

INDEKS FILE (TEKS) MENGGUNAKAN SWISH-E

By: Maulana Hizbullah (1708107010055) – Penuluruhan Informasi – 2020
Informatika – Universitas Syiah Kuala

SWISH-e adalah sebuah *tool* yang digunakan untuk meng-indeks teks dalam berbagai format, seperti PDF, html, txt, XML, PostScript, dll. Selain meng-indeks, *tool* ini juga dapat mencari dokumen berdasarkan *query* yang di-*input*-kan. *Tool* ini memiliki banyak fitur/kemampuan yang dimilikinya atau untuk lebih jelasnya dapat dilihat pada laman “<https://www.esa.org/tiee/search/html/readme.html>”

Penulis membuat skrip python untuk mengunduh (*crawling*) berita pada website:

- AntaraNews
- KompasNews
- VivaNews

Berita yang terkumpul Dapat dilihat pada *folder* ./file. Secara keseluruhan berita yang diunduh sebanyak 10K. Berita-berita ini nantinya dijadikan bahan *text indexing*. Berikut cuplikan skrip *crawling* berita:

a. AntaraNews

```
[-] ▶ + Python 3.8.10
import re
import mechanize
import os
import sys
import urllib.request

[-] ▶ + Python 3.8.10
br = mechanize.Browser()
linklist=[]
kamus={}
values = 1
count = 0

day = 1
dayend = 31
month = 2
year = 2019
tahun = year
bulan = month
tanggal = day
halistart = 1
halend = 15

[-] ▶ + Python 3.8.10
print("Start to execute...")
for tanggal in range(tanggal, dayend):
    print("Crawling data on:", tahun, "/", bulan, "/", tanggal, "[Page:", halistart, "-", halend-1])
    i=1
    for i in range(halistart, halend):
        #print("Include page", i)
        if tanggal < 10:
            r = br.open("https://www.antarane.ws.com/index/* str(tahun)+ '-' + str(bulan)+ '-' + str(tanggal) + '/' + str(i))
        else:
            r = br.open("https://www.antarane.ws.com/index/* str(tahun)+ '-' + str(bulan)+ '-' + str(tanggal) + '/' + str(i))
        if bulan < 10:
            r = br.open("https://www.antarane.ws.com/index/* str(tahun)+ '-' + str(bulan)+ '-' + str(tanggal) + '/' + str(i))
        else:
            r = br.open("https://www.antarane.ws.com/index/* str(tahun)+ '-' + str(bulan)+ '-' + str(tanggal) + '/' + str(i))
        for link in br.links():
            linklist.append(link.url)
        for links in linklist:
            if links not in kamus and re.search("/berita/",links):
                kamus[links]=values
                count+=1
                #print(kamus[links])
            else:
                pass
    print("Executed successfully")
```

```

[-] ▶ == MI
print(komus)

[-] ▶ == MI
f = open("dataAntara.txt","w")
for x in komus:
    f.write(x + "\n")
f.close()

[-] ▶ == MI
f = open("dataAntara.txt","r")
count = 1

for link in f:
    print(link)
    urllib.request.urlretrieve(link, "Antarajenis-" + str(count) + ".html")
    count+=1
    print("Downloading file: ", count)
print("Finish : ", count)
f.close()

[-] ▶ == MI

```

b. KompasNews

```

# CrawlerKompas.py
import re
import mechanize
import os
import sys
import urllib.request

br = mechanize.Browser()
linklist=[]
kamus={}
values = 1
count = 0

day = 1
dayend = 31
month = 4
year = 2019
tahun = year
bulan = month
tanggal = day
halstart = 1
halend = 30

print("Start to execute...")
for tanggal in range(tanggal, dayend):
    print("Crawling data on:", tahun,"/", bulan,"/", tanggal, "[Page", halstart,"-", halend-1)
    i=1
    for i in range(halstart, halend):
        #print("Include page", i)
        if tanggal < 10:
            r = br.open("https://news.kompas.com/search/"+ str(tahun)+ "-" + str(bulan)+ "-" + str(tanggal) + "/" + str(i))
        else:
            r = br.open("https://news.kompas.com/search/"+ str(tahun)+ "-" + str(bulan)+ "-" + str(tanggal) + "/" + str(i))
        if bulan < 10:
            r = br.open("https://news.kompas.com/search/"+ str(tahun)+ "-" + str(bulan)+ "-" + str(tanggal) + "/" + str(i))
        else:
            r = br.open("https://news.kompas.com/search/"+ str(tahun)+ "-" + str(bulan)+ "-" + str(tanggal) + "/" + str(i))
        for link in br.links():
            linklist.append(link.uri)
        for links in linklist:
            if links not in kamus and re.search("read/",links):
                kamus[links]=values
                count+=1
                #print(kamus[links])
            else:
                pass
    print("Executed successfully")

