



### **Hak cipta dan penggunaan kembali:**

Lisensi ini mengizinkan setiap orang untuk menggubah, memperbaiki, dan membuat ciptaan turunan bukan untuk kepentingan komersial, selama anda mencantumkan nama penulis dan melisensikan ciptaan turunan dengan syarat yang serupa dengan ciptaan asli.

### **Copyright and reuse:**

This license lets you remix, tweak, and build upon work non-commercially, as long as you credit the origin creator and license it on your new creations under the identical terms.

## **BAB III**

### **PELAKSANAAN KERJA MAGANG**

#### **3.1. Kedudukan dan Koordinasi**

Praktek kerja magang dilaksanakan pada Universitas Multimedia Nusantara di *Lab Big Data*, berlokasi di gedung C 503. Deskripsi pekerjaan magang ini sebagai *Text Mining* yang dibimbing langsung oleh Bapak Yustinus Eko selaku sekretaris program studi sistem informasi dan supervisor. Selama praktek kerja magang di Universitas Multimedia Nusantara, dikoordinasikan langsung dengan tim dari program studi ilmu komunikasi dan program studi sistem informasi mengenai beberapa kegiatan, seperti *fact checking* mengenai debat pemilihan presiden 2019 di saat siaran *live* berlangsung akan dilakukan pencarian sumber berita yang relevan melalui *keyword* yang disusun oleh tim Ilkom.

#### **3.2. Tugas yang dilakukan**

Pada kerja magang ini terdapat beberapa *script* yang sudah dibuat oleh rekan magang sebelumnya, sehingga penulis dapat langsung menggunakan dan melakukan pengembangan program yang sudah ada untuk meningkatkan kemudahan penggunaan. Berikut merupakan hasil yang dilakukan selama 2 bulan kerja magang di Universitas Multimedia Nusantara:

##### **3.2.1. Fact Checking Debat Capres Ke- 4**

Persiapan *fact checking* untuk debat capres ke-4 ini dilakukan sedikit *briefing* bersama tim Ilkom dan tim SI sebelum debat capres dimulai. Bersama

dengan Bapak Bintang, selaku dosen dari tim Ilkom dan anggota-anggota tim Ilkom lainnya menjelaskan rundown yang harus dilakukan saat *live* akan berlangsung.

Bapak Bintang menjelaskan bahwa tim Ilkom akan berada di ruangan B6 yaitu *News Room* dan tim dari SI bertempat di *Lab Big Data* di gedung C 503. Tim Ilkom akan memberikan keyword setiap sesi dari debat capres, yang disusun oleh tim Ilkom di pimpin oleh Mario selaku representatif. Apabila *keyword* yang disusun sudah siap untuk dicari sumber beritanya oleh tim SI maka si Mario akan memberitahu kepada tim SI bahwa keyword *ready* melalui Whatsapp dan proses pengerjaan dari persiapan *keyword* dan hasil pencarian akan dikomunikasikan melalui Google Drive yang telah di sediakan oleh tim Ilkom.

Jadwal debat capres ke-4 diadakan pada tanggal 30 Maret 2019. Sebelum kegiatan *fact checking* dimulai, dikarenakan hari pertama penulis melakukan kerja magang ini. Penulis diberikan sedikit *tutorial* oleh mantan rekan magang tim SI sebelumnya, mengenai cara kerja setiap program yang mereka buat untuk melakukan *fact checking* seperti program *linkkeyword.py* untuk mencari sumber-sumber artikel Kompas yang relevan pada database yang telah di *scraping*, cara mengambil dan memperbaharui database artikel Kompas dengan menggunakan perintah *wget --recursive "link website"* sebagai contoh *wget --recursive <https://nasional.kompas.com/>* akan mengambil seluruh artikel berita dengan kategori berita nasional Kompas pada gambar 3.1. *Terminal Linux dan Command Wget.*

```

pc-bigdata-1@pc-bigdata-1:~$ cd Desktop/artikel2019
pc-bigdata-1@pc-bigdata-1:~/Desktop/artikel2019$ wget --recursive https://nasional.kompas.com/
--2019-06-16 21:12:43-- https://nasional.kompas.com/
Resolving nasional.kompas.com (nasional.kompas.com)... 52.84.251.60, 52.84.251.119, 52.84.251.39, ...
Connecting to nasional.kompas.com (nasional.kompas.com)|52.84.251.60|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: unspecified [text/html]
Saving to: 'nasional.kompas.com/index.html'

nasional.kompas.com  [ <=>          ] 126,62K  --.-KB/s  in 0,02s

2019-06-16 21:12:44 (5,33 MB/s) - 'nasional.kompas.com/index.html' saved [129664]

Loading robots.txt; please ignore errors.
--2019-06-16 21:12:44-- https://nasional.kompas.com/robots.txt
Reusing existing connection to nasional.kompas.com:443.
HTTP request sent, awaiting response... 200 OK
Length: 325 [text/plain]
Saving to: 'nasional.kompas.com/robots.txt'

nasional.kompas.com 100%[=====]      325  --.-KB/s  in 0s

2019-06-16 21:12:44 (20,0 MB/s) - 'nasional.kompas.com/robots.txt' saved [325/325]

--2019-06-16 21:12:44-- https://nasional.kompas.com/twitter.com
Reusing existing connection to nasional.kompas.com:443.
HTTP request sent, awaiting response... 404 Not Found

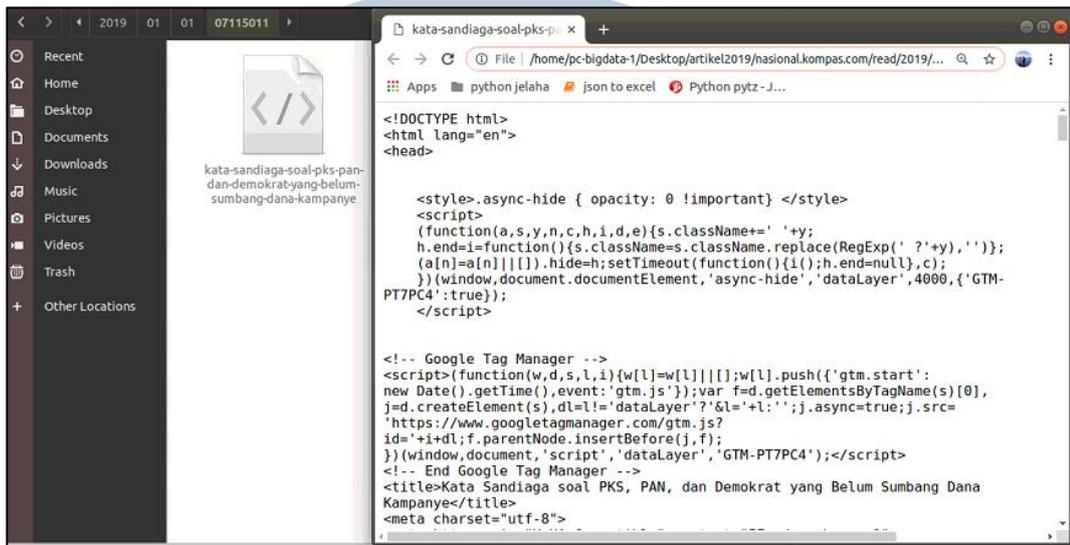
```

**Gambar 3.1. Terminal Linux dan Command Wget**

Hasil dari perintah `wget --recursive` “link website” tersebut akan menyimpan data-data artikel dengan format *HTML document* pada gambar 3.3. *HTML Document*. Kemudian, *HTML document* tersebut di ekstrak isinya dengan menggunakan *script* `Scrap.py` untuk mengambil data-data yang diperlukan dari file *HTML document* dan data tersebut akan lebih terstruktur yang disimpan ke dalam bentuk format `CSV` isi yang didapatkan seperti judul berita, link berita, link foto, *thumb*, isi konten berita, tanggal *publish*, nama situs, jenis artikel, nama pengarang dan nama penerbit dapat dilihat pada gambar 3.4. Hasil `Scrap.py`. Hasil dari `CSV` tersebut sebagai database sumber berita yang akan digunakan dalam kegiatan *fact checking*.



