

데이터 조회 프로그램  
SKN Family AI Camp 2기 : 최종 프로젝트 2조



프로젝트 주제	경제 지식이 없는 아동 및 청소년을 대상으로 경제 지식을 쉽게 설명하는 AI 인플루언서 블로그 관리 Tool
원본 데이터 (텍스트)	<b>[ 한국은행 제공 경제교육용 PDF 문서 ]</b> - <u>알기 쉬운 경제 이야기 (일반인용)</u> - <u>경제금융용어 700선</u>

알기 쉬운 경제 이야기는 경제 교양서처럼 서술되었고, 경제금융용어 700선은 사전처럼 서술되었습니다.  
따라서 각각의 데이터는 형식 및 용도 면에서 큰 차이가 있으므로 데이터 고유 특성을 고려하여 전처리 방식을 결정해야 합니다.

데이터 전처리 과정	<b>[ 알기 쉬운 경제 이야기 (일반인용) ]</b> 1. 날짜 부분에 붙은 '용어 설명'과 '페이지 번호' 제거된 부분을 이미지 형태로 추출합니다. 2. OCR을 사용하여 이미지를 텍스트로 추출합니다.  <b>[ 경제금융용어 700선 ]</b> 1. pdf 라이브러리를 사용하여 텍스트 추출합니다. 2. 페이지 번호 및 사전 인덱스(한글 자음) 삭제합니다.
------------	---

DB 사용 용도	<b>[ Vector DB ]</b> AI 인플루언서 블로그 포스팅 작성 시, 제공 정보의 정확도를 높이기 위해 RAG를 사용합니다. 이를 위해 전처리가 끝난 각각의 pdf 파일을 서로 다른 Vector DB에 넣습니다. 두 종류의 블로그 글 스타일(일반 경제 교양 포스팅, 오늘의 경제 단어 포스팅)을 고려하여 위와 같은 방향으로 접근했습니다.  <b>[ RDB ]</b> AI 인플루언서 블로그 포스팅 및 회원 정보를 RDB에 저장합니다. 자주 바뀌는 성질의 데이터는 아니라고 판단하여 RDB에 저장하기로 결정했습니다.
----------	---

## Vector DB 구조

	ids	documents
0	01c4613f-fb75-4514-9805-e06b4dae62c9	하기 때문에 정책수립자가 당초 생각했던 결과가 나타나 지 않는
1	01f9c3ae-b58d-4a6c-827a-ddc2ef3469a2	사람을 뜻하는 것이 아니고 매월 15일이 속한 1주말 동안에 수입
2	02440f51-ca21-4f79-8ad8-d12e5bcb1d12	물 것 없이 가격과 생산량을 결정할 수 있기 때문입니다. 기업이
3	027e1a16-334f-4f05-9e39-082ad0258f85	나 현실에서는 어느 누구도 고소득으로 남에 대해 소유권을 가지
4	032f7a68-b749-4d30-a651-de037976266e	자가 확대되고 있다면 향후 경제가 좋아질 것이라고 시장이 기대
5	03baa389-d981-44bb-b12c-74ed73fa98ba	합니다. 미국 연방준비은행에서는 나우캐스트(nowcast)란 이름
6	04515a77-3419-4b75-8350-3835f14dd6ce	음을 나타냅니다. 이러한 물가지수는 상품별로 중요한 정도에 따
7	054e293b-b14d-42b4-849f-c59685934fb3	준다면 아무 이득도 가져다주지 못할 것이기 때문입니다. 경쟁기
8	05ea1712-7669-4bda-8ae3-2254ba7ceb93	때문에 제값물가와 지수물가 사이에 차이가 나기도 합니다. 이
9	06e0344d-6d37-48f8-9488-41153b129f2b	이처럼 한 나라의 경제에서 중요한 위치를 차지하는 임금은 어쩔
10	09765a1b-de9d-4b30-96e5-c76811209eca	얼굴 환경문제는 사람들이 문명사회를 이루어 주변의 환경을 이
11	097d496d-c77c-48b8-8287-5d3845c27242	로 생산한 재화와 서비스의 가치한 각 생산단계에서 추가된 가치
12	0a3ef704-32c7-461a-9146-c259ff2045e7	다. 우선 제1안정을 위해 정부가 어느 정도 개입할 것인지에 대
13	0aa11ca4-cdab-4e12-df94-58bd6d91753c	통의 크기는 기회비용으로 측정할 수 있는데 한 국가에서 다른 지
14	0afb6116-574a-42f7-b474-2db840dd8b08	과확기기술자 인시사, 변호사와 같은 소위 지식노동자들의 지적노
15	08ced623-370b-df8e-a59c-f87361ebe144	의한 결과라고 비유했습니다. 일반적으로 경제성장은 자본과 노
16	0cccd687-b1f0-4252-a1d2-1cf065a79fb7	것입니다. 지리산 속의 반달가슴곰, 왜 팔팔했을까? 우리 주위에
17	0d4cf5f1-4058-43a5-af1c-9273107ab639	높은 것이 일반적입니다. 공급경쟁은 독점에 비해 많지만 경쟁시장
18	1080d91e-3b04-4e7b-9551-47f2d270eee5	정보가 불완전한 상황에서는 여러 가지 방법으로 필요한 정보를
19	14b2a59b-d228-45d3-ab21-e102aafdf192	를 수 있습니다. 쉬운 예로 국내 경제에서 '갑'이라는 사람이 상대
20	15b3cf0b-3db6-4f75-b9c0-408c32e3f9a4	이 중에서도 가장 중요한 목표는 물가안정이라고 할 수 있습니다
21	16dc3273-9e91-4927-a5ad-01842bbb405e	주거나 예금을 받을 것과 마찬가지로 중앙은행이 금융기관에 부

