70%인 경우, 노란색으로, DCR < 50%인 경우, 빨간색을 DCR 값의 배경색으로 하고, 100%를 기준으로 상기 DCR 값만큼 슬라이드바 형태로 표시하는 웹사이트 프레임워크 처리부를 포함한다.

발명의 효과

[0022] 전술한 구성에 의하여, 본 발명은 개방 공공 데이터 포털(Open Government Data Portal, ODGP)에서 데이터의 품질을 실시간으로 시각화하여 개방 공공 데이터의 DCR 결과를 모니터링하여 개방 공공 데이터의 품질을 개선할 수 있는 효과가 있다.

도면의 간단한 설명

[0023] 도 1은 본 발명의 실시예에 따른 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치의 구성을 나타낸 도면이다.

도 2는 본 발명의 실시예에 따른 개방 공공 데이터 인터페이스 서버의 내부 구성을 간략하게 나타낸 블록도이다.

도 3은 본 발명의 실시예에 따른 개방 공공 데이터 포털에 있는 DCR을 체크하기 위한 프레임워크를 나타낸 도면이다.

도 4는 본 발명의 실시예에 따른 개방 공공 데이터 포털에서 이용 가능한 DCR을 체크하는 Django 기반의 프레임워크를 나타낸 도면이다.

도 5는 본 발명의 실시예에 따른 API 목록과 CSV 형식의 API의 일례를 나타낸 도면이다.

도 6은 본 발명의 실시예에 따른 전국 및 수도권 지역 데이터 파일의 분류를 나타낸 도면이다.

도 7은 본 발명의 실시예에 따른 전국 및 수도권 지역의 평균 DCR를 나타낸 도면이다.

도 8은 본 발명의 실시예에 따른 전국, 인천, 서울, 경기 지역의 일평균 DCR를 나타낸 도면이다.

도 9는 본 발명의 실시예에 따른 개방 공공 데이터 포털의 일례를 나타낸 도면이다.

도 10은 본 발명의 실시예에 따른 개방 공공 데이터의 일례를 나타낸 도면이다.

도 11은 본 발명의 실시예에 따른 전체에 대해 나열된 파일의 DCR을 보여주는 대시보드를 나타낸 도면이다.

도 12는 본 발명의 실시예에 따른 개방 공공 데이터의 DCR 값을 기준으로 정렬된 파일을 보여주는 대시보드를 나타낸 도면이다.

도 13은 본 발명의 실시예에 따른 행이 비어 있는 파일의 헤더를 개수와 함께 표시하는 대시보드를 나타낸 도면이다.

발명을 실시하기 위한 구체적인 내용

[0024] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.

[0025] 개방 공공 데이터 포털(Open Government Data Portal, ODGP)은 다양한 소스에서 기존 데이터 세트를 수집하고, 이를 사용자 친화적인 대시보드에 게시하여 사용자가 표준화된 파일 형식(CSV, XLSX, JSON)으로 API(응용 프로그래밍 인터페이스)를 통해 보고, 다운로드하고 검색할 수 있는 웹 기반 시스템이다.

[0026] OGDP를 높은 신뢰성으로 만드는 주요 특징은 품질, 완전성, 접근성, 사용성과 이해성, 적시성, 가치와 유용성, 세분성과 비교 가능성이다. 이러한 모든 특성 외에도 데이터의 품질은 OGDP의 성공에 중요한 역할을 한다.

[0027] 고품질 데이터는 최고 성능을 달성할 가능성을 높일 수 있기 때문이다. DAMA(International Data Management Association)에서 데이터 품질을 보장하기 위해 정의한 6가지 특성은 완전성, 고유성, 적시성, 유효성, 정확성 및 일관성이다.

[0028] 완전성은 소스에서 null이 아닌 값의 수와 보편적 관계의 크기 사이의 비율을 나타낸다. 정확도는 데이터가 정확하고, 인증되고, 신뢰할 수 있는 정도를 나타낸다. 일관성은 데이터가 동일한 형식으로 표시되고, 이전 데이터와 호환되는 정도를 나타낸다. 유효성은 데이터가 정의의 구문(형식, 유형, 범위)을 준수하는 정도를 나타낸다. 고유성은 데이터가 중복되지 않는 정도를 나타낸다. 적시성은 데이터의 사용 기간이 당면한 작업에 적절한 정도를 나타낸다.

[0029] OGDP에 저장된 데이터의 품질이 좋지 않은 문제를 해결하기 위해서 OGDP에 저장된 데이터의 품질을 자동으로 검사하는 프레임워크가 필요하다.

[0031] 도 1은 본 발명의 실시예에 따른 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치의 구성을 나타낸 도면이다.

[0032] 본 발명의 실시예에 따른 개방 공공 데이터의 품질을 알려주는 개방 공공 데이터 관리 장치(100)는 공공 데이터 제공 서버(200), 개방 공공 데이터 인터페이스 서버(300) 및 사용자 단말기(400)를 포함한다.

[0033] 공공 데이터 제공 서버(200)는 개방형 공공 데이터를 제공하는 다양한 기관들 각각에 대응하는 서버로서, 복수의 서버(210, 220, 230 등)를 포함할 수 있다.

[0034] 예를 들어, 공공 데이터를 제공하는 기관은 국가행정기관, 자치행정조직, 교육행정조직, 입법조직, 사법조직, 헌법조직, 교육기관, 공공기관, 위원회 및 경제자유구역청 등이 있을 수 있으며, 세부적으로 국토교통부, 경찰청, 국민건강보험공단, 소상공인시장진흥공단, 해양수산부, 한국수자원공사, 농림축산식품부, 한국농수산식품교육문화정보원, 산림청, 국립수목원 등이 있을 수 있다.

[0035] 일예로, 각 기관들이 제공하는 공공 데이터를 간단히 살펴보면 다음과 같다.

[0036] 국토교통부는 건축정보 관련 공공 데이터로서, 건축물대장(기본개요, 주택가격, 부속지번 등), 건물에너지(수용가별 에너지사용량), 건물 유지점검(점검접수보고, 점검결과 등) 등을 제공할 수 있다. 또한, 국토교통부는 부동산정보 관련 공공 데이터로서, 부동산 지적정보(법정구역정보, 연속지적도형정보, 토지등급, 대지권등록정보 등), 건물정보(GIS건물통합정보) 등을 제공할 수 있다.

[0037] 경찰청은 교통사고정보 관련 공공 데이터로서, 교통사망사고 위치정보, 보행자 무단횡단 사고 다발지역 정보, 스쿨존 내 어린이사고 다발지역 정보 등을 제공할 수 있다.

[0038] 국민건강보험공단은 국민건강정보 관련 공공 데이터로서, 진료내역정보(서식코드, 진료과목코드, 요양일수, 등), 약품처방내역정보(1회투약량, 총 투여일수, 단가, 금액 등), 건강검진정보(신장, 체중, 허리둘레, 혈압, 혈당, 콜레스테롤, 시력, 구강, 흡연, 음주 등) 등을 제공할 수 있다.

[0039] 상기의 예들은 본원의 이해를 돕기 위한 하나의 실시예일 뿐 이에 한정된 것은 아니며, 다른 일 실시예는, 공공 데이터 제공 서버(200)를 통해 제공되는 공공데이터로서, 교육, 국토관리, 공공행정, 재정금융, 산업고용, 사회복지, 식품건강, 문화관광, 보건의료, 재난안전, 교통물류, 환경기상, 과학기술, 농축수산, 통일외교안보, 법률 관련 정보 등이 포함될 수 있으며, 또한, GPS 정보, 버스 정보, 지하철 도착 시간 정보, 기상 정보, 지리정보, 부동산 정보, 관광정보, 문화정보, 보건의료정보, 시설 정보, 주차장 정보 등 국가의 모든 기관이 만들어내는 정보나 자료 등이 포함될 수 있다

[0040] 개방 공공 데이터 인터페이스 서버(300)는 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 계산하고, ODGP에서 데이터의 품질을 실시간으로 시각화하여 개방 공공 데이터의 DCR 결과를 모니터링하여 개방 공공 데이터의 품질을 개선할 수 있다.

[0041] 사용자 단말기(400)는 개방 공공 데이터 인터페이스 서버(300)로부터 개방 공공 데이터의 품질을 정량적으로 나타내는 데이터 완전성 비율(Data Completeness Ration, DCR)을 알려주는 개방 공공 데이터 포털(Open Government Data Portal, ODGP)을 웹 페이지 형태로 제공받는다.

[0043] 도 2는 본 발명의 실시예에 따른 개방 공공 데이터 인터페이스 서버의 내부 구성을 간략하게 나타낸 블록도이고, 도 3은 본 발명의 실시예에 따른 개방 공공 데이터 포털에 있는 DCR을 체크하기 위한 프레임워크를 나타낸 도면이고, 도 4는 본 발명의 실시예에 따른 개방 공공 데이터 포털에서 이용 가능한 DCR을 체크하는 Django 기반의 프레임워크를 나타낸 도면이다.

