

2021-winter-intern

2 월 7 일 업무 진행 상황

기존 결과에서 인용횟수, 출원날짜, google patent link 와 특허 pdf file link 를 추가하였다.

Main

Developer notes

Simple search

Advanced search

Apple search

Advanced search service

You can search patents with keywords in google advanced patent

Input all values you want (connect with , (ex.autonomous driving, deep learning))

Keywords:

Inventor:

assignee:

country:

→ Keywords : autonomous driving

assignee : stradvision 으로 세부검색한 결과는 아래와 같다.

Advanced patent search result							
Rank	Patent No.	Inventor	Assignee	Pub. No.	Google Patent Link	Pub. No.	Patent Link
1	Method for monitoring wheel spin of vehicle and wheel spin monitor using the same	20170000074	이명환	2018.01.08	US 20180000074	2018.01.08	https://patents.google.com/patent/US20180000074
2	Method for monitoring wheel spin of vehicle and wheel spin monitor using the same	20180100007	이명환	2018.01.08	US 20180100007	2018.01.08	https://patents.google.com/patent/US20180100007
3	Method and device for wheel spin path planning of autonomous driving through information fusion by using V2X communication and image processing	20180100008	이명환	2018.01.08	US 20180100008	2018.01.08	https://patents.google.com/patent/US20180100008
4	Method and device for optimized resource allocation in autonomous driving on the basis of reinforcement learning using deep brain data, radar, and camera data	20180100009	이명환	2018.01.08	US 20180100009	2018.01.08	https://patents.google.com/patent/US20180100009
5	Method and device for adaptive resource allocation in autonomous driving on the basis of reinforcement learning using deep brain data, radar, and camera data	20180100010	이명환	2018.01.08	US 20180100010	2018.01.08	https://patents.google.com/patent/US20180100010
6	Method and device for generating image data and for using the learning information for driving information in autonomous driving environment and testing method and testing device using the same	20180100011	이명환	2018.01.08	US 20180100011	2018.01.08	https://patents.google.com/patent/US20180100011
7	Method and device for creating traffic scenario with dynamic adaptation on virtual driving environment for testing, validating and training autonomous vehicles	20180100012	이명환	2018.01.08	US 20180100012	2018.01.08	https://patents.google.com/patent/US20180100012
8	Method and device for attention-driven resource allocation by using reinforcement learning and V2X communication in mobility without safety of autonomous driving	20180100013	이명환	2018.01.08	US 20180100013	2018.01.08	https://patents.google.com/patent/US20180100013
9	Learning method and learning device for general fusion in adaptive allocation required to make model of dynamic extraction and information required by sensor to thereby improve neural network for supporting autonomous driving and testing method and testing device using the same	20180100014	이명환	2018.01.08	US 20180100014	2018.01.08	https://patents.google.com/patent/US20180100014

→ 해당 결과는 google patent 의 세부검색 결과로 먼저 테스트 해보았다.

현재 google patent, patently apple 의 모든 자율주행 관련 특허를 하나의 데이터 베이스 추합하고 있는 중이다.

이 후 몽고 DB 에 있는 데이터 베이스 중복제거 작업을 수행할 예정이다.

2 월 3 일 업무 진행 상황

다음은 28 일에 설계한 크롤링 로봇의 동작 과정이다.

- 먼저 patently apple 에 존재하는 자율주행에 관한 특허들을 크롤링한 결과이다.
140 여개가 추출되었으며 동적 크롤링을 진행하였기에 16 분이 소요되었다.

```
db에 예들을 저장

url = 'https://www.patentlyapple.com/patently-apple/autonomous-vehicle-technology/'
df = applepatent(url)
df.to_db('autonomous driving',df,my_client)

[5] ✓ 16m 10.3s Python

... C:\Users\junkil\AppData\Local\Temp\ipykernel_36416\4025898212.py:502: DeprecationWarning: executable_path has been deprecated, please pass in a Service
object
driver = webdriver.Chrome('../chromedriver')

Maximum page is 5

C:\Users\junkil\AppData\Local\Temp\ipykernel_36416\4025898212.py:540: DeprecationWarning: executable_path has been deprecated, please pass in a Service
object
driver = webdriver.Chrome('../chromedriver')
```

- 구글 특허 세부검색을 통하여 db 에 데이터를 저장한다.

```
db에 키워드들을 저장 (리스트를 수정해서 돌리면 됨)

keyword = 'autonomous driving'
inventor = ''
assignee_list = ['stradvision','Waymo LLC','Nvidia','argo AI','baidu','cruise','motional','Mobileye','Aurora','zoox','nuro','ford','intel','Hyundai']
country = ''
for assignee in assignee_list:
    df = detail_search(keyword, inventor, assignee, country)
    df.to_db('autonomous driving',df,my_client)
```

다만 90 분정도 진행되다가 connection error 가 발생해서 ford 부터 다시 실행하였다.

```
ford 중간에 끊김 (connection error)
그 이후로 실행

keyword = 'autonomous driving'
inventor = ''
assignee_list = ['ford','intel','Hyundai Motor Company','volkswagen','yandex','daimler','bosch','aptiv','toyota','renault','nissan','mitsubishi motc']
country = ''
for assignee in assignee_list:
    df = detail_search(keyword, inventor, assignee, country)
    df.to_db('autonomous driving',df,my_client)

[26] ✓ 126m 51.3s Python

... C:\Users\junkil\AppData\Local\Temp\ipykernel_36416\2423918927.py:315: DeprecationWarning: executable_path has been deprecated, please pass in a Service
object
driver = webdriver.Chrome('../chromedriver')
```

총합 3 시간 30 분 정도 크롤링에 걸렸다.

다만 한가지 의아한 점은 반복문을 1 개만 시행했을 때에는 1000 개의 검색결과를 추출하는데 2 분 40 여초 정도만 걸린것을 확인할 수 있었다.

```

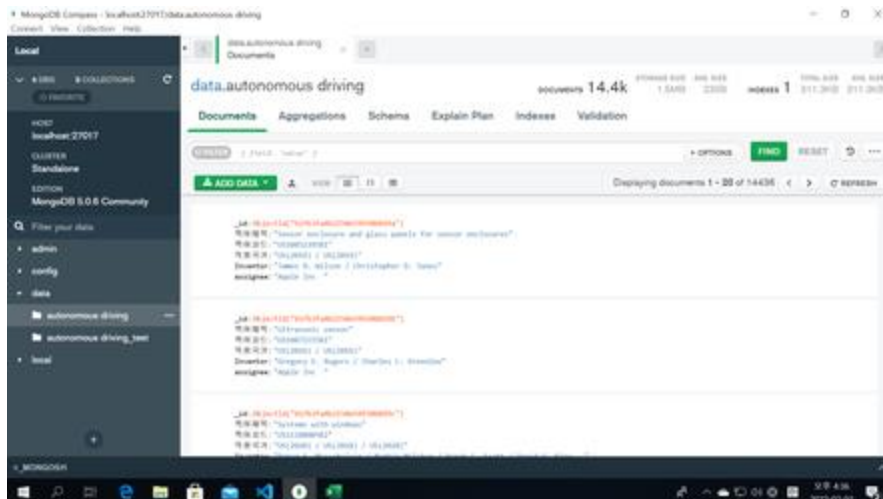
keyword = 'autonomous driving'
inventor = ''
assignee_list = ['volkswagen']
country = ''
for assignee in assignee_list:
    df = detail_search(keyword, inventor, assignee, country)
    df_to_db('autonomous driving_test',df,my_client)

[31] ✓ 2m 41.2s Python
... C:\Users\junkil\AppData\Local\Temp\ipykernel_36416\3201541700.py:315: DeprecationWarning: executable_path has been deprecated, please pass in a Service
object
driver = webdriver.Chrome('./chromedriver')

```

stradvision	24sec	12페이지
intel	2m 37sec	100페이지
volkswagen	2m 41sec	100페이지

마찬가지로 상당히 짧은 시간에 크롤링이 진행된것을 볼 수 있다. 실제로 해당 함수에서는 정적인 데이터만을 가져왔기 때문에 2 분 40 여초 시간이 걸려야 했으나 상대적으로 오랜 시간이 걸린것을 확인할 수 있다.



크롤링이 모두 끝난 이후 데이터베이스에 14436 개의 데이터가 존재하는 것을 확인할 수 있다.

다만, 이 때에 각각의 특허 링크에 접속하지 않았기 때문에 가져온 데이터는 한정적이다.

이는 추후 우리가 설계한 로봇의 2 번째 기능을 고안하면서 추가할 계획이다.

db 에 쌓인 데이터를 바탕으로 유의미한 정보를 추출해낼 것이다.

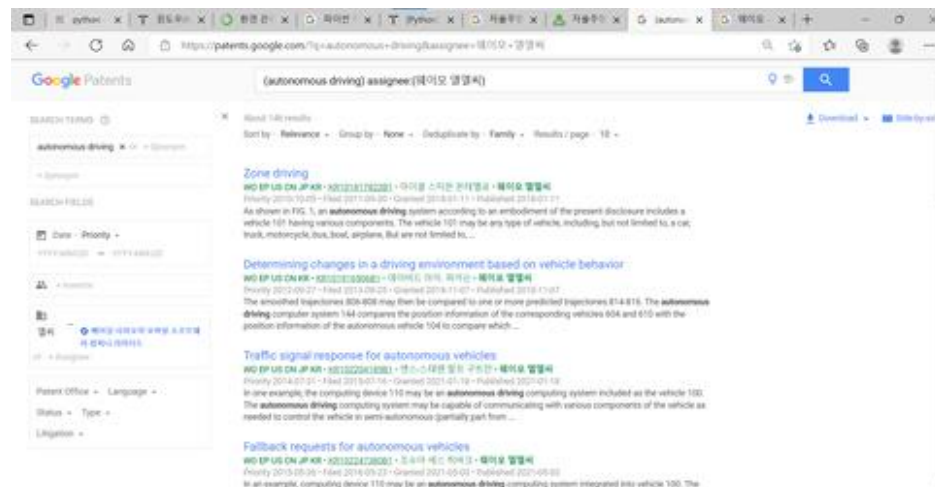
우리는 그 지표로 다른 특허에서 인용된 횟수와 출원 날짜를 고민해보았다.

이는 더 고민해보아야 할 것이다.

특허 페이지에 접속해 해당 정보들을 더 가져올 수 있도록 프로그램을 설계해야 할 것이다.

1 월 28 일 업무 진행 상황

우리가 만들 크롤링 로봇에 대해서 설명하고자 한다.



먼저 구글 patent 의 세부검색기능을 이용하여 자율주행관련된 기업에서 autonomous driving 가 포함된 특허들을 추출해내고자 한다.

가이드하우스 인사이트 자율주행 기술 종합 순위

순위	2019년	2020년	2021년
1	웨이모(구글)	웨이모(구글)	웨이모(구글)
2	크루즈(GM)	포드	엔비디아
3	포드	크루즈(GM)	아르고AI(포드·폭스바겐)
4	애플	바이두	바이두
5	인텔-모빌아이	인텔-모빌아이	크루즈
6	폭스바겐	현대차그룹-애플	모셔널(현대차그룹-애플)
7	다임러-보쉬	폭스바겐	모빌아이
8	바이두	안텍스	오로라
9	도요타	죽스	죽스
10	르노-닛산-미쓰비시	다임러-보쉬	뉴로

자료:가이드하우스 인사이트

현재 생각하고 있는 기업들은 기술 종합 순위에 존재하는 기업들로 가능하다면 해당 사이트에서 매년 기술 종합 순위가 업데이트 된다면

db 에 데이터를 계속 쌓아나가는 방식으로 설계하고 싶다.

이 때 db 에 데이터를 쌓아갈 때에 크롤링해서 정보를 얻어내는 곳은 모두 google patent 에서 얻어내도록 설계하였기에 동일한 데이터가 db 에 쌓이게 될 경우

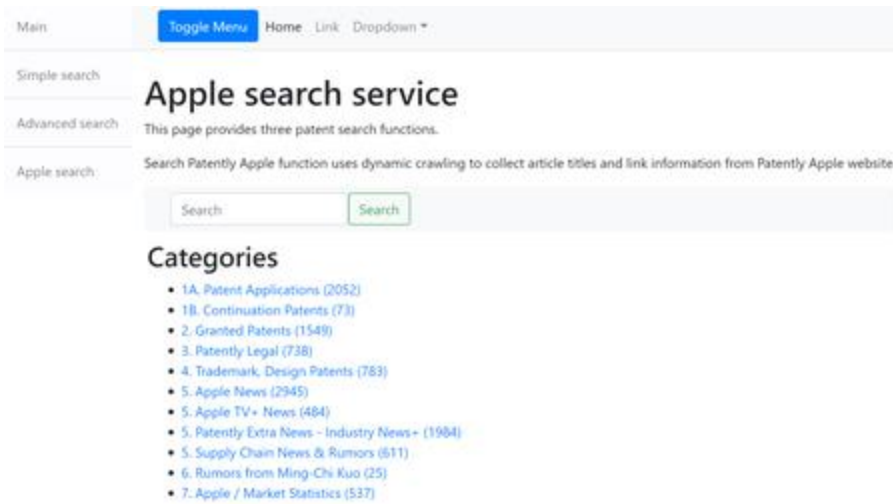
해당 데이터의 형태는 동일하다. 그렇기에 이를 고려해서 프로그램을 설계한다면 문제를 해결하기 쉬울 것이다.

또한, patently apple 을 통하여 차량 기술에 대한 특허들을 크롤링해 db 에 저장하는 것을 현재의 목표로 하고 있다.

이 때 크롤링 할 때에는 추후에 데이터베이스에서 로봇 2 가 필터링해내기 위해서 각각 특허의 링크내용을 포함하고 있어야 한다.