## < 알기 쉬운 경제 이야기 (일반인용) Vector DB >

	ids	documents
0	00077fe2-bc05-4220-9be7-122598773a7	거나(call option) 팔 수 있는 권리(put option)를 말한다. 통화옵션은 거래소나 장외,
1	0117b01e-ae00-4a36-93d1-b0ea15fbbfb0	원이고, 기초가 격은 생산자 가격 800원에서 정부에 납부하는 세금 100원을 차감하
2	01e905f5-589c-4bad-92f0-2924459fa000	재고지수 현실 경제에서는 경기의 총고 나 불황에 따라 재고의 증가는 감소 현상이
3	021cf88a-2cc4-4880-94d7-39395969c1ab	런 상황에서 판매자가 구매자에게 물품이 없는 물건을 시장에 아예 내놓지 않게
4	0253dc40-5ee9-4aac-8dc0-f36737079d5	말한다. 총 자금이 거래되는 시장을 페더럴펀드시장(federal funds market)이라고
5	02d5cc9b-1b6f-4399-b1f8-6bcd7ecc5d20	일 목적으로 시행되고 한다. 우리나라에서는 1953년의 제1차 통화조치에 따라 1
6	02da62d5-c98a-4493-ac19-c1d2d01307d9	은행(G-SIB)에는 2022년 1월부터 시스템적 중요도에 따라 추가 레버리지 비율이 부
7	03402b1b-9c30-4716-957b-591022a6e1fc	규제나 간섭에서 벗어나 자유롭게 거래를 할 수 있다는 장점이 있고 이러한 자율은
8	03678255-1183-473a-bb3c-ed62f6c0b4b1	무한대(∞)가 된다. 공급탄력성은 가격의 상승과 하락에 따라 공급량을 조절하는 것
9	03c7367a-75ec-40c9-a10b-a54748126bf0	로 자산을 운용하는 상업은행과 대비되는 투자은행은 자기자본금을 활용하여 글!
10	0497f2c3-d674-405b-a8f4-45813d80bbdb	(Common Equity Tier 1) 보통주자본(CE1; Common Equity Tier 1)은 보통주, 자본잉
11	04e57bc9-3ae0-404a-a47f-ab90a5fd6961	통 주요 기관투자의 장기 무이표채에 대한 수요를 충족시킬 수 있다. 그리고 이러
12	057602c6-ccc1-4108-b0c5-d2536e21bdd8	SR 가중치) / 필요안정자금 조달금액(자산원형액 × RSR 가중치) × 100이다. 본가
13	060a9a07-d364-4795-b13d-77b8c1b584d4	물상실 인장, 관세환급금 등의 장점이 있다. 내국 신용장장은 내 물품공급자는 물론
14	065f8833-9a82-4db5-b671-9b4486c8bd58	물가인의 책임경제체제를 확립하고자 하는데 그 주된 목적이 있다. 평가방식으로;
15	0678f8d2-e0bf-4d0e-956f-b2770e4aa286	적으로 이루어지고 있는 반면 규제 감독이 금융기관의 국적 또는 지역기반을 중
16	06a75df9-745c-494b-a46f-0bcc16a20fa1	에 대한 권리가 이전함으로써 실용증권 이동에 따른 본질 도난의 위험성, 사무처리
17	0761a2f4-5063-4f77-8f50-d3ce115a9f38	를 미 달러화로 기록통화라는 규범반체제도 조정과는 고정환율제도를 통해 사
18	08351bca-4dff-4de8-9c4d-c4f259bc1967	부가지는 일반적으로 프로토콜, 부로, 할당, 미디어, 속도 등의 변환을 전송을 부
19	087aa785-1242-4e7f-8a1f-324522d6d192	검색어: 추가지수선물거래 주당은이익(EPS) 주당순이익(EPS; Earning Per Share)은
20	0884f5b3-3460-4c5b-9f64-2de4959b22ac	무, 수출상임외 회피책에 기여하는 정도를 나타낸다. 산업연관도를 이용하면 수
21	08c20555-7a6f-42ba-a375-30e82aea997	(SNA), 공공사용비율(SUT) 국민부담률(국민부담률(total tax revenue as percentage of
22	09a96272-b681-4c56-8f46-5a3e932ec85e	소입의 거래비용을 보면 기업 자금, 환리스크의 관리, 결제일의 조정, 금리차익 획득

## < 경제금융용어 700선 Vector DB >

## Django 내부 RDB 구조

