# Al Newsletter (No. 6)

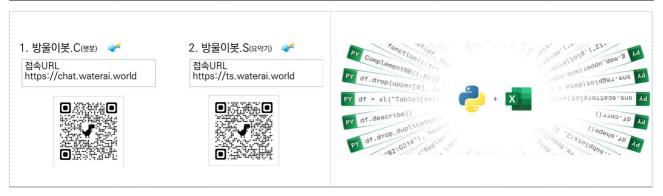
연구관리처 AI연구센터 / (2023년 10월 25일)

K-water연구원 AI연구센터에서 정기적으로 발간하는 뉴스레터입니다. AI 뉴스, Hands-on 프로젝트, 팁 등을 다양한 내용과 난이도로 담았습니다.

※ 코드 및 뉴스 등 외부링크가 다수 포함되어 있으므로 **인터넷 환경 PC에서 사용 권장** 

# >> 6호 주요 콘텐츠

#AI연구센터 #Hello, K-water #직접 돌려보기 #AI 2학년 4반



방울이봇 시리즈(2종) 출시!!

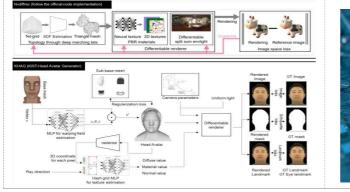
엑셀 x 파이썬 콜라보





제3회「K-water 대국민 AI 경진대회」실시

크롬브라우저를 통해서 Youtube 영상을 간단히 요약





컴퓨터비전(CV) 국제대회 입상 & 성과 논문 게재

[Tips & Info] AI 반도체 & 파이썬 가상환경 만들기

모든 실습 예제와 링크는 "AI연구센터" 누리집(Homepage, GitHub)에 게시되어 활용 가능합니다.

### □ (신기술) 반도체면 반도체지... AI 반도체는 뭐람?

- AI반도체(NPU<sup>Neural Processing Unit</sup>)는 말 그대로 AI 전용 반도체입니다. NPU가 등장하게 된 배경에는 ChatGPT와 같은 거대 AI 모델의 등장이 있습니다. 거대모델은 사이즈가 크기 때문에 적정 시간 내에 수많은 데이터를 학습하고 추론하여 결과를 처리하는 것이 중요합니다. 이에 따라 비교적 처리 속도가 빠른 GPU에 대한 수요가 폭발적으로 증가하였고 품귀현상이 발생하기도 했습니다.
- 그러나 GPU(와 CPU)는 애초에 목적이 AI를 위한 것이 아닙니다. 그러다 보니 AI 개발만을 용도로 사용하기에는 처리 성능과 비용, 전력 소모 등에서 비효율적인 부분이 많습니다. 이러한 문제를 해결하기 위해 등장한 것이 바로 AI 반도체, NPU입니다.
- 아래 표를 보면, AI 반도체(NPU)는 인공지능 개발에 최적화되어 있다는 것을 알 수 있습니다. 기존 CPU(GPU)의 기능은 여러 가지를 처리할 수 있게 만들었다면. NPU는 오로지 인공지능에만 집중했습니다.



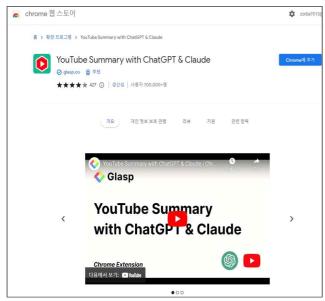
○ 향후 AI 반도체는 데이터 센터(Data Center)와 같은 고성능 서버에서 활용되다 점차 자동차, 스마트폰 등에 탑재되는 디바이스용으로 사용범위가 확대될 것으로 전망됩니다. 가트너에 의하면 시장규모 또한 커질 것으로 예상되는데, 2023년에는 343억 달러(약 40조 원), 2030년에는 전체 시스템 반도체 시장의 31.3%를 점유할 거라고 합니다.

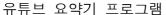
### □ (어플리케이션) Youtube 요약기

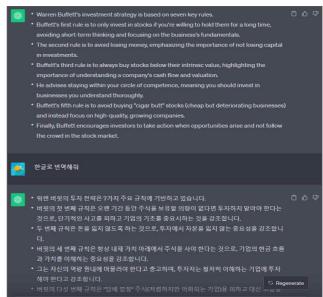
- 오늘 소개할 앱은 "Youtube summary with ChatGPT & Claude" 라는 크롬 확장프로그램입니다.
- 유튜브는 엔터테인먼트의 영역을 넘어 정보 탐색에도 활용되고 있습니다. 기존의 서적, 논문 등의 텍스트와는 다르게 동영상을 통해 빠르고 쉽게 이해할 수 있다는 점이 유튜브의 큰 장점이죠. 그러나 이제는 이것마저 시간이 아깝다! 하는 사용자들을 위해 동영상을 요약하는 앱이 나왔습니다.
- 인터넷망 환경에서 사용 가능하며, 방법은 아래와 같습니다.
  - ① 확장프로그램 설치 : 크롬 앱스토어에서 검색하거나 아래 URL을 이용해 프로그램을 설치합니다.