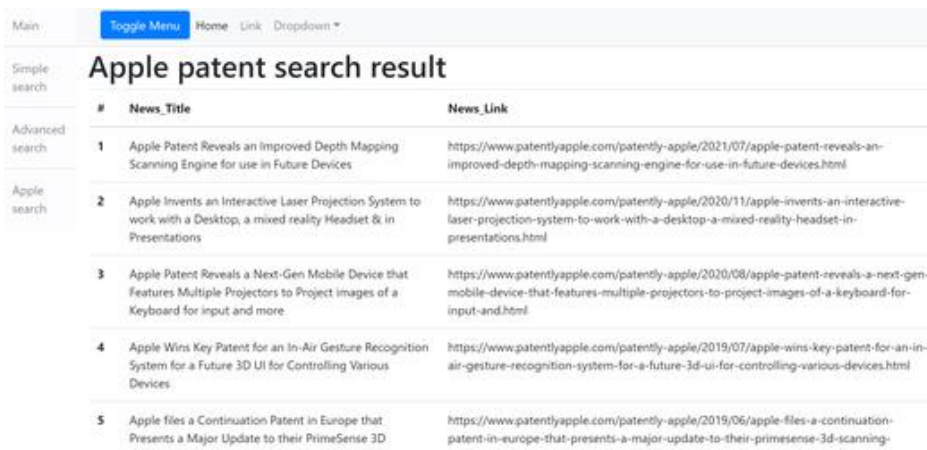
1 월 27 일 업무 진행 상황

- Apple patent archives 검색 기능 설정



간단하게 Apple search service 라는 창을 만들었다. Categories 는 patently apple 사이트에 있는 archives 를 모두 가져 온 것이다.

해당 archives filed 를 누르게 되면 그에 해당하는 특허 제목과 링크를 추출하도록 하였다.



특허 번호와 assignee, inventor 등 모든 정보를 가져오는 것은 더 많은 시간이 소요될 것 같다.

- 추가적으로 해야 할 업무

→ 위의 데이터베이스를 MongoDB 에 주기적으로 업로드 할 것

→ MongoDB 의 데이터들을 filtering 하는 기능

1 월 26 일 업무 진행 상황



기존 정의한 google patent 함수를 사용하여 반환된 데이터 프레임을 json file 로 변환하였다.

```
@app.route("/mongo")
def index():
    client = MongoClient('mongodb://localhost:27017/')
    db = client.my_database
    collection = db.my_collection
    results = collection.find()
    return render_template('index.html', data=results)

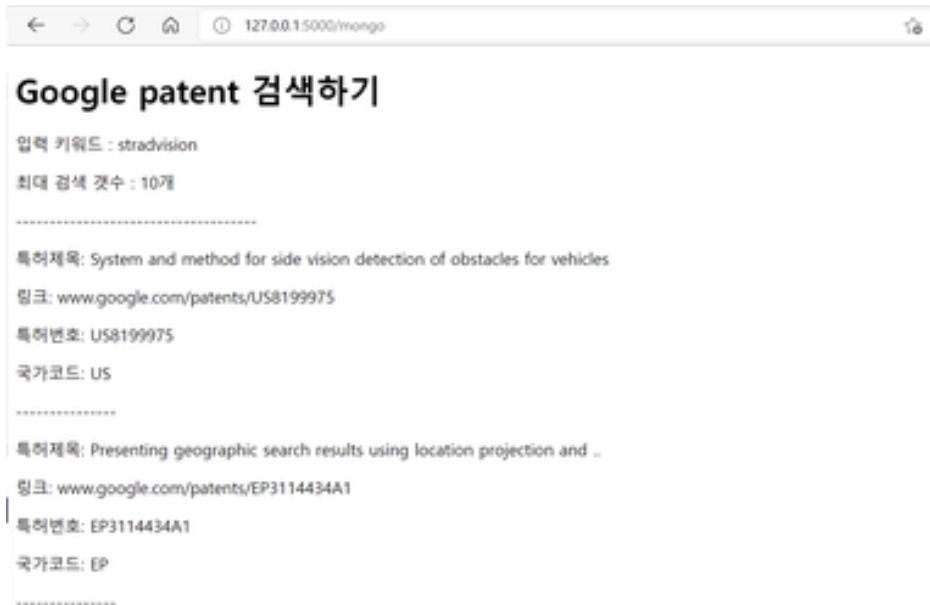
if __name__ == "__main__":
    app.run(debug=True)
```

```
* Detected change in 'c:\\Users\\LEGION\\Desktop\\projects\\myproject\\app.py', reloading
* Restarting with stat
* Debugger is active!
* Debugger PIN: 956-077-675
* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)
```

my_database 의 하위폴더 my_collection 에 있는 json 파일을 불러 와서 data 에 저장하였다.

html 구조로 변환하여 출력되도록 하였다.

- Mongo DB 와 python Flask 연동



- 추가적으로 해야 할 업무들

→ Action Diagram 에 따라 웹 페이지 상에서 입력이 가능하도록 해야함

→ 입력 및 결과에 대한 기본적인 웹 페이지 틀 구상하기

1 월 25 일 업무 진행 상황

- flask 를 이용한 웹페이지 만들어보기



하위 페이지로 우리가 만든 함수들로 구성하면 되겠다는 생각이 든다.

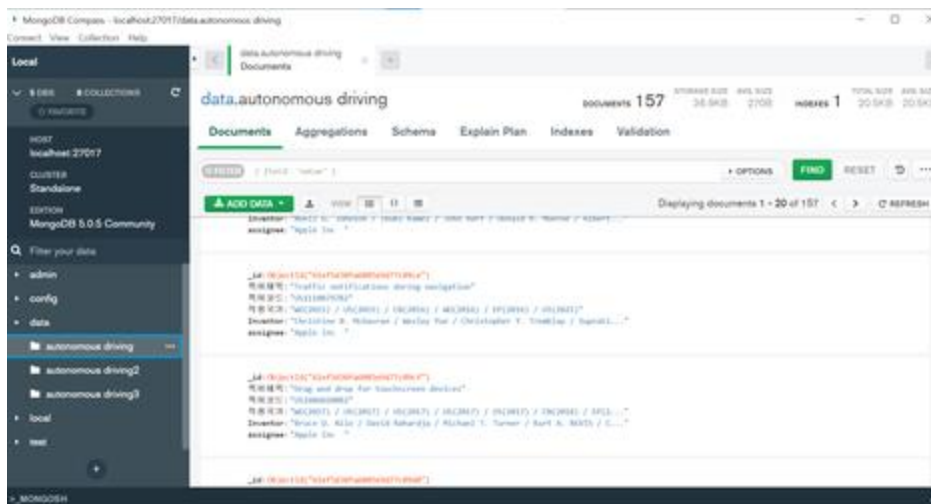
- 데이터 프레임을 db 에 저장하는 함수 선언

```
def df_to_db(keyword, df, client):
    items = df.to_dict("records")
    mydb = client['data']
    mycol = mydb[keyword]

    mycol.insert_many(items)

df_to_db('autonomous driving3',df,my_client)
```

해당 함수를 활용하여 데이터프레임의 저장이 가능한 것을 확인하였다.

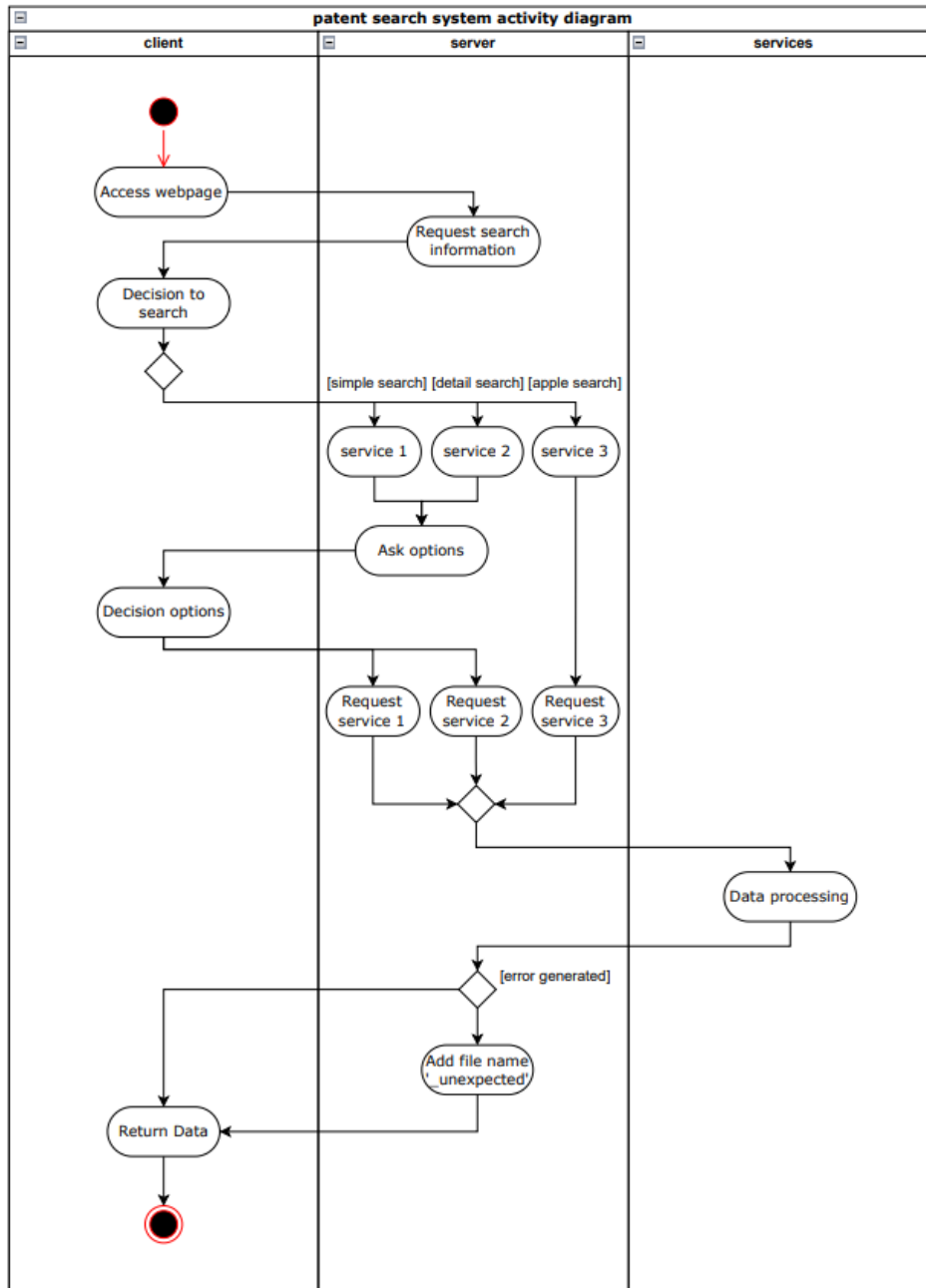


이 때 데이터프레임을 저장할 때의 이름은 보통 어떤 것으로 지정하는지는 더 찾아봐야 할 것 같다.

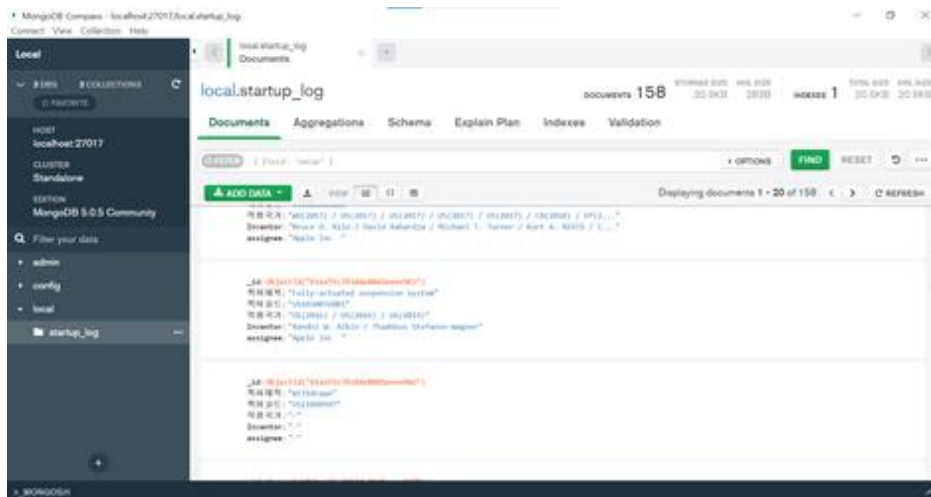
(데이터프레임을 생성한 시간 or 입력 키워드의 요약정보)

1 월 24 일 업무 진행 상황

- activity diagram 작성



- mongoDB 설치 및 csv 파일 업로드해보기



설치 및 사용법만 보았을 뿐 어떻게 사용하는지는 차차 관련서적과 인터넷을 통해 찾아볼 예정

1 월 20 일 업무 진행 상황

- 자율주행 관련 특허로 추출되는 데이터를 csv 형태로 변환하였다.

특허재목	특허코드	적용국가	Inventor	assignee
1 Passenger vehicle and door structure	US10309132B1	US(2017) / US(2017)	Noell G. Johnson / Ibuki Kamei / John F Apple Inc	
2 Traffic notifications during navigation	US11100797B2	WO(2015) / US(2015)	Christine B. McGavran / Wesley Yue / C Apple Inc	
3 Drag and drop for touchscreen devices	US10860200B2	WO(2017) / US(2017)	Bruce D. Nilo / David Rahardja / Michar Apple Inc	
4 Fully-actuated suspension system	US10300760B1	US(2016) / US(2016)	Randol W. Aikin / Thaddeus Stefanov-V Apple Inc	
5 Withdrawn	US11090997	-	-	-
6 System and method for visual communication of an operational status	US10053001B1	US(2016) / US(2016)	Bartholomeus C. Nabbe / Byron B. Han Apple Inc	
7 Compact optical packaging of LIDAR systems using diffractive structures	US20200348418A1	CN(2020) / US(2020)	Andrew J. Sutton / Michael C. Wharton Apple Inc	
8 Extendable bumpers for vehicles	US10336290B1	US(2016) / US(2016)	I Juan A. Lazaro / Erik Mikael Fermer / Al Apple Inc	
9 Body structure ventilation	US10787060B1	US(2017) / US(2017)	Johannes Mayr / John M. Kearney / Rei Apple Inc	
10 Electronic devices with millimeter wave antennas and metal housings	US10862195B2	DE(2019) / US(2015)	Yuehui Ouyang / Yi Jiang / Matthew A. Apple Inc	
11 Systems and methods for resolving hidden features in a field of vi	US20200342623A1	WO(2020) / WO(2020)	Christy F. Cull / Evan C. Cull Apple Inc	
12 Efficient connection management in a SAS target	US10089041B2	US(2016) / US(2012)	Arie Peled Apple Inc	
13 Device, Method, and Graphical User Interface for Presenting Vehicular No	US20180334175A1	WO(2018) / EP(2018)	Andre Boule / Stephen Chick / Thomas Apple Inc	
14 Systems with windows	US11180005B2	US(2020) / US(2018)	Peter F. Masschelein / Martin Melcher / Apple Inc	

- 자율주행 외에도 archives 에 다른 category 에도 동일하게 적용하여 함수를 사용할 수 있다.

