

영등포구 청년을 위한

# 생성형 AI 활용 데이터 시각화 교육

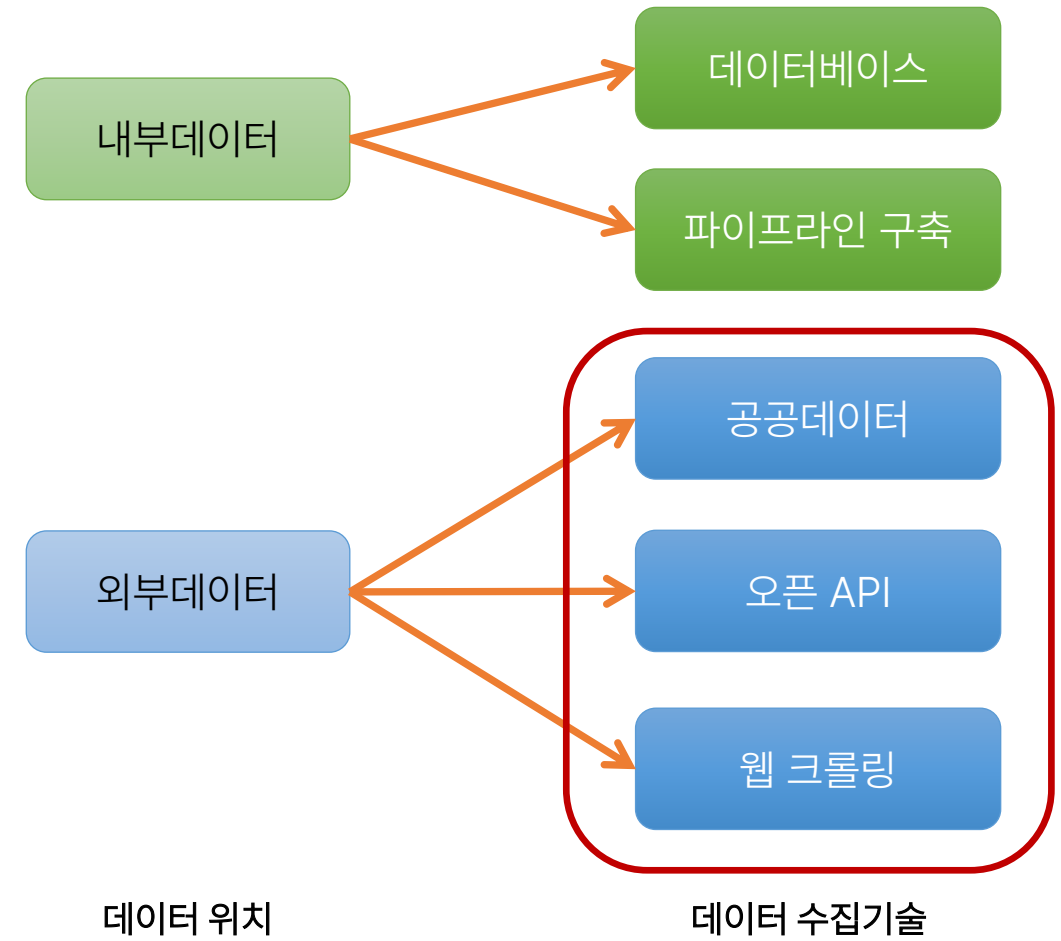
데이터 수집



## 데이터 수집 방법

### ▪ 어디서, 어떻게 데이터를 수집할 것인가?

- 온라인을 통한 데이터 수집(공공데이터, 크롤링, API 등)
- 설문 조사 등을 통한 데이터 수집(구글 설문지, 오프라인 설문지 등)
- 이미 가지고 있는 데이터(회사 내부 데이터, 고객 데이터 등)
- 데이터 수집, 저장을 위한 개발이 필요(로그 데이터 등)