(https://chrome.google.com/webstore/detail/youtube-article-summary-p/nmmicjeknamkfloonkhhcjmomieiodli?hl=ko)

- ② 유튜브에서 요약하고 싶은 동영상을 선택합니다.
- ③ 동영상 우측 상단에 "Transcript & Summary"를 펼치면, 전체 스크립트를 확인할 수 있습니다.
- ④ "View Al Summary" 버튼을 누르면 ChatGPT가 요약해줍니다. 기본설정이 영어이므로 한국어 요약을 원할 경우 "한국어로 번역해줘"와 같은 프롬프팅을 하거나 번역기(파파고, 구글번역기, DeepL 등)을 사용합니다.
- ⑤ 상세한 스크립트는 가장 오른쪽의 복사 버튼을 눌러 복사하여 (번역기를 통해) 전체 내용을 확인할 수 있습니다.







동영상 내용이 요약된 모습

# 2. "10줄 코딩" Hands-on Al Project

#openpyxl 사용법 #python excel #python excel 다루기 #파이썬 업무자동화 엑셀

# □ openpyxl을 활용해 파이썬으로 엑셀 다루기 실습

○ 파이썬으로 엑셀을 다룰 때 가장 많이 사용하는 openpyxl 패키지에 대해 알아보겠습니다. openpyxl은 엑셀 파일(.xlsx, .xlsm, .xltx, xltm 등)을 읽고 쓰는 기능을 제공합니다. 엑셀은 회사에서 업무를 할 때 정말 많이 사용하는 오피스프로그램이니 알아두시면 많은 도움이 될 거라 생각합니다! 아래의 실습 코드를 가지고 설명해 드릴 테니 차근차근 따라와 주세요.

○ 우선, 실습 작업을 위해 openpyxl 패키지를 설치하고, 라이브러리를 불러옵니다.

```
# 패키지를 설치합니다.
pip install openpyxl
                         # 라이브러리를 불러옵니다.
import openpyxl
from random import *
                         # random에서 함수들을 불러옵니다.
```

- 통합문서(Work Book) 생성: 엑셀 파일이 있어야만 openpyxl을 사용할 수 있는 것은 아닙니다. Workbook()클래스로 통합문서를 만들 수 있습니다.
  - openpyxl.Workbook()함수를 호출하면. 안이 비어 있는 새로운 워크북 객체가 생성됩니다. 그러나 이 타이밍에서 새로운 엑셀 "파일"이 생성됐다고 할 수 없습니다. 반드시 아래에서 소개할 save 메소드를 이용해서 저장까지 해야 합니다.
  - 워크북은 항상 하나 이상의 워크시트로 생성됩니다. Workbook.active 속성을 이용해 워크 시트 객체를 만들 수 있습니다. 참고로 처음 워크북을 만들면 'Sheet'라는 이름을 가진 시트가 하나 생성되며, 이 첫 번째 시트를 활성화하는 것이 기본값으로 설정됩니다.

```
# 엑셀파일 만들기
                             # 워크북을 생성합니다.
wb = openpyxl.Workbook()
                             # 워크시트 객체를 생성합니다.
ws=wb.active
ws.title = 'test_sheet'
                             # 워크시트의 이름을 변경합니다.
# append 함수로 한꺼번에 1줄씩 데이터 넣기
ws.append(['이름', '국어', '영어', '수학', '총점']) # Head로 사용될 행을 입력합니다.
for i in ['둘 리','고길동','마이콜']: # 3개 행에 이름, 국·영·수 각 과목
                                                # 3개 행에 이름, 국·영·수 각 과목 점수를
                                                  난수로 생성합니다.
# randint(start, stop) 함수로 정수난수를 생성합니다.
ws.append([i, randint(60, 100), randint(60, 100), randint(60, 100), ''])
ws.append(['평
               균'])
                         # 5행에 '평균'을 입력
wb.save('test.xlsx')
                         # 생성한 워크북에 파일명('test.xlsx')을 주어 저장합니다.
                         # 엑셀파일을 닫습니다.
wb.close()
```