Inductive Power Transf	US20210036553A1	WO(2020) / US(2018) / Aiguo Hu / Jeffrey D Apple Inc
Transient Power Management Circu	US20210075214A1	US(2020) / US(2020) / Samuel B. Schaevitz / Apple Inc
Inductive module	US20180233951A1	CN(2017) / US(2017) / Karl Ruben F. Larsson / Apple Inc
System and method for enhanced high throughput (eht) st	US20210076437A1	CN(2020) / US(2020) / Jarkko L. KNECKT / Apple Inc
Withdrawn	US80090973	-
Laser system	US10074958B2	WO(2016) / JP(2014) / Osamu Wakabayashi / Gigaphoton Inc
Method for manufacturing graphene using cover member	US10035708B2	WO(2014) / KR(2013) / Kilwon Cho / Hyojin Academy Industry F
Systems and methods for object detection by radio frequ	US20200300996A1	US(2019) / US(2019) / Berke Cetinoneri / Apple Inc
Withdrawn	US70117754	-
Multi-layer transmitter coil arrangement for wireless chargi	US20180091000A1	WO(2017) / US(2017) / Eric S. Jol / Christopher Apple Inc
Withdrawn	US80090730	-
Detection of coil coupling in an inductive charging system	US10110051B2	US(2017) / US(2014) / Todd K. Moyer / Ch Apple Inc
Wireless charging systems for electronic device	US20190363565A1	US(2019) / US(2019) / Christopher S. Graha / Apple Inc
Thermally optimized rx wireless charger for small rx devic	US20210036557A1	CN(2019) / US(2019) / Grant S. Haug / Chri Apple Inc
Configurable wireless transmitter device	US10158244B2	WO(2016) / EP(2016) / Christopher M. Pinc / Apple Inc
Withdrawn	US80233951	-
Inductive charging between electronic devices	US10886771B2	WO(2019) / US(2015) / Darshan R. Kasar / C Apple Inc
Electronic Devices with Broadband Ranging Capabiliti	US20200182964A1	CN(2019) / US(2019) / Carlo Di Nallo / Simc Apple Inc

- 사용자가 입력한 "autonomous driving" 키워드 검색에 따른 특허정보를 추출하는 기능도 가능하도록 하였다. (다른 키워드를 검색하더라도 동일하게 적용가능하다.)

applepatent_keyword()					
✓ 3m 40.6s Python					
Maximum page is 3					
	특허제목	특허코드	적용국가	Inventor	assignee
0	Automated capture of image data for points of ...	US10810444B2	WO(2016) / US(2016) / WO(2016) / US(2020)	Benjamin B. Lyon / Ahmad Al-Dahle / Xinyu Xu	Apple Inc
1	Method for representing virtual information in...	US10453267B2	WO(2018) / DE(2009) / JP(2010) / US(2010) / WO...	Peter Meier / Michael Kuhn / Frank Angermann	Apple Inc
2	Body structure ventilation	US10787060B1	US(2017) / US(2017)	Johannes Mayr / John M. Kearney / Reinhard HOE...	Apple Inc
3	Intention Recogniti	US20200302194A1	WO(2020) / US(2016) / CN(2016) / WO(2016) / CN...	Ahmad Al-Dahle / Bruno A. HEDSEL	Apple Inc
4	Radar system including dual receive array	US10446938B1	US(2016) / US(2016)	Jian Wang / Gregory E. Rogers	Apple Inc
5	Vehicle video system	US10810443B2	WO(2017) / WO(2017) / DE(2017) / CN(2017) / US...	Jared A. CRAWFORD / Yingen Xiong / Marco Pontil	Apple Inc
6	Systems for improving side-mirror functionalit...	US10486600B1	US(2016) / US(2016)	David G. Havskjold / Arthur Y. Zhang / Hyungry...	Apple Inc
7	System and method for road friction coefficient	US10442439B1	US(2017) / US(2017)	Youngwoo Seo / Randol Alkin / Kibata Gun	Apple Inc

1 월 19 일 업무 진행 상황

현재 알고리즘

- 필터 1 실행 -> 페이지 접속
추출되는 데이터의 형태 : 00,000,00
- 아무런 정보를 얻어내지 못한 링크들로부터 필터 2 실행 -> 페이지 접속
추출되는 데이터의 형태 : 20190 or 20200 or 20180

- 아무런 정보를 얻어내지 못한 링크들로부터 필터 3 실행 -> 페이지 접속
추출되는 데이터의 형태 : href = uspto

<https://www.patentlyapple.com/patently-apple/2020/12/a-supply-chain-report-claims-that-apples-vehicle-related-project-titan-is-moving-to-the-next-phase.html>
<https://www.patentlyapple.com/patently-apple/2020/05/apple-won-60-patents-today-covering-a-future-sliding-door-hinge-system-for-a-vehicle-and-a-future-helicopter.html>
<https://www.patentlyapple.com/patently-apple/2020/05/apple-wins-a-project-titan-patent-illustrating-next-gen-seat-belt-systems-that-includes-roof-and-lapbel-airbags.html>
<https://www.patentlyapple.com/patently-apple/2019/07/apple-invents-next-gen-augmented-reality-maps-requiring-smartglasses-or-iphone-with-an-camera-on-dash.html>
<https://www.patentlyapple.com/patently-apple/2019/06/a-project-titan-patent-reveals-autonomous-vehicles-that-can-communicate-with-surrounding-vehicles-and-pedestrians.html>
<https://www.patentlyapple.com/patently-apple/2019/05/apple-invents-a-vehicle-radar-system-that-reads-road-signs-and-warns-of-road-hazards-in-low-visibility-conditions-like-a-snowstorm.html>
<https://www.patentlyapple.com/patently-apple/2019/03/apples-project-titan-team-invents-an-advanced-nighttime-sensing-system-with-3x-the-power-of-traditional-headlights.html>
<https://www.patentlyapple.com/patently-apple/2019/03/a-project-titan-patent-reveals-how-headlight-tail-light-elements-will-house-camera-for-autonomous-mode-driving.html>
<https://www.patentlyapple.com/patently-apple/2018/10/a-project-titan-patent-focuses-on-a-new-anti-glare-system-to-protect-drivers-eyes-from-the-sun-and-lights-of-all-kinds.html>
<https://www.patentlyapple.com/patently-apple/2018/10/an-apple-patent-discovered-in-europe-today-covers-range-of-view-and-data-comparison-systems-for-driving-titan.html>
<https://www.patentlyapple.com/patently-apple/2018/10/apple-reveals-more-details-about-their-air-windshield-system-that-could-extend-to-driving-simulator.html>
<https://www.patentlyapple.com/patently-apple/2018/09/a-project-titan-invention-autonomous-vehicle-navigation-system-with-intention-recognition.html>
<https://www.patentlyapple.com/patently-apple/2018/09/european-apple-patent-describes-a-driver-accessing-and-starting-a-vehicle-with-an-iphone-or-apple-watch.html>
<https://www.patentlyapple.com/patently-apple/2018/09/apple-reinvents-a-vehicles-headlight-system-that-focuses-on-illuminating-objects-advances-air-windshield.html>
<https://www.patentlyapple.com/patently-apple/2018/07/like-a-science-fiction-novel-apple-reveals-how-your-personal-driving-profile-will-control-your-autonomous-car.html>

<https://www.patentlyapple.com/patently-apple/2018/06/apple-invents-an-autonomous-vehicle-guidance-system-that-recognizes-intent-signals-from-authorized-drivers.html>
<https://www.patentlyapple.com/patently-apple/2018/06/new-apple-invention-covers-inspection-and-primal-neural-networks-for-controlling-future-autonomous-vehicles.html>
<https://www.patentlyapple.com/patently-apple/2018/06/apple-was-granted-a-patent-this-week-for-the-roof-body-structure-of-a-vehicle.html>
<https://www.patentlyapple.com/patently-apple/2018/03/apples-most-advanced-project-titan-invention-surfaces-detailing-gesture-controls-for-a-heads-up-display.html>
<https://www.patentlyapple.com/patently-apple/2018/02/a-new-project-titan-patent-surfaces-in-europe-covering-interchangeable-heads-up-display-interfaces-more.html>
<https://www.patentlyapple.com/patently-apple/2017/12/the-first-apple-patent-regarding-autonomous-vehicle-navigation-was-published-today-by-the-us-patent-office.html>

필터 3의 알고리즘을 만들기 위하여 필터에 걸러지지 않은 남은 링크들을
 직접살펴보았으며, 정보가 추출되지 않은 링크들은

대체로 같은 특허에 관한 기사형태로, 다른 기사를 다시 설명하는 정보였다.

이 때 필터 1,2,3 에서 접속하는 페이지들이 동일한 페이지임

=> 때문에 더 오래걸림 (현재 필터 1 : 14 분, 필터 2 : 10 분, 필터 3 : 4 분정도 소요)

이를 보완하기 위한 작업이 필요 (보완시 14 분정도 소요될 것으로 예상)

※해당 시간은 patently apple 의 archives 중 vehicle technology 의 정보들을 크롤링한
시간으로 5 페이지 크롤링하는데 걸리는 시간이다

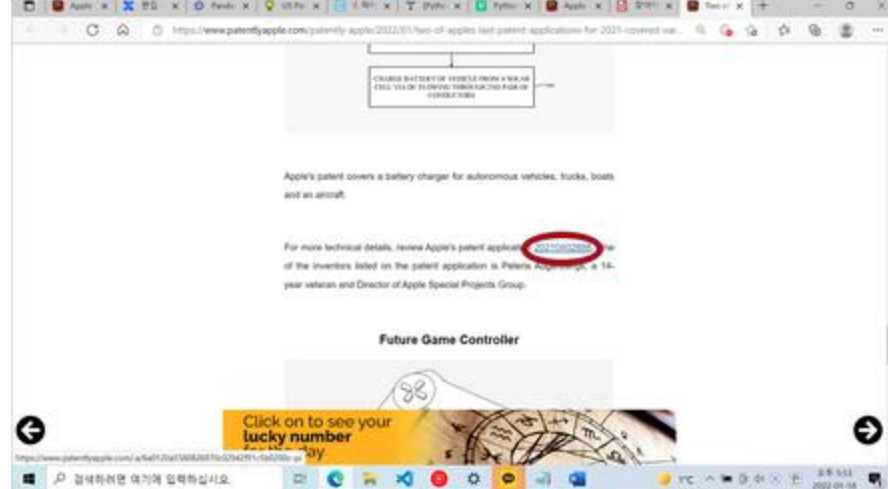
필터 1 을 통하여 추출된 특허 정보들

	특허제목	특허코드	적용국가	Inventor	assignee
0	Exterior lighting and warning system	US11104267B1	US(2018) / US(2016) / US(2018)	Clarisse Mazuir / Arthur Y. Zhang / Albert J. ...	Apple Inc.
1	System with windows	US10730368B1	US(2018) / US(2018) / US(2020)	Peter F. Masschelein / Martin Melcher / Derek ...	Apple Inc.
2	Climate control system with slit-vent fluid de...	US10752082B1	US(2017) / US(2017)	John M. Kearney / Clarisse Mazuir / Arthur Y. ...	Apple Inc.
3	Redundant vehicle actuator system	US11046330B1	US(2017) / US(2017)	Diomidis KATZOURAKIS / Hulbert Mees / Robert M. ...	Apple Inc.
4	Climate control	US10875380B2	WO(2015) / WO(2015) / US(2015) / EP(2015) / CN...	Derek A. Faust / Emery A. Sanford / James Foster	Apple Inc.
...
89	Active near-field communication device facilit...	US11005533B2	US(2019) / US(2019) / US(2021)	Vincent Chauvin / Peter M. Agboh / Vusthla Sun...	Apple Inc.
90	Gesture based control of autonomous vehicles	US10913463B2	WO(2017) / US(2017) / CN(2017) / WO(2017) / EP...	Scott M. Herz / Karlin Bark / Nguyen-Cat Le	Apple Inc.
91	Extendable bumpers for vehicles	US10343634B1	US(2017) / US(2017)	Erik Mikael Ferner / Alice C. Watts / Joel Fre...	FUNCTION ENGINEERING Inc/ Apple Inc.
92	Glass fastening and sealing systems	US10562274B1	US(2017) / US(2017) / US(2020)	David E. Kingman / Derek C. Scott / Donald R. ...	Apple Inc.
93	Fully-actuated suspension system	US10300760B1	US(2016) / US(2016) / US(2019)	Randol W. Alkin / Thaddeus Stefanov-Wagner	Apple Inc.