[0044] 본 발명의 실시예에 따른 개방 공공 데이터 인터페이스 서버(300)는 데이터 수신부(310), API 저장부(320), 자동 데이터 다운로드부(330), DCR 계산부(340), 데이터베이스부(342), 제어부(350), 웹사이트 프레임워크 처리부(360), 디스플레이부(370)를 포함한다.

[0045] 데이터 수신부(310)는 공공 데이터 제공 서버(200)로부터 개방 공공 데이터 및 개방 공공 데이터의 등록 요청 신호를 수신할 수 있다.

[0046] 애플리케이션 프로그래밍 인터페이스 저장부(320)는 공공 데이터 제공 서버(200)의 개방 소스 데이터 플랫폼으로부터 수신한 개방 공공 데이터 $\mathcal{N} = \{N_1, N_2 \ \ldots, N_M\}$ 의 애플리케이션 프로그래밍 인터페이스(Application Programming Interface, API) $\mathcal{A} = \{ a_1, a_2, \ldots, a_M\}$를 추출하여 저장한다.

[0047] API는 컴퓨터나 컴퓨터 프로그램 사이의 연결이다. 일종의 소프트웨어 인터페이스이며 다른 종류의 소프트웨어에 서비스를 제공한다.

[0048] 애플리케이션 프로그래밍 인터페이스 저장부(320)에 저장된 API는 Python, Java, Ruby 등과 같은 스크립팅 언어(Scripting Language)를 사용하여 관련 데이터 파일을 자동으로 다운로드하는데 사용한다.

[0049] 자동 데이터 다운로드부(330)는 길이가 M인 애플리케이션 프로그래밍 인터페이스 저장부(320)에 저장된 API $\mathcal{A}$ 를 사용하여 M개의 데이터 파일 $\mathcal{F} = \{f_1, f_2, \ldots, f_M\}$ 목록을 다운로드한다.

[0050] 자동 데이터 다운로드부(330)는 표준 데이터 이름 $\mathcal{N}$ 과 이에 대응하는 API $\mathcal{A}$를 사용하여 표준 파일 포맷인 데이터 파일 $f_i \ (i \in \{1,2,3,\ldots,M\})$ 을 다운로드한다.

[0051] 자동 데이터 다운로드부(330)는 다운로드된 파일 $\mathcal{F}$ 을 저장한다.

[0052] 애플리케이션 프로그래밍 인터페이스 저장부(320)에서 수신한 데이터와 자동 데이터 다운로드부(330)에서 다운로드한 데이터의 알고리즘을 정리하면 다음과 같다.

Algorithm 1. Automatic data download

Input: API database $\mathcal{D}$ containing API $\mathcal{A} = \{a_1, a_2, \ldots, a_M\}$ and its corresponding data file name $\mathcal{N} = \{N_1, N_2 \ \ldots, N_M\}$

Output: Downloaded data file $\mathcal{F} = \{f_1, f_2, \ldots, f_M\}$

Initialize: $API_response$ = [], f = []

Automatically download data file $\mathcal{F}$: To download the data file f_i from its respective API a_i

1. **for** i in range(M) **do**
2. $\quad API_response_i \leftarrow$ requests.get(a_i)
3. $\quad f_i \leftarrow$ open(N_i .(standard_file_format) from preferred system location)
4. $\quad f_i$.write($API_response_i$.content)
5. $\quad f_i$.close()
6. **endfor**

[0053]

[0054] 데이터 완전성 비율(Data Completeness Ration, DCR) 계산부(340)는 표준 파일 포맷의 파일 $\mathcal{F}$를 다운로드한 후, 이러한 파일의 데이터를 추가로 분석하여 파일의 지역 기반의 DCR를 계산한다. 여기서, 지역은 국가의 행정 구역과 국가 전체를 나타낸다.

[0055] DCR 계산부(340)에서 파일의 지역 기반 DCR을 결정하는 알고리즘 2는 다음과 같다.

Algorithm 2. Automatic region-based DCR calculation

Input: Data file $\mathcal{F} = \{f_1, f_2, \ldots, f_M\}$, list of unique regions in the file $\mathcal{L} = \{l_1, l_2, \ldots, l_L\}$, list of overall and mandatory fields

Output: Region-based DCR stored in the data file $\mathcal{R} = \{r_1, r_2, \ldots, r_M\}$

Initialize: N, DCR, O, M = []

Automatically find region-based DCR : To find DCR in the data file f_i

1. **for** i in range(M) **do**
2. $\quad O_{f_i} \leftarrow$ Count (data points of overall fields in the file f_i)
3. $\quad M_{f_i} \leftarrow$ Count (data points of mandatory fields in the file f_i)
4. $\quad$ **for** j in range(L) **do**
5. $\quad\quad N_{f_i}^{l_j,overall} \leftarrow$ Count (overall field null data points in f_i for region L_j)
6. $\quad\quad N_{f_i}^{l_j,mandatory} \leftarrow$ Count (mandatory field null data points in f_i for region L_j)
7. $\quad\quad DCR_{f_i}^{l_j,overall} \leftarrow \left(1 - \frac{N_{f_i}^{l_j,overall}}{O_{f_i}}\right) \times 100\ \%$
8. $\quad\quad DCR_{f_i}^{l_j,mandatory} \leftarrow \left(1 - \frac{N_{f_i}^{l_j,mandatory}}{M_{f_i}}\right) \times 100\ \%$
9. $\quad$ **endfor**
10. $\quad$ **save** the data frame containing $DCR_{f_i}^{l_j,overall}$, and $DCR_{f_i}^{l_j,mandatory}$ as r_i in the preferred system location
11. **endfor**

[0056]

[0057] 데이터 완전성 비율 계산부(340)는 파일 $\mathcal{F}$에 리스트된 특정 지역 $\mathcal{L} = \{l_1, l_2, \ldots, l_L\}$을 기초로 위의 알고리즘 2를 이용하여 파일의 지역 기반 데이터 완전성 비율(DCR)을 계산한다.

[0058] 데이터 완전성 비율(DCR)은 파일의 데이터 필드에 데이터가 차지하는 비율을 계산하여 개방 공공 데이터의 품질을 정량적으로 나타내고, 전체 데이터 필드와 필수 데이터 필드의 경우를 고려하여 계산된다. 전체 데이터 필드(Overall Data Field)는 파일 $\mathcal{F}$에서 컬럼명(Columm Name)의 리스트를 나타내고, 파일 $\mathcal{F}$의 필수 데이터 필드(Mandatory Data Field)는 사전 정의되어 개방 공공 데이터 포털(Open Government Data Portal, ODGP)의 데이터 표준화 정책 섹션에 표시된다.

[0059] 여기서, 필수 데이터 필드는 전체 데이터 필드에서 필수로 입력되어야 하는 사항을 나타낸다.

[0060] 데이터 완전성 비율 계산부(340)는 데이터 파일 $\mathcal{F} = \{f_1, f_2, \ldots, f_M\}$ 과, 파일에서 특정 지역 $\mathcal{L} = \{l_1, l_2, \ldots, l_L\}$ 의 리스트와, 전체 데이터 필드와 필수 데이터 필드의 리스트를 입력받고, 하기의 수학식 1과 수학식 2를 이용하여 지역 기반 DCR을 포함시킨 데이터 프레임 $\mathcal{R} = \{r_1, r_2, \ldots, r_M\}$ 을 출력한다.

[0061] 데이터 완전성 비율 계산부(340)는 수신한 데이터 파일 f_i 에서 지역 기반의 데이터 완전성 비율(DCR)을 자동적으로 찾아낸다.

[0062] 데이터 완전성 비율 계산부(340)는 파일 f_i 에 있는 전체 필드의 데이터 포인트의 개수(O_{f_i})를 카운트한다.

[0063] 데이터 완전성 비율 계산부(340)는 파일 f_i 에 있는 필수 필드의 데이터 포인트(Null Data Point)의 개수(M_{f_i})를 카운트한다.

[0064] 데이터 완전성 비율 계산부(340)는 지역 L_j 에 대한 파일 f_i 의 전체 필드의 널 데이터 포인트(Null Data Point)의 개수($N_{f_i}^{l_j,overall}$)를 카운트한다.

[0065] 데이터 완전성 비율 계산부(340)는 지역 L_j 에 대한 파일 f_i 의 필수 필드의 널 데이터 포인트의 개수($N_{f_i}^{l_j,mandatory}$)를 카운트한다.

[0066] 데이터 완전성 비율 계산부(340)는 위에 계산된 데이터 포인트의 개수, 널 데이터 포인트의 개수를 하기의 수학식 1과 수학식 2에 대입하여 전체 데이터 필드의 제1 데이터 완전성 비율($DCR_{f_i}^{l_j,overall}$)과 필수 데이터 필드의 제2 데이터 완전성 비율($DCR_{f_i}^{l_j,mandatory}$)을 계산한다.