```

```
CrawlerKompas.pyntb x
>> print(kampus)

[-] >+= Mi
print(kampus)

[-] >+= Mi
f = open("dataKompas.txt", "w")
for x in kampus:
    f.write(x + "\n")
f.close()

[-] >+= Mi
f = open("dataKompas.txt", "r")
count = 3945

for link in f:
    print(link)
    urllib.request.urlretrieve(link, "Kompas-" + str(count) + ".html")
    count+=1
    print("\nDownloading file: ", count)
print("Finish : ", count)
f.close()
```

c. VivaNews

```
F CrawlerVivaNews.pynt X
import re
import mechanize
import os
import sys
import urllib.request

br = mechanize.Browser()
linklist=[]
kamus={}
values = 1
count = 0

day = 1
dayend = 31
month = 6
monthend = 8
year = 2019
tahun = year
bulan = month
tanggal = day
halstart = 1
halend = 4

print("Start to execute...")
for tanggal in range(tanggal, dayend):
    print("Crawling data on:", tahun,"/", bulan,"/", tanggal)
    i=1
    for i in range(month, monthend):
        #print("Include page", i)
        if tanggal < 10:
            r = br.open("https://www.viva.co.id/index/berita/all/"+ str(tahun)+ "/" + str(bulan)+ "/" + str(tanggal))
        else:
            r = br.open("https://www.viva.co.id/index/berita/all/"+ str(tahun)+ "/" + str(bulan)+ "/" + str(tanggal))
        else:
            if bulan < 10:
                r = br.open("https://www.viva.co.id/index/berita/all/"+ str(tahun)+ "/" + str(bulan)+ "/" + str(tanggal))
            else:
                r = br.open("https://www.viva.co.id/index/berita/all/"+ str(tahun)+ "/" + str(bulan)+ "/" + str(tanggal))

        for link in br.links():
            linklist.append(link.url)

        for links in linklist:
            if links not in kamus and re.search("berita",links):
                kamus[links]=values
                count+=1
                #print(kamus[links])
            else:
                pass
    print("Executed successfully")
```

```
F CrawlerVivaNews.pynt X

print(kamus)

f = open("dataVivaNews3.txt","w")
for x in kamus:
    f.write(x + "\n")
f.close()

f = open("dataVivaNews3.txt","r")
count = 2287

for link in f:
    print(link)
    urllib.request.urlretrieve(link, "VivaNews-" + str(count) + ".html")
    count+=1
    print("Downloading file: ", count)
print("Finish : ", count)
f.close()
```

Setelah *crawling*, dilanjutkan dengan proses *indexing*:

1. Update terminal: **\$ sudo apt update**

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
Microsoft Windows [Version 10.0.19042.572]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\Users\hizbu\PI>bash
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ sudo apt update
[sudo] password for maulanahz:
Get:1 http://security.ubuntu.com/ubuntu focal-security InRelease [107 kB]
Hit:2 http://archive.ubuntu.com/ubuntu focal InRelease
Get:3 http://archive.ubuntu.com/ubuntu focal-updates InRelease [111 kB]
Get:4 http://archive.ubuntu.com/ubuntu focal-backports InRelease [98.3 kB]
Get:5 http://security.ubuntu.com/ubuntu focal-security/main amd64 Packages [350 kB]
Get:6 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 Packages [637 kB]
Get:7 http://security.ubuntu.com/ubuntu focal-security/main amd64 c-n-f Metadata [5176 B]
Get:8 http://archive.ubuntu.com/ubuntu focal-updates/main Translation-en [160 kB]
Get:9 http://archive.ubuntu.com/ubuntu focal-updates/main amd64 c-n-f Metadata [10.8 kB]
Get:10 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 Packages [686 kB]
Get:11 http://archive.ubuntu.com/ubuntu focal-updates/universe Translation-en [131 kB]
Get:12 http://archive.ubuntu.com/ubuntu focal-updates/universe amd64 c-n-f Metadata [12.9 kB]
Fetched 2310 kB in 14s (165 kB/s)
Reading package lists... Done
Building dependency tree
Reading state information... Done
114 packages can be upgraded. Run 'apt list --upgradable' to see them.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$
```

2. Install SWISH-e: `$ sudo apt-get install swish-e`

Cek Versi SWISH-e: `$ swish-e -V`

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
Reading state information... Done
114 packages can be upgraded. Run 'apt list --upgradable' to see them.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ sudo apt-get install swish-e
Reading package lists... Done
Building dependency tree
Reading state information... Done
swish-e is already the newest version (2.4.7-6build2).
0 upgraded, 0 newly installed, 0 to remove and 114 not upgraded.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ swish-e -V
SWISH-E 2.4.7
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$
```