필터 2 를 통하여 추출된 특허 정보들을 추가한 데이터프레임

	특허제목	특허코드	적용국가	Inventor	assignee
0	Exterior lighting and warning system	US11104267B1	US(2018) / US(2016) / US(2018)	Clarisse Mazuir / Arthur Y. Zhang / Albert J. ...	Apple Inc.
1	System with windows	US10730368B1	US(2018) / US(2018) / US(2020)	Peter F. Masschelein / Martin Melcher / Derek ...	Apple Inc.
2	Climate control system with slit-vent fluid de...	US10752082B1	US(2017) / US(2017)	John M. Kearney / Clarisse Mazuir / Arthur Y. ...	Apple Inc.
3	Redundant vehicle actuator system	US11046330B1	US(2017) / US(2017)	Diomidis KATZOURAKIS / Hulbert Mees / Robert M. ...	Apple Inc.
4	Climate control	US10875380B2	WO(2015) / WO(2015) / US(2015) / EP(2015) / CN...	Derek A. Faust / Emery A. Sanford / James Foster	Apple Inc.
...
122	Color filter array for machine vision syst	US20180088586A1	WO(2017) / US(2017) / WO(2017)	Lucian Ion / Mahesh Krishnamurthy	Apple Inc.
123	Camera shades	US20180091717A1	WO(2017) / US(2017) / WO(2017)	Lucian Ion	Apple Inc.
124	Augmented virtual display	US20180089900A1	US(2017) / US(2017) / US(2017) / US(2019) / US...	Mark B. Rober / Sawyer I. Cohen / Daniel Kurz ...	Apple Inc.
125	Traffic direction gesture recognition	US20180082134A1	WO(2017) / US(2017) / WO(2017)	Sayanar V. Sivaraman	Apple Inc.
126	Vehicle Video System	US20180082135A1	WO(2017) / WO(2017) / DE(2017) / CN(2017) / US...	Jared A. CRAWFORD / Yingen Xiong / Marco Pottel	Apple Inc.






```
new_list = []
for v in list2:
    if v not in new_list:
        new_list.append(v)
print(new_list)

[ '20210402888', '20210402290', '20210229630', '20210218994', '20210160817', '20210135507', '20210089041', '20200348143', '20200049837', '20200348418',
'20200342623', '20200314614', '20200258400', '20200255002', '20200214148' ]

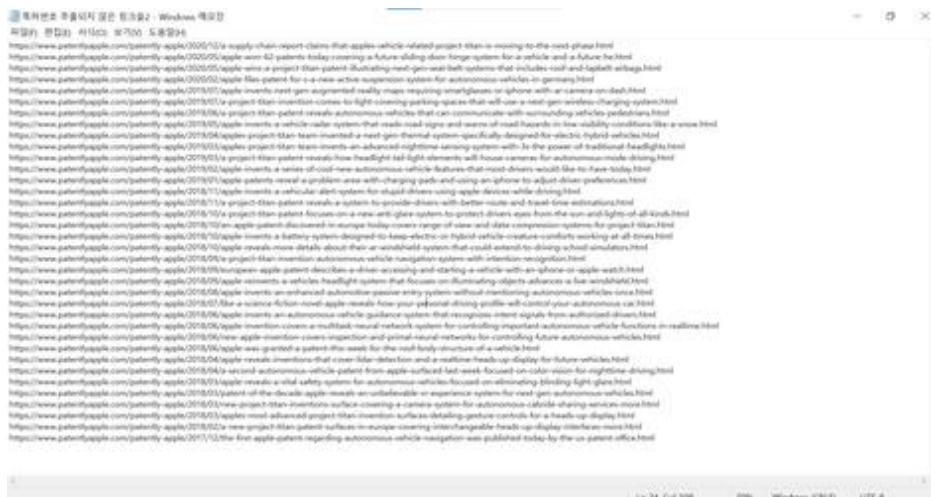
print(len(new_list))

df = pd.read_table('특허번호 추출되지 않은 링크들2.txt', header=None, sep=" ")
print(len(df))
df[0][0]
print(type(df))
```

추출되지 않은 링크들은 55 개에서 36 개로 줄었다.

하지만 처음 00,000,000 의 형태로 걸러졌어야할 링크들에서 걸러지지 못한 링크들이 존재했다.

때문에 해당 문제가 발생한 이유에 대해서 찾아봐야 할 것 같다.

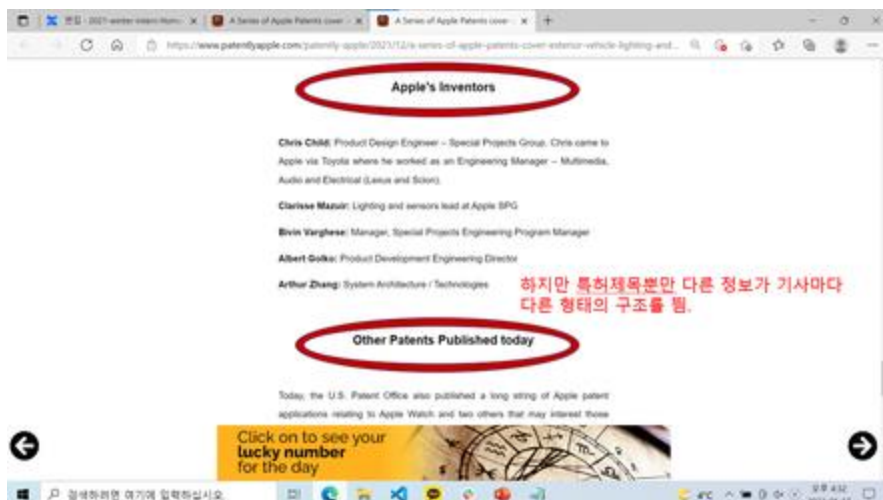


1 월 17 일 업무 진행 상황

특허번호가 추출되지 않은 링크들

[illegible]

14 일자 업무 결과를 살펴보면 특허코드가 94 개 추출된 것을 확인할 수 있다. 이는 00,000,000 형태의 특허번호만을 추출하여 변환한 결과이다.



다만 실제 사이트에는 중간에 특허 제목이 언급되어있는 경우도 존재한다. 위의 2 사진은 동일한 링크에서의 결과이다.

제목과 다른 정보들을 구별해낼 수 있는 방법이 필요하고, 이 외의 형태들일 경우 특허 정보를 검색할 방안도 필요하다.

현재 고안하고 있는 해결방안은 다음과 같다.

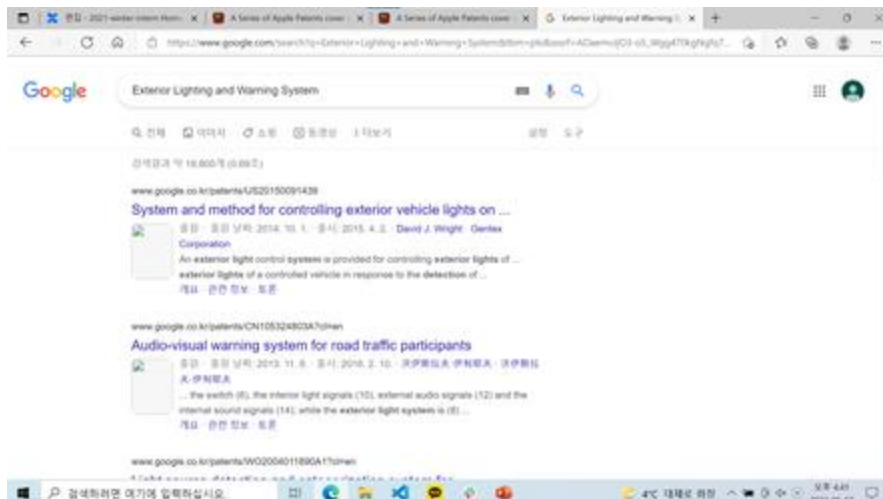
- 해결방안의 알고리즘

- 빨간색 동그라미의 모든 정보를 가져온 이후 여러 개의 요소가 존재하는 것들은 제거한다. (Inventors 와 같은 다른 기사들에서의 공통적인 내용이 삭제될 것으로 예상)

다만 걱정되는 사항은 한 개의 특허에 관해 여러 개의 기사가 존재하는 경우에는 데이터가 손실될 수 있다는 것이다.

(이는 프로그래밍과정 중에서 삭제되는 요소들을 print 를 통해 직접 확인해서 제목이 존재하는지 확인해 볼 것임)

- 얻어낸 제목을 통하여 검색할 때에는 동적 크롤링인 세부검색을 사용해야 한다.



위의 사진은 google patent 에서 제목을 검색한 결과이다. 일치하는 검색결과가 나타나지 않는 것을 확인할 수 있으며,



세부 검색을 통하여 첫 페이지에 제목이 가장 유사한 특허를 찾아내면 해결가능할 듯 하다.

다만 그 외에 txt 파일에 존재하는 링크들에는 다양한 예외들이 존재한다.

그렇기 때문에 모든 특허들에 대한 정보가 필요하다면 파이썬을 통하여 얻어내지 특허들은 직접 링크에 접속해 수작업으로 마무리해야할 것 같다.

1 월 14 일 업무 진행 상황

- apple patent 의 특허 정보수집

... Maximum page is 5

	News Title	News Link
0	Apple won two Project Titan patents today cove...	https://www.patentlyapple.com/patently-apple/2...
1	Two of Apple's last Patent Applications for 20...	https://www.patentlyapple.com/patently-apple/2...
2	Apple wins a Pair of Project Titan Patents cov...	https://www.patentlyapple.com/patently-apple/2...
3	Apple has won a patent for a Suspension System...	https://www.patentlyapple.com/patently-apple/2...
4	Apple wins a Project Titan Patent relating to ...	https://www.patentlyapple.com/patently-apple/2...
...
136	Patent of the Decade: Apple Reveals an Unbelie...	https://www.patentlyapple.com/patently-apple/2...
137	New Project Titan Inventions Surface covering ...	https://www.patentlyapple.com/patently-apple/2...
138	Apple's Most Advanced Project Titan Invention ...	https://www.patentlyapple.com/patently-apple/2...
139	A New Project Titan Patent Surfaces in Europe ...	https://www.patentlyapple.com/patently-apple/2...
140	The First Apple Patent Regarding Autonomous Ve...	https://www.patentlyapple.com/patently-apple/2...

141 rows x 2 columns

- 해당 링크로부터 US 특허코드 추출

```
> print(len(driving_ptcode))
print(driving_ptcode)

134]

... 94
['1104267', '10730368', '10752082', '11046338', '10875380', '10053001', '10970518', '10465314', '10351162', '10769525', '10594850', '10314116',
'10293718', '10308290', '11130391', '10309132', '10555133', '10661748', '11214100', '11173766', '10661669', '10814690', '11092689', '11021126',
'10446938', '11056928', '11163157', '10788316', '11006257', '10328897', '11100797', '10948582', '11124035', '10906370', '10108202', '10862195',
'10916056', '10569672', '10919520', '10871555', '10860034', '10908328', '10479413', '10955531', '10661764', '10625580', '10891745', '10922886',
'10486608', '11092690', '10222529', '10787192', '10189434', '11003147', '10415286', '11046143', '10852553', '10656777', '10558037', '10630127',
'10860142', '11180005', '11090997', '10662676', '10336290', '10798673', '10247816', '10482669', '10878699', '11179991', '10889261', '11097658',
'10286761', '10866414', '10919482', '10872228', '10827564', '10757373', '11104291', '10988142', '10712446', '11214997', '10412334', '10627248',
'10372132', '10928597', '10860200', '10843535', '10442439', '11005533', '10913463', '10343634', '10562274', '10300760']
```

1 월 13 일 업무 진행 상황

- 동적크롤링을 활용한 특허 관련 기사 제목과 링크추출

Search_patent_news 사용예제

```
[34] Apple_patently = Search_patent_news()
```

```
[35] display(Apple_patently)
```

	News_Title	News_Link
0	Apple Won 43 Patents Today Covering a Future A...	https://www.patentlyapple.com/patently-apple/2...
1	A new Project Titan Patent from Apple covers a...	https://www.patentlyapple.com/patently-apple/2...
2	Apple Wins Patent for Handling Emergency Situa...	https://www.patentlyapple.com/patently-apple/2...
3	Apple won three 'Project Titan' patents today ...	https://www.patentlyapple.com/patently-apple/2...
4	Apple Files Patent for s a new 'Active Suspens...	https://www.patentlyapple.com/patently-apple/2...
5	Apple Wins Patent for Vehicle side mirrors tha...	https://www.patentlyapple.com/patently-apple/2...
6	The Rumor of Apple Acquiring drive.ai earlier ...	https://www.patentlyapple.com/patently-apple/2...
7	Apple Won 52 Patents Today covering a Futurist...	https://www.patentlyapple.com/patently-apple/2...
8	Google to Drop out of Making Autonomous Cars a...	https://www.patentlyapple.com/patently-apple/2...
9	A Series of Apple Patents cover exterior Vehic...	https://www.patentlyapple.com/patently-apple/2...
10	A Core Engineer who worked on both NeXT Comput...	https://www.patentlyapple.com/patently-apple/2...
11	Apple Wins Two Project Titan Patents covering ...	https://www.patentlyapple.com/patently-apple/2...
12	Apple has Won Two 'Project Titan' Patents cove...	https://www.patentlyapple.com/patently-apple/2...
13	Apple Wins a new Series of Project Titan Paten...	https://www.patentlyapple.com/patently-apple/2...
14	Apple's CEO Confirms Project Titan is the 'Mot...	https://www.patentlyapple.com/patently-apple/2...
15	Intel on Target to Deliver a 10nm Discreet Gra...	https://www.patentlyapple.com/patently-apple/2...
16	Xiaomi Rockets to Second Place Globally in Q2-...	https://www.patentlyapple.com/patently-apple/2...
17	A new Supply Chain report Claims that Foxconn ...	https://www.patentlyapple.com/patently-apple/2...
18	Quick Notes from the Wall Street Journal Inter...	https://www.patentlyapple.com/patently-apple/2...
19	Apple Partner Foxconn and Zhejiang Geely have f...	https://www.patentlyapple.com/patently-apple/2...
20	The NTBS Slammed Tesla today for the death of ...	https://www.patentlyapple.com/patently-apple/2...
21	A Group of Apple Employees have Openly Rebelle...	https://www.patentlyapple.com/patently-apple/2...
22	Apple Won a Trio of Project Titan Patents toda...	https://www.patentlyapple.com/patently-apple/2...
23	As Expected, Apple Launched Podcasts Subscript...	https://www.patentlyapple.com/patently-apple/2...

→ Autonomous driving 키워드를 입력했을때의 검색된 기사 제목과 링크를 추출한 결과이다.

