



챗GPT 러닝데이 & MS애저톤  
#무료 #유튜브라이브

# Open API를 활용한 연구원의 업무 효율화



활용

4월 25일(화) 19:00

MS MVP AI **이제현**

한국에너지기술연구원  
에이아이프렌즈학회

# 연구원의 삶

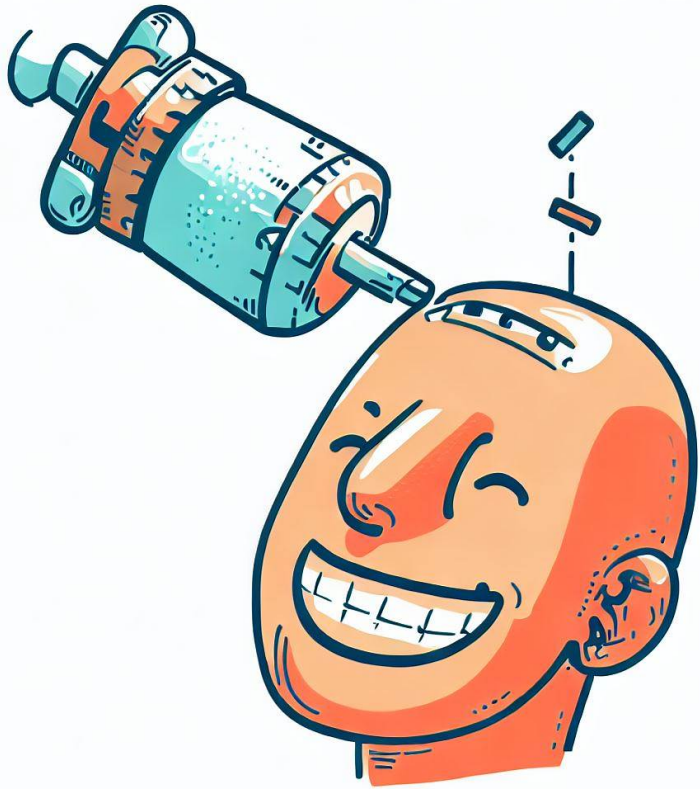


**식사**



**안주**

# 이러고 싶다



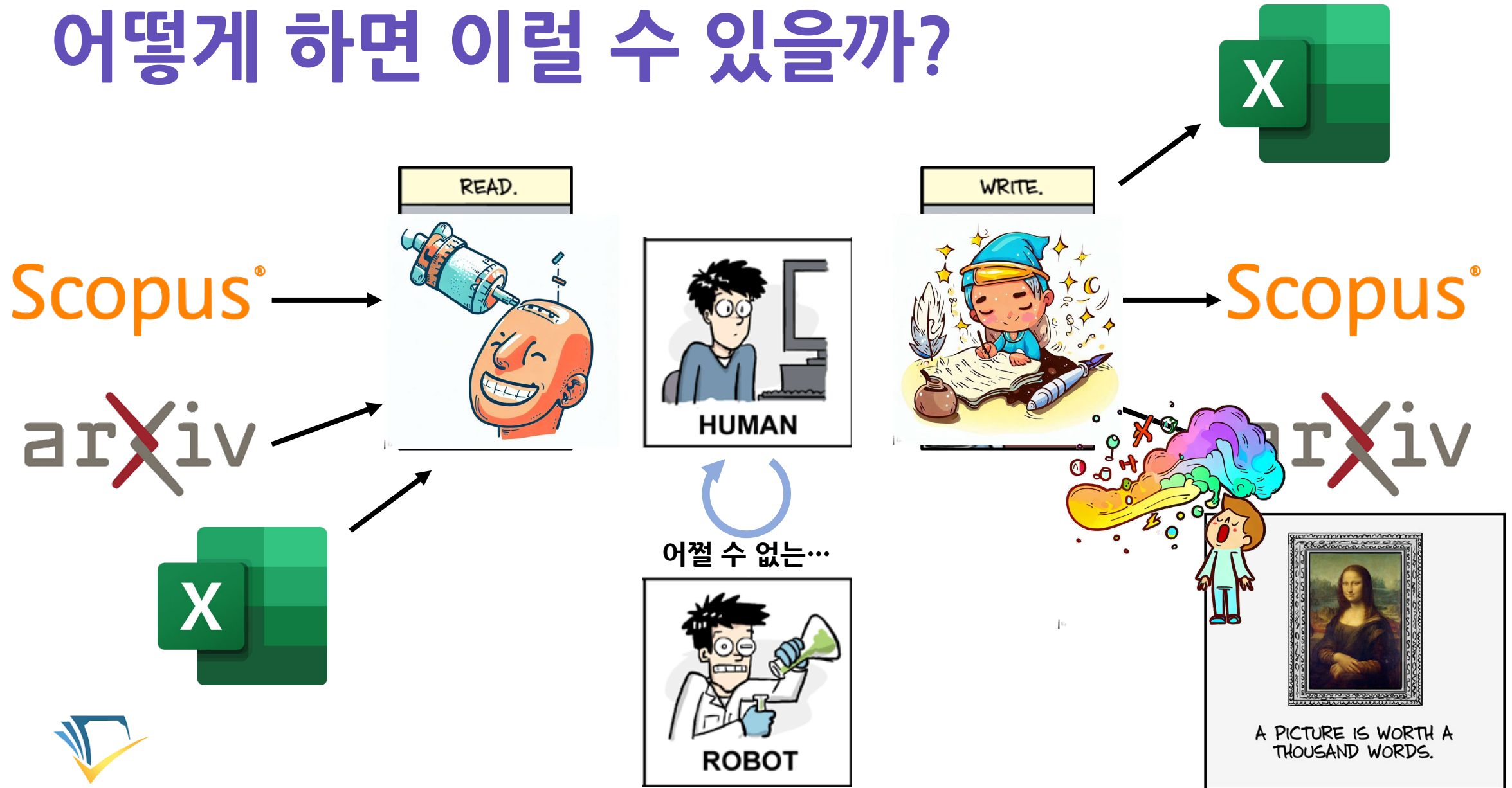
Bing image creator,

“cartoon of a machine injecting new knowledges on a smiling person's head”

“a kid-style cartoon of sleeping man, while a magic quill is transferring his knowledge onto a sheet of paper”

“a kid-style cartoon of man generating a rainbow-colorful chemical illustrations from his deep exhale breath”

# 어떻게 하면 이럴 수 있을까?



# Open API : Scientific Literatures

- **Scopus**

- <https://dev.elsevier.com/>

- **ScienceDirect**

- <https://www.elsevier.com/solutions/sciencedirect/librarian-resource-center/api>

- **Semantic Scholar**

- <https://www.semanticscholar.org/product/api>

The image shows two browser windows. The left window is the Elsevier Developer Portal, displaying the 'Scopus APIs' page. The right window is the ScienceDirect API documentation page. Overlaid on these is a diagram comparing Elsevier and Other publishers. The diagram is a 2x2 grid with dashed blue lines. The top-left quadrant is labeled 'Elsevier' and 'Abstract-only'. The top-right quadrant is labeled 'Other publishers' and 'Scopus'. The bottom-left quadrant is labeled 'Full-text' and 'ScienceDirect'. The bottom-right quadrant is empty. A URL <https://dev.elsevier.com/support.html> is visible at the bottom of the diagram.

The image shows the Semantic Scholar Academic Graph API Overview page. The page features a search bar at the top with the text 'Search over 207 million papers from all fields of science'. Below the search bar is the heading 'Semantic Scholar Academic Graph API - Overview' and a large blue heading 'Providing a reliable source of scholarly data for developers'. Underneath, it says 'Build projects that accelerate scientific progress with the Semantic Scholar Academic Graph API' and includes a 'View Core Use Cases' button. At the bottom, there is a graphic of a blue and yellow abstract design.

# Open API : Summarization & Translation

## Summarization APIs

Browse the best premium and free APIs on the world's largest API Hub. Read about the latest API news, tutorials, SDK documentation, and API examples. RapidAPI offers free APIs all within one SDK. One API key. One dashboard.