**Gambar 3.2. Kategori Artikel Kompas**



Gambar 3.3. HTML Document

The image shows a spreadsheet application displaying a table with the following columns: title, link, photo, thumb, content, date, site\_name, section\_name, editor, and author. The table contains 45 rows of data, including titles like "Mandi Pemandi Kita Membuang?", "Sepekan Ini, Al-Azhar Tetap Prinsip Kurangnya", and "Tahun Ini, Mandi Syariat Fokus Empat".

no	title	link	photo	thumb	content	date	site_name	section_name	editor	author
1	Mandi Pemandi Kita Membuang?	https://nasional.kompas.com/read/2019/03/01/mandi-pemandi-kita-membuang			Kembali tidak diperbolehkan? JAKARTA, KOMPAS.com-01/03/19 15:37:31	Money	Spend Smart	Barabang Pityo Jatmiko	Muti Ali Lingga	
2	Sepekan Ini, Al-Azhar Tetap Prinsip Kurangnya	https://nasional.kompas.com/read/2019/03/01/sepekan-ini-al-azhar-tetap-prinsip-kurangnya			Ambil saja pencekukan? JAKARTA, KOMPAS.com-01/03/19 14:48:19	Money	Whats New	Barabang Pityo Jatmiko	Muti Ali Lingga	
3	Tahun Ini, Mandi Syariat Fokus Empat	https://nasional.kompas.com/read/2019/03/01/tahun-ini-mandi-syariat-fokus-empat			Flur yang saat ini tergep? JAKARTA, KOMPAS.com-01/03/19 14:23:35	Money	Whats New	Barabang Pityo Jatmiko	Ambareng Naika Kemala Mowata	
4	Pengadilan Eropa Mengonfirmasi	https://nasional.kompas.com/read/2019/03/01/pengadilan-eropa-mengonfirmasi			Untuk konstitusi emas? JAKARTA, KOMPAS.com-01/03/19 14:07:00	Money	Whats New	Sakina Ekharma Dah Setiawan	Muti Ali Lingga	
5	Sudut 4.75 Jula Wakil Paksi Lapor	https://nasional.kompas.com/read/2019/03/01/sudut-4-75-jula-wakil-paksi-lapor			Hingga Senin (11/3/2019). JAKARTA, KOMPAS.com-01/03/19 13:41:37	Money	Whats New	Sakina Ekharma Dah Setiawan	Yoga Sukmana	
6	Trial Lapor SPT Pajak	https://nasional.kompas.com/read/2019/03/01/trial-lapor-spt-pajak			Demis bagi siapa saja? JAKARTA, KOMPAS.com-01/03/19 13:21:20	Money	Whats New	Eterangga Durmana	Yoga Sukmana	
7	Prata Sandiaga Shapiro Taken Core	https://nasional.kompas.com/read/2019/03/01/prata-sandiaga-shapiro-taken-core			Konstitusi itu bukan orang? JAKARTA, KOMPAS.com-01/03/19 13:14:59	Money	Whats New	Sakina Ekharma Dah Setiawan		
8	INACIA Hanya Awak Tak Berencana	https://nasional.kompas.com/read/2019/03/01/inacia-hanya-awak-tak-berencana			Mendat INACIA, keabakan? JAKARTA, KOMPAS.com-01/03/19 13:02:16	Money	Whats New	Sakina Ekharma Dah Setiawan	Eka Nurul Ulya	
9	Chipping Terasaki Digital	https://nasional.kompas.com/read/2019/03/01/chipping-terasaki-digital			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 12:44:10	Money	Whats New	Eterangga Durmana	Ambareng Naika Kemala Mowata	
10	Emponan Airlines Kerdakaan	https://nasional.kompas.com/read/2019/03/01/emponan-airlines-kerdakaan			Pendapatan disebarkan ke? NEW YORK, KOMPAS.com-01/03/19 12:30:00	Money	Whats New	Sakina Ekharma Dah Setiawan	Ambareng Naika Kemala Mowata	
11	Pulutan di 8 Utefly - Kapan Tak	https://nasional.kompas.com/read/2019/03/01/pulutan-di-8-utefly-kapan-tak			Jaga, tidak dipan dibenar? SEORANG pra baq yang? JAKARTA, KOMPAS.com-01/03/19 12:25:27	Money	Whats New	Henu Mangalato	Mahroka K Dulu	
12	Chiquan Airlines Patakan Tak	https://nasional.kompas.com/read/2019/03/01/chiquan-airlines-patakan-tak			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 12:14:13	Money	Whats New	Sakina Ekharma Dah Setiawan	Eka Nurul Ulya	
13	Dukung Kerdakaan PT SMI	https://nasional.kompas.com/read/2019/03/01/dukung-kerdakaan-pt-smi			Hai terasah diawal? JAKARTA, KOMPAS.com-01/03/19 11:31:08	Money	Whats New	Eterangga Durmana	Eka Nurul Ulya	
14	Kerendakan Pengawasan	https://nasional.kompas.com/read/2019/03/01/kerendakan-pengawasan			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 11:20:36	Money	Whats New	Eterangga Durmana		
15	Maw Terasah Diawal	https://nasional.kompas.com/read/2019/03/01/maw-terasah-diawal			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 11:02:42	Money	Whats New	Spend Smart	Eterangga Durmana	
16	Message China Untuk	https://nasional.kompas.com/read/2019/03/01/message-china-untuk			Bagi yang ingin mendat? KOMPAS.com - Bag A? 2019-03-11 09:33:03	Money	Whats New	Eterangga Durmana	Ambareng Naika Kemala Mowata	
17	Dampak Fosfor	https://nasional.kompas.com/read/2019/03/01/dampak-fosfor			Begalunya bantuan? JAKARTA, KOMPAS.com-01/03/19 09:58:50	Money	Whats New	Sakina Ekharma Dah Setiawan		
18	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			UMI adalah program? KOMPAS.com - Kaitu? 2019-03-11 09:34:00	Money	Whats New	Label		
19	Pilih Mana	https://nasional.kompas.com/read/2019/03/01/pilih-mana			Diawal satu? Jakarta, KOMPAS.com - Mancu? 2019-03-11 07:31:19	Money	Whats New	Eterangga Durmana	Muti Ali Lingga	
20	Mendat Risi	https://nasional.kompas.com/read/2019/03/01/mendat-risi			Diawal satu? Jakarta, KOMPAS.com - Mancu? 2019-03-11 07:31:19	Money	Whats New	Eterangga Durmana	Muti Ali Lingga	
21	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 07:00:00	Money	Whats New	Eterangga Durmana		
22	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
23	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
24	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
25	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
26	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
27	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
28	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
29	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
30	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
31	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
32	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
33	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
34	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
35	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
36	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
37	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
38	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
39	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
40	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
41	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
42	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
43	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
44	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		
45	Ini Konsep	https://nasional.kompas.com/read/2019/03/01/ini-konsep			Ini konsepnya jayanya? ANDES ABABA, KOMPAS.com-01/03/19 06:37:00	Money	Whats New	Eterangga Durmana		

Gambar 3.4. Hasil Scrap.py

Setelah tutorial singkat selesai oleh mantan rekan kerja SI bernama Yulius David dan memantau kinerja pertama penulis saat siaran live debat capres ke-4 di mulai. Tugas dari tim SI adalah menggunakan script Python yang telah dibuat sebelumnya untuk setiap keyword yang diberikan tim Ilkom

akan dilakukan pencarian sumber-sumber beritanya. Terdapat 5 sesi pada debat capres, tim dari SI akan memberikan tanggapan kepada tim Ilkom hasil dari *script* yang berupa *link* artikel tersebut dan menentukan apakah pernyataan dari calon presiden sesuai fakta atau tidak. Mario akan membagikan 2 *link online document* dari Google Drive, link pertama merupakan *form* untuk *claim* untuk tim Ilkom memberikan *keyword* kepada tim SI dengan *link* <https://docs.google.com/spreadsheets/d/1gUIFqlaYE5MaxeRubd9zQ5cry7Z1fACxFisYMujqTn4/edit?usp=sharing> pada gambar 3.5. *Form Keyword Claim*.