3. Membuat file konfigurasi SWISH-e untuk *indexing*: `$ touch crawl.conf`

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
Microsoft Windows [Version 10.0.19042.572]
(c) 2020 Microsoft Corporation. All rights reserved.

C:\Users\hizbu\PI>bash
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ touch crawl.conf
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$
```

4. Menulis file konfigurasi yang telah dibuat sebelumnya: `$ pico crawl.conf`

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
GNU nano 4.8                                crawl.conf
#directory yang diindex
IndexDir ./file

#file yang diindex, ekstensi .html
IndexOnly .html

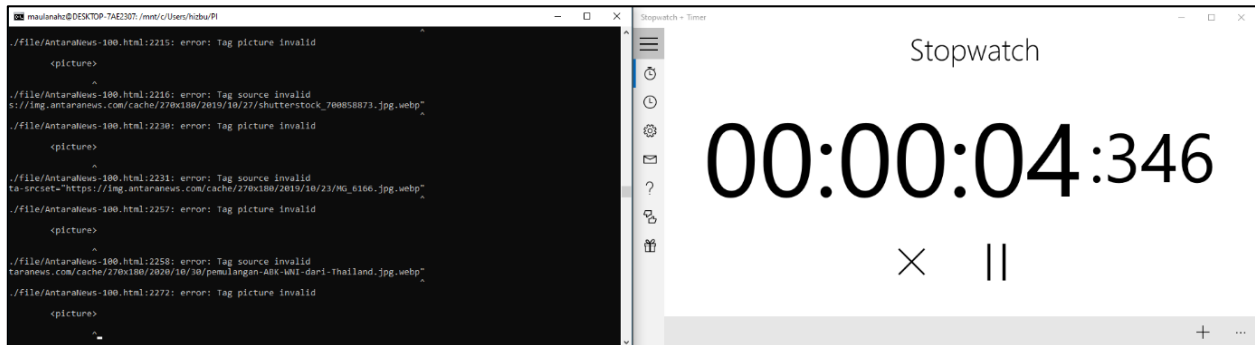
#mengambil konten html
DefaultContents HTML*
StoreDescription HTML* <body> 200000

#create file index
IndexFile ./result

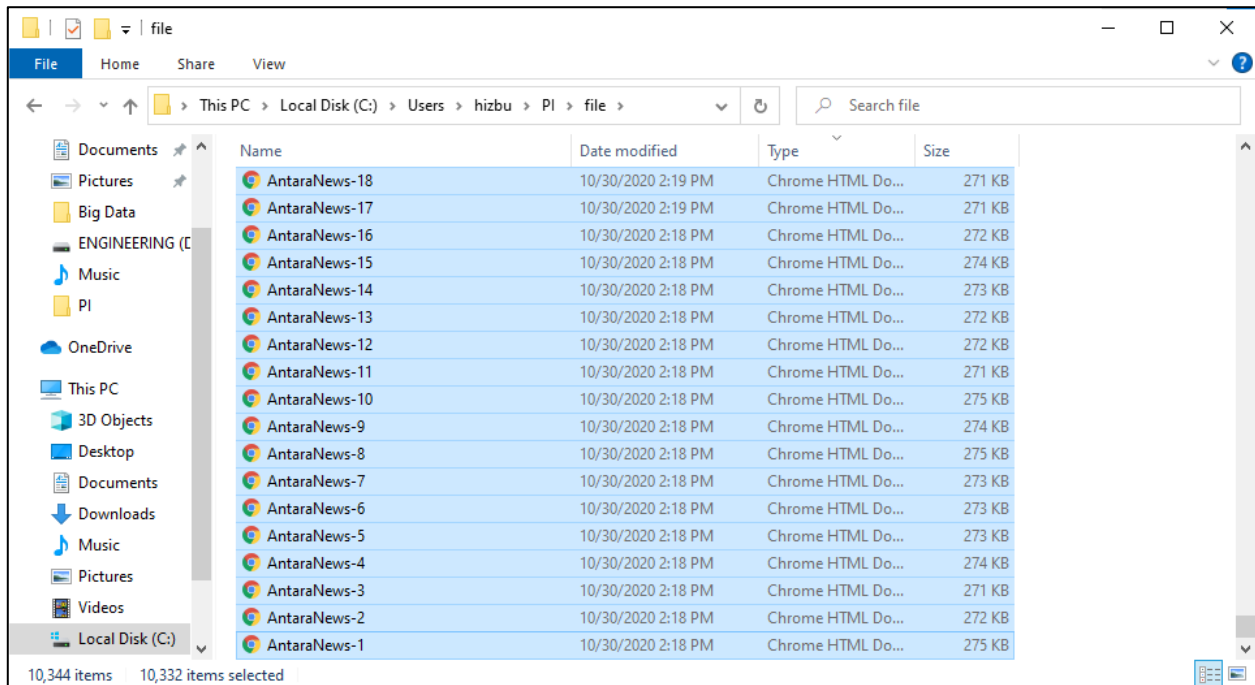
^G Get Help      ^O Write Out    ^W Where Is     ^K Cut Text     ^J Justify      ^C Cur Pos      M-U Undo
^X Exit          ^R Read File    ^\ Replace      ^U Paste Text   ^T To Spell     ^_ Go To Line     M-E Redo
```