→ 동적 크롤링을 활용하였기 때문에 정적 크롤링을 사용한 이전의 Search 함수와 달리 최대 검색 갯수를 입력 parameter 로 인가하지 않았다.

```
try:
    if(k==0):
        driver.find_element_by_xpath(
    if(k > 0):
        driver.find_element_by_xpath(
except NoSuchElementException :
    print("Maximum page is ", k+1)
    break
```

→ NoSuchElementException 예외처리를 하여 Next page 가 존재하지 않는 경우 함수를 종료하도록 하였다.

→ 따라서 검색 키워드에 따른 최대 페이지까지 모든 정보를 수집가능하다.

autonomous driving 의 경우 Maximum page 는 3 이며 총 24 개의 정보가 검색 되었다.

```
[31] Apple_patently = Search_patent_news()
```

```
검색할 키워드를 입력하세요: autonomous driving
/usr/local/lib/python3.7/dist-packages/ipykern
/usr/local/lib/python3.7/dist-packages/ipykern
/usr/local/lib/python3.7/dist-packages/ipykern
/usr/local/lib/python3.7/dist-packages/ipykern
Maximum page is 3
```

1 월 12 일 업무 진행 상황

동적크롤링과 정적크롤링의 시간차이

```
url = 'https://patents.google.com/patent/KR2020027885A/en'
data = requests.get(url)
soup = BeautifulSoup(data.content, 'html.parser')
title = soup.select_one('body > search-app > article > h1')
print(title.text)

✓ 0.5s Python
```

KR2020027885A - Method and device for generating image data set to be used for learning cnn capable of detecting obstruction in autonomous driving circumstance, and testing method, and testing device using the same

- Google Patents

위의 사진은 정적인 데이터만 가져오는 크롤링 방식이다 0.5 초가 걸린것을 확인할 수 있다.

```
# from selenium import webdriver
# import time

# driver = webdriver.Chrome('./chromedriver')

url = 'https://patents.google.com/patent/KR2020027885A/en'
driver.get(url)
html = driver.page_source
soup = BeautifulSoup(html, 'html.parser')

title = soup.select('#title')
for t in title:
    print(t.text)

# driver.close

[P] ✓ 3.9s Python
```

... Title

Method and device for generating image data set to be used for learning cnn capable of detecting obstruction in autonomous driving circumstance, and testing method, and testing device using the same

위의 사진은 selenium 을 활용하여 직접 웹페이지에 접속해서 데이터를 가져오는 방식이다. 3.9 초가 걸린것을 확인할 수 있다.

최대한 정적인 웹페이지에서 가져오도록 코드를 작성한다면 프로그램이 더 빠르게 동작할 것이다

```
query - Windows 메모장
파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)
input = (page number, keywords, inventor, assignee, country)
autonomous driving, deep learning, CNN
김계현
stradvision
KR,US,JP
(end)
```

Ln 6, Col 6 100% Windows (CRLF) UTF-8

세부사항의 input query 에 따라 동작하는 함수를 선언하였다.


```
search_details 사용예제

patent_name, patent_code, inventor, assignee, country = search_details()

# 리스트가 올바른지 확인하기 위한 검증코드
for p in patent_name:
    print(p)
print(len(patent_name))
```

search_details 를 이용하여 list 를 생성하고,

```
search_details_df 사용예제

output = search_details_df(patent_name, patent_code, inventor, assignee, country)

display(output)
```

	특허제목	특허코드	inventor	assignee	적용국가
0	... set to be used for learning cnn capable of ...	KR20200027885A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
1	Learning method and learning device for integr...	KR20200095333A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
2	Learning method and learning device for runtim...	KR20200093424A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
3	Method and device for segmenting image to be u...	KR20200091317A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
4	Learning method and learning device for segmen...	KR20200091781A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
...
59	Method and device for generating deceivable co...	KR20200095337A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
60	... whether to switch mode of vehicle from manu...	KR20200094643A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
61	Learning method and learning device for reduci...	KR20200093417A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR

search_details_df 을 이용해 데이터프레임으로의 변환이 가능하다.

여기에서 논문 제목이 길다면 ...으로 표기되는데 이는 또 다른 페이지에 접속해서 제목을 가져와야한다.

때문에 더 오랜시간이 소요되고 이는 optional 로 함수를 정의하였다.

```
change_ptname 사용예제

output2 = change_ptname(output)

display(output2)
```

	특허제목	특허코드	inventor	assignee	적용국가
0	Method and device for generating image data se...	KR20200027885A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
1	Learning method and learning device for integr...	KR20200095333A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
2	Learning method and learning device for runtim...	KR20200093424A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
3	Method and device for segmenting image to be u...	KR20200091317A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
4	Learning method and learning device for segmen...	KR20200091781A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
...
59	Method and device for generating deceivable co...	KR20200095337A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR
60	Learning method and learning device for deter...	KR20200094643A	김계현	주식회사 스트라드비전	EP, US, CN, JP, KR

change_ptname 함수를 사용하여 해결할 수 있고

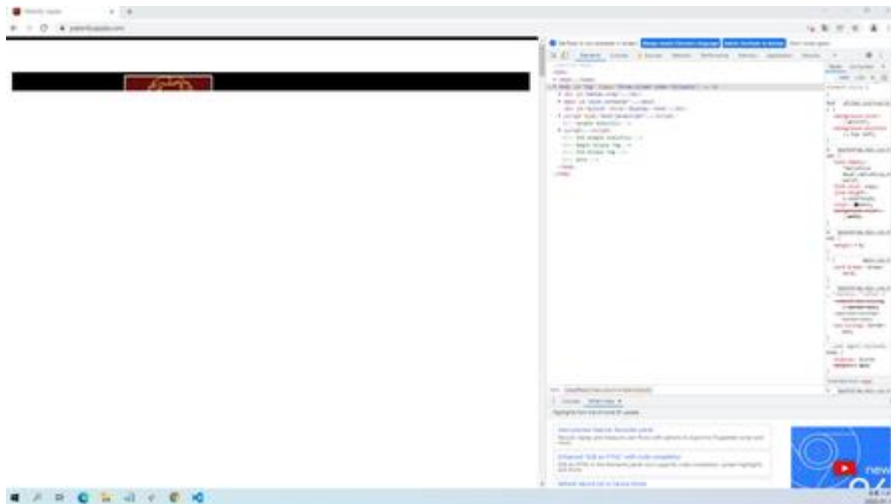
```
print(output['특허지도'])[0])
```

✓ 0.4s Python

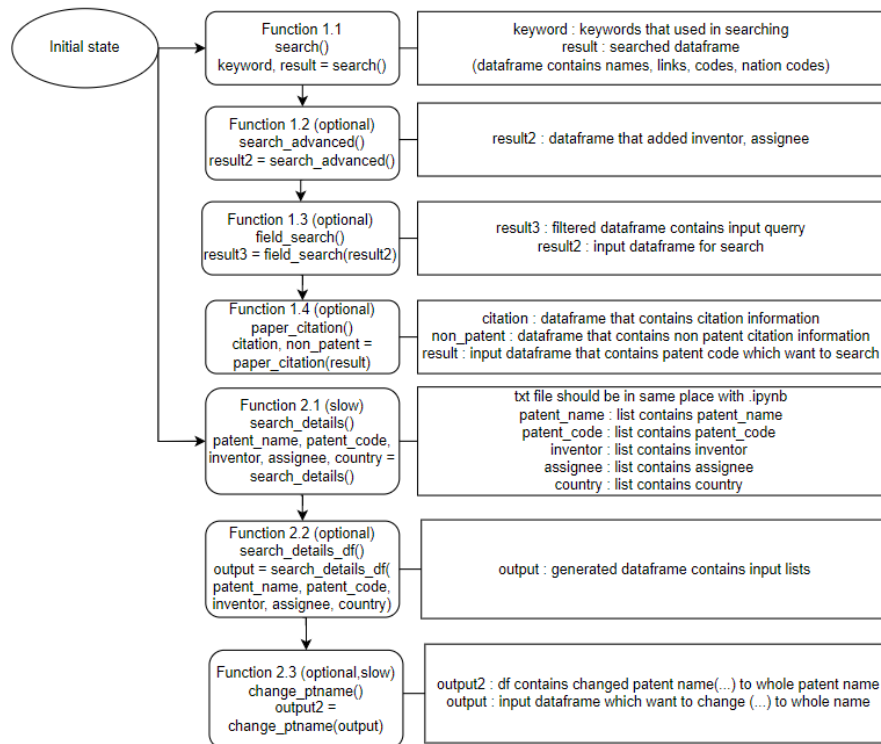
Method and device for generating image data set to be used for learning cnn capable of detecting obstruction in autonomous driving circumstance, and testing method, and testing device using the same

전체 제목으로 변화였음을 확인할 수 있다.

– Patentlyapple.com



javascript 를 꺾을 때에는 아무런 페이지가 나오지않는 것을 볼 수 있다. 때문에 동적크롤링을 사용해야한다.



현재 진행 상황 flow chart

function1 은 조금 더 단순한 기능을 수행한다. 정적인 웹페이지 구조만을 가져오기 때문에 더 빠른 연산속도인 것을 확인할 수 있다.

function2 는 더 복잡한 기능을 수행한다. 셀레니움을 활용하여 직접 웹페이지에 접속해 동적인 웹페이지 구조를 가져온다.

때문에 function1 에서는 페이지수를 지정해주었지만 function2 에서는 검색되는 모든 정보를 저장할 수 있지만, 그만큼 오래 걸린다.

또한 구글에서는 검색결과가 1000 개를 넘어갈 경우에는 실제로 검색결과가 더 존재하지만 보여주지 않는다.

따라서 1000 개가 넘어가는 경우에는 input query 를 수정하여 검색결과를 줄일 필요가 있다.

1 월 11 일 업무 진행 상황

- Family to family 을 제외한 patent citation 데이터 추출
- Non patent citations 데이터 추가

▼ paper_citation 사용예제

citation : 입력한 논문을 인용한 논문들의 데이터프레임

```
[ ] citation, Non_Patent = paper_citation(result)
```

특허번호 입력 :CN105844239A

```
[ ] display(citation)
```

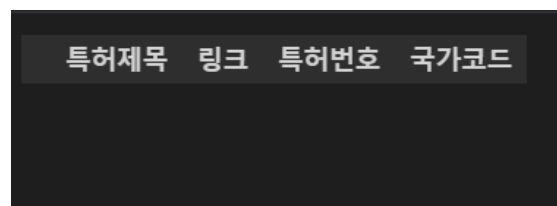
	Publication number	Priority date	Publication date	Assignee	Title
0	CN103218608A	2013-04-19	2013-07-24	中国科学院自动化研究所	一种网络暴力视频的识别方法
1	CN103473555A	2013-08-26	2013-12-25	中国科学院自动化研究所	基于多视角多示例学习的恐怖视频场景识别方法
2	CN105005772A	2015-07-20	2015-10-28	北京大学	一种视频场景检测方法

```
[ ] display(Non_Patent)
```

	Non-Patent Citations title
0	YQWANG2006: "SPP_mask阅读报告", 《百度文库》
1	ZHANGWEN XU 等: ""A Discriminative CNN Video Re...
2	ZUXUAN WU 等: ""Modeling Spatial-Temporal Clues...

→ paper citation 함수를 사용하여 citation, 과 Non_patent 라는 데이터 프레임을 반환하도록 하였다. Non-patent citation 은 Title 내용만 있기 때문에 데이터 프레임을 각각 분할 생성하였다.

1 월 10 일 업무 진행 상황



vs 코드에서 html 에서 필요한 정보를 가져오지 못하는 문제가 존재

	특허제목	링크	특허번호	국가코드
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.co.kr/patents/WO2017079522A1	www.google.co.kr/patents/WO2017079522A1	www
1	Method for increasing object detection rate or...	www.google.co.kr/patents/US9230193	www.google.co.kr/patents/US9230193	www
2	Proposition de région pour des régions d'image...	www.google.co.kr/patents/WO2017139927A1	www.google.co.kr/patents/WO2017139927A1	www
3	Method for detecting texts included in an imag...	www.google.co.kr/patents/US9524430	www.google.co.kr/patents/US9524430	www
4	Three-dimensional (3d) convolution with 3d bat...	www.google.co.kr/patents/WO2017031088A1	www.google.co.kr/patents/WO2017031088A1	www
5	Pedestrian notifications	www.google.co.kr/patents/US8954252	www.google.co.kr/patents/US8954252	www
6	Deepstereo: learning to predict new views from...	www.google.co.kr/patents/WO2016183464A1	www.google.co.kr/patents/WO2016183464A1	www
7	Distance measurement device and environment ma...	www.google.co.kr/patents/CN103154666A	www.google.co.kr/patents/CN103154666A	www
8	Driving support method and apparatus	www.google.co.kr/patents/EP1878618A2	www.google.co.kr/patents/EP1878618A2	www
9	Method, apparatus and computer readable record...	www.google.co.kr/patents/WO2013048162A2	www.google.co.kr/patents/WO2013048162A2	www

해결하였으나 replace()함수가 동작하지 않아서 직접 문자열을 slicing 해야할듯

→해당문제인줄알았으나 www.google.co.kr 과 www.google.com 주소가 두개 존재하기 때문에 발생하던 문제

	특허제목	링크	특허번호	국가코드
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1	017079522A1	01
1	Method for increasing object detection rate or...	www.google.com/patents/US9230193	230193	23
2	Method for detecting texts included in an imag...	www.google.com/patents/US9524430	524430	52
3	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1	017139927A1	01
4	Pedestrian notifications	www.google.com/patents/US8954252	954252	95
5	Three-dimensional (3d) convolution with 3d bat...	www.google.com/patents/WO2017031088A1	017031088A1	01
6	Deepstereo: learning to predict new views from...	www.google.com/patents/WO2016183464A1	016183464A1	01
7	Driving support method and apparatus	www.google.com/patents/EP1878618A2	878618A2	87
8	Distance measurement device and environment ma...	www.google.com/patents/CN103154666A	03154666A	03
9	Method, apparatus and computer readable record...	www.google.com/patents/WO2013048162A2	013048162A2	01