- Text Analysis**  
sentiment-analysis, text-summarization, language-detection, article-extraction, named-entity-recognition, extract-text from documents  
👍 9.9 ⌚ 3,705 ms ✓ 87%
- TLDRThis**  
Summarize any URL or text using state-of-the-art abstractive and extractive summarization models.  
👍 9.7 ⌚ 6,190 ms ✓ 93%
- TextGears**  
Spelling and grammar checker with automatic correction. Text summarization and keyword extraction. Language  
👍 9.7 ⌚ 2,494 ms ✓ 100%
- News Article Data Extract and Summarization**  
Extract data from online news & articles. Get full metadata with content.  
👍 9.5 ⌚ 1,028 ms ✓ 100%
- Text Summarization**  
Text Summarization API provides professional text summarizer service which is based on advanced Natural Language Processing and  
👍 9.7 ⌚ 313 ms ✓ 100%
- GPT Summarization**  
Summarize text using an abstractive summarizer based on the GPT machine learning model.  
👍 9.4 ⌚ 19,826 ms ✓ 100%
- TextAnalysis**  
TextAnalysis API provides customized Text Analysis,Text Mining and Text Processing Services like Text Summarization,  
👍 9.2 ⌚ 342 ms ✓ 100%
- Summarization**  
Summarization is MeaningCloud's solution to extract a summary for a given document, selecting the most relevant sentences  
👍 9.2 ⌚ 933 ms ✓ 100%
- TLDR Text Analysis**  
TLDR (Too Long Didn't Read) is a Text Analysis API that allows you to extract summaries and ranked keywords from articles on  
👍 9.2 ⌚ 2,081 ms ✓ 99%
- Summarize Texts**  
Use One AI's NLP models to automatically summarize texts. Configure this API in [the Language Studio] (<https://studio.oneai.com/?>)  
👍 9.3 ⌚ 675 ms ✓ 100%
- AssemblyAI Speech-to-Text**  
The Top-Rated API for Speech-to-Text
- CryptolInfo**  
We collect news from more than 30 crypto/financial sources and process them using neural networks. We estimate news sentiment  
👍 9.2 ⌚ 342 ms ✓ 100%
- Integrate AI**  
Unleash the power of GPT-4 and simplify integration with standard JSON responses.
- Pros and Cons**  
Summarize textual content (ie: reviews and more) into a powerful and highly compact format, Pros and Cons, using Fakespot's Generative Pre-  
👍 9.2 ⌚ 342 ms ✓ 100%
- News Article Data Extract and Summarization**  
Extract data from online news & articles. Get full metadata with content.







# Open API : Summarization & Translation





## Translator APIs

Browse the best premium and free APIs on the world's largest API Hub. Read about the latest API news, tutorials, SDK documentation, and API examples. RapidAPI offers free APIs all within one SDK. One API key. One dashboard.