[결과] 워크북 'test.xlsx'의 워크시트 'test\_sheet'에 학생들의 과목별 점수 출력

1	Α	В	С	D	E
1	이름	국어	영어	수학	총점
2	둘 리	100	96	72	
3	고길동	81	85	77	
4	마이콜	88	61	69	
5	평 균				
7		test_sheet	(+)		

- **엑셀파일 읽고 쓰기**: 워크시트의 키값이나(예: 'A1', 'D5'), Worksheet.cell() 메서드로 행 및 열 표기법을 이용해 셀에 접근하여 읽거나 쓸 수 있습니다. 또한 슬라이싱을 통해 셀 범위(Range)에 접근할 수도 있습니다.
  - openpyxl.load\_workbook()을 통해 이미 만들어진 통합 문서를 새로 열 수 있습니다.
  - 워크시트 슬라이싱을 이용해 원하는 행들이나 열들을 골라 순환문을 실행할 수 있습니다.
  - Worksheet.cell 속성을 사용하려면 행, 열에 대한 위치값을 각각 알면 됩니다. 가령 위 결과에서 둘리의 영어성적은 Cell(2,3)라고 표현할 수 있습니다. Cell() 내부의 숫자는 행row=2(행), column=3(열)을 의미합니다. 만약, 엑셀의 고유 인덱스인 Range로 표기하려면 "C2"라고 표기할 수 있습니다.
  - 실제 셀에 입력된 값을 읽거나 해당 셀 위치에 값을 쓰려면 Cell함수나 Range인덱스의 .value속성을 이용해 접근할 수 있습니다.

```
# 엑셀파일 열기
wb = openpyxl.load_workbook('test.xlsx') # 파일명을 인수로 하여, 워크북의 객체를 반환
ws = wb['test sheet']
                                    # 워크시트 객채를 생성합니다.
# 슬라이싱과 for문을 이용해 행단위, 열단위를 읽거나 씁니다.

      row_range=ws[2:4]
      # 2번째 줄에서 4번째 줄까지 가져오기

      for r in row_range:
      # 행단위에서 각 셀 값을 얻는다.

   sum_range=r[1].coordinate +':'+ r[3].coordinate
   # 각 행의 셀들은 0부터 인덱싱 합니다. (아래 r[1]은 2행 2열, 즉 'B2'셀)
   # .coordinate 속성: 현재 작업중인 좌표 위치를 알아냅니다. (아래 결과는 'B2:B4')
   # cell(행, 열) 함수를 이용해 원하는 셀에 값을 씁니다.
   # 단일 값을 직접 입력하거나, 엑셀함수를 문자열로 입력해 계산할 수도 있습니다.
   ws.cell(r[0].row, 5).value = '=sum(' + sum_range + ')' # 합계(sum)를 계산합니다.
   # 셀 값 쓰기2: ws.cell([0].row, 5, '=sum(' + sum_range + ')')
   # 셀 값 쓰기3: ws['B5']='=sum(' + sum_range + ')
col_range=ws['B:E']
for c in col_range:
                          # B열에서 E열까지 가져오기
                          # 열단위에서 각 셀 값을 얻는다.
   # 각 열의 셀들은 0부터 인덱싱 합니다. (아래 c[3]은 4행 2열, 즉 'B4'셀)
   ave_range=c[1].coordinate +':'+ c[3].coordinate
                                                         # 'B2:D2'
   ws.cell(5, c[0].column).value = '=average(' + ave_range + ')' # 평균(average) 계산
# 엑셀 파일 저장하고 닫기
wb.save("test.xlsx")
wb.close()
```

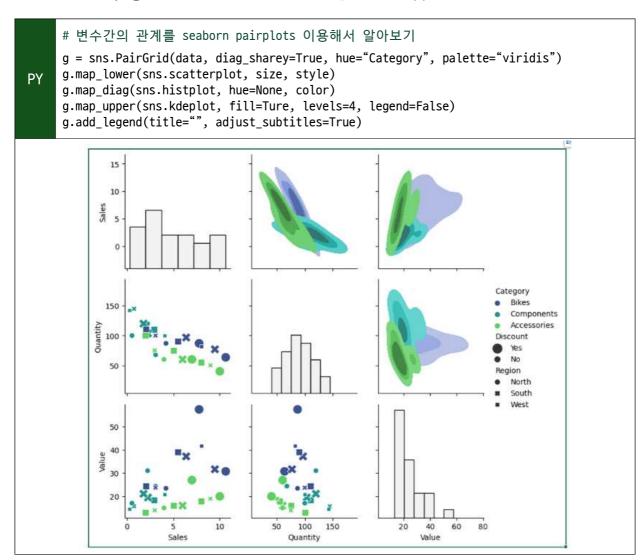
#### [결과] 학생별 과목 총점, 과목별 학생 평균을 계산

1	Α	В	C	D	E
1	이름	국어	영어	수학	총점
2	둘 리	100	96	72	268
3	고길동	81	85	77	243
4	마이콜	88	61	69	218
5	평 균	89.7	80.7	72.7	243
4	F	test_sheet	(+)		