5. Menjalankan SWISH-e: `$ swish-e -c crawl.conf`

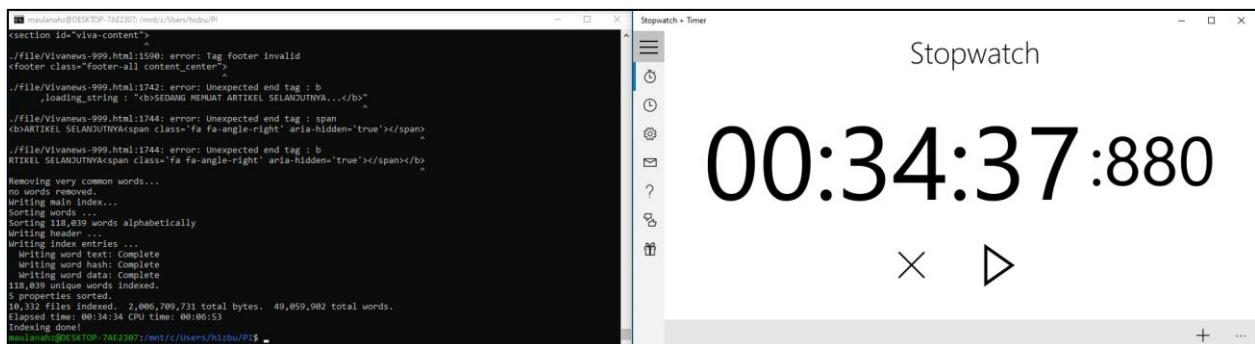
Proses *indexing* di SWISH-e dihitung menggunakan *stopwatch*.



Dapat dilihat *file* hasil *crawling* yang diindeks berjumlah 10.322 *files*.



Dari hasil pengukuran waktu dengan *stopwatch*, dihasilkan lama proses *indexing* yaitu 34 menit dengan jumlah 10K *files*.



Hasil *indexing* dengan SWISH-e.