 **DeepL Translator** 



The world's most accurate and nuanced machine translation

🌐 9.9 🕒 2,309 ms ✓ 100%

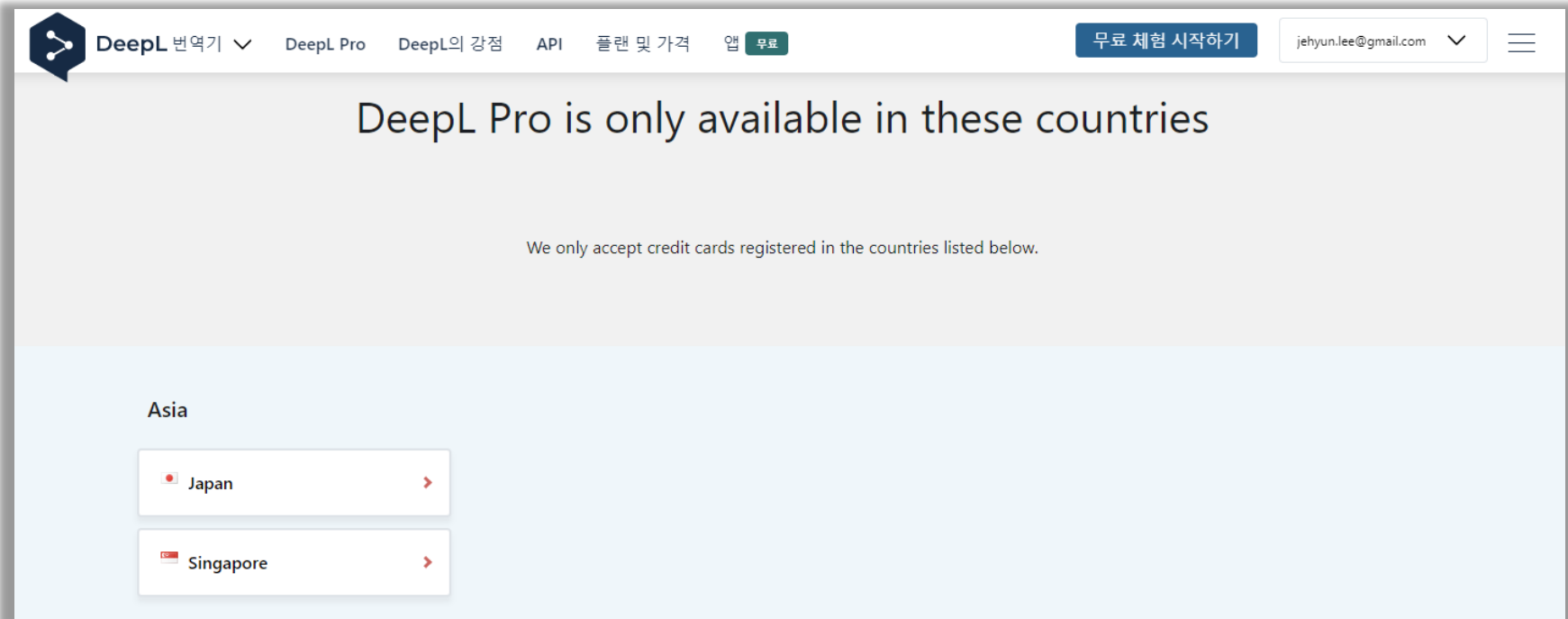
 **Google Translator** 

Google Translate is a multilingual neural machine translation service developed by Google to translate text

🌐 9.7 🕒 158 ms ✓ 100%

 **Long Translator** 

Translate text into other languages. Can translate long texts, fast processing, cost-effective.




DeepL 번역기 **DeepL Pro** DeepL의 장점 API 플랜 및 가격 앱 **무료** [무료 체험 시작하기](#) jehyun.lee@gmail.com

### DeepL Pro is only available in these countries



We only accept credit cards registered in the countries listed below.

**Asia**



- Japan
- Singapore

 **Translator** 



Translate in over 100 languages, and even exclude it from translation. Demo: <https://leafy-biscuit-ccb9a4.netlify.app/> Translate

 **Text to Speech** 

Convert text into a human-sounding voice instantly. Currently, you can choose from over 200 different voices.

 **Fast DeepL Translator** 

Neural Translate. HTML / Text / JSON. Cheaper than original, same quality. DeepL Translator API allows developers to access the

 **Indic-Translator** 

Indic Translator is an advance AI based Translator Indian languages translator API. It has outperformed Google and other translators

# Open API : ETRI

ETRI AI API, Data : <https://aiopen.etri.re.kr/>

AI API-DATA

서비스 안내

API 개발가이드

API 데모

모델 및 데이터

고객지원

API Key



사전 준비사항

사이트 소개

정책 및 약관

## » 오픈API 목록

기술명	API명	1일 허용량
언어 분석 기술 (문어)	<ul style="list-style-type: none"><li>· 형태소 분석 API</li><li>· 동음이의어 분석 API</li><li>· 의존 구문분석 API</li><li>· 개체명 인식 API</li><li>· 다의어 분석 API</li><li>· 의미역 인식 API</li></ul>	5,000건/일 (1회 사용 시 입력은 1만 글자 이하)
언어 분석 기술 (구어)	<ul style="list-style-type: none"><li>· 형태소 분석 API</li><li>· 개체명 인식 API</li></ul>	5,000건/일 (1회 사용 시 입력은 1만 글자 이하)
어휘관계 분석 기술	<ul style="list-style-type: none"><li>· 문장 패러프레이즈 인식 API</li><li>· 동음이의어 정보 API</li><li>· 어휘간 유사도 분석 API</li><li>· 상호참조 해결 API</li><li>· 어휘 정보 API</li><li>· 다의어 정보 API</li><li>· 개체 연결 API</li></ul>	5,000건/일
질의응답 기술	<ul style="list-style-type: none"><li>· 질문분석 API</li><li>· 위키백과 QA API</li><li>· 행정문서 QA API</li><li>· 기계독해 API</li><li>· 법률 QA API</li></ul>	5,000건/일 (1회 사용 시 입력은 1만 글자 이하)
	<ul style="list-style-type: none"><li>· 한국어 인식 API</li><li>· 중국어 인식 API</li><li>· 영어 인식 API</li><li>· 일본어 인식 API</li></ul>	1,000건/일 (최대 20초/건당)