#### # Python in Excel

#### □ ♣+■, 액셀에서 파이썬을?

- 드디어 액셀에서 파이썬을 사용할 수 있게 되었습니다!!!
  - \* 참고) Hands-On은 반대입니다. (파이썬에서 액셀 다루기)
- 파이썬에서 데이터분석, 시각화, 머신러닝 등 다양한 일을 할 수 있지만, 여전히 액셀이 많이 사용되고 익숙하죠. 작업을 하다보면 이 둘을 적당히 섞어서 사용할 수 있으면 좋겠다는 생각이 듭니다. 이런 불편을 알았는지 마이크로소프트에서 액셀에서 파이썬 코딩을 할 수 있는 기능(베타버전)이 나왔습니다! (\*^▽^\*)
- 아직은 베타버전이라 Microsoft365 참가자 프로그램 가입 후 사용 가능합니다. 문법은 셀 클릭 후 =PY(코딩)을 입력합니다. 간단하죠?
- 파이썬 기본문법뿐만 아니라 시각화(seaborn, Matplotlib), 모델링(scikit-learn, statsmodels) 등을 위한 라이브러리도 제공한다고 합니다.



#### □ (파이썬) Youtube Download

○ 파이썬으로 유튜브 영상을 다운로드 할 수 있습니다.

#라이브러리 설치 pip install pytube

import pytube #import

pytube.YouTube(input("Enter the URL: ").streams.first().download()) #URL넣으면 다운 끝!

○ 영상에 대한 정보 가져오기도 가능합니다.

from pytube import YouTube

url = "https://youtu.be/UAUHviUQ9cA?si=LMudXb2k\_0JGbilT" # 소양강댐 50주년 캠패인 영상

yt = YouTube(url)

print("제목 : ", yt.title) print("길이 : ", yt.length) print("게시자 : ", yt.autho #제목 #영상길이 #게시자

----(기시사: ", yt.author) print("게시날짜: ", yt.nubi--print("조취스 ', yt.publish\_date) #영상 게시 날짜 print("조회수 : ", yt.views)
print("키워드 : ", yt.keywords) #조회수 #키워드 추출 print("설명 : ", yt.description) #영상 설명 추출

[결과]

제목 : [소양강댐 50주년 캠페인 CF] K-water와 함께 희망의 미래로!

길이 : 40

게시자 : 한국수자원공사 K-water Official

게시날짜: 2023-09-13 00:00:00

조회수 : 123 키워드:[] 설명 : None

#### □ (가상환경 시리즈) 1탄. 가상환경이 뭔데?

- AI기초 시리즈(1~4호)에 이어 이번에는 가상환경 시리즈로 돌아왔습니다. 이번 호는 가상환경이 무엇인지, 그리고 왜 사용하는지에 대해 알아보겠습니다.
- 가상환경은 독립적인 실행환경을 의미 합니다. 따라서 가상환경끼리는 서로에게 영향을 미치지 않으며, 서로 다른 버전과 패키지를 가질 수 있습니다.
- 왜 필요할까요? 옆에 그림을 볼까요? 이처럼 버전이나 라이브러리 서로 다른 프로젝트를 관리하기 편하기 때문입니다. 파일별로 폴더를 나눠서 관리하는 거랑 비슷하네요!
- 다음 호(7호)에서는 가상환경 만드는 방법에 대해 알아보겠습니다.



# 4. Meanwhile, in K-water AI Lab.

# 방울이봇 등장

# 제3회 AI경진대회

# AI 국제 경진대회 입상

## □ (홍보) 🦖 방울이봇 시리즈 출시

- 두둥! Al연구센터에서는 ChatGPT와 같은 Chatbot을 업무에 활용하기 위해 방울이봇 시리즈 방울이봇.C와 방울이봇.S를 개발하여 베타버전을 사내에 공개했습니다.
- **방울이봇.C**(Chatbot)는 **챗봇** 형태로 알리오에 공시된 사규(108개) 및 간단한 질문에 대해 질의 응답할 수 있습니다. 아래 예시를 참고해주세요! 더불어 방울이봇.C에는 간단한 심리 상담 기능도 있습니다.