SEGMENT 1	SEGMENT 2	SEGMENT 3
	Indonesia 1100 bahasa daerah	Pemerintah bangun Divisi 3, Komando AU
	Pemerintah sudah bangun badan pembinaan ideologi pancasila	2019 pemerintah gelar pasukan di Natuna
	4,5 tahun jokowi dituduh PKI	Indonesia memiliki Tank Harimau dan Kap
	4,5 tahun 23 lembaga dibubarkan	11 titik radar maritim sudah terkoneksi
	perpajakan sekarang sudah online	107 Triliun Anggaran pertahanan
		2019 Anggaran pertahanan terbesar ked
		2018 menengahi pengungsi di Rakhine St
		Indonesia mendamaikan faksi-faksi konflik
		Indonesia penduduk muslim terbesar di du
		investasi dan perdagangan bertandanga
		Australia terima produk Indonesia melalui

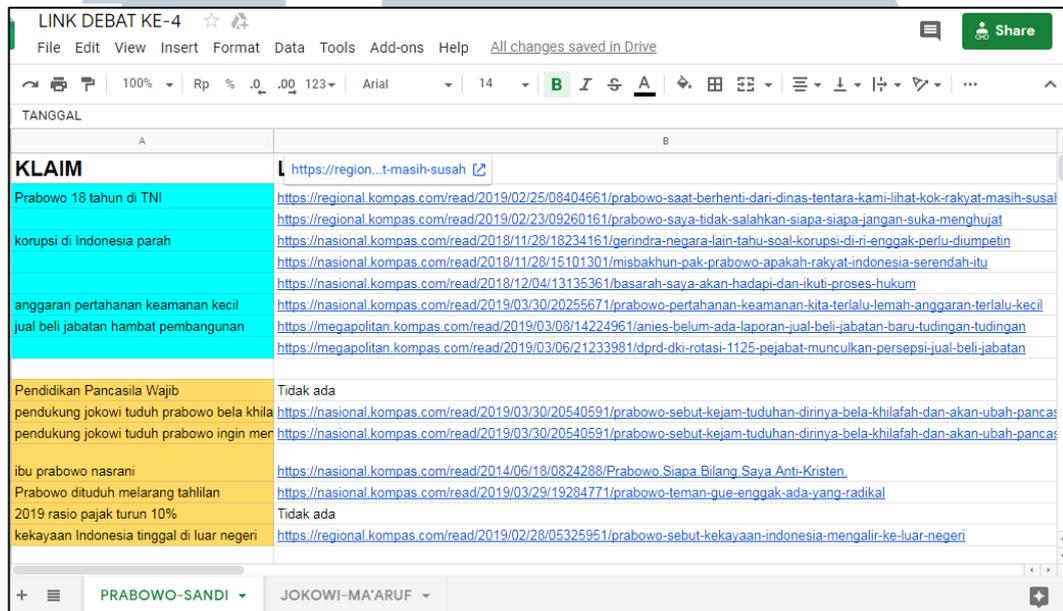
**Gambar 3.5. Form Keyword Claim**

*Link* kedua merupakan *form* untuk *link* berita yang ditemukan dan digunakan oleh tim SI untuk memasukkan *data link* beserta tanggal dari artikel

[https://docs.google.com/spreadsheets/d/1fyJ6aCAYGbZVyr9oB-G\\_6v-](https://docs.google.com/spreadsheets/d/1fyJ6aCAYGbZVyr9oB-G_6v-Hgrh74VET6NwnHH-mLFI/edit?usp=sharing)

[Hgrh74VET6NwnHH-mLFI/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1fyJ6aCAYGbZVyr9oB-G_6v-Hgrh74VET6NwnHH-mLFI/edit?usp=sharing) pada gambar 3.6. *Form Result*

*Keyword.*



TANGGAL	
<b>KLAIM</b>	<a href="https://region...t-masih-susah">https://region...t-masih-susah</a>
Prabowo 18 tahun di TNI	<a href="https://regional.kompas.com/read/2019/02/25/08404661/prabowo-saat-berhenti-dan-dinas-tentara-kami-lihat-kok-rakyat-masih-susah">https://regional.kompas.com/read/2019/02/25/08404661/prabowo-saat-berhenti-dan-dinas-tentara-kami-lihat-kok-rakyat-masih-susah</a>
korupsi di Indonesia parah	<a href="https://regional.kompas.com/read/2019/02/23/09260161/prabowo-saya-tidak-salahkan-siapa-siapa-jangan-suka-menghujat">https://regional.kompas.com/read/2019/02/23/09260161/prabowo-saya-tidak-salahkan-siapa-siapa-jangan-suka-menghujat</a>
	<a href="https://nasional.kompas.com/read/2018/11/28/18234161/gerindra-negara-lain-tahu-soal-korupsi-di-ri-enggak-perlu-diumpetin">https://nasional.kompas.com/read/2018/11/28/18234161/gerindra-negara-lain-tahu-soal-korupsi-di-ri-enggak-perlu-diumpetin</a>
	<a href="https://nasional.kompas.com/read/2018/11/28/15101301/misbakhun-pak-prabowo-apakah-rakyat-indonesia-serendah-itu">https://nasional.kompas.com/read/2018/11/28/15101301/misbakhun-pak-prabowo-apakah-rakyat-indonesia-serendah-itu</a>
	<a href="https://nasional.kompas.com/read/2018/12/04/13135361/basarah-saya-akan-hadapi-dan-ikuti-proses-hukum">https://nasional.kompas.com/read/2018/12/04/13135361/basarah-saya-akan-hadapi-dan-ikuti-proses-hukum</a>
anggaran pertahanan keamanan kecil	<a href="https://nasional.kompas.com/read/2019/03/30/20255671/prabowo-pertahanan-keamanan-kita-terlalu-lemah-anggaran-terlalu-kecil">https://nasional.kompas.com/read/2019/03/30/20255671/prabowo-pertahanan-keamanan-kita-terlalu-lemah-anggaran-terlalu-kecil</a>
jual beli jabatan hambat pembangunan	<a href="https://megapolitan.kompas.com/read/2019/03/08/14224961/anies-belum-ada-laporan-jual-beli-jabatan-baru-tudingan-tudingan">https://megapolitan.kompas.com/read/2019/03/08/14224961/anies-belum-ada-laporan-jual-beli-jabatan-baru-tudingan-tudingan</a>
	<a href="https://megapolitan.kompas.com/read/2019/03/06/21233981/dprd-dki-rotasi-1125-pejabat-munculkan-persepsi-jual-beli-jabatan">https://megapolitan.kompas.com/read/2019/03/06/21233981/dprd-dki-rotasi-1125-pejabat-munculkan-persepsi-jual-beli-jabatan</a>
Pendidikan Pancasila Wajib	Tidak ada
pendukung jokowi tuduh prabowo bela khilafah	<a href="https://nasional.kompas.com/read/2019/03/30/20540591/prabowo-sebut-kejam-tuduhan-dirinya-bela-khilafah-dan-akan-ubah-pancas">https://nasional.kompas.com/read/2019/03/30/20540591/prabowo-sebut-kejam-tuduhan-dirinya-bela-khilafah-dan-akan-ubah-pancas</a>
pendukung jokowi tuduh prabowo ingin merubah	<a href="https://nasional.kompas.com/read/2019/03/30/20540591/prabowo-sebut-kejam-tuduhan-dirinya-bela-khilafah-dan-akan-ubah-pancas">https://nasional.kompas.com/read/2019/03/30/20540591/prabowo-sebut-kejam-tuduhan-dirinya-bela-khilafah-dan-akan-ubah-pancas</a>
ibu prabowo nasrani	<a href="https://nasional.kompas.com/read/2014/06/18/0824288/Prabowo_Siapa_Bilang_Saya_Anti-Kristen">https://nasional.kompas.com/read/2014/06/18/0824288/Prabowo_Siapa_Bilang_Saya_Anti-Kristen</a>
Prabowo dituduh melarang tahlilan	<a href="https://nasional.kompas.com/read/2019/03/29/19284771/prabowo-teman-gue-enggak-ada-yang-radikal">https://nasional.kompas.com/read/2019/03/29/19284771/prabowo-teman-gue-enggak-ada-yang-radikal</a>
2019 rasio pajak turun 10%	Tidak ada
kekayaan Indonesia tinggal di luar negeri	<a href="https://regional.kompas.com/read/2019/02/28/05325951/prabowo-sebut-kekayaan-indonesia-mengalir-ke-luar-negeri">https://regional.kompas.com/read/2019/02/28/05325951/prabowo-sebut-kekayaan-indonesia-mengalir-ke-luar-negeri</a>