# 실습 Google Colab

- <https://bit.ly/3HaUIgy>



230415\_애저톤.ipynb ☆

파일 수정 보기 삽입 런타임 도구 도움말 모든 변경사항이 저장됨

☰ 목차

🔍 라이브러리 설치

기능 구현

{x}

논문 검색 (semantic scholar)

📁

논문 출판 동향 (Matplotlib)

초록 요약, 질의 응답 (langchain)

일괄 초록 요약, 질의 응답  
(pd.Series.apply)

한글 번역 (Google Translator)

한글 번역 (DeepL)

일괄 방법론, 메시지 번역  
(pd.Series.apply)

보고서 생성 (python-docx)

📄 섹션

+ 코드 + 텍스트

## 라이브러리 설치

```
[1] 1 !pip install openai # OpenAI와 상호작용에 사용
    2 !pip install langchain # 자연어 처리에 사용
    3 !pip install tiktoken # 텍스트 토큰화에 사용
    4 !pip install googletrans==4.0.0-rc1 # 번역에 사용
    5 !pip install python-docx # 보고서 생성에 사용
```

```
Requirement already satisfied: aiohttp>=1.1.2 in /usr/local/lib/python3.9/dist-packages (from aiohttp->openai) (1.3.1)
Requirement already satisfied: async-timeout<5.0,>=4.0.0a3 in /usr/local/lib/python3.9/dist-packages (from aiohttp->openai) (4.0.3)
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: langchain in /usr/local/lib/python3.9/dist-packages (0.0.148)
Requirement already satisfied: tenacity<9.0.0,>=8.1.0 in /usr/local/lib/python3.9/dist-packages (from langchain) (8.2.2)
Requirement already satisfied: dataclasses-json<0.6.0,>=0.5.7 in /usr/local/lib/python3.9/dist-packages (from langchain) (0.5.7)
Requirement already satisfied: pydantic<2,>=1 in /usr/local/lib/python3.9/dist-packages (from langchain) (1.10.7)
Requirement already satisfied: async-timeout<5.0.0,>=4.0.0 in /usr/local/lib/python3.9/dist-packages (from langchain) (4.0.3)
Requirement already satisfied: PyYAML>=5.4.1 in /usr/local/lib/python3.9/dist-packages (from langchain) (6.0)
Requirement already satisfied: openapi-schema-pydantic<2.0,>=1.2 in /usr/local/lib/python3.9/dist-packages (from langchain) (1.2.2)
Requirement already satisfied: numpy<2,>=1 in /usr/local/lib/python3.9/dist-packages (from langchain) (1.24.1)
```



# 1. 년도별 출판 건수 시각화



publication trends per year



All

News

Images

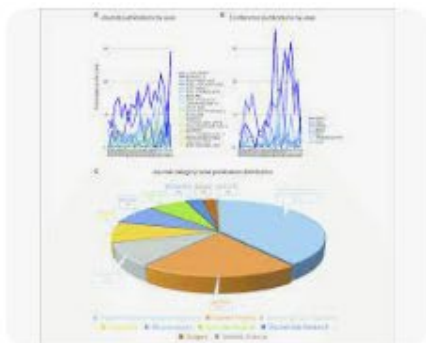
Books

Videos

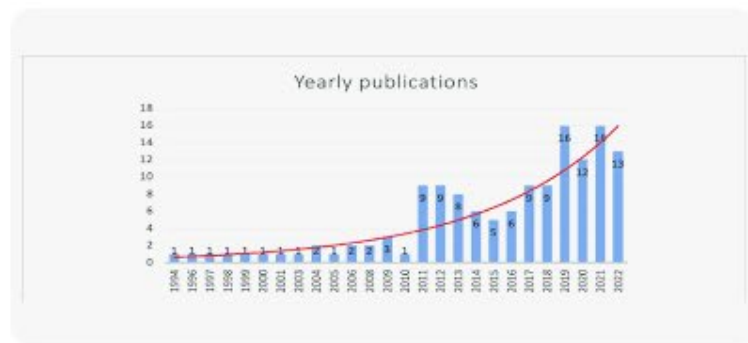
More

Tools

Collections SafeSearch



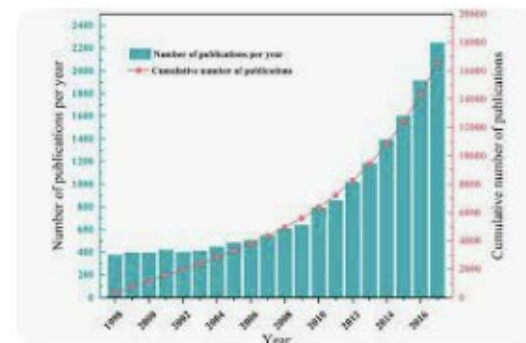
ResearchGate  
Publication trends. Distribution o...