[0067] 데이터 완전성 비율 계산부(340)는 각 파일 f_i 의 전체 데이터 필드의 제1 데이터 완전성 비율과 필수 데이터 필드의 제2 데이터 완전성 비율을 포함한 데이터 프레임을 DCR 계산부(340)의 특정 위치에 저장한다.

수학식 1

[0068]
$$DCR_{f_i}^{l_j,overall} = \left(1 - \frac{N_{f_i}^{l_j,overall}}{O_{f_i}}\right) \times 100\ \%$$

수학식 2

$$DCR_{f_i}^{l_j,mandatory} = \left(1 - \frac{N_{f_i}^{l_j,mandatory}}{M_{f_i}}\right) \times 100\ \%$$

[0069]

[0070] 웹사이트 프레임워크 처리부(360)는 데이터 시각화를 위해 오픈 소스 웹사이트 프레임워크를 제공하고, 데이터 완전성 비율 계산부(340)로부터 제1 데이터 완전성 비율과 제2 데이터 완전성 비율을 포함한 각 파일의 데이터 프레임을 수신한다. 웹 개발 프레임워크의 품질을 결정하는 기능은 AJAX 지원, 클라우드 컴퓨팅, Comet 지원, 사용자 정의 오류 메시지, 사용자 정의 및 확장성, 디버깅, 문서화, 양식 유효성 검사, HTML5 지원, JavaScript 기반 프레임워크 지원, 개체 관계형 매핑, 병렬 렌더링이다.

[0071] 웹사이트 개발을 위한 강력한 백엔드가 있는 오픈 소스 프레임워크 중 일부는 Ruby on Rails, Cake PHP, ASP.NET, Django, Laravel 등이다.

[0072] 웹사이트 프레임워크 처리부(360)는 오픈 소스 웹 서버를 구현하여 개방 공공 데이터 포털(Open Government Data Portal, ODGP)을 생성하여 사용자 단말기(400)로 제공한다. 오픈 소스 웹 서버는 World Wide Web을 통해 웹 페이지를 제공하도록 설계된 공개 도메인 소프트웨어이다.

[0073] 웹사이트 프레임워크 처리부(360)는 ODGP 사이트에서 사용자가 시각화를 위해 웹 브라우저에 웹사이트 URL(Uniform Resource Locator)을 입력할 수 있다. 웹 브라우저에 URL이 입력되면 웹사이트를 호스팅하는 서버인 웹사이트 프레임워크 처리부(360)는 리소스 또는 웹 페이지에 대한 URL 요청을 수신하고, 이에 대한 응답으로 웹 페이지를 사용자 단말기(400)의 브라우저로 전송한다.

[0074] ODGP 사이트는 도 4와 같이, OGDP에서 사용할 수 있는 데이터의 DCR을 자동으로 평가하는 Django 기반 프레임워크를 제공한다. Django 기반 프레임워크는 파이썬(Python)으로 만들어진 무료 오픈소스 웹 애플리케이션 프레임워크(web application framework)이다. 프레임워크는 애플리케이션 개발에 바탕이 되는 템플릿과 같은 역할을 하는 클래스들과 인터페이스 집합이다. 도 4는 도 3의 프레임워크를 구현한 것이다.

[0075] 웹사이트 프레임워크 처리부(360)는 웹 서버 소프트웨어와 파이썬으로 작성된 웹 응용 프로그램 간의 표준 인터페이스인 WSGI(Web Server Gateway Interface)를 제공한다. 표준 인터페이스는 여러 웹 서버에서 WSGI를 지원하는 응용 프로그램을 쉽게 사용할 수 있도록 한다. WSGI는 웹 서버와 웹 애플리케이션의 인터페이스를 위한 파이썬 프레임워크이다. 웹 서버(Web Server)는 HTTP를 통해 웹 브라우저에서 요청하는 HTML 문서나 오브젝트(이미지 파일 등)을 전송해주는 서비스 프로그램을 나타낸다. 웹 서버는 웹 브라우저와 같은 클라이언트로부터 HTTP 프로토콜로 요청을 받아 HTML 문서 등과 같은 정적 웹 페이지를 응답해준다.

[0077] 도 5는 본 발명의 실시예에 따른 API 목록과 CSV 형식의 API의 일례를 나타낸 도면이다.

[0078] 도 5의 (a)와 같이, API 저장부(320)는 공공 데이터 제공 서버(200)의 개방 공공 데이터 포털에 접속하여 표준 형식의 데이터를 담고 있는 API 목록을 수집한다.

[0079] API 목록은 도 5의 (a)와 같이, 플랫폼의 메타 데이터 섹션에서 찾을 수 있다. 도 5의 (a)는 공공 데이터 제공 서버(200)의 개방 공공 데이터 포털(웹 페이지)에 표시된 다양한 표준화된 형식으로 데이터를 다운로드하기 위한 API 목록을 보여주는 메타 데이터이다.

[0080] 도 5의 (a)는 "국가별 공연 정보 표준 자료"를 다양한 규격으로 다운로드할 수 있는 API 목록을 포함하는 메타 데이터 섹션을 보여준다.

[0081] API 저장부(320)는 API 목록에 해당하는 메타 데이터를 CSV 형식으로 다운로드할 수 있는 API(도 5의 (b))로 저장한다. 도 5의 (b)는 API 저장부(320)에 저장된 데이터 목록이다.

[0082] 자동 데이터 다운로드부(330)는 알고리즘 1을 기반으로 Python 코드를 작성하고, API 저장부(320)에 나열된 API를 사용하여 파일을 자동으로 다운로드한다.

[0083] DCR 계산부(340)는 다운로드된 파일을 Python으로 분석하여 알고리즘 2를 기반으로 지역별 DCR를 계산한다.

[0084] 데이터베이스부(342)는 전체 데이터 필드와 필수 데이터 필드를 포함하고, 지역 기반 DCR를 포함한 데이터 프레임을 저장하고 있다.

[0085] 웹사이트 프레임워크 처리부(360)는 데이터베이스부(342)에 저장된 데이터 프레임을 이용하여 동적 웹사이트인 OGDP의 생성을 단순화하는 Python 기반 웹 프레임워크를 제공한다.

[0087] 도 6은 본 발명의 실시예에 따른 전국 및 수도권 지역 데이터 파일의 분류를 나타낸 도면이고, 도 7은 본 발명의 실시예에 따른 전국 및 수도권 지역의 평균 DCR를 나타낸 도면이고, 도 8은 본 발명의 실시예에 따른 전국, 인천, 서울, 경기 지역의 일평균 DCR를 나타낸 도면이다.

[0088] 도 6에 도시된 바와 같이, 제어부(350)는 데이터베이스부(342)를 분석하여 DCR 개수의 다른 값에 대한 전국 및 수도권 지역 데이터 파일의 분류하여 디스플레이부(370)를 통해 출력한다. 제어부(350)는 완전성 비율 계산부(340)에서 계산한 DCR 값을 기준으로 DCR ≥ 90%인 경우, 파란색으로, 70% ≤ DCR < 90%인 경우, 녹색으로, 50% ≤ DCR < 70%인 경우, 노란색으로, DCR < 50%인 경우, 빨간색으로 히스토그램을 생성하여 디스플레이부(370)를 통해 표시한다.

[0089] 도 7에 도시된 바와 같이, 제어부(350)는 데이터베이스부(342)를 분석하여 전국 및 수도권 지역에 대하여 전체 데이터 필드 및 필수 데이터 필드를 고려한 평균 DCR를 계산하여 디스플레이부(370)를 통해 출력한다.

[0090] 도 8에 도시된 바와 같이, 제어부(350)는 데이터베이스부(342)를 분석하여 전국, 인천, 서울, 경기 지역의 일평균 DCR를 계산하여 디스플레이부(370)를 통해 출력한다.

[0092] 도 9는 본 발명의 실시예에 따른 개방 공공 데이터 포털의 일례를 나타낸 도면이고, 도 10은 본 발명의 실시예에 따른 개방 공공 데이터의 일례를 나타낸 도면이고, 도 11은 본 발명의 실시예에 따른 전체에 대해 나열된 파일의 DCR을 보여주는 대시보드를 나타낸 도면이다.

[0093] 웹사이트 프레임워크 처리부(360)는 개방 공공 데이터 포털(Open Government Data Portal, ODGP)을 생성하여 대시보드 홈페이지 형태로 시각화하여 디스플레이부(370)를 통해 출력한다. 여기서, 개방 공공 데이터 포털은 'cityview.inu.ac.kr'이라는 도메인 이름으로 나타낸다.

[0094] OGDP는 개방 공공 데이터를 다양한 형태의 대시보드를 제공할 수 있다.

[0095] 대시보드는 웹에서 한 화면에 다양한 정보를 중앙 집중적으로 관리하고 찾을 수 있도록 하는 사용자 인터페이스를 나타낸다.

[0096] 웹사이트 프레임워크 처리부(360)는 ODGP의 '개방 공공 데이터' 버튼 신호를 수신하면, 개방 공공 데이터를 도 10과 같은 항목이 있는 테이블 형태의 페이지로 생성하여 디스플레이부(370)를 통해 출력한다.