**Gambar 3.6. Form Result Keyword**

Setelah *form result keyword* telah berhasil ditemukan sumber-sumber berita yang disusun oleh tim Ilkom berdasarkan debat kedua pasangan calon presiden terkait klaimnya, maka tim Ilkom akan memastikan isi sumber berita yang dipublikasi oleh Kompas sesuai dengan fakta yang dilontarkan saat *live* debat berlangsung.

### 3.2.2. Scarping Artikel Kompas.com

*Scarping* artikel Kompas biasanya dilakukan 1-2 hari menjelang debat capres akan dimulai, bertujuan untuk mendapatkan hasil berita yang *up-to-date*

pada database. Berikut ini tahapan untuk melakukan *scarping* artikel Kompas.com yang penulis gunakan masih *script* dari mantan rekan magang.

Pertama penulis melakukan *update* isi artikel Kompas untuk tiap kategori beritanya, yaitu dengan menggunakan command `wget --recursive "link website"`. Untuk melakukan *scarping* ini dapat dilakukan secara bersamaan dengan membuka *terminal* Linux lebih dari satu, kemudian memasukkan command sebagai contoh `wget --recursive https://nasional.kompas.com/`. Kategori artikel Kompas terdiri dari 7 kategori yaitu, nasional, edukasi, ekonomi, megapolitan, regional, sains, dan tekno, dapat dilihat pada gambar 3.2. Kategori Artikel Kompas. Setiap isi dari kategori merupakan *HTML document* yang berisi konten berita artikel pada gambar 3.3. *HTML document*.

Setelah selesai *download* dengan perintah `wget`, langkah berikutnya agar data lebih terstruktur rapi dan mudah dibaca, dengan menggunakan *script* `scrap.py` dapat di lihat pada gambar 3.7. Isi `Scrap.py`. Hasil yang diperoleh saat `script.py` dijalankan pada *terminal* Linux dengan menggunakan *command* Python "`namascript.py`" contoh Python `Scrap.py`. Hasil yang diperoleh ialah sebuah *file* dalam bentuk CSV dapat dilihat pada gambar 3.4. Hasil `Scrap.py`.

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A



```
from bs4 import BeautifulSoup
import csv
import os
from datetime import datetime

path = '/home/pc-bigdata-1/Desktop/artikel2019/ekonomi.kompas.com/read'
for root, dirs, files in os.walk(path):
    for file in files:
        with open(os.path.join(root, file)) as myfile:
            soup = BeautifulSoup(myfile, "html.parser")
            title = soup.find('meta', property='og:title').get("content", None)
            links = soup.findAll("link", rel='canonical')
            for link in links:
                newlink = link['href']
            photo = soup.find('meta', property='og:image').get("content", None)
            description = soup.find('meta', property='og:description').get("content", None)
            content = soup.find("div", {"class": "read_content"})
            date = soup.find('meta', attrs={'name': 'content_PublishedDate'}).get("content", None)
            site_name = soup.find('meta', attrs={'name': 'content_category'}).get("content", None)
            section_name = soup.find('meta', attrs={'name': 'content_subcategory'}).get("content", None)
            editor = soup.find('meta', attrs={'name': 'content_editor'}).get("content", None)
            author = soup.find('meta', attrs={'name': 'content_author'}).get("content", None)

            with open('ekonomikompastes.csv', 'a') as csv_file:
                writer = csv.writer(csv_file)
                writer.writerow([title, newlink, photo, description, content, date, site_name, section_name,
                                editor, author])
```

**Gambar 3.7. Isi Scrap.py**

Maksud dari *coding* di atas adalah dengan menaruh lokasi *directory* pada *folder* “read” saja maka secara otomatis akan menelusuri setiap *folder* yang ada di dalam *folder* “read” tersebut dengan menggunakan *os.walk*. Setiap *file directory* yang ditemukan *os.walk* akan melakukan perintah *BeautifulSoup*. *BeautifulSoup* berguna untuk mempermudah pengambilan *tag* pada sebuah *html* karena bentuk *html* lebih berantakan. Dicarilah *tag title, link, photo, thumb, description, content, date, site\_name, section\_name, editor, dan author* satu persatu. Setelah itu, membuat *file output* bernama “ekonomikompastes.csv” dengan *format* “a” yang berfungsi *file output* akan terus dilanjutkan ke baris selanjutnya tanpa menimpa *file* tersebut dan dimasukkan *tag* yang telah dicari ke dalam 1 baris per 1 *html* dapat dilihat pada gambar 3.4. Hasil Scrap.py. Setelah semua kategori Kompas sudah di scrap.py, maka seluruh kategori digabung menjadi 1 *file csv* atau *database* besar secara

*manual copy paste* karena hanya ada 7 *file csv*. *File csv* ini akan digunakan untuk debat capres selanjutnya.

### **3.2.3. Briefing Magang**

Penulis diminta oleh Bapak Yustinus Eko selaku supervisor untuk rapat mengenai *progress* pekerjaan dan pembagian kerja saat magang, meminta saran mengenai penyajian data twitter agar lebih mudah di baca, diminta mencari sebuah program untuk melakukan *scraping* Instagram, meminta saran-saran penyajian data Instagram yang diperlukan dari hasil *scraping* dan saran-saran persiapan dan perbaikan untuk laporan dan presentasi hasil analisa twitter yang diperoleh untuk disampaikan kepada tim *fact checking*, meminta saran penulisan laporan teknis mengenai analisa Kompas Twitter.

### **3.2.4. Pengembangan Script Python Linkkeyword.py**

Dikarenakan *script* yang saat ini kurang efisien maka dilakukan pengembangan terhadap *script*. Perubahan yang dilakukan dari *script* sebelumnya antara lain, dapat menentukan berapa banyak jumlah keyword yang ingin dicari dalam 1 sekali program di jalankan tanpa harus merubah isi koding setiap ingin mencari keyword tertentu dapat dilihat pada gambar 3.11. *Execute Python Linkkeywordth1*. Jika keyword yang dicari tidak sesuai atau tidak ada memungkinkan data tersebut memang tidak ada atau *keyword* yang dimasukkan kurang tepat. Perubahan yang terjadi pada *script* dapat dilihat pada gambar di bawah ini.

```

linkkeyword(ori).py
~/Desktop/Tugas Rapi/data artikel program

import csv
data = []
a = 1

with open('/home/pc-bigdata-1/Desktop/artikel2019/gabunganekonominasionalregional.csv', newline='') as csvfile:
    spamreader = csv.reader(csvfile, delimiter=',')
    for row in spamreader:
        data.append(row)

keyword1 = 'proyek'
keyword2 = 'strategis'
keyword3 = 'nasional'
#keyword4 = 'Angkasa pura'
#keyword5 = '2014'
#keyword6 = 'bubar'
#keyword7 = ''
namaoutputfile = 'hasil.csv'
datawrite = []

with open(namaoutputfile, 'w', newline='') as csvfile:
    writer = csv.writer(csvfile, delimiter=',')
    for a in range(1,10343):
        if keyword1 in data[a][5]:
            #if keyword2 or keyword3 or keyword4 or keyword5 in data[a][5]:
            #if keyword2 in data[a][5]:
            #if keyword3 in data[a][5]:
            #if keyword4 in data[a][5]:
            #if keyword5 in data[a][5]:
            #if keyword6 in data[a][5]:

            print ('Link: ' + data[a][1] + ' Date: ' + data[a][2])
            datawrite = data[a][1],data[a][2]
            writer.writerow(datawrite)
            a += 1