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
<section id="viva-content">
^
./file/Vivanews-999.html:1590: error: Tag footer invalid
<footer class="footer-all content_center">
^
./file/Vivanews-999.html:1742: error: Unexpected end tag : b
,loading_string : "<b>SEDANG MEMUAT ARTIKEL SELANJUTNYA...</b>"
^
./file/Vivanews-999.html:1744: error: Unexpected end tag : span
<b>ARTIKEL SELANJUTNYA<span class='fa fa-angle-right' aria-hidden='true'></span>
^
./file/Vivanews-999.html:1744: error: Unexpected end tag : b
RTIKEL SELANJUTNYA<span class='fa fa-angle-right' aria-hidden='true'></span></b>
^
Removing very common words...
no words removed.
Writing main index...
Sorting words ...
Sorting 118,039 words alphabetically
Writing header ...
Writing index entries ...
Writing word text: Complete
Writing word hash: Complete
Writing word data: Complete
118,039 unique words indexed.
5 properties sorted.
10,332 files indexed. 2,006,709,731 total bytes. 49,059,902 total words.
Elapsed time: 00:34:34 CPU time: 00:06:53
Indexing done!
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI$
```

6. Melakukan proses *Search Word*

- search word **Boolean Query**

\$ swish-e -f result -m 10 -w Jakarta AND bandung : menghasilkan 608 hit.

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
Writing header ...
Writing index entries ...
Writing word text: Complete
Writing word hash: Complete
Writing word data: Complete
118,039 unique words indexed.
5 properties sorted.
10,332 files indexed. 2,006,709,731 total bytes. 49,059,902 total words.
Elapsed time: 00:34:34 CPU time: 00:06:53
Indexing done!
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI$ swish-e -f result -m 10 -w jakarta AND bandung
# SWISH format: 2.4.7
# Search words: jakarta AND bandung
# Removed stopwords:
# Number of hits: 608
# Search time: 0.003 seconds
# Run time: 0.014 seconds
1000 ./file/Vivanews-1442.html "Bor Raksasa Proyek Kereta Cepat Jakarta-Bandung Siap Beroperasi" 72576
955 ./file/AntaraNews-1320.html "Gubernur prihatin banjir sebabkan korban jiwa di Bandung - ANTARA News" 277792
950 ./file/AntaraNews-1039.html "Polres bandung adakan lomba kreatifitas untuk bobotoh Persib - ANTARA News" 278757
940 ./file/AntaraNews-2822.html "Syuting "Dilan 1991", Ira Wibowo Jakarta-Bandung PP tiap hari - ANTARA News" 277172
927 ./file/AntaraNews-1339.html "Warga Bandung yang hanyut saat banjir ditemukan tewas - ANTARA News" 277703
925 ./file/Kompas-4048.html "Tanggul Jebol, Siswa SDN Aji Tunggul Bandung Terjebak di Sekolah" 230247
917 ./file/AntaraNews-1266.html "Persiwa ubah strategi saat hadapi Persib Bandung - ANTARA News" 277760
912 ./file/AntaraNews-1877.html "BJB Pakuan buka peluang ke final usai jegal Popsivo - ANTARA News" 280093
910 ./file/AntaraNews-26.html "Persib tunggu surat resmi PSSI soal kompetisi lanjut Februari 2021 - ANTARA News" 279148
905 ./file/AntaraNews-1789.html "BJB putri dan Pertamina putra optimistis raih hasil maksimal - ANTARA News" 279042
.
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI$
```

\$ swish-e -f result -m 10 -w Jakarta OR bandung : menghasilkan 8765 hit.