[0097] 도 10은 대한민국의 표준 공공 데이터 정보를 보여주는 대시보드를 나타낸 것이다.

[0098] 개방 공공 데이터는 파일 이름, 다운로드 링크, 데이터 완전성 비율(DCR)(10), 지도, 행이 비어 있는 파일의 헤더를 개수와 함께 표시하는 널 아이템 헤더(Null Item Headers)(20)를 행별로 이루어져 있다. 다운로드 링크는 선택 신호를 수신하면, 전체 데이터 필드를 포함한 데이터 프레임이 엑셀 파일로 연결된다.

[0099] 도 10 및 도 11에 도시된 바와 같이, DCR 값은 전체 데이터 필드의 제1 데이터 완전성 비율(수학식 1)을 나타낸다. 필수 데이터 필드의 제2 데이터 완전성 비율(수학식 2)을 나타내려면, 필수 데이터 필드의 DCR 값을 표시하는 항목을 클릭하면 된다.

[0100] 도 11은 전체에 대해 나열된 파일의 DCR(30)을 보여주는 대시보드이다. 여기서, DCR(30)은 변경 사항을 쉽게 식별할 수 있도록 다양한 배경색으로 표시된다.

[0101] 도 6에 도시된 바와 같이, 웹사이트 프레임워크 처리부(360)는 파일의 DCR(30)이 ≥ 90%인 경우, 가장 완전한 것으로 적용하고, DCR이 < 50%인 경우 가장 낮은 완성도가 적용된다.

[0102] 도 11에 도시된 바와 같이, 웹사이트 프레임워크 처리부(360)는 아래의 표 1과 같은 규칙에 따라 배경색을 다르

게 하여 DCR 값을 바 형태로 표시한다.

[0103] DCR(30)의 배경색은 아래의 표 1과 같은 규칙에 따라 변경된다.

표 1

[0104]

Condition	Background Color
DCR ≥ 90%	Blue
70% ≤ DCR < 90%	Green
50% ≤ DCR < 70%	Yellow
DCR < 50%	Red
No Data	Grey

[0105] 웹사이트 프레임워크 처리부(360)는 완전성 비율 계산부(340)에서 계산한 DCR 값을 기준으로 DCR ≥ 90%인 경우, 파란색으로, 70% ≤ DCR < 90%인 경우, 녹색으로, 50% ≤ DCR < 70%인 경우, 노란색으로, DCR < 50%인 경우, 빨간색을 DCR 값의 배경색으로 하고, 100%를 기준으로 DCR 값만큼 슬라이드바 형태로 표시한다.

[0106] 도 7은 지역별 93개 표준 파일의 평균 DCR을 보여준다. 그 결과, 전국(필수) 지역이 94.87%로 가장 높은 평균 DCR을 보인 반면, 전국(전체) 지역이 86.63%로 가장 낮은 평균 DCR을 보임을 알 수 있다.

[0107] 93개 표준화된 파일에서 필수 데이터 필드에 비해서 비필수 데이터 필드가 불완전한 것으로 유추할 수 있다.

[0108] 도 8은 전국, 인천, 서울, 경기 지역의 전체 및 의무적 사례에 대한 1일 평균 DCR을 나타낸 것이다.

[0109] 웹사이트 프레임워크 처리부(360)는 DCR(30)의 결과가 개방 공공 데이터 포털(Open Government Data Portal, ODGP)을 사용하여 실시간으로 시각화한다.

[0110] 웹사이트 프레임워크 처리부(360)는 ODGP에 표시된 DCR 결과를 모니터링하여 개방 공공 데이터의 품질을 개선할 수 있다.

[0112] 도 12는 본 발명의 실시예에 따른 개방 공공 데이터의 DCR 값을 기준으로 정렬된 파일을 보여주는 대시보드를 나타낸 도면이고, 도 13은 본 발명의 실시예에 따른 행이 비어 있는 파일의 헤더를 개수와 함께 표시하는 대시보드를 나타낸 도면이다.

[0113] 도 12에 도시된 바와 같이, 웹사이트 프레임워크 처리부(360)는 ODGP에 저장된 파일의 DCR 값을 기준으로 DCR 값을 오름차순으로 정렬하고, 정렬한 DCR 값에 대응하는 개방 공공 데이터를 ODGP에 업로드하여 디스플레이부(370)를 통해 출력한다. DCR 값은 ODGP에 저장된 파일의 품질을 정량화하는데 사용한다.

[0114] ODGP에 저장된 파일은 API를 통해 다운로드할 수 있고, 테이블 형식의 데이터 파일에 적용할 수 있다.

[0115] 도 13에 도시된 바와 같이, 웹사이트 프레임워크 처리부(360)는 ODGP의 '널 아이템 헤더' 버튼 신호를 수신하면, 행이 비어 있는 파일의 헤더를 개수(Null Value)와 함께 테이블 형태의 페이지로 생성하여 디스플레이부(370)를 통해 출력한다.

[0116] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

부호의 설명

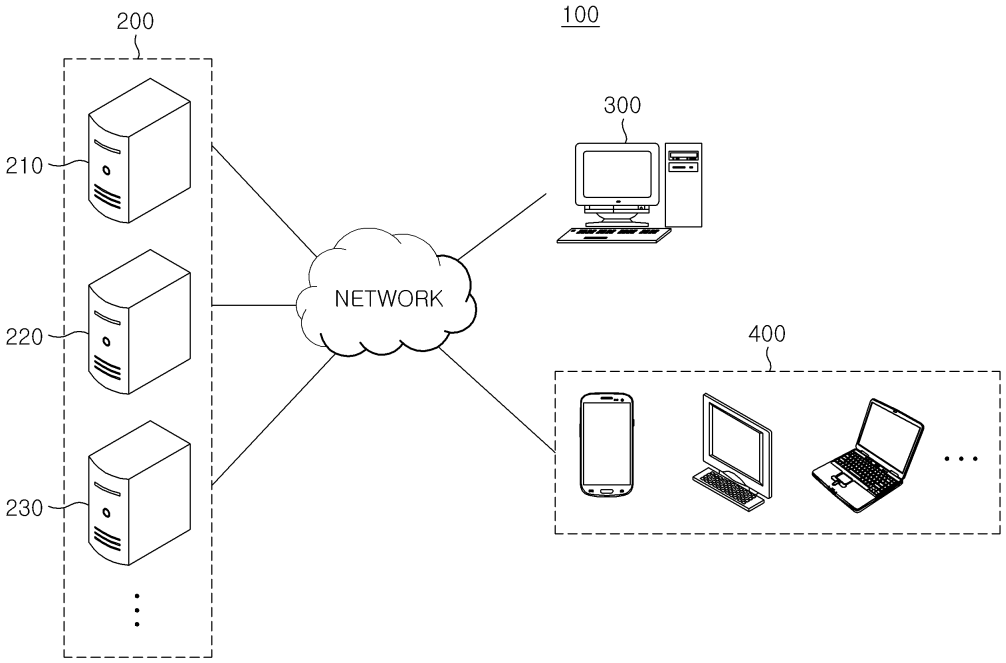
[0117] 100: 개방 공공 데이터 관리 장치 200: 공공 데이터 제공 서버

300: 개방 공공 데이터 인터페이스 서버 310: 데이터 수신부

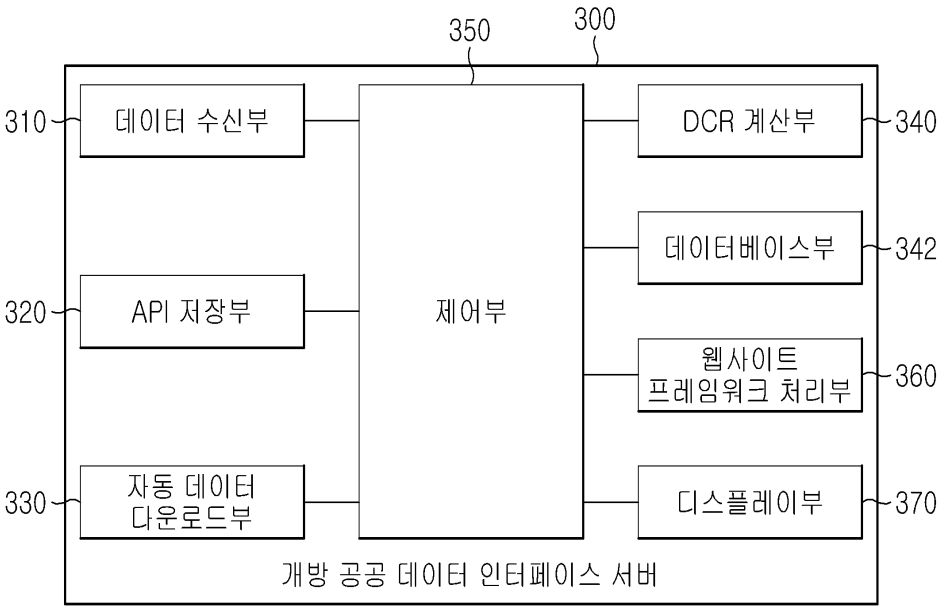
320: API 저장부
330: 자동 데이터 다운로드부
340: DCR 계산부
342: 데이터베이스부
350: 제어부
360: 웹사이트 프레임워크 처리부
370: 디스플레이부
400: 사용자 단말기