## 크롤링

- 크롤링 – 웹 크롤러가 일정 규칙으로 웹페이지를 브라우징 하는 것
- 스크래핑 – 웹페이지 상에서 원하는 데이터를 추출하는 행위
- 크롤링을 허용한다고 해도, 해당 내용의 저작권까지 자유로운 것은 아니기 때문에 논문, 사진, 기사 등의 데이터를 활용할 때는 주의
- 무단으로 크롤링하는 것은 불법이 될 수 있음
- 해당 사이트의 약관이나 '사이트주소/robots.txt'를 확인 후 크롤링 권장
  - <https://www.saramin.co.kr/robots.txt>

## 크롤링

### ▪ User Agent(사용자 에이전트)

- 사용자를 대신하여 일을 수행하는 소프트웨어 에이전트
- 사용자가 어떤 플랫폼으로 접속했는지를 알려줌(모바일, 데스크탑 등)
- <http://www.useragentstring.com/> 접속

[Home](#) | [List of User Agent Strings](#) | [Links](#) | [API](#) |

---

### User Agent String explained :

Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/79.0.3945.88 Safari/537.36

Copy/paste any user agent string in this field and click 'Analyze'

Analyze

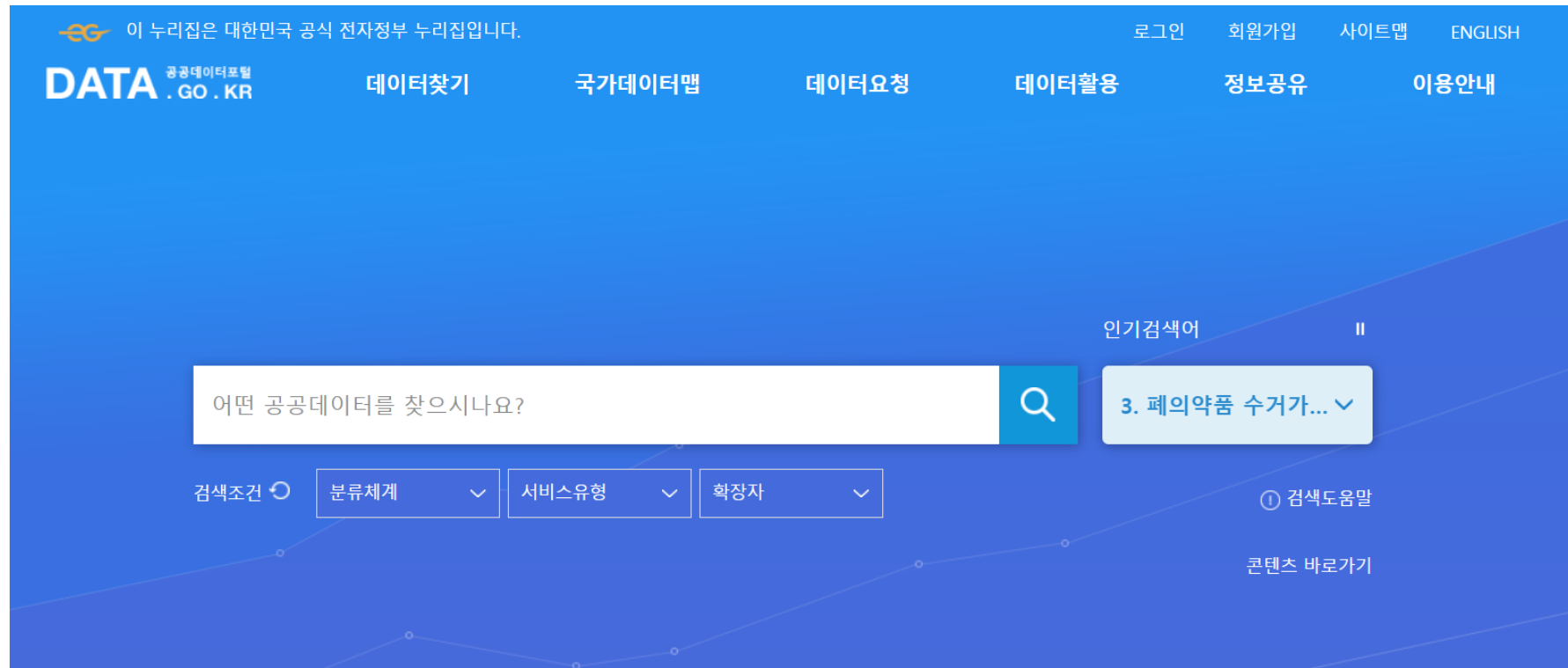
 <b>Chrome 79.0.3945.88</b>	
<b>Mozilla</b>	MozillaProductSlice. Claims to be a Mozilla based user agent, which is only true for Gecko browsers like Firefox and Netscape. For all other user agents it means 'Mozilla-compatible'. In modern browsers, this is only used for historical reasons. It has no real meaning anymore.