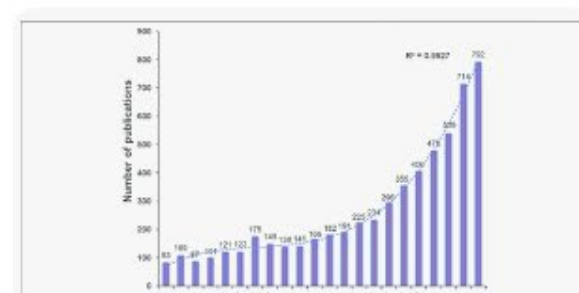
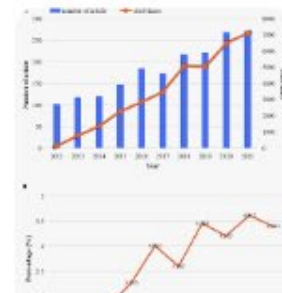
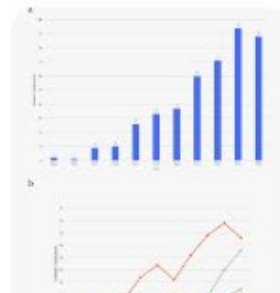
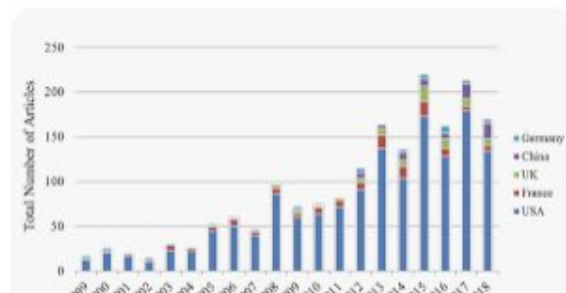
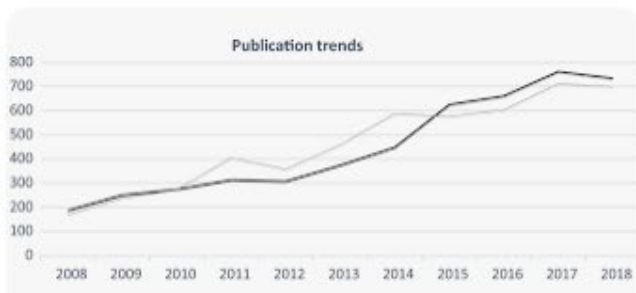
ResearchGate  
Publication trends per year. | Download ...



ResearchGate  
depicts the number of publications per ...



ResearchGate  
Number of publications per year and the ...

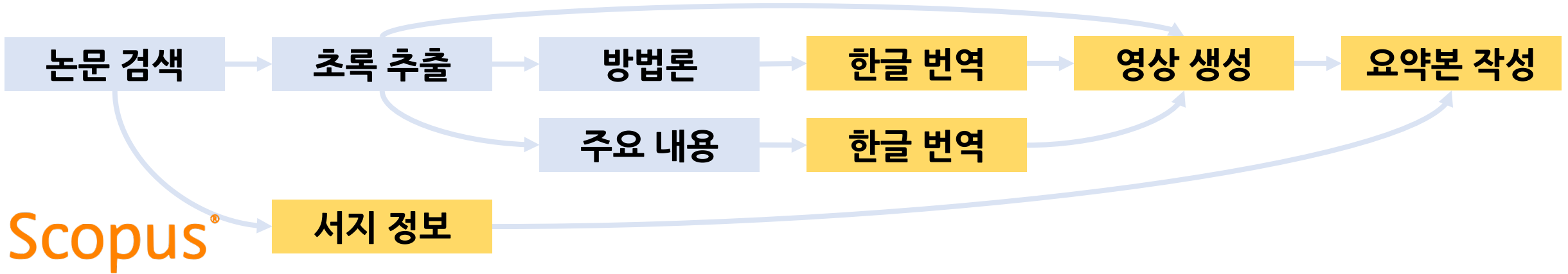


# 1. 연도별 출판 건수 시각화

semantic scholar: deep+learning @2023.04.25



# 2. 논문 요약본 생성



Scopus®  
라이선스 필요

  
Semantic Scholar  
라이선스 불필요

  
Semantic Scholar



 OpenAI  
DALL·E 2



~~arXiv~~  
라이선스 불필요



# 2. 논문 요약본 생성

## 3. "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning"

Computational Intelligence and Neuroscience (2021)