도면

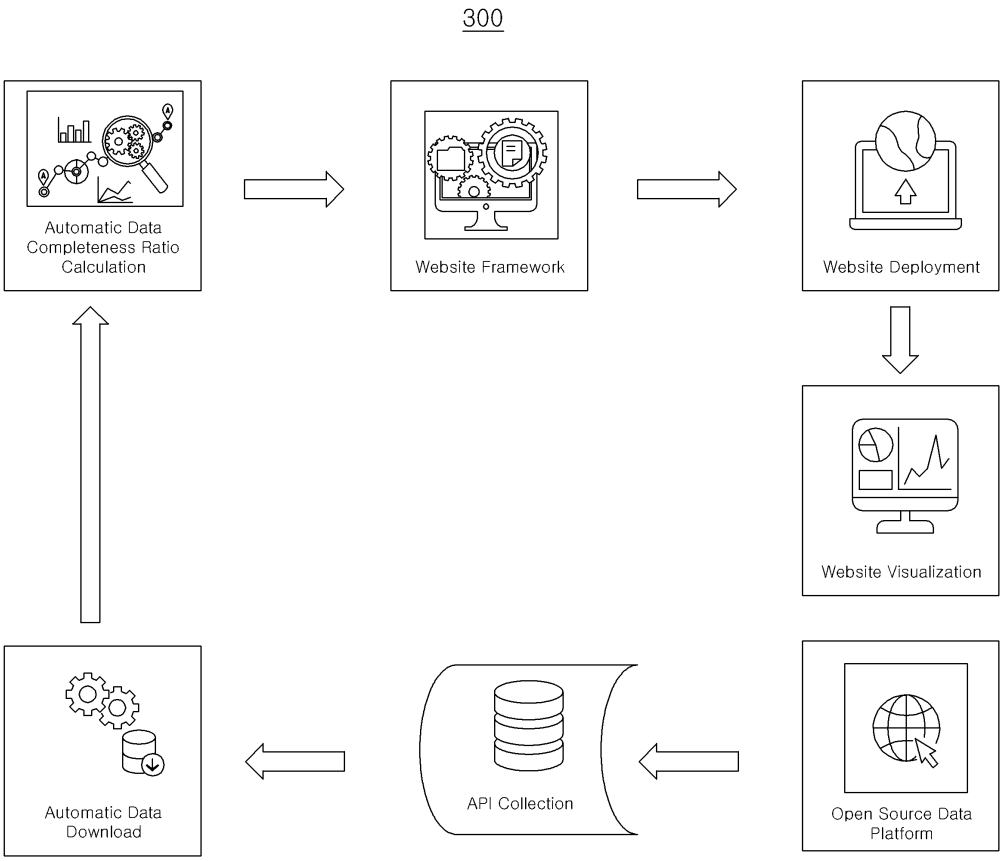
도면1



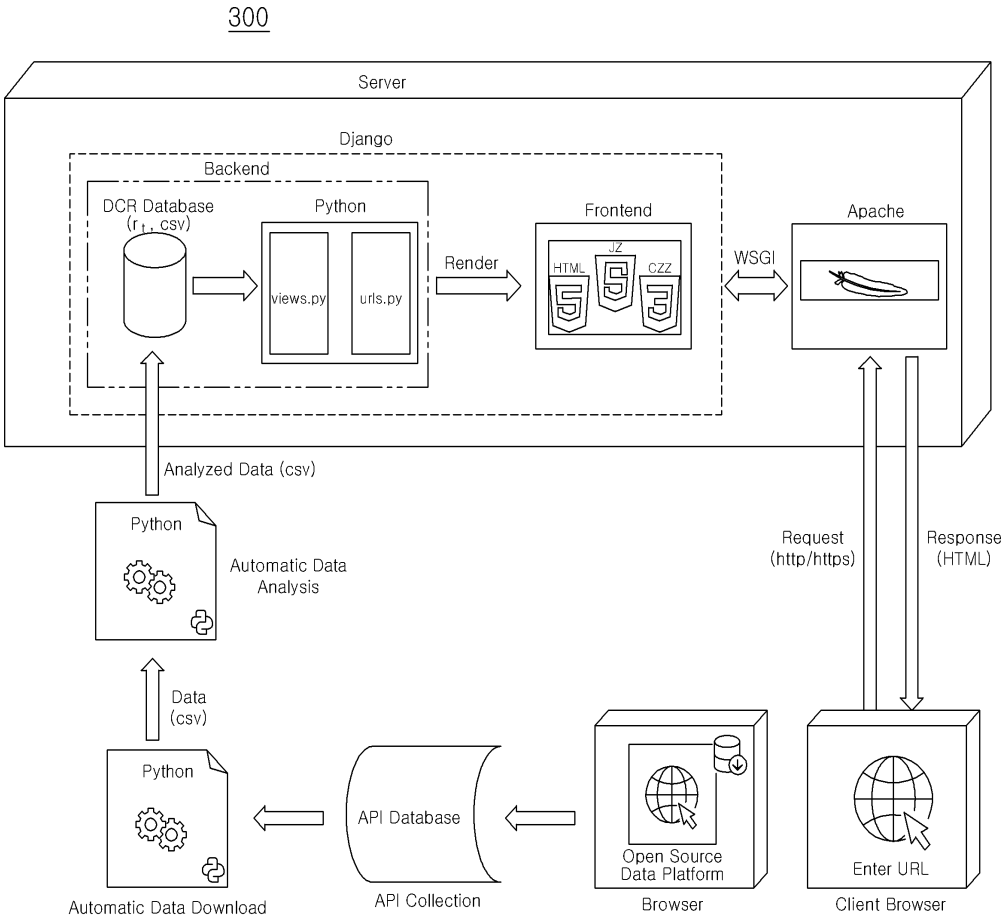
도면2



도면3


300
Automatic Data Completeness Ratio Calculation
Website Framework
Website Deployment
Website Visualization
Open Source Data Platform
API Collection
Automatic Data Download

도면4


300
Server
Django
Backend
DCR Database
(r t , csv)
Python
views.py
urls.py
Render
Frontend
HTML
JZ
CZZ
WSGI
Apache
Analyzed Data (csv)
Python
Automatic Data
Analysis
Request
(http/https)
Response
(HTML)
Data
(csv)
Python
Automatic Data Download
API Database
API Collection
Open Source
Data Platform
Browser
Enter URL
Client Browser

도면5

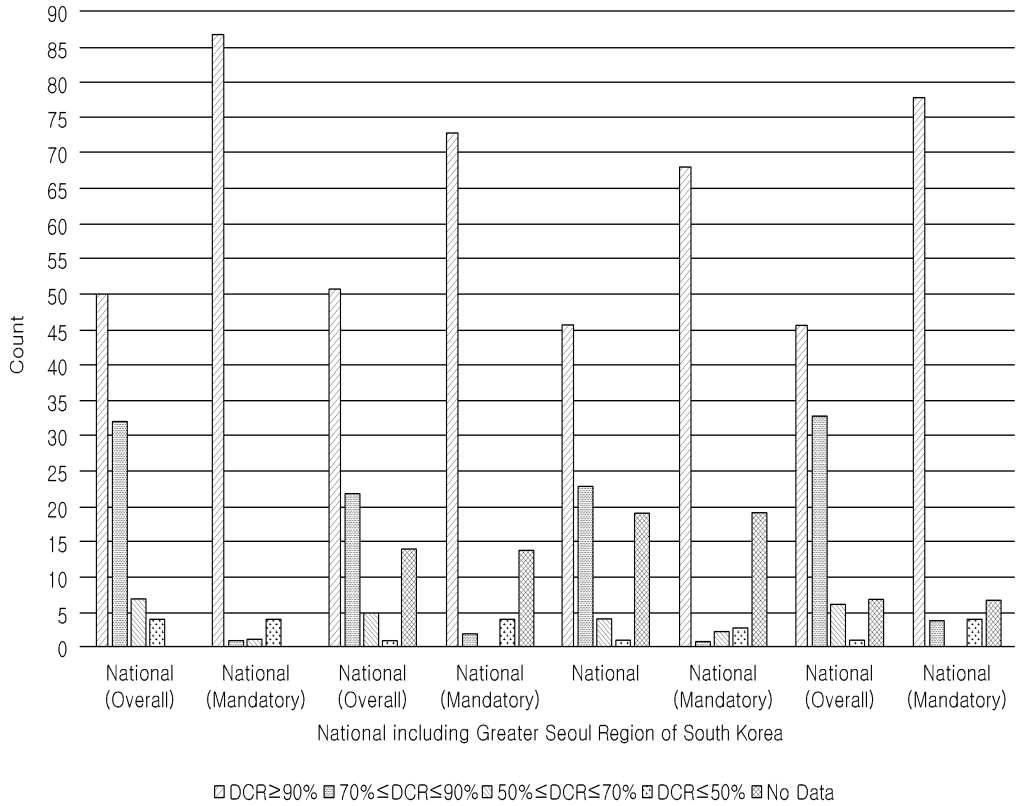
{"name": "National performance event information standard data","description":"Provides per-formance event information
(performance information, venue, schedule, fee information, etc.).
Public data open standard data attribute information (expression format/unit) etc.), please refer to the [Public Data Open Standards] Notice.
(Information Sharing>Resources>Laws (No-tice)/Guidelines).
The standard data registered by each institution is collected and provided. The update cycle of standard data is different for each file.
(Data registered by institutions is collected and the update time for each file is different)",
"url":"https://www.data.go.kr/data/15013106/standard.do","keywords": ["Art, Performance, Event"],
"license":"https://data.go.kr/ugs/selectPortalPolicyView.do","distribution":
[{"encodingFormat":"CSV","contentUrl": "https://www.data.go.kr/download/15013106/standard.do?dataType=csv","@type":"DataDownload"},
{"encodingFormat":"XML","contentUrl":" https://www.data.go.kr/download/15013106/standard.do?dataType=xml","@type":"DataDownload"},
{"encodingFor-mat":"JSON","contentUrl":"https://www.data.go.kr/download/15013106/standard.do ?dataType=json","@type":"DataDownload"},
{"encodingFor-mat":"RDF","contentUrl":"https://www.data.go.kr/download/15013106/standard.do? dataType=rdf","@type":"DataDownload"},
{"encodingFor-mat":"XLS","contentUrl":"https://www.data.go.kr/download/15013106/standardExcel.do", "@type":"DataDownload"}],
"@context":"https://schema.org","@type":"Dataset"}