## 크롤링 실습

- 로또 1등 배출점 데이터 분석해보기
  - <https://www.dhlottery.co.kr/store.do?method=topStore&pageGubun=L645>
- 로또 1등 당첨이 많은 소재지 파악하기
- 자동/수동/반자동 중에 어떤 방식으로 구매하는 것이 1등 당첨이 높을지 확인하기

## 공공데이터 포털

- 공공기관에서 생성 또는 취득하여 관리하고 있는 공공데이터를 제공하는 사이트
  - <https://www.data.go.kr>



## 서울열린데이터광장

### ■ 서울시와 관련된 다양한 공공데이터를 제공하는 사이트

- <https://data.seoul.go.kr>

서울특별시

전기자동차 안전운행 교육 전기자동차 화재는 멀리, 예방은 가까이

서울소식

응답소

정보공개

분야별정보

서울 열린데이터 광장

공공데이터

통계

서울빅데이터

소식&참여

이용안내

로그인 | 회원가입 | 사이트맵

찾고 싶은 데이터를 입력해 주세요.

Q 검색

지하철

인구

노인

공공자전거

시설

2024년 7월

서울시  
등록 인구현황



2024년 7월 기준 서울시 등록 인구현황

1/6 < || >

서울시민을 위한 공공데이터

열린데이터광장은 서울시 및 자치구, 그 산하기관에서 보유하고 있는 공공데이터를 지속적으로 발굴·개방하고 있습니다.