```

Gambar 3.8. Sebelum Modifikasi *Linkkeyword.py*

```

*testlinkkeywordth1.py
~/Desktop/Tugas Rapi

import csv
import re
data = []
jumlah = []
keyword = []
kata = 1
a = 0
row_count= 0
no = 1

with open('/home/pc-bigdata-1/Desktop/testing/gabunganhasil.csv', newline='') as csvfile:
#with open('/home/pc-bigdata-1/Desktop/artikel2019/gabunganekonominasionalregional.csv', newline='') as csvfile:#membaca data csv
    spamreader = csv.reader(csvfile, delimiter=',')
    for row in spamreader:
        data.append(row)

with open('/home/pc-bigdata-1/Desktop/testing/gabunganhasil.csv', newline='') as csvfile: #menghitung jumlah data.csv
    spamreader = csv.reader(csvfile, delimiter=',')
    for row in spamreader:
        row_count = sum(1 for row in spamreader)

print ("Jumlah Data Berita yang tersedia: " + str(row_count))

jumlah = input("Masukkan Jumlah Keyword yang ingin di cari : ")
for i in range (int(jumlah)) :
    keyword = input("Masukkan Keyword" + str(kata) + " : ")
    kata += 1

namaoutputfile = 'hasil.csv'
datawrite = []

with open (namaoutputfile, 'w', newline='') as csvfile:
    writer = csv.writer(csvfile, delimiter=',')
    for a in range (1,row_count):
        if re.search(keyword, data[a][5], re.IGNORECASE):
            print (str(no) + '=>Link: ' + data[a][1] + '\n -->Date: ' + data[a][2] + '\n ==>Title: ' + data[a][0])
            datawrite = data[a][1],data[a][2]
            writer.writerow(datawrite)
            a += 1
            no += 1

```

Gambar 3.9. Setelah Modifikasi *Linkkeywordth1.py*

M U L T I M E D I A  
N U S A N T A R A

```

pc-bigdata-1@pc-bigdata-1:~/Desktop/Tugas Rapi/data artikel program$ python linkkeyword(orl\).py
Link: https://regional.kompas.com/read/2019/03/11/10373931/pemprov-jatim-ambil-langkah-cepat-tang
ani-banjir-trenggalek Date: 03/11/19 10:37 AM
Link: https://nasional.kompas.com/read/2019/03/10/10143281/jokowi-tinjau-proyek-pengendali-banjir
-sungai-citarum-yang-pernah-mangkrak Date: 03/10/19 10:14 AM
Link: https://nasional.kompas.com/read/2019/03/10/10143281/jokowi-tinjau-proyek-pengendali-banjir
-sungai-citarum-yang-pernah-mangkrak Date: 03/10/19 10:14 AM
Link: https://nasional.kompas.com/read/2019/03/10/08255711/ini-kegiatan-jokowi-selama-kunjungan-k
e-bandung Date: 03/10/19 08:25 AM
Link: https://money.kompas.com/read/2019/03/09/194836426/menakar-proyeksi-pemerintah-untuk-meneka
n-harga-karet Date: 03/09/19 07:48 PM
Link: https://megapolitan.kompas.com/read/2019/03/08/14025411/gerai-pelayanan-publik-dibuka-di-pl
aza-cibubur-bisa-urus-izin-usaha Date: 03/08/19 02:02 PM
Link: https://regional.kompas.com/read/2019/03/05/14334201/pasca-serangan-kkb-600-prajurit-tni-dl
siagikan-jaga-keamanan-pembangunan-di Date: 03/05/19 02:33 PM
Link: https://regional.kompas.com/read/2019/02/27/05411471/edy-rahmayadi-apbd-sumut-2023-diprojek
si-rp-18-triliun-bukan-wacana Date: 02/27/19 05:41 AM
Link: https://regional.kompas.com/read/2019/02/26/16481951/percepat-pembangunan-di-pesisir-selata
n-pemprov-jatim-bangun-bandara-dan-tol Date: 02/26/19 04:48 PM
Link: https://ekonomi.kompas.com/read/2019/02/26/123531626/kemenkeu-genjot-inovasi-teknologi-dan-
kolaborasi-blu Date: 02/26/19 12:35 PM

```

**Gambar 3.10. Execute Python Linkkeyword.py**

```

pc-bigdata-1@pc-bigdata-1:~/Desktop/Tugas Rapi$ python testlinkkeywordth1.py
Jumlah Data Berita yang tersedia: 31723
Masukkan Jumlah Keyword yang ingin di cari : 2
Masukkan Keyword1 : proyek nasional
Masukkan Keyword2 : program nasional
1=>link: https://edukasi.kompas.com/read/2018/06/06/13144481/kesejahteraan-guru-tidak-berdampak-p
ada-kualitas-pendidikan
-->Date: 06/06/18 01:14 PM
==>Title: Kesejahteraan Guru Tidak Berdampak pada Kualitas Pendidikan - Kompas.com
2=>link: https://edukasi.kompas.com/read/2018/05/05/09030001/program-keahlian-ganda-jalan-untuk-s
ukseskan-revitalisasi-smk
-->Date: 05/05/18 09:03 AM
==>Title: Program Keahlian Ganda Jalan untuk Sukseskan Revitalisasi SMK
3=>link: https://ekonomi.kompas.com/read/2018/11/28/150000226/pgn-saka-dan-skk-migas-bangun-fasil
itas-atr-bersih-di-gresik
-->Date: 11/28/18 03:00 PM
==>Title: PGN SAKA dan SKK MIGAS Bangun Fasilitas Air Bersih di Gresik - Kompas.com
4=>link: https://ekonomi.kompas.com/read/2018/04/10/111300326/ini-cara-industri-ritel-ikuti-perke
mbangan-zaman
-->Date: 04/10/18 11:13 AM
==>Title: Ini Cara Industri Ritel Ikuti Perkembangan Zaman
5=>link: https://megapolitan.kompas.com/read/2017/10/21/21052181/mahesh-mengaku-sudah-lama-merela
kan-tanahnya-untuk-mrt-asal

```

**Gambar 3.11. Execute Linkkeywordth1.py**

Pada gambar 3.10 dan 3.11 mengenai eksekusi dapat terlihat perbedaan

dimana tujuan hasil dari kedua program tersebut memberikan hasil yang sama.

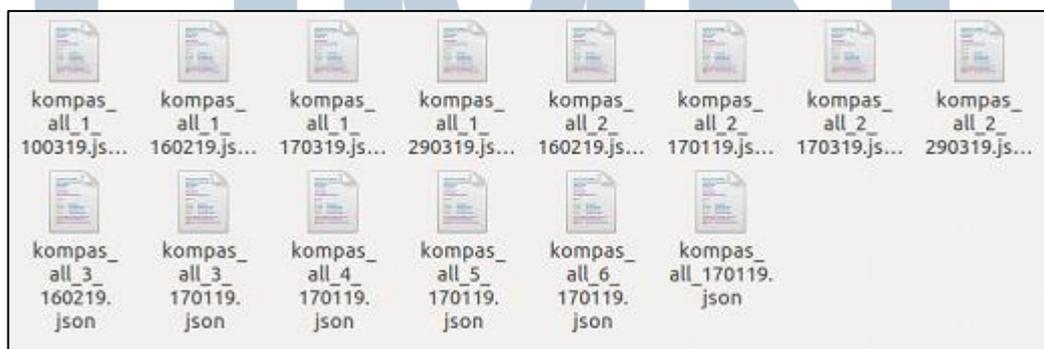
Perbedaan antar eksekusi program tersebut terlihat dari cara penggunaan melalui

*terminal* Linux dimana *script* yang telah dimodifikasi menampilkan hasil yang

lebih interaktif dan mudah digunakan serta mempercepat pencarian sumber data saat *fact checking* berlangsung.