(a)

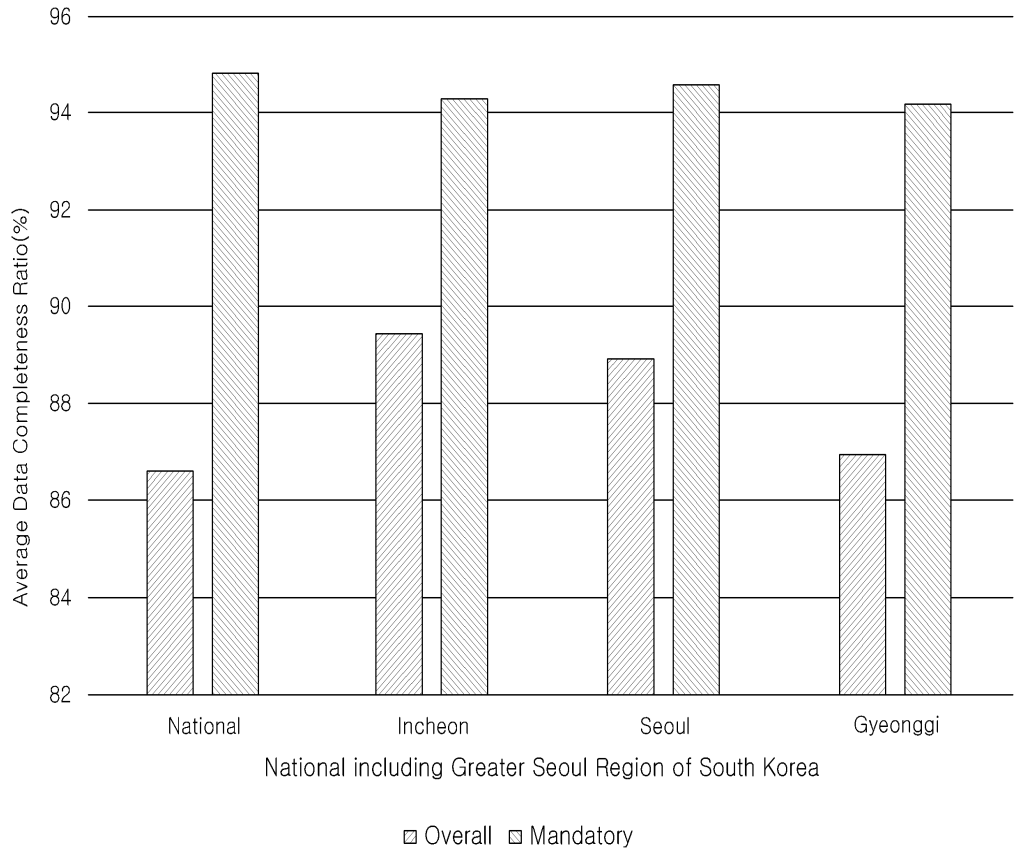
SN	English Name	API
1	Standard Data for Protection Zones for the Elderly and Disabled Nationwide	https://www.data.go.kr/download/15034532/standard.do?dataType=csv (accessed on 12 July 2021)
2	National Performance Event Information Standard Data	https://www.data.go.kr/download/15013106/standard.do?dataType=csv (accessed on 12 July 2021)
3	National Library Standard Data	https://www.data.go.kr/download/15013109/standard.do?dataType=csv (accessed on 12 July 2021)
4	National Automobile Maintenance Company Standard Data	https://www.data.go.kr/download/15028204/standard.do?dataType=csv (accessed on 12 July 2021)
5	National Lifelong Learning Course Standard Data	https://www.data.go.kr/download/15013110/standard.do?dataType=csv (accessed on 12 July 2021)
⋮		
91	National Elementary School Commuting Area Standard Data	https://www.data.go.kr/download/15021149/standard.do?dataType=csv (accessed on 12 July 2021)
92	National Elementary and Secondary School Location Standard Data	https://www.data.go.kr/download/15021148/standard.do?dataType=csv (accessed on 12 July 2021)
93	National Recreational Forest Standard Data	https://www.data.go.kr/download/15013111/standard.do?dataType=csv (accessed on 12 July 2021)

(b)