공공데이터  
7,919

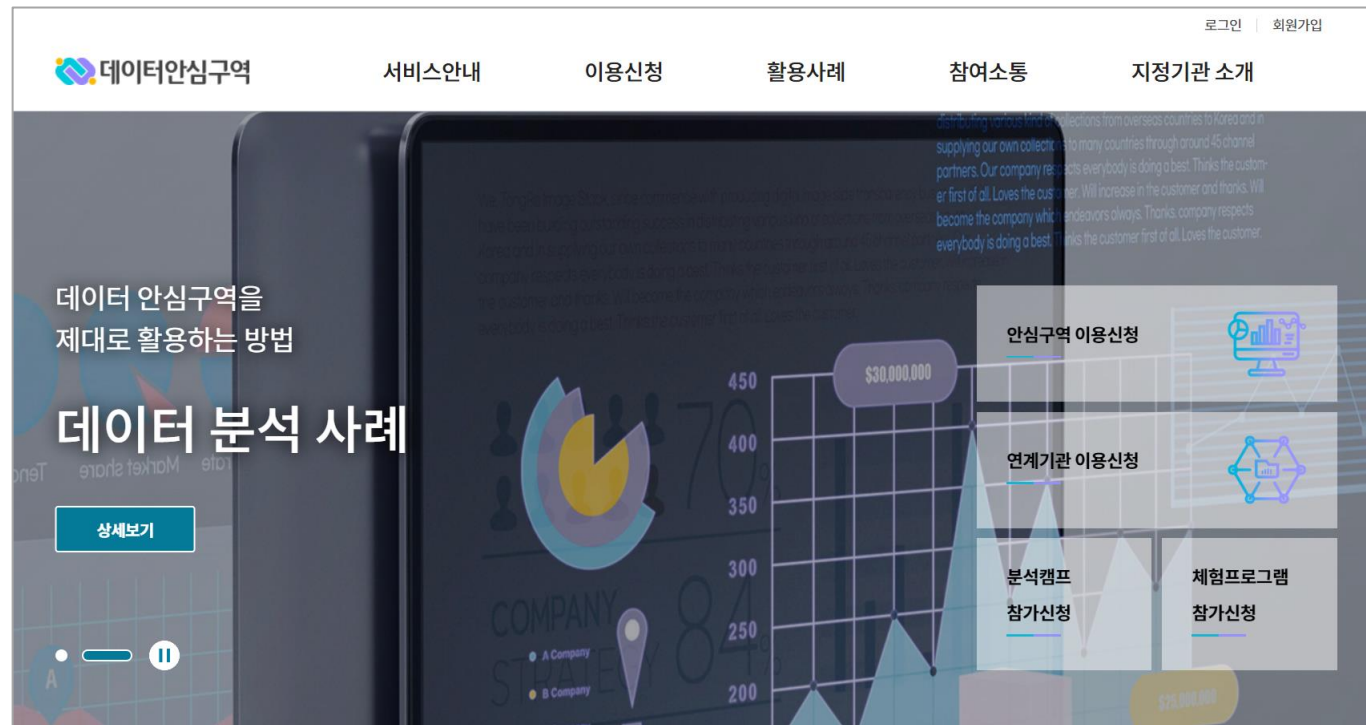
서비스  
16,026

AI학습데이터  
17

OpenAPI	5,592	SHEET	7,315
FILE	905	LOD	105
CHART	1,867	MAP	126
LINK	116		

## 데이터안심구역

- 미개방 데이터를 안정적으로 분석&활용 할 수 있는 플랫폼
  - <https://dsz.kdata.or.kr/>
- 쉽게 접할 수 없는 공공기관 및 민간기업의 미개방 데이터를 누구나 안전한 분석환경에서 무료로 활용 가능





## 오픈API 실습

- 공공데이터 포털(<https://www.data.go.kr>) 회원가입 & 로그인 진행
- [데이터찾기] 메뉴에서 [데이터 목록] 선택
- 파일데이터, 오픈API등의 형태로 데이터셋 제공

전체(92,167건)
파일데이터(68,938건)
오픈 API(11,488건)
표준데이터셋218개(9,957건)
연계데이터(1,784건)

수정일자순 ▼
10개씩 ▼
정렬

파일데이터 (68,938건)

사회복지
자치행정기관
미리보기

CSV
JSON + XML
인천광역시 강화군\_노인복지시설 현황
New

인천광역시 강화군 관내 노인복지시설 현황 데이터입니다.연번, 종류(노인복지법에 따른 구분), 유형, 시설명, 주소, 위도, 경도, 정원(수용인원), 데이터기준일자 항목을 제공하고 있습니다...

제공기관 인천광역시 강화군 수정일 2024-08-27 조회수 1 다운로드 0 키워드 노인복지시설,경로당,노인요양시설

다운로드

환경기상
자치행정기관
미리보기

CSV
JSON + XML
충청북도 단양군\_사업장폐기물배출자 신고현황
Update

충청북도 단양군 사업장폐기물배출자 신고현황 데이터로 폐기물구분, 상호, 사업자등록번호, 전화번호, 신고일, 폐기물 종류, 운반자, 처리업소명, 처리방법, 사업장도로명주소, 데이터 ...

제공기관 충청북도 단양군 수정일 2024-08-27 조회수 1968 다운로드 1267 키워드 사업장명,폐기물종류,배출량

다운로드

환경기상
자치행정기관
미리보기

CSV
JSON + XML
충청북도 단양군\_음식물류폐기물 다량배출사업장 현황
Update

충청북도 단양군 음식물류폐기물 다량배출사업장 현황 데이터로 사업장명, 사업장주소, 사업장전화번호 등의 항목을 포함함.