### 3.2.5. Mendownload *File* Twitter

Mendownload *file* Twitter yang di berikan oleh Bapak Yustinus Eko untuk di analisa dari *server*, dengan menggunakan perintah `scp -r`, sebelum itu kita perlu melakukan login ke server dengan perintah `ssh` di *terminal* Linux, dengan ID penulis yaitu [vege@172.17.2.200](mailto:vege@172.17.2.200). Kemudian untuk melakukan duplikasi folder dari *server* ke lokal komputer dengan *command* `scp -r vege@172.17.2.200: /folder /local` direktori atau jika ingin melakukan duplikasi *file* tertentu di *server* dapat dengan menggunakan *command* contoh `scp vege@172.17.2.200:"letak folder"/"file.format(*.json, *.txt dan lain-lain) /"direktori lokal"/"file.format"` merupakan tujuan direktori lokal. Hasil yang diperoleh dari *server* yaitu *data* mengenai Twitter Kompas dalam bentuk *format* JSON dapat di lihat pada gambar di bawah ini.

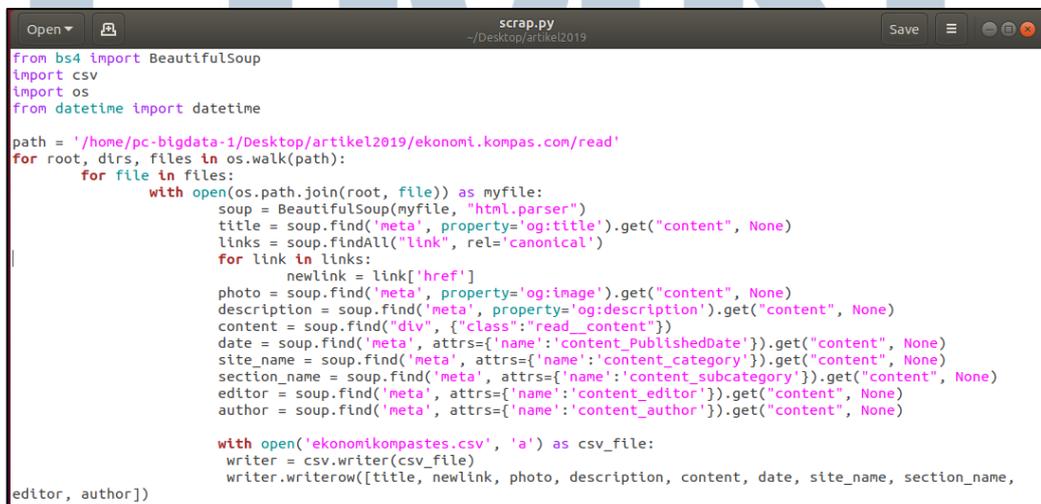


Gambar 3.2. *File* Kompas Twitter Dari Server

CENTRO  
MULTIMEDIA  
NUSANTARA

### 3.2.6. Pengembangan Script Python Scrapkompas.py

Dikarenakan *script* yang saat ini kurang efisien maka dilakukan pengembangan terhadap *script*. Perubahan yang dilakukan terhadap *script* pada saat dijalankan antara lain, menggunakan 1 *script* untuk melakukan *scraping* 7 kategori berita tanpa harus merubah isi koding sebagai contoh direktori pembacaan *file* dan penamaan *output file* dan jika *scraping* dilakukan berulang maka akan terjadi duplikasi isi yang sama. *Script* baru terdiri 2 *file output*, yaitu *output* pertama hasil scrap sudah dapat langsung menggabungkan seluruh data dengan 7 kategori menjadi 1 *file csv* secara otomatis tanpa harus dilakukan secara manual lagi, dan output kedua hasil scrap output per kategori artikel Kompas. Kemudian *script* baru sudah dapat menghapus isi data yang sama atau data redundansi akibat program scrap yang dijalankan lebih dari 1 kali. *Script* baru sudah terdapat jenis-jenis kategori yang ingin di *scraping* secara langsung di gunakan untuk memudahkan pengguna jika pengguna ingin melakukan pembaharuan artikel Kompas.



```
Open | scrap.py | Save |
~/Desktop/artikel2019

from bs4 import BeautifulSoup
import csv
import os
from datetime import datetime

path = '/home/pc-bigdata-1/Desktop/artikel2019/ekonomi.kompas.com/read'
for root, dirs, files in os.walk(path):
    for file in files:
        with open(os.path.join(root, file)) as myfile:
            soup = BeautifulSoup(myfile, "html.parser")
            title = soup.find('meta', property='og:title').get("content", None)
            links = soup.findAll("link", rel="canonical")
            for link in links:
                newlink = link['href']
            photo = soup.find('meta', property='og:image').get("content", None)
            description = soup.find('meta', property='og:description').get("content", None)
            content = soup.find("div", {"class": "read__content"})
            date = soup.find('meta', attrs={'name': 'content_PublishedDate'}).get("content", None)
            site_name = soup.find('meta', attrs={'name': 'content_category'}).get("content", None)
            section_name = soup.find('meta', attrs={'name': 'content_subcategory'}).get("content", None)
            editor = soup.find('meta', attrs={'name': 'content_editor'}).get("content", None)
            author = soup.find('meta', attrs={'name': 'content_author'}).get("content", None)

            with open('ekonomikompastes.csv', 'a') as csv_file:
                writer = csv.writer(csv_file)
                writer.writerow([title, newlink, photo, description, content, date, site_name, section_name,
                                editor, author])
```

Gambar 3.13. Script Scrap.py Sebelum Modifikasi

```

Open  testscrapkompas.py  Save
from bs4 import BeautifulSoup
import csv
import os
import pandas as pd
from datetime import datetime
import numpy as np
labelpilihan = []

print(" Label" + " " + "Angka\n")
print("1.edukasi" + " " + "====> 1")
print("2.ekonomi" + " " + "====> 2")
print("3.megapolitan" + " " + "====> 3")
print("4.nasional" + " " + "====> 4")
print("5.regional" + " " + "====> 5")
print("6.sains" + " " + "====> 6")
print("7.tekno" + " " + "====> 7")

labelpilihan = input("Masukkan Angka yang ingin di scraping : ") #dibuat menggunakan swtich case tiap labelpilihan

base_path = "/home/pc-bigdata-1/Desktop/artikel2019"
save_path = "/home/pc-bigdata-1/Desktop/testing"

read_kategori = {'1':'edukasi.kompas.com/read', '2':'ekonomi.kompas.com/read', '3':'megapolitan.kompas.com/read', '4':'nasional.kompas.com/read', '5':'regional.kompas.com/read', '6':'sains.kompas.com/read', '7':'tekno.kompas.com/read'}

save_kategori = {'1':'edukasikompashasil.csv', '2':'ekonomikompashasil.csv', '3':'megapolitankompashasil.csv', '4':'nasionalkompashasil.csv', '5':'regionalkompashasil.csv', '6':'sainskompashasil.csv', '7':'teknokompashasil.csv'}

if labelpilihan in read_kategori :
    with open(os.path.join(save_path, save_kategori[labelpilihan]), 'w') as csv_file:
        writer = csv.writer(csv_file)

writer.writerow(["Judul", "Link", "Tanggal_Publikasi", "Deskripsi", "Foto", "Konten", "Nama_Situs", "Kategori_Berita", "Editor", "Author"])
print("in progress")
path = os.path.join(base_path, read_kategori[labelpilihan]) #path edukasi.kompas.com di rubah
for root, dtrs, files in os.walk(path):
    for file in files:
        with open(os.path.join(root, file)) as myfile:
            soup = BeautifulSoup(myfile, "html.parser")
            title = soup.find('meta', property='og:title').get("content", None)
            links = soup.findAll("link", rel='canonical')
            for link in links:
                newlink = link['href']
            photo = soup.find('meta', property='og:image').get("content", None)
            description = soup.find('meta', property='og:description').get("content", None)
            content = soup.find("div", {"class":"read_content"})
            date = soup.find('meta', attrs={'name':'content_PublishedDate'}).get("content", None)
            site_name = soup.find('meta', property='og:site_name').get("content", None)
            content_category = soup.find('meta', attrs={'name':'content_category'}).get("content", None)
            editor = soup.find('meta', attrs={'name':'content_editor'}).get("content", None)