보완할 수 있는 코드가 필요해보임

1 월 6 일 업무 진행 상황

▼ search 함수 사용예제

result : 생성하고자 하는 데이터 프레임

✓
13
초

```
[19] keyword, result = search()
```

최대 검색갯수 입력(숫자) : 10
키워드 입력(끝내고싶으면 end 입력) : CNN
키워드 입력(끝내고싶으면 end 입력) : deep learning
키워드 입력(끝내고싶으면 end 입력) : stradvision
키워드 입력(끝내고싶으면 end 입력) : end

	특허제목	링크	특허번호	국가코드	적용국가	Inventor	assignee
0	Method and system for vision-centric deep-lear...	www.google.com/patents/US9760806	US9760806	US	US(2016) / CN(2017)	Guanghan Ning / Haohong Wang / Wenqiang Bo / X...	TCL Research America Inc
1	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US	US(2015) / JP(2016) / EP(2016)	José A. Rodríguez-Serrano / Albert Gordo Solde...	Xerox Corp
2	Regionlets with shift invariant neural pattern...	www.google.com/patents/US9202144	US9202144	US	US(2014)	Xiaoyu Wang / Yuanqing Lin / Will Zou / Miao SUN	NEC Corp
3	Hierarchical deep convolutional neural network...	www.google.com/patents/US20160117587	US20160117587	US	US(2014) / CN(2015) / WO(2015) / EP(2015) / KR...	Zhicheng Yan / Robinson Piramuthu / Vignesh JA...	eBay Inc
4	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US	US(2016)	Jian Sun / Ross Girshick / Shaoqing Ren / Kaim...	Microsoft Technology Licensing LLC
5	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282	US9633282	US	US(2015)	Arjun Sharma / Pramod Sankar Kompalli	Xerox Corp
6	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315	US20170169315	US	US(2016)	Gonzalo Vaca Castano / Syed Zain Masood / Step...	Sighthound Inc
7	Context-based priors for object detection in I...	www.google.com/patents/US20170011281	US20170011281	US	US(2015) / KR(2016) / JP(2016) / EP(2016) / CN...	Daniel Hendricus Franciscus DIKMAN / Regan BL...	Qualcomm Inc
8	Hierarchical deep convolutional neural network	www.google.com/patents/WO2016069581A1	WO2016069581A1	WO	US(2014) / CN(2015) / WO(2015) / EP(2015) / KR...	Zhicheng Yan / Robinson Piramuthu / Vignesh JA...	
9	Neural network image representation	www.google.com/patents/US9396415	US9396415	US	IL(2014) / US(2015) / US(2016)	Michael Chertok / Alexander LORBERT / Adi Pinhas	Superfish Ltd

키워드 CNN, deep learning, stradvision 검색 결과에 대해 field 검색을 수행하도록 하였다.

field 를 잘못 검색 할 경우 오류검출문과 함께 다시 입력하도록 기능을 수정하였음.

field_search 사용예제

```
[19] result3 = field_search(result2)
```

검색가능한 field들 : ['특허제목', '링크', '특허번호', '국가코드', '적용국가', 'Inventor', 'assignee']
검색할 field를 입력하세요(끝내고싶으면 end 입력) : 국가번호
검색할 field를 다시 입력하세요(끝내고싶으면 end 입력) : 국가번호
검색할 키워드를 입력하세요 : US
검색할 field를 입력하세요(끝내고싶으면 end 입력) : assignee
검색할 키워드를 입력하세요 : Xerox Corp
검색할 field를 입력하세요(끝내고싶으면 end 입력) : end

```
[20] display(result3)
```

	특허제목	링크	특허번호	국가코드	적용국가	Inventor	assignee
1	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US	US(2015) / JP(2016) / EP(2016)	José A. Rodríguez-Serrano / Albert Gordo Solde...	Xerox Corp
5	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282	US9633282	US	US(2015)	Arjun Sharma / Pramod Sankar Kompalli	Xerox Corp

→ field 를 잘못 입력 할 시, '검색할 field 를 다시 입력하세요' 경고문과 함께 다시 입력가능하도록 하였음.

→ field_search 사용예제에서 국가코드가 US 이면서 Current_assignee 가 Xerox Corp 인 특정 필드 검색을 하였고 해당 결과가 잘 나타난 것을 확인 하였다.

1 월 5 일 업무 진행 상황

생성된 데이터셋 내에서 검색기능을 수행해 검색된 데이터들을 새로운 데이터프레임으로 만들어주는 기능

field_search 라는 함수이름으로 정의되어서 검색필드를 먼저 지정하고 해당 내용을 검색하면 포함된 내용이 들어간 행 추출

field_search 사용예제



	특허제목	링크	특허번호	국가코드	작품국가	Inventor	assignee
1	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US	US(2015) / JP(2016) / EP(2016)	José A. Rodríguez-Serrano / Albert Gordo Solde...	Xerox Corp
2	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US	US(2016)	Jian Sun / Ross Girshick / Shaoqing Ren / Kaim...	Microsoft Technology Licensing LLC
3	Cross-trained convolutional neural networks us...	www.google.com/patents/US20170032222	US20170032222	US	US(2015)	Arjun Sharma / Pramod Sankar Kompalli	Xerox Corp
5	Deeply learned convolutional neural networks L...	www.google.com/patents/US20170169315	US20170169315	US	US(2016)	Gonzalo Vaca Castano / Syed Zain Masood / Step...	Sighthound Inc
6	Multi-task deep convolutional neural networks ...	www.google.com/patents/US9286524	US9286524	US	US(2015)	Xue Mei / Danil V. Prokhorov / Jun Li / Dachen...	University of Technology Sydney/ Toyota Motor...

< documentation >

result3 = 검색된 새로운 데이터 프레임

result2 = 검색할 데이터 프레임

(이 기능은 search_advanced 를 사용하지 않아도 사용할 수 있도록 정의하였음)

현재 사용 방식은 (검색할 field 지정 → 키워드 입력)의 과정을 반복해서 해당하는 데이터들을 데이터 프레임으로 생성

다만 검색 field 를 지정할 때 아래 사진처럼 field 를 잘못 입력할 경우에는 이미 많이 검색했더라도 error 가 발생해 새로 입력해야하는 불편함이 존재함.

```
result3 = field_search(result2)

검색할 field를 입력하세요(물냉고있으면 end 입력) : 기LUCI
검색할 키워드를 입력하세요 : goot

KeyError                                Traceback (most recent call last)
/usr/local/lib/python3.7/dist-packages/pandas/core/indexes/base.py in get_loc(self, key, method, tolerance)
    2897         try:
    2898             return self._engine.get_loc(casted_key)
    2899         except KeyError as err:
    2900             except KeyError as err:

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: '기LUCI'

The above exception was the direct cause of the following exception:

KeyError                                Traceback (most recent call last)
3 frames
/usr/local/lib/python3.7/dist-packages/pandas/core/indexes/base.py in get_loc(self, key, method, tolerance)
    2898         return self._engine.get_loc(casted_key)
    2899         except KeyError as err:
    2900             raise KeyError(key) from err
    2901
    2902         if tolerance is not None:

KeyError: '기LUCI'
```

이를 해결하기위해 while 문을 사용해 '키워드를 다시 입력하세요' 라는 경고문구를 출력할 예정

<진행되어야할 업무>


- 경고문구 출력
- search 함수 return 이 2 개인데 간소화할 방안


1 월 4 일 업무 진행 상황


특허 발명자(inventor) 와 소유자정보(Current assignee) 검색기능 추가

US20170011281A1

United States

 Download PDF

 Find Prior Art

 Similar

Inventor: Daniel Hendricus Franciscus DIJKMAN, Regan Blythe TOWAL, Venkata Sreekanta Reddy ANNAPUREDDY

Current Assignee : Qualcomm Inc

위의 inventor, current Assignee 만 추출하여 만든 데이터 프레임 형태는 아래와 같다.

	Inventor	assignee
0	Daniel Hendricus Franciscus DIJKMAN/Regan Blyt...	Qualcomm Inc

- Inventor 와 assignee 의 정보를 추가한 데이터 프레임 생성

▼ search_advanced 함수 사용예제

result2 : 적용국가가 추가된 데이터 프레임
result : search 함수를 통해 생성된 데이터 프레임

```
[6] result2 = search_advanced(result)
```

```
[7] display(result2)
```

	특허제목	링크	특허번호	국가코드	적용국가	Inventor	assignee
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1	WO2017079522A1	WO	US(2016) / WO(2016) / DE(2016) / JP(2016)	Wongun CHOI / Yuanqing Lin / Yu XIANG / Silvio...	
1	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US	US(2015) / JP(2016) / EP(2016)	José A. Rodríguez-Serrano / Albert Gordo Solde...	Xerox Corp
2	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US	US(2016)	Jian Sun / Ross Girshick / Shaoqing Ren / Kaim...	Microsoft Technology Licensing LLC
3	Cross-trained convolutional neural networks us...	www.google.com/patents/US20170032222	US20170032222	US	US(2015)	Ajrun Sharma / Pramod Sankar Kompalli	Xerox Corp
4	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1	WO2017139927A1	WO	US(2016) / CN(2016) / WO(2016)	Anbang YAO / Tao Kong / Yurong Chen	
5	Fully convolutional pyramid networks for pedes...	www.google.com/patents/WO2017101036A1	WO2017101036A1	WO	EP(2015) / WO(2015) / US(2015)	Anbang YAO / Ruoyan WANG / Yurong Chen	
6	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315	US20170169315	US	US(2016)	Gonzalo Vaca Castano / Syed Zain Masood / Step...	Sighthound Inc
7	Multi-task deep convolutional neural networks ...	www.google.com/patents/US9286524	US9286524	US	US(2015)	Xue Mei / Danil V. Prokhorov / Jun Li / Dachen...	University of Technology Sydney/ Toyota Motor...
8	Lane boundary detection using images	www.google.com/patents/US9081385	US9081385	US	US(2012)	David Ian Ferguson / Turgay Senlet	Google LLC/ Waymo LLC
9	一种优化人工神经网络的方法 A method of optimizing artificia...	www.google.com/patents/CN107239829A	CN107239829A	CN	US(2016) / CN(2016) / CN(2017)	魏强	Xilinx Technology Beijing Ltd

→ assignee 열의 공백은 해당 특허 assignee 가 없음을 의미한다.

1 월 3 일 업무 진행 상황

- 링크를 추출할 때 검색창에서 특정 언어의 결과로 출력되는 오류 수정

0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1? cl=fr
1	Object detection and classification in images	www.google.com/patents/US20170206431
2	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1? cl=fr
3	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792
4	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282
5	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315
6	Face hallucination using convolutional neural ...	www.google.com/patents/US9405960
7	Hierarchical deep convolutional neural network	www.google.com/patents/WO2016069581A1? cl=en

해당 사진처럼 언어에 관한 검색어 설정이 ?cl=이후에 나타난다

	특허제목	링크	특허번호	국가코드
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1	WO2017079522A1	WO
1	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US
2	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1	WO2017139927A1	WO
3	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US
4	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282	US9633282	US
5	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315	US20170169315	US
6	Face hallucination using convolutional neural ...	www.google.com/patents/US9405960	US9405960	US
7	Hierarchical deep convolutional neural network	www.google.com/patents/WO2016069581A1	WO2016069581A1	WO
8	Hierarchical deep convolutional neural network...	www.google.com/patents/US20160117587	US20160117587	US
9	Context-based priors for object detection in i...	www.google.com/patents/US20170011281	US20170011281	US

오류를 해결한 이후의 결과는 위와 같다

✓ 1초 [60] citation = paper_citation(result)

특허번호 입력 : US9286524
<https://www.google.com/patents/US9286524>
 Name: 링크, dtype: object
 <Response [404]>

✓ 0초 ▶ display(citation)

Publication number	Priority date	Publication date	Assignee	Title
--------------------	---------------	------------------	----------	-------

→ 특허번호 입력에 따라 링크를 새로 생성하여 Url 404 error 문제 해결하였음.

• 인용정보 데이터프레임 생성

▾ paper_citation 사용예제

citation : 입력한 논문을 인용한 논문들의 데이터프레임

✓ 6분 [154] citation = paper_citation(result)

특허번호 입력 : #02017139927A1

✓ 0초 [152] display(citation)

	Publication number	Priority date	Publication date	Assignee	Title
0	CN104794527B	2014-01-20	2018-03-27	富士通株式会社	Disaggregated model construction method and eq...
1	CN104036323B	2014-06-26	2016-11-09	叶茂	A kind of vehicle checking method based on con...
2	CN104517103A	2014-12-26	2015-04-15	广州中国科学院先进技术研究所	Traffic sign classification method based on de...
3	CN104573731B	2015-02-06	2018-03-23	厦门大学	Fast target detection method based on convolut...
4	CN104809443B	2015-05-05	2018-12-28	上海交通大学	Detection method of license plate and system b...
5	CN104866868B	2015-05-22	2018-09-07	杭州朗和科技有限公司	Metal coins recognition methods based on deep ...
6	CN105138973B	2015-08-11	2018-11-09	北京天诚盛业科技有限公司	The method and apparatus of face authentication
7	CN105120130B	2015-09-17	2018-06-29	京东方科技集团股份有限公司	A kind of image raising frequency system, its ...

Patent Citations (13)

90%

Publication number	Priority date	Publication date	Assignee	Title
US20080304740A1 *	2007-06-06	2008-12-11	Microsoft Corporation	Salient Object Detection
US20110182469A1 *	2010-01-28	2011-07-28	Nec Laboratories America, Inc.	3d convolutional neural networks for automatic human action recognition
CN104217225A *	2014-09-02	2014-12-17	中国科学院自动化研究所	A visual target detection and labeling method
US20150078654A1 *	2013-09-13	2015-03-19	Interra Systems, Inc.	Visual Descriptors Based Video Quality Assessment Using Outlier Model
CN104572804A *	2013-10-24	2015-04-29	Tcl集团股份有限公司	Video object retrieval system and method
Family To Family Citations				
CN104794527B *	2014-01-20	2018-03-27	富士通株式会社	Disaggregated model construction method and equipment based on convolutional neural networks
CN104036323B *	2014-06-26	2016-11-09	叶茂	A kind of vehicle checking method based on convolutional neural networks
CN104517103A *	2014-12-26	2015-04-15	广州中国科学院先进技术研究所	Traffic sign classification method based on deep neural network
CN104573731B *	2015-02-06	2018-03-23	厦门大学	Fast target detection method based on convolutional neural networks
CN104809443B *	2015-05-05	2018-12-28	上海交通大学	Detection method of license plate and system based on convolutional neural networks
CN104866868B *	2015-05-22	2018-09-07	杭州朗和科技有限公司	Metal coins recognition methods based on deep neural network and device
CN105138973B *	2015-08-11	2018-11-09	北京天诚盛业科技有限公司	The method and apparatus of face authentication
CN105120130B *	2015-09-17	2018-06-29	京东方科技集团股份有限公司	A kind of image raising frequency system, its training method and image raising frequency method

해당 정보는 특허 링크에 접속해 paper citation 중 family to family citation 부분이다.