제공기관 충청북도 단양군 수정일 2024-08-27 조회수 915 다운로드 237 키워드 사업장명,사업장주소,사업장전화번호

다운로드

전체(92,167건)
파일데이터(68,938건)
오픈 API(11,488건)
표준데이터셋218개(9,957건)
연계데이터(1,784건)

수정일자순 ▼
10개씩 ▼
정렬

오픈 API (11,488건)

재난안전
국가행정기관
미리보기

JSON
행정안전부\_긴급재난문자
New

재난 및 민방공 상황발생으로 인한 인명·재산 피해가 예상되는 때, 그 피해를 예방하기 위하여 사용기관에서 발송한 재난문자 정보를 제공합니다. 공공데이터포털(data.go.kr)에서 재...

제공기관 행정안전부 수정일 2024-08-27 조회수 1 활용신청 0 키워드 재해문자,경보발령,발령내용

바로가기

산업고용
공공기관
미리보기

JSON
한국석유관리원\_수소충전소\_부대시설
Update

수소충전소 내 세차장 여부, 경정비 여부, 편의점 여부, 화장실 여부, 대기실 여부, 카페 여부 등 부대시설 정보 데이터

제공기관 한국석유관리원 수정일 2024-08-27 조회수 1 활용신청 1 키워드 수소,하일,수소유통

활용신청

산업고용
공공기관
미리보기

JSON
한국석유관리원\_수소충전소\_실시간정보
Update

수소트루트래일러 압력, 완충가능한 수소차량 대수, 수소차량 대기차수, 수소차량 혼잡상태 등 수소충전소 실시간 정보 데이터

제공기관 한국석유관리원 수정일 2024-08-27 조회수 1 활용신청 1 키워드 수소,수소유통정보시스템,실시간

활용신청

# 오픈API 실습

## 파일 데이터의 경우 CSV 파일 다운로드를 통해 이용 가능

서울특별시 영등포구\_의류수거함 위치현황

서울특별시 영등포구 의류수거함 위치현황입니다.

제공데이터: 행정동, 도로명주소, 지번주소, 위도, 경도, 데이터기준일자

0

0

관심

파일데이터

오픈API

주천데이터

공공데이터활용지원센터는 공공데이터포털에 개방되는 3단계 이상의 오픈 포맷 파일데이터를 오픈 API(RestAPI 기반의 JSON/XML)로 자동변환하여 제공합니다.

오픈 API를 활용하기 위해서는 공공데이터포털 회원 가입 및 활용신청이 필요하며, 활용 관련 문의는 공공데이터활용지원센터로 연락주시기 바라며,

데이터 자체에 대한 문의는 아래 제공기관의 관리부서 전화번호로 연락주시기 바랍니다.

파일데이터는 로그인 없이 다운로드를 통해 이용하실 수 있습니다.