```

**Gambar 3.14. Setelah Modifikasi Scrap.py**

```

        author = soup.find('meta', attrs={'name':'content_author'}).get("content", None)
        content = content.text.strip().replace("\n", ' ')

kebutuhan
        with open(os.path.join(save_path, save_kategori[labelpilihan]), 'a') as csv_file:# ubah bagian sin tsesuai
            writer = csv.writer(csv_file)
            writer.writerow([title, newlink, date, photo, description, content, site_name, content_category,
editor, author])

        with open("gabunganhasil.csv", 'a+') as csv_fille:
            writer = csv.writer(csv_fille)
            writer.writerow([title, newlink, date, photo, description, content, site_name, content_category,
editor, author])
        print("scrap selesai" + save_kategori[labelpilihan])

#removeduplikat masing-masing file label csv
baca_csv = pd.read_csv(os.path.join(save_path, save_kategori[labelpilihan]))
isi = pd.DataFrame(baca_csv, columns=["Judul", "Link", "Tanggal_Publikasi", "Deskripsi", "Foto", "Konten", "Nama_Situs", "Kategori_Berita", "Editor", "Author"])
hapus_duplikat = isi.drop_duplicates(subset='Link', keep="first")
path_hasil = os.path.join(save_path, save_kategori[labelpilihan])
hasilcsv = hapus_duplikat.to_csv(path_hasil, index=False)
print("Menghapus duplikat selesai" + save_kategori[labelpilihan])

#removeduplikat gabungan file csv
baca_csv = pd.read_csv(os.path.join(save_path, 'gabunganhasil.csv'), header=None, names=['Judul', 'Link', 'Tanggal_Publikasi', 'Deskripsi', 'Foto', 'Konten', 'Nama_Situs', 'Kategori_Berita', 'Editor', 'Author'])
isi = pd.DataFrame(baca_csv, columns=["Judul", "Link", "Tanggal_Publikasi", "Deskripsi", "Foto", "Konten", "Nama_Situs", "Kategori_Berita", "Editor", "Author"])
hapus_duplikat = isi.drop_duplicates(subset='Link', keep="first")
path_hasil = os.path.join(save_path, 'gabunganhasil.csv')
hasilcsv = hapus_duplikat.to_csv(path_hasil, index=False)
print("Menghapus duplikat gabunganhasil selesai")

```

**Gambar 3.15 Setelah Modifikasi Scrap.py Lanjutan**

Hasil yang di dapatkan pada gambar di atas berupa data yang terstruktur dan mudah dibaca dalam bentuk CSV. Isi data CSV tersebut berisi judul berita, *link* berita, tanggal publikasi, deskripsi, *link* foto, isi berita, nama situs, kategori berita, nama *editor*, dan nama *author* seperti gambar dibawah ini.