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
950 ./file/AntaraNews-1039.html "Polres bandung adakan lomba kreatifitas untuk bobotoh Persib - ANTARA News" 278757
940 ./file/AntaraNews-2822.html "Syuting "Dilan 1991", Ira Wibowo Jakarta-Bandung PP tiap hari - ANTARA News" 277172
927 ./file/AntaraNews-1339.html "Warga Bandung yang hanyut saat banjir ditemukan tewas - ANTARA News" 277703
925 ./file/Kompas-4048.html "Tanggul Jebol, Siswa SDN Aji Tunggal Bandung Terjebak di Sekolah" 230247
917 ./file/AntaraNews-1266.html "Persiwa ubah strategi saat hadapi Persib Bandung - ANTARA News" 277760
912 ./file/AntaraNews-1877.html "BJB Pakuan buka peluang ke final usai jegal Popsivo - ANTARA News" 280093
910 ./file/AntaraNews-26.html "Persib tunggu surat resmi PSSI soal kompetisi lanjut Februari 2021 - ANTARA News" 279148
905 ./file/AntaraNews-1789.html "BJB putri dan Pertamina putra optimistis raih hasil maksimal - ANTARA News" 279042
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ swish-e -f result -m 10 -w jakarta OR bandung
# SWISH format: 2.4.7
# Search words: jakarta OR bandung
# Removed stopwords:
# Number of hits: 8765
# Search time: 0.004 seconds
# Run time: 0.014 seconds
1000 ./file/Vivanews-1442.html "Bor Raksasa Proyek Kereta Cepat Jakarta-Bandung Siap Beroperasi" 72576
953 ./file/AntaraNews-1320.html "Gubernur prihatin banjir sebabkan korban jiwa di Bandung - ANTARA News" 277792
950 ./file/AntaraNews-1039.html "Polres bandung adakan lomba kreatifitas untuk bobotoh Persib - ANTARA News" 278757
938 ./file/AntaraNews-2822.html "Syuting "Dilan 1991", Ira Wibowo Jakarta-Bandung PP tiap hari - ANTARA News" 277172
926 ./file/AntaraNews-1339.html "Warga Bandung yang hanyut saat banjir ditemukan tewas - ANTARA News" 277703
925 ./file/Kompas-4048.html "Tanggul Jebol, Siswa SDN Aji Tunggal Bandung Terjebak di Sekolah" 230247
917 ./file/AntaraNews-1266.html "Persiwa ubah strategi saat hadapi Persib Bandung - ANTARA News" 277760
912 ./file/AntaraNews-1877.html "BJB Pakuan buka peluang ke final usai jegal Popsivo - ANTARA News" 280093
910 ./file/AntaraNews-26.html "Persib tunggu surat resmi PSSI soal kompetisi lanjut Februari 2021 - ANTARA News" 279148
905 ./file/AntaraNews-1789.html "BJB putri dan Pertamina putra optimistis raih hasil maksimal - ANTARA News" 279042
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$
```

\$ swish-e -f result -m 10 -w Jakarta NOT bandung : menghasilkan 7863 hit.

```
maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
950 ./file/AntaraNews-1039.html "Polres bandung adakan lomba kreatifitas untuk bobotoh Persib - ANTARA News" 278757
938 ./file/AntaraNews-2822.html "Syuting "Dilan 1991", Ira Wibowo Jakarta-Bandung PP tiap hari - ANTARA News" 277172
926 ./file/AntaraNews-1339.html "Warga Bandung yang hanyut saat banjir ditemukan tewas - ANTARA News" 277703
925 ./file/Kompas-4048.html "Tanggul Jebol, Siswa SDN Aji Tunggal Bandung Terjebak di Sekolah" 230247
917 ./file/AntaraNews-1266.html "Persiwa ubah strategi saat hadapi Persib Bandung - ANTARA News" 277760
912 ./file/AntaraNews-1877.html "BJB Pakuan buka peluang ke final usai jegal Popsivo - ANTARA News" 280093
910 ./file/AntaraNews-26.html "Persib tunggu surat resmi PSSI soal kompetisi lanjut Februari 2021 - ANTARA News" 279148
905 ./file/AntaraNews-1789.html "BJB putri dan Pertamina putra optimistis raih hasil maksimal - ANTARA News" 279042
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ swish-e -f result -m 10 -w jakarta NOT bandung
# SWISH format: 2.4.7
# Search words: jakarta NOT bandung
# Removed stopwords:
# Number of hits: 7863
# Search time: 0.003 seconds
# Run time: 0.014 seconds
1000 ./file/Kompas-4362.html "Tarif LRT Jakarta Flat Rp 5.000" 224735
970 ./file/Kompas-4323.html "Kantongi Sertifikasi dan Rekomendasi, LRT Jakarta Siap Jalani Simulasi" 225921
950 ./file/Kompas-4310.html "Tinggal Tentukan Tanggal, LRT Jakarta Siap Beroperasi..." 228014
937 ./file/Kompas-2649.html "Camat: Gambir Jadi Percontohan "Jakarta Satu" karena Kompleksitasnya" 224546
933 ./file/AntaraNews-3041.html "Cepat dan nyaman, MRT untuk masyarakat Jakarta yang lebih dinamis - ANTARA News" 299411
921 ./file/Kompas-4224.html "Anies Targetkan Lebih dari 65.000 Penumpang Naik MRT Jakarta Per Hari" 228689
919 ./file/Vivanews-2908.html "Rute Pendek, LRT Jakarta Penting Engga? [SUARA JAKARTA] - Vlix.id" 39478
917 ./file/AntaraNews-23.html "Hujan ringan bayangi wilayah Jakarta pada Jumat pagi hingga sore - ANTARA News" 278209
913 ./file/Kompas-3870.html "Cara Anies Cegah Becak dari Luar Daerah Masuk Jakarta" 224031
909 ./file/Kompas-4365.html "Anies Sebut Indeks Pembangunan Manusia Jakarta Tembus 80 pada 2018" 227244
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$
```

- search word **Keyword-Base Query**

\$ swish-e -f result -m 10 -w Jakarta bandung : menghasilkan 608 hit.