csv

서울특별시 영등포구\_의류수거함 위치현황

다운로드

파일데이터 정보

메타데이터 다운로드

데이터 항목(컬럼) 정보

데이터 개신요청

오류신고 및 문의

행정동	도로명주소	지번주소	위도	경도	데이터기준일자
영등포본동	서울특별시 영등포구 도신로51길 9	서울특별시 영등포구 신길동 186-230	37.51246	126.9134	2023-09-01
영등포본동	서울특별시 영등포구 영신로9라길 6	서울특별시 영등포구 영등포동 631-8	37.51284	126.9059	2023-09-01
영등포본동	서울특별시 영등포구 신길로61길 17	서울특별시 영등포구 영등포동 592-70	37.51467	126.9093	2023-09-01
영등포본동	서울특별시 영등포구 신길로 276	서울특별시 영등포구 영등포동 585-25	37.51611	126.9127	2023-09-01
영등포본동	서울특별시 영등포구 영등포로60길 29	서울특별시 영등포구 신길동 35	37.5168	126.915	2023-09-01
영등포본동	서울특별시 영등포구 영등포로60길 17	서울특별시 영등포구 신길동 31-20	37.51681	126.9164	2023-09-01
영등포본동	서울특별시 영등포구 영등포로62길 34	서울특별시 영등포구 영등포동 576-9	37.5164	126.9144	2023-09-01
영등포본동	서울특별시 영등포구 신길로62나길 9	서울특별시 영등포구 신길동 184-13	37.51476	126.916	2023-09-01
영등포본동	서울특별시 영등포구 신길로62길 37	서울특별시 영등포구 신길동 187-14	37.51429	126.9155	2023-09-01
영등포본동	서울특별시 영등포구 신길로60다길 11	서울특별시 영등포구 신길동 189-25	37.5138	126.914	2023-09-01
영등포본동	서울특별시 영등포구 신길로60길 25	서울특별시 영등포구 신길동 186-182	37.51304	126.9131	2023-09-01
영등포본동	서울특별시 영등포구 신길로62길 17	서울특별시 영등포구 신길동 185-1	37.51485	126.9137	2023-09-01
영등포본동	서울특별시 영등포구 도신로65길 5	서울특별시 영등포구 신길동 169-10	37.51295	126.9169	2023-09-01
영등포본동	서울특별시 영등포구 도신로53길 14	서울특별시 영등포구 신길동 159-1	37.512	126.915	2023-09-01
영등포본동	서울특별시 영등포구 도신로 191	서울특별시 영등포구 신길동 219-27	37.51141	126.9123	2023-09-01
영등포본동	서울특별시 영등포구 신길로60길 32	서울특별시 영등포구 신길동 186-268	37.51326	126.9137	2023-09-01
영등포본동	서울특별시 영등포구 영신로20가길 11	서울특별시 영등포구 영등포동 592-57	37.51468	126.909	2023-09-01
영등포본동	서울특별시 영등포구 영등포로60길 25	서울특별시 영등포구 신길동 34-5	37.51683	126.9155	2023-09-01
영등포본동	서울특별시 영등포구 영등포로60길 7	서울특별시 영등포구 신길동 46-9	37.51617	126.9167	2023-09-01

- 10 -

## 오픈API 실습

- [https://www.data.go.kr/data/15059126/openapi.do?recommendDataYn=Y#tab\\_layer\\_detail\\_function](https://www.data.go.kr/data/15059126/openapi.do?recommendDataYn=Y#tab_layer_detail_function)
- 활용 신청

오픈API 상세

XML

JSON

도로교통공단\_사망교통사고정보서비스

활용신청

사망 교통사고에 대한 개별 정보를 제공(위치 데이터 포함)

9

0

관심

OpenAPI 정보

메타데이터 다운로드

오픈API 에러코드

데이터 개선요청

오류신고 및 문의

분류체계	공공질서및안전 - 안전관리	제공기관	도로교통공단
관리부서명	데이터융합처	관리부서 전화번호	033-749-5266
API 유형	REST	데이터포맷	JSON+XML
활용신청	1958	키워드	사망 교통사고, 개별정보, 위치좌표
등록일	2015-12-29	수정일	2024-07-08
비용부과유무	무료	신청가능 트래픽	개발계정 : 10,000 / 운영계정 : 활용사례 등록시 신청하면 트래픽 증가 가능
심의유형	개발단계 : 자동승인 / 운영단계 : 자동승인		
이용허락범위	이용허락범위 제한 없음		
참고문서	기술문서 도로교통공단_사망교통사고정보.hwp		

활용목적 선택

\*표시는 필수 입력항목입니다.

\*활용목적

☐ 웹 사이트 개발
 ☐ 앱개발 (모바일, 솔루션 등)
 ☐ 기타
 ☐ 참고자료
 ☒ 연구(논문 등)

API 공부

6/250

첨부파일

파일 선택

Drag & Drop으로 파일을 선택 가능합니다.

상세기능정보 선택

<input checked="" type="checkbox"/>	상세기능	설명	일일 트래픽
<input checked="" type="checkbox"/>	사망교통사고정보 Rest 조회	사망교통사고정보 Rest 조회	10000
<input checked="" type="checkbox"/>	사망교통사고정보 WMS 조회	사망교통사고정보 WMS 조회	10000

라이선스 표시

\*이용허락범위

이용허락범위 제한 없음

☒ 동의합니다.

취소

활용신청

# 감사합니다.