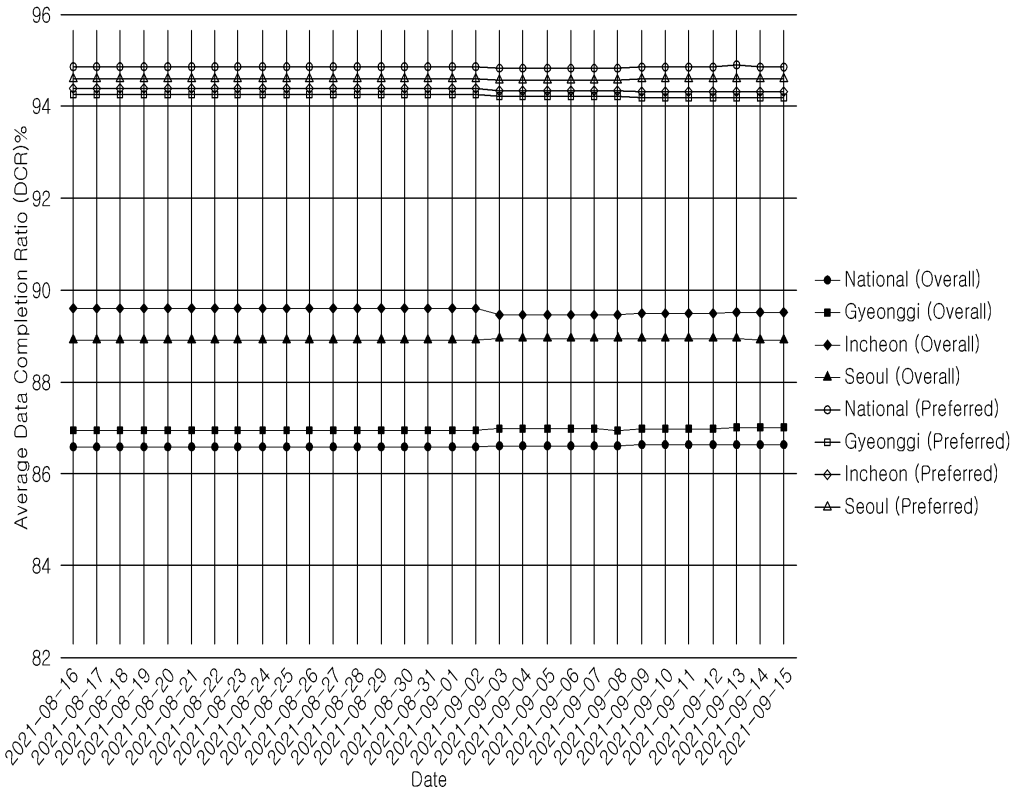
도면6



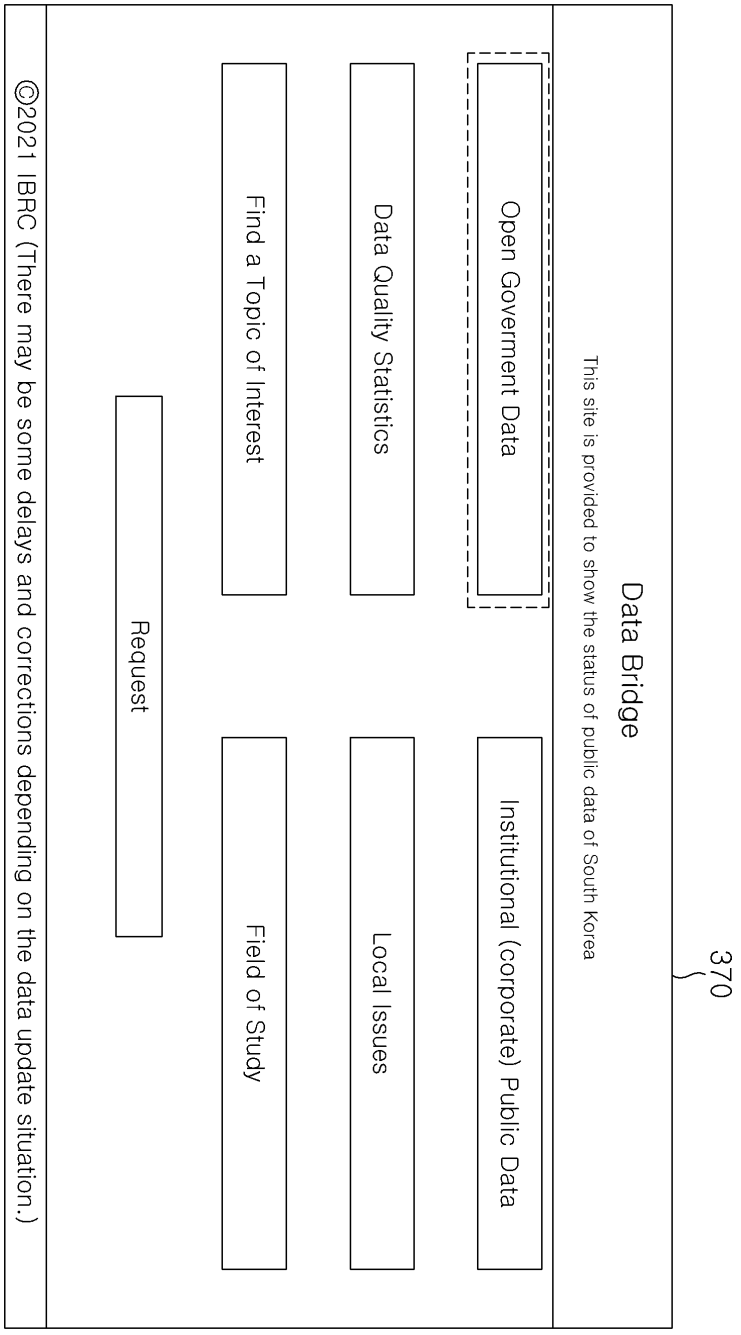
도면7


96
94
92
90
88
86
84
82
Average Data Completeness Ratio(%)
National
Incheon
Seoul
Gyeonggi
National including Greater Seoul Region of South Korea
Overall
Mandatory

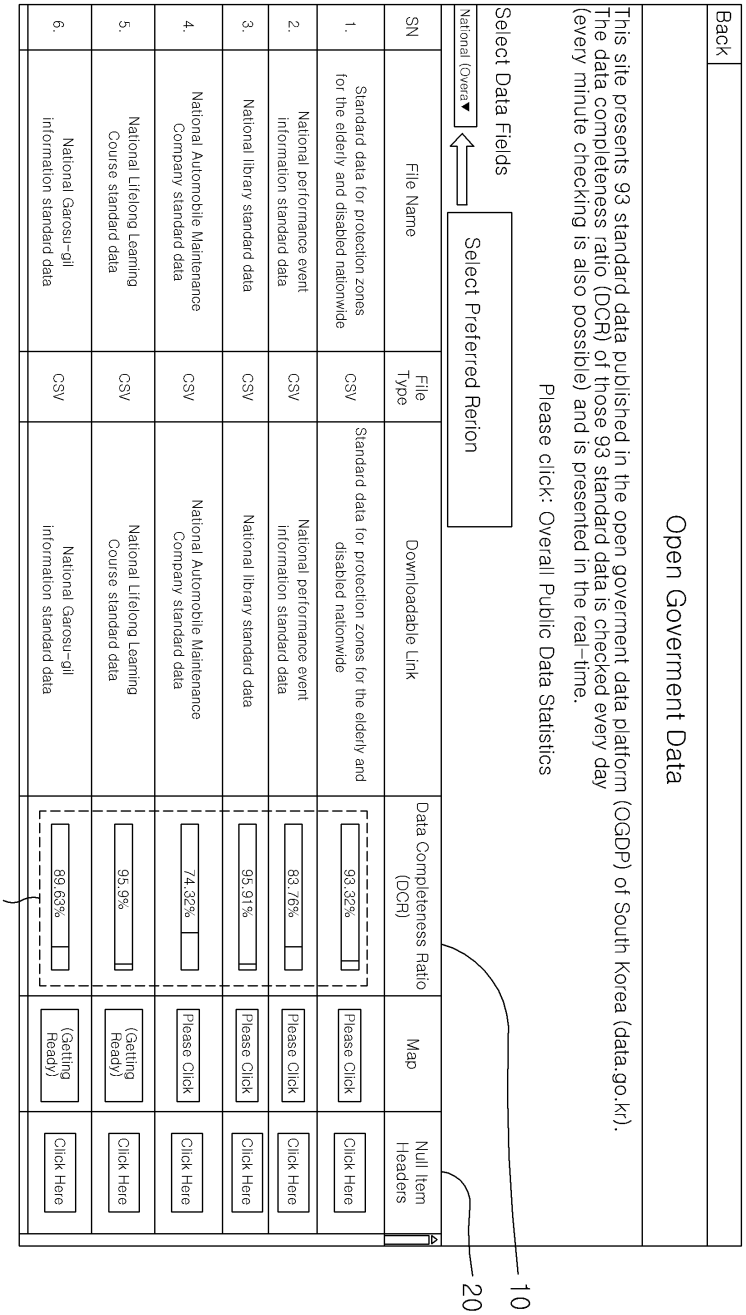
도면8


96
94
92
90
88
86
84
82
Average Data Completion Ratio (DCR)%
2021-08-16
2021-08-17
2021-08-18
2021-08-19
2021-08-20
2021-08-21
2021-08-22
2021-08-23
2021-08-24
2021-08-25
2021-08-26
2021-08-27
2021-08-28
2021-08-29
2021-08-30
2021-08-31
2021-09-01
2021-09-02
2021-09-03
2021-09-04
2021-09-05
2021-09-06
2021-09-07
2021-09-08
2021-09-09
2021-09-10
2021-09-11
2021-09-12
2021-09-13
2021-09-14
2021-09-15
Date
National (Overall)
Gyeonggi (Overall)
Incheon (Overall)
Seoul (Overall)
National (Preferred)
Gyeonggi (Preferred)
Incheon (Preferred)
Seoul (Preferred)

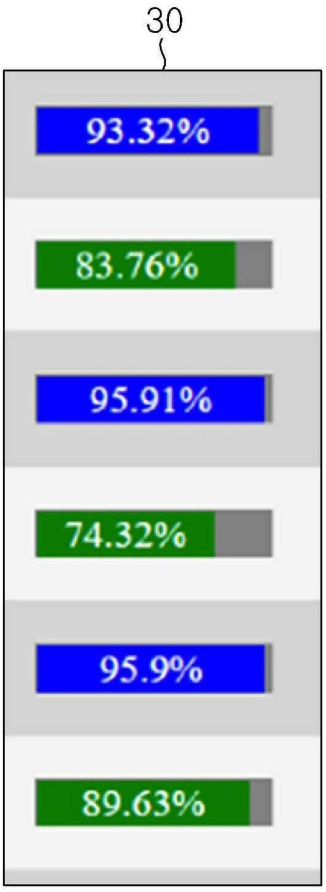
도면9


Data Bridge
This site is provided to show the status of public data of South Korea
370
Open Goverment Data
Institutional (corporate) Public Data
Data Quality Statistics
Local Issues
Find a Topic of Interest
Field of Study
Request
©2021 IBRC (There may be some delays and corrections depending on the data update situation.)