질문해도 <b>.</b>	질문 불가능한 것	
사규 관련 질문	간단한 질문	정보 탐색
<ul> <li>▶재택당직은 언제 할 수 있어?</li> <li>▶설계심의위원회는 몇 명으로 구성해야 해?</li> <li>▶ K-water의 ESG에 대해 알려줘</li> </ul>	▶ 근의 공식에 대해 알려줘 ▶ pH가 수질에 영향을 주나? ▶ 환경과 녹색의 관계는?	▶K-water CEO의 이름은? ▶K-water AI연구센터란?

○ **방울이봇.S**(Summary)는 요약기로, 신문기사·논문·긴 문서 등의 텍스트를 효율적으로 요약할 수 있습니다. 입력 형태와 선택하는 언어 모델은 아래 표를 참고해주세요.

입력 형태			
직접 입력	WEB		
▶텍스트를 직접 입력할 수 있습니다.	▶웹페이지 URL을 넣을 수 있습니다. ▶주로 인터넷 기사(네이버, CNN 등)를 요약해 서 보고 싶을 때 활용하면 좋습니다.		
논문 - DOI	Upload File		
<ul><li>► 논문 DOI(논문 고유 ID)를 넣을 수 있습니다.</li><li>► 현재는 MDPI 저널의 DOI만 입력 가능합니다. (www.mdpi.com)</li></ul>	▶컴퓨터에 있는 파일을 넣을 수도 있습니다. ▶ pdf, hwp, docs, txt 형식만 가능합니다.		

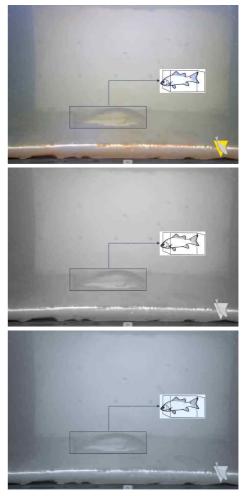
선택할 수 있는 언어 모델					
기본형태(GPT-3.5)	다국어(mT5)	영어(T5)			
▶ 기본형태입니다. ▶ 어떤 언어를 넣어도 무난하게 번역/요약해줍니다.	▶외국어일 경우 추천합니다. ▶중국어, 일본어, 프랑스어 등	▶영어 사용 시 추천합니다.			

#### □ (홍보) 제3회 대국민 K-water AI 경진대회 개최

- 지난 2년간 열린 K-water Al 경진대회가 올해도 개최될 예정입니다!
- 이번 주제는「어종(魚種) 식별 및 분류 알고리즘 개발」입니다. 제공되는 데이터는 라벨링 된 낙동강 하굿둑 물고기 영상입니다. 해당 영상을 보고 어떤 물고기인지 식별하는 대회입니다.
- AI에 관심 있는 사람은 누구나 참여할 수 있으며 상금은 총 1,000만원입니다! 10월 17일부터 참가 신청을 받아 11월 13일(월)까지 진행되니 관심과 참여 부탁드립니다.





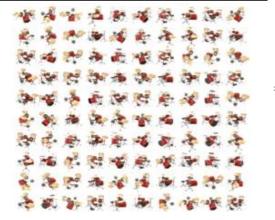


### □ 컴퓨터비전(CV) 국제대회 입상 및 결과 논문\* 게재

- \* ICCV(International Conference on Computer Vision)
- AI연구센터는 물인프라안전연구소와의 협업으로 디지털 트윈과 메타버스의 가속화를 위해 NeRF(Neural Radiance Fields) 기술을 연구하고 있습니다. NeRF는 아래 그림과 같이 2D 사진 몇 장으로 3D 모델을 만들어주는 렌더링 기술입니다.

여러 각도에서 찍은 사진들(2D)

3D로 변환(영상)





- 연구중인 NeRF 기술로 올해 영국 Imperial College London 주관으로 진행된 국제대회\*에 참가했으며, 전세계에서 참여한 100팀 중 최종 6위로 입선했습니다. 대회 결과는 컴퓨터비전(CV) 분야에 있어 세계적인 학술대회 ICCV 2023 Workshop Paper에 게재되었습니다.
  - \* 주제 : To NeRF or not to NeRF: A View Synthesis Challenge for Human Heads



#### \* 논문주소 :

https://openaccess.thecvf.com/content/ICCV2023W/RHWC/papers/Jang\_VSCHH\_2023\_A\_Benchmark\_for\_the \_View\_Synthesis\_Challenge\_of\_ICCVW\_2023\_paper.pdf