어느 데이터들이 필요한지 모르기에 우선 이 정도로 마무리하였다.

12 월 29 일 업무 진행 상황

- 적용국가 추출하는 데이터프레임 및 함수 생성

search_advanced 함수 사용예제

result2 : 적용국가가 추가된 데이터 프레임

result : search 함수를 통해 생성된 데이터 프레임

```
[6] result2 = search_advanced(result)
```

```
[7] display(result2)
```

	특허제목	링크	특허번호	국가코드	적용국가	Inventor	assignee
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1	WO2017079522A1	WO	US(2016) / WO(2016) / DE(2016) / JP(2016)	Wongun CHOI / Yuanqing Lin / Yu XIANG / Silvio...	
1	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US	US(2015) / JP(2016) / EP(2016)	José A. Rodríguez-Serrano / Albert Gordo Solde...	Xerox Corp
2	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US	US(2016)	Jian Sun / Ross Girshick / Shaoqing Ren / Kaim...	Microsoft Technology Licensing LLC
3	Cross-trained convolutional neural networks us...	www.google.com/patents/US20170032222	US20170032222	US	US(2015)	Ajrun Shama / Pramod Sankar Kompalli	Xerox Corp
4	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1	WO2017139927A1	WO	US(2016) / CN(2016) / WO(2016)	Anbang YAO / Tao Kong / Yurong Chen	
5	Fully convolutional pyramid networks for pedes...	www.google.com/patents/WO2017101036A1	WO2017101036A1	WO	EP(2015) / WO(2015) / US(2015)	Anbang YAO / Ruoyan WANG / Yurong Chen	
6	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315	US20170169315	US	US(2016)	Gonzalo Vaca Castano / Syed Zain Masood / Step...	Sighthound Inc
7	Multi-task deep convolutional neural networks ...	www.google.com/patents/US9286524	US9286524	US	US(2015)	Xue Mei / Danil V. Prokhorov / Jun Li / Dachen...	University of Technology Sydney/ Toyota Motor...
8	Lane boundary detection using images	www.google.com/patents/US9081385	US9081385	US	US(2012)	David Ian Ferguson / Turgay Senlet	Google LLC/ Waymo LLC
9	一种优化人工神经网络的方法 A method of optimizing artificia...	www.google.com/patents/CN107239829A	CN107239829A	CN	US(2016) / CN(2016) / CN(2016) / CN(2017)	路强	Xilinx Technology Beijing Ltd

- 데이터 프레임을 입력으로 특허번호를 검색했을 때, 인용문을 출력하는 기능 추가

0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1?cl=fr
1	Object detection and classification in images	www.google.com/patents/US20170206431
2	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1?cl=fr
3	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792
4	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282
5	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315
6	Face hallucination using convolutional neural ...	www.google.com/patents/US9405960
7	Hierarchical deep convolutional neural network	www.google.com/patents/WO2016069581A1?cl=en

문제상황 : 위의 기능을 수행하는 함수를 만들기위해 데이터프레임에서 가져온 링크로 함수를 실행시킬 경우에 링크에 실제로는 접속가능하나 urlopen 했을 경우 404 에러가 존재한다고 나타남.

	특허제목	링크	특허번호	국가코드
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1	WO2017079522A1	WO
1	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US
2	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1	WO2017139927A1	WO
3	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US
4	Cross-trained convolutional neural networks us...	www.google.com/patents/US9633282	US9633282	US
5	Deeply learned convolutional neural networks (...)	www.google.com/patents/US20170169315	US20170169315	US
6	Face hallucination using convolutional neural ...	www.google.com/patents/US9405960	US9405960	US
7	Hierarchical deep convolutional neural network	www.google.com/patents/WO2016069581A1	WO2016069581A1	WO
8	Hierarchical deep convolutional neural network...	www.google.com/patents/US20160117587	US20160117587	US
9	Context-based priors for object detection in i...	www.google.com/patents/US20170011281	US20170011281	US

아래의 링크일 경우에는 기능을 수행할 수 있음.

```
✓ [60] citation = paper_citation(result)
1초
특허번호 입력 : US9286524
https://www.google.com/patents/US9286524
Name: 링크, dtype: object
<Response [404]>

✓ ▶ display(citation)
0초
```

Publication number	Priority date	Publication date	Assignee	Title
--------------------	---------------	------------------	----------	-------

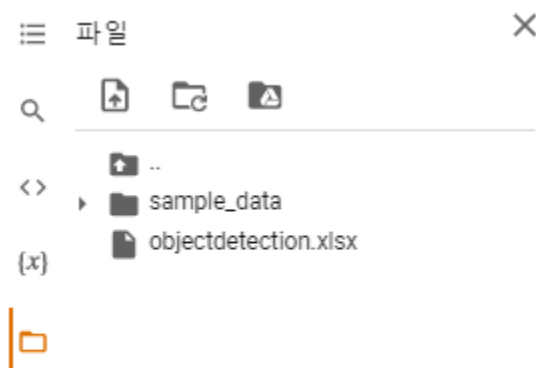
그렇기 때문에 도출해낼 수 있는 해결방안은 다음과 같으며, 이 중에서 편한 방법으로 해결할 예정

해결방안 1 : 현재 키워드의 정보는 search 함수에만 존재하기 때문에 키워드의 정보를 가져와서 링크를 생성

해결방안 2 : 인용 데이터를 간단한 링크에서 추출

12 월 28 일 업무 진행 상황

- 키워드 입력에 따른 파일명 지정



입력 한 키워드가 그대로 파일명으로 저장되도록 하였음. 파일명이 objectdetection.xlsx로 저장되었음.

, 키워드 입력(끝내고싶으면 end 입력) :object
 키워드 입력(끝내고싶으면 end 입력) :detection
 키워드 입력(끝내고싶으면 end 입력) :end
 objectdetection.xlsx

	특허제목	링크	특허번호	국가코드	적용국가
0	Object detection apparatus and object detectio...	www.google.com/patents/WO2017094891A1?cl=en	WO2017094891A1?cl=en	WO	NaN
1	Method and device for object detection in the ...	www.google.com/patents/US7791527	US7791527	US	NaN
2	Object detection Device	www.google.com/patents/US20100026491	US20100026491	US	NaN
3	Object detection system	www.google.com/patents/US20020014955	US20020014955	US	NaN
4	Object detection device, object detection meth...	www.google.com/patents/WO2015041081A1?cl=en	WO2015041081A1?cl=en	WO	NaN
...
195	Fast object detection for augmented reality sy...	www.google.com/patents/US20060233423	US20060233423	US	NaN
196	Foreign object detection and removal system fo...	www.google.com/patents/US20100048269	US20100048269	US	NaN
197	Context aware moving object detection	www.google.com/patents/US20130176430	US20130176430	US	NaN
198	Foreign object detection apparatus	www.google.com/patents/US7567088	US7567088	US	NaN
199	Nonlinear system identification for object det...	www.google.com/patents/US20140167704	US20140167704	US	NaN

200 rows × 5 columns

• 특허 인용 관련 정보 추출 기능

US6954744B2
(en)

2001-08-29
2005-10-11
Honeywell International, Inc.
Combinatorial approach for supervised neural network learning

AU2003301143A1
(en)

*

2002-12-19
2004-07-14
Drexel University
Apparatus and method for removing non-discriminatory indices of an indexed dataset

US20070094214A1
(en)

2005-07-15
2007-04-26
Li Eric Q
Parallelization of bayesian network structure learning

Patent Citations 에 대한 Publication number, Priority date, Publication date, Assignee, Title 의 정보를 추출하는데 성공함.

하지만 각각의 정보를 따로 추출하여 데이터 프레임 형태로 저장하기 위해선 추가적인 html 구조 분석이 필요해보임.

- 최대검색갯수 지정받아 데이터 저장

```
[92] result = search()

검색갯수 입력(숫자) : 10
키워드 입력(끝내고싶으면 end 입력) :stradvision
키워드 입력(끝내고싶으면 end 입력) :end
```

```
[93] # result 데이터 프레임 시각화
display(result)
# result 데이터 프레임 엑셀로 변환
result.to_excel('Search Results.xlsx')
```

	특허제목	링크	특허번호	국가코드
0	Réseaux neuronaux convolutionnels sensibles à ...	www.google.com/patents/WO2017079522A1?cl=fr	WO2017079522A1?cl=fr	WO
1	Pedestrian notifications	www.google.com/patents/US8954252	US8954252	US
2	Method for increasing object detection rate or...	www.google.com/patents/US9230193	US9230193	US
3	Method for detecting texts included in an imag...	www.google.com/patents/US9524430	US9524430	US
4	Proposition de région pour des régions d'image...	www.google.com/patents/WO2017139927A1?cl=fr	WO2017139927A1?cl=fr	WO
5	Three-dimensional (3d) convolution with 3d bat...	www.google.com/patents/WO2017031088A1?cl=en	WO2017031088A1?cl=en	WO
6	Deepstereo: learning to predict new views from...	www.google.com/patents/WO2016183464A1?cl=en	WO2016183464A1?cl=en	WO
7	Driving support method and apparatus	www.google.com/patents/EP1878618A2?cl=en	EP1878618A2?cl=en	EP
8	Distance measurement device and environment ma...	www.google.com/patents/CN103154666A?cl=en	CN103154666A?cl=en	CN
9	Object Detection Using Cascaded Convolutional ...	www.google.com/patents/US20160307074	US20160307074	US

• 적용국가 추출기능

```
[87] df_app = []
for link in result['링크']:
    df_app.append(application('http://'+link))
    # print(application('http://'+link))
print(df_app)

[['US(2016) / JP(2016) / WO(2016) / DE(2016)'], ['US(2015) / JP(2016) / EP(2016)'], ['US(2016)'], ['US(2015)'], ['US(2016)']]
```

현재 2 차원 리스트이기 때문에 원래 데이터프레임에 concat 가 되지않는 문제가 있음.

2 차원 리스트 안에 있던 요소들을 join 함수를 이용하여 하나의 요소로 합쳤으나, 2 차원 리스트로 남아있는 문제가 존재함.

파이썬 리스트에 대한 이해가 더 있어야 해결가능할 것으로 보임.

12 월 27 일 업무 진행 상황

- 인턴쉽 진행자료 문서화

Stradvision winter internship report

링크 : <https://colab.research.google.com/drive/1gxrKS6glVcW-cA5XGzCJ0snvU3DVjLH4?usp=sharing>

- 분석 코드들 함수화

함수 정의

- 페이지 추출함수 (주소를 입력해 해당 페이지의 제목, 링크, 특허번호, 국가코드를 추출) : pageExtraction
- 적용국가 추출함수 (주소를 입력해 해당 주소에 접속, 적용국가 추출) : application
- 키워드 검색해 데이터 프레임을 반환하는 함수 : search

(함수 사용예제)

search 함수 사용예제

```
[9] result = search()
```

```
키워드 입력(끝내고싶으면 end 입력) :stradvision  
키워드 입력(끝내고싶으면 end 입력) :deep learning  
키워드 입력(끝내고싶으면 end 입력) :CNN  
키워드 입력(끝내고싶으면 end 입력) :end
```

```
# result 데이터 프레임 시각화  
display(result)  
# result 데이터 프레임 엑셀로 변환  
result.to_excel('Search Results.xlsx')
```

	특허제목	링크	특허번호	국가코드
0	Method and system for vision-centric deep-lear...	www.google.com/patents/US9760806	US9760806	US
1	Regionlets with shift invariant neural pattern...	www.google.com/patents/US9202144	US9202144	US
2	Similarity-based detection of prominent object...	www.google.com/patents/US20170083792	US20170083792	US
3	Hierarchical deep convolutional neural network...	www.google.com/patents/US20160117587	US20160117587	US
4	Object detection and classification in images	www.google.com/patents/US20170206431	US20170206431	US
...
84	显著性信息取得装置以及显著性信息取得方法	www.google.com/patents/CN106296638A?cl=zh	CN106296638A?cl=zh	CN
85	Gemeinsames Tiefenschätzen und semantisches Be...	www.google.com/patents/DE102016005407A1?cl=de	DE102016005407A1?cl=de	DE
86	两层文本识别系统和方法	www.google.com/patents/CN101496036B?cl=zh	CN101496036B?cl=zh	CN
87	脸部识别方法和设备	www.google.com/patents/CN106295496A?cl=zh	CN106295496A?cl=zh	CN
88	神经网络训练与构建方法和装置以及目标检测方法和装置	www.google.com/patents/CN106295678A?cl=zh	CN106295678A?cl=zh	CN

89 rows × 4 columns

추가되어야할 내용들

1. 데이터프레임을 입력으로 적용국가를 추출하는 함수
2. 파이썬을 처음 사용하는 사람을 위해 엑셀변환 간소화
3. 엑셀 파일에 키워드 정보를 포함하는 방안 고안
4. 저장할 키워드 개수를 입력받아서 검색실행하도록 고안

12 월 24 일 업무 진행 상황

- 특허 적용국가 및 연도 추출
- 추출한 정보들을 엑셀에 데이터 프레임 형태 저장
- google patent 에 검색어 키워드를 직접 입력하는 대신에 프로그램에 입력해서 검색해주는 기능추가
- 키워드들을 입력했을 때 검색 갯수를 보여주고, 사용자가 타당하다 판단되면 추출하도록 하려 했으나 실제 구글에 검색되는 것은 cache memory 로 할당하였기에 완전히 파악 불가능 하였음
- 여러 페이지에 대해 데이터 추출 가능하도록 함