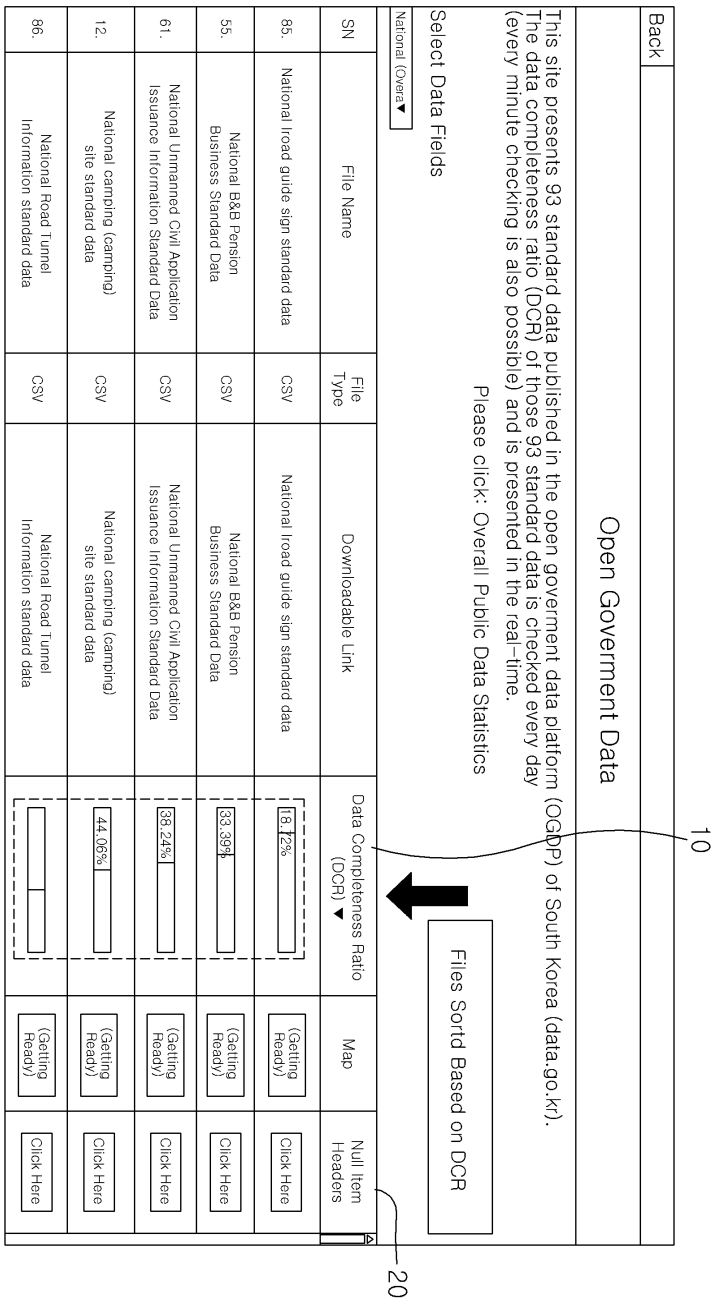
도면10


Back
Open Goverment Data
This site presents 93 standard data published in the open goverment data platform (OGDP) of South Korea (data.go.kr).
The data completeness ratio (DCR) of those 93 standard data is checked every day
(every minute checking is also possible) and is presented in the real-time.
Please click: Overall Public Data Statistics
Select Data Fields
National (Overa
Select Preferred Rerion
SN
File Name
File Type
Downloadable Link
Data Completeness Ratio (DCR)
Map
Null Item Headers
1.
Standard data for protection zones for the elderly and disabled nationwide
CSV
Standard data for protection zones for the elderly and disabled nationwide
93.32%
Please Click
Click Here
2.
National performance event information standard data
CSV
National performance event information standard data
83.76%
Please Click
Click Here
3.
National library standard data
CSV
National library standard data
95.91%
Please Click
Click Here
4.
National Automobile Maintenance Company standard data
CSV
National Automobile Maintenance Company standard data
74.32%
Please Click
Click Here
5.
National Lifelong Learning Course standard data
CSV
National Lifelong Learning Course standard data
95.9%
(Getting Ready)
Click Here
6.
National Garosu-gil information standard data
CSV
National Garosu-gil information standard data
89.63%
(Getting Ready)
Click Here
10
20
30

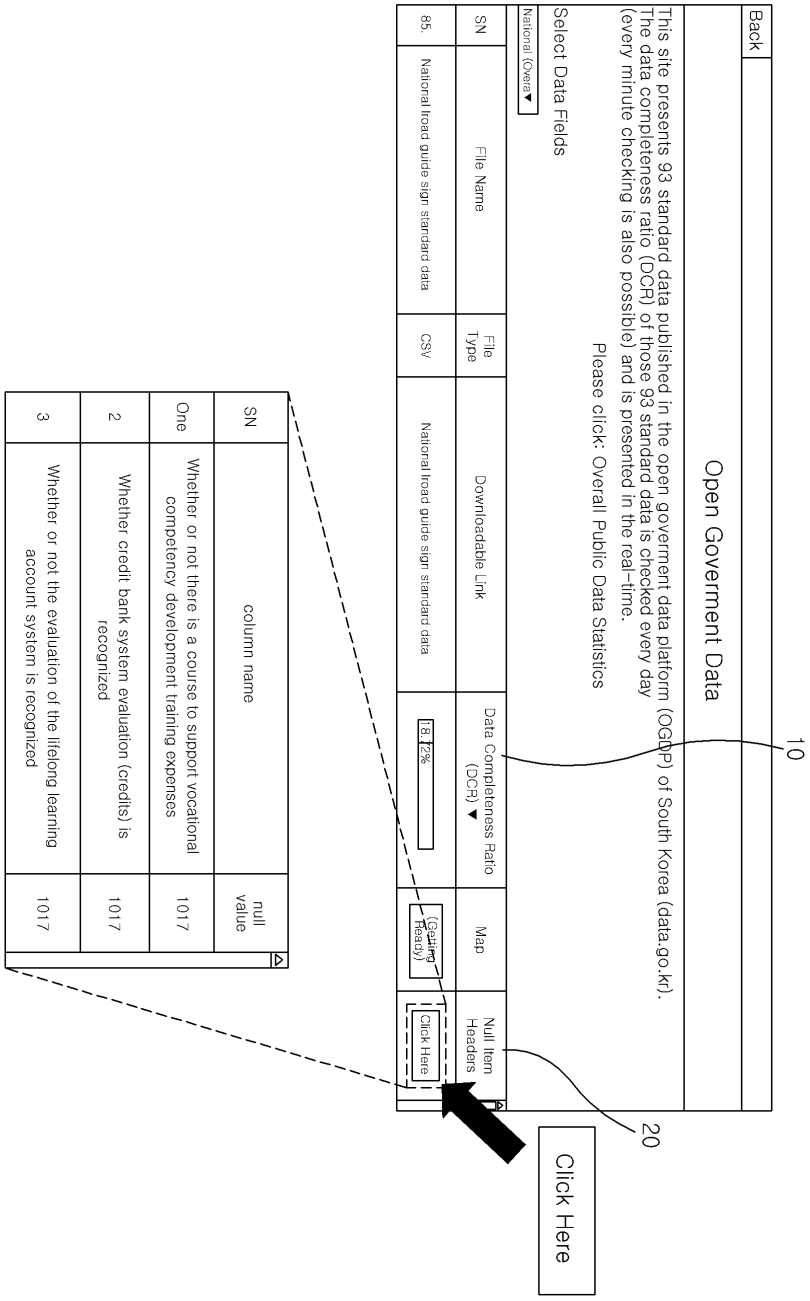
도면11



도면12


Back
Open Goverment Data
10
This site presents 93 standard data published in the open goverment data platform (OGDP) of South Korea (data.go.kr).
The data completeness ratio (DCR) of those 93 standard data is checked every day
(every minute checking is also possible) and is presented in the real-time.
Please click: Overall Public Data Statistics
Files Sortd Based on DCR
Select Data Fields
National (Overa
20
SN
File Name
File Type
Downloadable Link
Data Completeness Ratio (DCR)
Map
Null Item Headers
85.
National road guide sign standard data
CSV
National road guide sign standard data
18.72%
(Getting Ready)
Click Here
55.
National B&B Pension Business Standard Data
CSV
National B&B Pension Business Standard Data
33.39%
(Getting Ready)
Click Here
61.
National Unmanned Civil Application Issuance Information Standard Data
CSV
National Unmanned Civil Application Issuance Information Standard Data
38.24%
(Getting Ready)
Click Here
12.
National camping (camping) site standard data
CSV
National camping (camping) site standard data
44.06%
(Getting Ready)
Click Here
86.
National Road Tunnel Information standard data
CSV
National Road Tunnel Information standard data
(Getting Ready)
Click Here

도면13


Back
Open Goverment Data
10
This site presents 93 standard data published in the open goverment data platform (OGDP) of South Korea (data.go.kr).
The data completeness ratio (DCR) of those 93 standard data is checked every day
(every minute checking is also possible) and is presented in the real-time.
Please click: Overall Public Data Statistics
Select Data Fields
National (Overa
SN
File Name
File Type
Downloadable Link
Data Completeness Ratio (DCR)
Map
Null Item Headers
85
National Road guide sign standard data
CSV
National Road guide sign standard data
18.12%
(Getting Ready)
Click Here
20
Click Here
SN
column name
null value
One
Whether or not there is a course to support vocational competency development training expenses
1017
2
Whether credit bank system evaluation (credits) is recognized
1017
3
Whether or not the evaluation of the lifelong learning account system is recognized
1017