- <https://www.semanticscholar.org/paper/31cf4c96c5dd4ac5a6bbb4ac7b6bab763651624a>

230415\_애저톤.ipynb ☆  
파일 수정 보기 삽입 런타임 도구 도움말 모든 변경사항이 저장됨

목차

- 1. 라이브러리 설치
- 2. 기능 구현
  - 2-1. 논문 검색 (semantic scholar)
  - 논문 출판 동향 (Matplotlib)
  - 2-2. 초록 요약, 질의 응답 (langchain)
  - 일괄 초록 요약, 질의 응답 (pd.Series.apply)
  - 2-3. 한글 번역 (Google Translator)
  - 2-4. 한글 번역 (DeepL)
  - 일괄 방법론, 메시지 번역 (pd.Series.apply)
- 3. 보고서 생성 (python-docx)
  - 3-1. hyperlink 삽입 코드
  - 3-2. 방법론과 메시지를 함축한 그림 생성 (DALLE)

### 3-3. MS Word 보고서 생성

섹션

+ 코드 + 텍스트

[46] 31 fig = get\_figure(methodology, message)

### 3-3. MS Word 보고서 생성

```
[48] 1 from docx import Document
2 from docx.shared import Inches
3 from docx.shared import Pt
4 from docx.shared import RGBColor
5 from docx.enum.text import WD_ALIGN_PARAGRAPH
6
7 # document generation
8 document = Document()
9
10 # Title
11 report_title = f"semantic scholar @ {datetime.strftime(datetime.now(), '%Y.%m.%d')}\n(keywords)"
12 document.add_heading(f"{report_title}", 0)
13 document.add_page_break()
14
15 # Each Article
16 for i in range(df_paper_abstract20.shape[0]):
17     # article data
18     url = df_paper_abstract20["uris"].iloc[i]
19     title = df_paper_abstract20["title"].iloc[i]
20     journal = df_paper_abstract20["journals"].iloc[i]
21     volume = df_paper_abstract20["volumes"].iloc[i]
22     pages = df_paper_abstract20["pages"].iloc[i]
23     abstract = df_paper_abstract20["abstracts"].iloc[i]
24
25     methodology_ko = df_paper_abstract20["methodology_ko"].iloc[i]
26     methodology = df_paper_abstract20["methodology"].iloc[i]
27     message_ko = df_paper_abstract20["message_ko"].iloc[i]
28     message = df_paper_abstract20["message"].iloc[i]
29
30     # title
31     title_ = document.add_paragraph(style="List Number").add_run(f"{title}")
32     title_.font.size = Pt(14)
33     title_.font.italic = True
34
35     # journal
36     if journal:
```



- 방법론 (국문) :** 이 글의 방법론은 다양한 머신러닝 알고리즘과 딥러닝을 UCI 머신러닝 심장 질환 데이터 세트에 적용하고, 격리 포레스트(Isolation Forest)를 사용하여 관련 없는 특징을 처리하고, 데이터를 정규화하고, 정확도 및 혼동 행렬을 사용하여 결과를 검증하는 것입니다.
- 방법론 (영문) :** The methodology in this article is to apply different machine learning algorithms and deep learning to the UCI Machine Learning Heart Disease dataset, handle irrelevant features using Isolation Forest, normalize the data, and validate the results using accuracy and confusion matrix.
- 주요내용 (국문) :** 이 글의 주요 메시지는 머신러닝 알고리즘과 딥러닝을 사용하여 UCI 머신러닝 심장병 데이터 세트를 분석하여 심장병을 예측하고 94.2%의 정확도로 유망한 결과를 얻을 수 있다는 것입니다.
- 주요내용 (영문) :** The main message of this article is that machine learning algorithms and deep learning can be used to analyze the UCI Machine Learning Heart Disease dataset in order to predict heart disease and achieve promising results with an accuracy of 94.2%.
- ABSTRACT :** The correct prediction of heart disease can prevent life threats, and incorrect prediction can prove to be fatal at the same time. In this paper different machine learning