키워드 입력(끝내고싶으면 end 입력) :stradvision
 키워드 입력(끝내고싶으면 end 입력) :deep learning
 키워드 입력(끝내고싶으면 end 입력) :end

	특허제목	링크	특허번호	국가코드	적용국가
0	Method and system for vision-centric deep-lear...	www.google.com/patents/US9760806	US9760806	US	NaN
1	Object detection using deep neural networks	www.google.com/patents/US9275308	US9275308	US	NaN
2	Image processing apparatus and method based on...	www.google.com/patents/US20170169313	US20170169313	US	NaN
3	Utilizing deep learning for boundary-aware ima...	www.google.com/patents/US20170287137	US20170287137	US	NaN
4	Deepstereo: learning to predict new views from...	www.google.com/patents/WO2016183464A1?cl=en	WO2016183464A1?cl=en	WO	NaN
...
134	Fussgängerdetektion und Bewegungsvorhersage mi...	www.google.com/patents/DE102017105903A1?cl=de	DE102017105903A1?cl=de	DE	NaN
135	Gemeinsames Tiefenschätzen und semantisches Be...	www.google.com/patents/DE102016005407A1?cl=de	DE102016005407A1?cl=de	DE	NaN
136	显著性信息取得装置以及显著性信息取得方法	www.google.com/patents/CN106296638A?cl=zh	CN106296638A?cl=zh	CN	NaN
137	脸部识别方法和设备	www.google.com/patents/CN106295496A?cl=zh	CN106295496A?cl=zh	CN	NaN
138	神经网络训练与构建方法和装置以及目标检测方法和装置	www.google.com/patents/CN106295678A?cl=zh	CN106295678A?cl=zh	CN	NaN

139 rows × 5 columns

12 월 23 일 업무 진행 상황

- google patent 에 검색어 키워드를 직접 입력하는 대신에 프로그램에 입력해서 검색해주는 기능추가

```
#-----특허 키워드 입력-----//
from urllib.parse import quote_plus
from bs4 import BeautifulSoup
plusUrl = input('특허 검색 창 :')
baseUr1 = 'https://www.google.com/search?q=sth&biw=910&bih=646&tbm=pts&ei='
nextUr1 = baseUr1.replace('sth',quote_plus(plusUr1))
ur1 = nextUr1
print(ur1)
```

특허 검색 창 : "stradvision" "deep learning"
<https://www.google.com/search?q=%22stradvision%22+%22deep+learning%22&biw=910&bih=646&tbm=pts&ei=>

- 데이터 프레임 생성

	논문제목	링크	특허번호	국가코드	적용국가
0	Engaging and disengaging for autonomous driving	www.google.com/patents/US20140156134	US20140156134	US	NaN
1	Engaging and disengaging for autonomous driving	www.google.com/patents/US9352752	US9352752	US	NaN
2	Remote operation of autonomous vehicle in unex...	www.google.com/patents/US20170045885	US20170045885	US	NaN
3	启用和停用自动驾驶 Enable and disable the autopilot	www.google.com/patents/CN104837705B?cl=en	CN104837705B?cl=en	CN	NaN
4	Engaging and disengaging for autonomous driving	www.google.com/patents/US9511779	US9511779	US	NaN
5	Engaging and disengaging for autonomous driving	www.google.com/patents/WO2014085380A1?cl=en	WO2014085380A1?cl=en	WO	NaN
6	Engaging and disengaging for autonomous driving	www.google.com/patents/US8825258	US8825258	US	NaN
7	Remote operation of autonomous vehicle in unex...	www.google.com/patents/US9494935	US9494935	US	NaN
8	Système de navigation et de commande pour des ...	www.google.com/patents/WO2008048707A2?cl=fr	WO2008048707A2?cl=fr	WO	NaN
9	Method for Interacting with an Object Displaye...	www.google.com/patents/US20160041624	US20160041624	US	NaN

• 데이터 프레임 엑셀로의 변환

	A	B	C	D	E	F
1		논문제목	링크	특허번호	국가코드	적용국가
2	0	Engaging and disengaging for autonomous driving	www.google.com/pater	US20140156134	US	
3	1	Engaging and disengaging for autonomous driving	www.google.com/pater	US9352752	US	
4	2	Remote operation of autonomous vehicle in unexpected	www.google.com/pater	US20170045885	US	
5	3	启用和停用自动驾驶 Enable and disable the autopilot	www.google.com/pater	CN104837705B?cl=en	CN	
6	4	Engaging and disengaging for autonomous driving	www.google.com/pater	US9511779	US	
7	5	Engaging and disengaging for autonomous driving	www.google.com/pater	WO2014085380A1?cl=	WO	
8	6	Engaging and disengaging for autonomous driving	www.google.com/pater	US8825258	US	
9	7	Remote operation of autonomous vehicle in unexpected	www.google.com/pater	US9494935	US	
10	8	Système de navigation et de commande pour des véhic	www.google.com/pater	WO2008048707A2?cl=	WO	
11	9	Method for Interacting with an Object Displayed on Dat	www.google.com/pater	US20160041624	US	
12						
13						
14						

추가 되어야할 업무들

- 적용국가 추출
- 1 페이지 이상의 자료가 나타날 경우 정보 추출

12 월 22 일 업무 진행 상황

- list 에서 1 페이지에 해`당하는 논문 제목 추출 (이 중에서 제목만을 찾아서 슬라이싱해야함) 이 때 class 의 이름에 해당하는 zBAuLc 로 시작하는 부분은

PC 마다 다른 것을 확인하였음) 때문에 모든 PC 에서 적용가능한 방안을 고려해야함

논문 제목 추출

```
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Engaging and disengaging for autonomous driving </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Engaging and disengaging for autonomous driving </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Remote operation of autonomous vehicle in unexpected environment </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Engaging and disengaging for autonomous driving </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Engaging and disengaging for autonomous driving </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Remote operation of autonomous vehicle in unexpected environment </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Engaging and disengaging for autonomous driving </div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Multi-task deep convolutional neural networks for efficient and ...</div></h3>
<h3 class="zBAuLc 197dzf"><div class="BNeawe vjwJb AP7Wnd">Method for Interacting with an Object Displayed on Data Eyeglasses </div></h3>
```

- 논문링크 추출(어제 상황에서 필요한 부분만을 추출하는데 성공함. 이는 그대로 data frame 에 저장해도 될 듯함)

논문 링크 추출

```
https://www.google.com/patents/related/US20140156134&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw077ehm4YGVt_5fJ-ogL19p
https://www.google.com/patents/related/US9352752&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw1-MIjYMIJJ3uT1dVq_HKCe
https://www.google.com/patents/related/US20170045885&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2US0g04Pc2_6xqvkyYUMGx
https://www.google.com/patents/related/US9511779&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw1bbDDEg8630WBS4cw9f_Xn
https://www.google.com/patents/related/US2014085380A1&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2LTrJC5ZgDBmTD_e93nME9
https://www.google.com/patents/related/CN104837705B&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3ewwF0n2Z10FaWIPy1RxrU
https://www.google.com/patents/related/US9494935&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3vXq1VJh0rJgBE8eMHVIRd
https://www.google.com/patents/related/US8825258&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3r1u81_2Zxod90s9UUnO0X
https://www.google.com/patents/related/US9286524&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3t6jpuT_AmtA79vVx9wvA9
https://www.google.com/patents/related/US20160041624&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2GDPHP9Ez1GqH3trEq2B7B
```

- 특허번호 추출(특허번호의 자리수가 다양함. slicing 범위를 알아야함 find 함수로 &sa=U&ved= 의 위치를 찾아내서 slicing 하면될듯함)

특허 번호 추출

```
US20140156134&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw077ehm4YGVt_5fJ-ogL19p
US9352752&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw1-MIjYMIJJ3uT1dVq_HKCe
US20170045885&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2US0g04Pc2_6xqvkyYUMGx
US9511779&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw1bbDDEg8630WBS4cw9f_Xn
US2014085380A1&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2LTrJC5ZgDBmTD_e93nME9
CN104837705B&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3ewwF0n2Z10FaWIPy1RxrU
US9494935&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3vXq1VJh0rJgBE8eMHVIRd
US8825258&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3r1u81_2Zxod90s9UUnO0X
US9286524&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw3t6jpuT_AmtA79vVx9wvA9
US20160041624&sa=U&ved=2ahUKEwjQts7-1_b0AhXUxIsBHxOHD4MQ0gJ6BAgFEAU&usq=A0vYaw2GDPHP9Ez1GqH3trEq2B7B
```

- 국가코드 추출(특허번호의 앞자리 2 자리 or 3 자리가 국가코드임. 특허번호 추출을 성공하면 이는 마찬가지로 될듯함)
- 적용국가 추출(적용국가는 검색했을 때에 논문 링크에 접속해야 나타나는 정보임. 때문에 또 정보를 가져올 방안을 고민해보기)
- 검색 갯수가 10 개를 넘어갈 경우도 고려해야 함. (위의 일련의 과정을 하나의 함수로 선언하여서, 검색갯수까지 실행되도록 코딩)

- 추출한 정보들을 엑셀에 저장하기 위한 데이터 프레임 형태로 만들기 (지금까지 만든 데이터들 저장해봄. 크기만 늘리면 계속 가능할 듯)

	논문제목	링크	특허번호
0	[[Engaging and disengaging for autonomous driv...	https://www.google.com/patents/related/US20140...	NaN
1	[[Engaging and disengaging for autonomous driv...	https://www.google.com/patents/related/US93527...	NaN
2	[[Remote operation of autonomous vehicle in un...	https://www.google.com/patents/related/US20170...	NaN
3	[[Engaging and disengaging for autonomous driv...	https://www.google.com/patents/related/US95117...	NaN
4	[[Engaging and disengaging for autonomous driv...	https://www.google.com/patents/related/WO20140...	NaN
5	[[启用和停用自动驾驶 Enable and disable the autopilot]]	https://www.google.com/patents/related/CN10483...	NaN
6	[[Remote operation of autonomous vehicle in un...	https://www.google.com/patents/related/US94949...	NaN
7	[[Engaging and disengaging for autonomous driv...	https://www.google.com/patents/related/US88252...	NaN
8	[[Multi-task deep convolutional neural network...	https://www.google.com/patents/related/US92865...	NaN
9	[[Method for Interacting with an Object Displa...	https://www.google.com/patents/related/US20160...	NaN

12 월 21 일 업무 진행 상황

- beuatifulSoup 을 이용해 google patent 에 검색했을 때 나타나는 데이터들을 불러오기

```

soup = BeautifulSoup(rebpage.content, "html.parser")
print(soup.prettify())

#논문 제목 추출
paper_name = soup.find_all('h3')
print(paper_name)

#논문 링크 추출
links = soup.find_all('a')
for link in links:
    if link['href'].startswith('/url?q='):
        print((link['href']).replace('/url?q=', ''))

```

```

</li>
<li class="NFsl">
<a href="/search?q=x22neural+networkx22+x22autonomousdrivingx22+x22stnradvisonx22&amp;bi=725&amp;bih=609&amp;ie=UTF-8&amp;tbs=pt:s&amp;source=Int&amp;tbs=pt:so:cn&amp;sa=x&amp;ved=0ahJKEwJbuZ25uT0AHUJEyW4UkrCIAQwUjFA">
中國
</a>
</li>
<li class="NFsl">
<a href="/search?q=x22neural+networkx22+x22autonomousdrivingx22+x22stnradvisonx22&amp;bi=725&amp;bih=609&amp;ie=UTF-8&amp;tbs=pt:s&amp;source=Int&amp;tbs=pt:so:de&amp;sa=x&amp;ved=0ahJKEwJbuZ25uT0AHUJEyW4UkrCIAQwUjFA">
德國
</a>
</li>
<li class="NFsl">
<a href="/search?q=x22neural+networkx22+x22autonomousdrivingx22+x22stnradvisonx22&amp;bi=725&amp;bih=609&amp;ie=UTF-8&amp;tbs=pt:s&amp;source=Int&amp;tbs=pt:so:ca&amp;sa=x&amp;ved=0ahJKEwJbuZ25uT0AHUJEyW4UkrCIAQwUjFA">
加拿大
</a>
</li>
</ul>
</div>
<div class="Pa9J5">
<div class="Rfa0fd" role="button" tabindex="0">
<div class="TfR0Uc">
所有專利狀態
</div>
</div>
<span class="OnIzt">
</span>
</div>
</div>

```

불러온 데이터들을 바탕으로 필요한 정보들 추출

- 논문 제목 추출

- 논문 링크 추출

```

#논문 제목 추출
paper_name = soup.find_all('h3')
print(paper_name)

#논문 링크 추출
links = soup.find_all('a')
for link in links:
    if link['href'].startswith('/url?q='):
        print(link['href'].replace('/url?q=', ''))

```

[illegible]

- 특허번호 추출
- 국가코드 추출
- 적용국가 추출
- 추출한 정보들을 엑셀에 저장하기 위한 데이터 프레임 형태로 만들기
- 만든 데이터 프레임을 pandas 를 이용하여 csv 파일로 변환

그 외의 적용하면 유용할 기능들 고안해보기

- google patent 에 검색어 키워드를 직접 입력하는 대신에 프로그램에 입력해서 검색해주는 기능
- 적용국가를 추출하기 위해서는 논문 링크의 데이터를 불러와서 해당 국가 코드 데이터를 얻어야 함
- 키워드들을 입력했을 때 검색 갯수를 보여주고, 사용자가 타당하다 판단되면 추출하는 기능
- pc 마다 검색할 때 html 에 나타나는 class 이름이 달랐기에 일반화 가능한 법 고안하기

만든 데이터 프레임을 pandas 를 이용하여 csv 파일로 변환

Team members

- Song Yeong Won - [Confluence \(atlassian.net\)](#)
 - Lee JunGi - [Confluence \(atlassian.net\)](#)
-