```

maulanahz@DESKTOP-7AE2307: /mnt/c/Users/hizbu/PI
950 ./file/Kompas-4310.html "Tinggal Tentukan Tanggal, LRT Jakarta Siap Beroperasi..." 228014
937 ./file/Kompas-2649.html "Camat: Gambir Jadi Percontohan "Jakarta Satu" karena Kompleksitasnya" 224546
933 ./file/AntaraNews-3041.html "Cepat dan nyaman, MRT untuk masyarakat Jakarta yang lebih dinamis - ANTARA News" 299
411
921 ./file/Kompas-4224.html "Anies Targetkan Lebih dari 65.000 Penumpang Naik MRT Jakarta Per Hari" 228689
919 ./file/Vivanews-2908.html "Rute Pendek, LRT Jakarta Penting Engga? [SUARA JAKARTA] - Vlix.id" 39478
917 ./file/AntaraNews-23.html "Hujan ringan bayangi wilayah Jakarta pada Jumat pagi hingga sore - ANTARA News" 278209
913 ./file/Kompas-3870.html "Cara Anies Cegah Becak dari Luar Daerah Masuk Jakarta" 224031
909 ./file/Kompas-4365.html "Anies Sebut Indeks Pembangunan Manusia Jakarta Tembus 80 pada 2018" 227244
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$ swish-e -f result -m 10 -w jakarta bandung
# SWISH format: 2.4.7
# Search words: jakarta bandung
# Removed stopwords:
# Number of hits: 608
# Search time: 0.003 seconds
# Run time: 0.014 seconds
1000 ./file/Vivanews-1442.html "Bor Raksasa Proyek Kereta Cepat Jakarta-Bandung Siap Beroperasi" 72576
955 ./file/AntaraNews-1320.html "Gubernur prihatin banjir sebabkan korban jiwa di Bandung - ANTARA News" 277792
950 ./file/AntaraNews-1039.html "Polres bandung adakan lomba kreatifitas untuk bobotoh Persib - ANTARA News" 278757
940 ./file/AntaraNews-2822.html "Syuting "Dilan 1991", Ira Wibowo Jakarta-Bandung PP tiap hari - ANTARA News" 277172
927 ./file/AntaraNews-1339.html "Marga Bandung yang hanyut saat banjir ditemukan tewas - ANTARA News" 277703
925 ./file/Kompas-4048.html "Tanggul Jebol, Siswa SDN Aji Tunggul Bandung Terjebak di Sekolah" 230247
917 ./file/AntaraNews-1266.html "Persiwa ubah strategi saat hadapi Persib Bandung - ANTARA News" 277760
912 ./file/AntaraNews-1877.html "BJB Pakuan buka peluang ke final usai jegal Popsivo - ANTARA News" 280093
910 ./file/AntaraNews-26.html "Persib tunggu surat resmi PSSI soal kompetisi lanjut Februari 2021 - ANTARA News" 2791
48
905 ./file/AntaraNews-1789.html "BJB putri dan Pertamina putra optimistis raih hasil maksimal - ANTARA News" 279042
.
maulanahz@DESKTOP-7AE2307:/mnt/c/Users/hizbu/PI$

```

7. Hasil Boolean Query dan Keyword-Base Query

Dari kata "jakarta bandung" menghasilkan **Keyword-Base Query** yang sama dengan hasil **Boolean Query "AND"**, karena menghasilkan hit yang sama yaitu 608. Hal ini menunjukkan proses *default keyword-base query* merupakan proses *Boolean query "AND"*.