	A	B	C	D	E	F	G	H	I	J	
	Judul	Link	Tanggal Publikasi	Deskripsi	Foto	Konten	Nama Situs	Kategori	Berita	Editor	Author
1	Mempertus Akses Pendidikan Tinggi, dir Bidkin	<a href="https://edukasi.kompas.com/read/2019/10/14/13.09.42">https://edukasi.kompas.com/read/2019/10/14/13.09.42</a>	10/14/13 09:42 AM	Program unggulan Direktur Oleh Dipko Santoso			KOMPAS	KOMPAS	edukasi	Latief	
2	Hebat... Film Animasi Karya Siswa SMK Diputar di	<a href="https://edukasi.kompas.com/read/2019/10/28/13.01.55">https://edukasi.kompas.com/read/2019/10/28/13.01.55</a>	10/28/13 01:55 PM	Diputar di bioskop sejak KJAKARTA, KOMPAS			KOMPAS	KOMPAS	edukasi	Latief	
3	Mendubai Jejakkan Kelelahan Pesanan Teknologi	<a href="https://edukasi.kompas.com/read/2019/03/20/18.06.29">https://edukasi.kompas.com/read/2019/03/20/18.06.29</a>	03/20/18 06:29 PM	Pada era digital saat ini suKOMPAS.com - Meri			KOMPAS	KOMPAS	edukasi	Dimas Wahyu	Mikhael Gevati
4	Siapkah Sekolah Membekali Siswa Menemukan	<a href="https://edukasi.kompas.com/read/2019/09/16/18.11.14">https://edukasi.kompas.com/read/2019/09/16/18.11.14</a>	09/16/18 11:14 AM	Pada era modern seperti KOMPAS.com - Digi			KOMPAS	KOMPAS	edukasi	Si Noviyanti	Mikhael Gevati
5	Corona, Pahalah Kecerdasan dan Karakter Anak	<a href="https://edukasi.kompas.com/read/2019/03/16/18.09.58">https://edukasi.kompas.com/read/2019/03/16/18.09.58</a>	03/16/18 09:58 PM	kecerdasan dan karakter KOMPAS.com - Kes			KOMPAS	KOMPAS	edukasi	Dimas Wahyu	Josephus Pitrus
6	Keska Anak Bertanya, Gadget Merawab	<a href="https://edukasi.kompas.com/read/2019/03/13/18.09.11">https://edukasi.kompas.com/read/2019/03/13/18.09.11</a>	03/13/18 09:11 AM	zaman sekarang KOMPAS.com - Sele			KOMPAS	KOMPAS	edukasi	Latief	
7	Siswa, Dendisi Dir Karakter Playdoh - Kompas	<a href="https://edukasi.kompas.com/read/2019/03/20/18.12.29">https://edukasi.kompas.com/read/2019/03/20/18.12.29</a>	03/20/18 12:29 PM	Wakil Ketua Sibero Cencu JAKARTA, KOMPAS			KOMPAS	KOMPAS	edukasi	Latief	
8	Teknologi, Pengasi Harfu Matematis	<a href="https://edukasi.kompas.com/read/2019/03/23/18.09.03">https://edukasi.kompas.com/read/2019/03/23/18.09.03</a>	03/23/18 09:03 AM	Siswa Indonesia bepeluaKOMPAS.com - Har			KOMPAS	KOMPAS	edukasi	Dimas Wahyu	Kurniasih Budi
9	Apa Jandrya Indonesia Tanpa Wirusaka?	<a href="https://edukasi.kompas.com/read/2019/03/23/18.08.45">https://edukasi.kompas.com/read/2019/03/23/18.08.45</a>	03/23/18 08:45 AM	Mau usaha? Tidak perlu taKOMPAS.com - Mar			KOMPAS	KOMPAS	edukasi	Dimas Wahyu	Josephus Pitrus
10	Coba Dili, "Elek Sampun" Gadget Untuk Anak	<a href="https://edukasi.kompas.com/read/2019/03/14/18.12.55">https://edukasi.kompas.com/read/2019/03/14/18.12.55</a>	03/14/18 12:55 PM	Penggunaan gadget oleh KOMPAS.com - Jac			KOMPAS	KOMPAS	edukasi	Si Noviyanti	Mikhael Gevati
11	Mengembangkan "Soft Skills" Anak-anak Generasi Z	<a href="https://edukasi.kompas.com/read/2019/03/19/18.02.24">https://edukasi.kompas.com/read/2019/03/19/18.02.24</a>	03/19/18 02:24 PM	Cara berit melalifika KOMPAS.com - Tit			KOMPAS	KOMPAS	edukasi	Latief	
12	Kemampuan Matematis Siswa Indonesia Memorib	<a href="https://edukasi.kompas.com/read/2019/03/21/18.09.21">https://edukasi.kompas.com/read/2019/03/21/18.09.21</a>	03/21/18 09:21 AM	Hasil Survei Programme foKOMPAS.com - Kep			KOMPAS	KOMPAS	edukasi	Latief	
13	Mahasiswa Politeknik Ciptakan Alat Pengupas Serap	<a href="https://edukasi.kompas.com/read/2019/01/20/18.09.03">https://edukasi.kompas.com/read/2019/01/20/18.09.03</a>	01/20/18 09:03 AM	Para buruh pengupas serapKOMPAS.com - Ser			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Auzi Amalia Domesti
14	Dongkrak Mani Tenaga Kerja, Pahlawan Vokasi Bur	<a href="https://edukasi.kompas.com/read/2019/01/26/18.07.43">https://edukasi.kompas.com/read/2019/01/26/18.07.43</a>	01/26/18 07:43 AM	Belalain, dan abadi perip JAKARTA, KOMPAS			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Kurniasih Budi
15	Siap siap... Ada Pendaftaran Beasiswa ke Belanda	<a href="https://edukasi.kompas.com/read/2019/01/23/18.07.27">https://edukasi.kompas.com/read/2019/01/23/18.07.27</a>	01/23/18 07:27 AM	Beberapa pendafarlan beaKOMPAS.com - Aw			KOMPAS	KOMPAS	edukasi	Latief	
16	Lewat Bahasa Inggris, Mahasiswa Politeknik Kemb	<a href="https://edukasi.kompas.com/read/2019/01/12/18.09.04">https://edukasi.kompas.com/read/2019/01/12/18.09.04</a>	01/12/18 09:04 AM	Mengembangkan wisata sKOMPAS.com - Kam			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Auzi Amalia Domesti
17	Lewat Beasiswa, Siasata Dukung Penerimaan Penr	<a href="https://edukasi.kompas.com/read/2019/01/25/18.09.10">https://edukasi.kompas.com/read/2019/01/25/18.09.10</a>	01/25/18 09:10 AM	Perlu adanya atongan royoK JAKARTA, KOMPAS			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Kurniasih Budi
18	Tahun Ajaran Baru, Sekolah Waqa Terapkan Kuriku	<a href="https://edukasi.kompas.com/read/2019/06/02/18.11.47">https://edukasi.kompas.com/read/2019/06/02/18.11.47</a>	06/02/18 11:47 PM	78.000 sekolah memesikaKOMPAS.com - Kem			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Kurniasih Budi
19	Penerapan Kurikulum 2013, Sekolah Mesti Dukung	<a href="https://edukasi.kompas.com/read/2019/06/30/18.10.32">https://edukasi.kompas.com/read/2019/06/30/18.10.32</a>	06/30/18 10:32 PM	Tahun ini semua sekolah hKOMPAS.com - Mer			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Kurniasih Budi
20	Wawasan XII UMN, Menyajikan Irisan Unggul	<a href="https://edukasi.kompas.com/read/2019/06/30/18.06.39">https://edukasi.kompas.com/read/2019/06/30/18.06.39</a>	06/30/18 06:39 PM	Kurikulum UMN sudah dirKOMPAS.com - Unr			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
21	Mendukung Kurikulum yang Sesungguhnya adalah	<a href="https://edukasi.kompas.com/read/2019/06/20/18.11.25">https://edukasi.kompas.com/read/2019/06/20/18.11.25</a>	06/20/18 11:25 PM	Kepala sekolah mesti merKOMPAS.com - Sos			KOMPAS	KOMPAS	edukasi	Kurniasih Budi	Kurniasih Budi
22	10 Universitas Terbaik ASEAN 2019, Di Mana Posi	<a href="https://edukasi.kompas.com/read/2019/06/22/18.03.16">https://edukasi.kompas.com/read/2019/06/22/18.03.16</a>	06/22/18 03:16 PM	ASEAN menempatkan 42 KOMPAS.com - QS			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
23	Jurusan dan Pekerjaan Paling Dibutuhkan 2020	<a href="https://edukasi.kompas.com/read/2019/06/20/18.03.46">https://edukasi.kompas.com/read/2019/06/20/18.03.46</a>	06/20/18 03:46 PM	Mengajukan, berdasarKOMPAS.com - Dis			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
24	Jelang Tahun Ajaran Baru, Ini 5 Tipe Meribak TK	<a href="https://edukasi.kompas.com/read/2019/06/20/18.06.56">https://edukasi.kompas.com/read/2019/06/20/18.06.56</a>	06/20/18 06:56 PM	Ada banyak peribanganKOMPAS.com - Lib			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
25	Beasiswa Bukalapak Segera Dituan. In Syarat dan	<a href="https://edukasi.kompas.com/read/2019/06/06/18.04.31">https://edukasi.kompas.com/read/2019/06/06/18.04.31</a>	06/06/18 04:31 PM	Bukalapak memberikan beKOMPAS.com - Buk			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
26	Mahasiswa dari 25 Universitas Ini Purwa Kesempaa	<a href="https://edukasi.kompas.com/read/2019/06/06/18.10.31">https://edukasi.kompas.com/read/2019/06/06/18.10.31</a>	06/06/18 10:31 AM	Sebanyak 25 universitas KOMPAS.com - JAK			KOMPAS	KOMPAS	edukasi	Inggrid Dwi Wedi Mela Amora	
27	Kesepakatan Guru Tidak Berdamak pada Kualita	<a href="https://edukasi.kompas.com/read/2019/06/06/18.01.14">https://edukasi.kompas.com/read/2019/06/06/18.01.14</a>	06/06/18 01:14 PM	Tribun ruah yang dibelaKOMPAS.com - Luh			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
28	10 SMP Kota Tembung Ringing Nila UN IPA Tere	<a href="https://edukasi.kompas.com/read/2019/06/16/18.03.21">https://edukasi.kompas.com/read/2019/06/16/18.03.21</a>	06/16/18 03:21 PM	SMP Negeri dan Swasta KOMPAS.com - Har			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
29	Masalah Sosial pada Anak Konepagan Gaywa	<a href="https://edukasi.kompas.com/read/2019/06/16/18.03.00">https://edukasi.kompas.com/read/2019/06/16/18.03.00</a>	06/16/18 03:00 PM	Anak rebel lebih mudah KOMPAS.com - Pa			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
30	Hari Ini Harapas. In 3 Kurikulum ukul Keluaran	<a href="https://edukasi.kompas.com/read/2019/06/29/18.11.43">https://edukasi.kompas.com/read/2019/06/29/18.11.43</a>	06/29/18 11:43 AM	Harapas menjadi momeKOMPAS.com - Tar			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Yohanes Enggar Hari
31	Ibu Guru Nur Fitriah Bisa Membangun Jembatan	<a href="https://edukasi.kompas.com/read/2019/06/29/18.08.58">https://edukasi.kompas.com/read/2019/06/29/18.08.58</a>	06/29/18 08:58 PM	Murid roud saya, sepeleKOMPAS.com - JAK			KOMPAS	KOMPAS	edukasi	Yohanes Enggar	Josephus Pitrus

Gambar 3.16. Hasil Struktur Scraping Artikel

### 3.2.7. Fact Checking Debat Capres Ke- 5

Penulis akan melanjutkan cara Kerja *fact checking* ini, seperti *fact checking* sebelumnya, tim Ilkom akan mengirimkan 2 *link* online document yang berisi *claim keyword* dan *link* berita yang diberikan oleh Mario melalui Whatsapp. Untuk *link claim keyword* dan *link* berita hasil pencarian dapat diakses pada *link* dibawah ini:

- A. [https://docs.google.com/spreadsheets/d/1ZM\\_XWYNM5ccrVH\\_hoGtXOaa--zxQRKDaAxtIJHa8ag/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1ZM_XWYNM5ccrVH_hoGtXOaa--zxQRKDaAxtIJHa8ag/edit?usp=sharing)
- B. <https://docs.google.com/spreadsheets/d/1iEucHgxDmNHwsn0jL99F-gV9-f0fthugxlqzIHOakCI/edit?usp=sharing>

Penulis akan memberikan satu contoh dengan menggunakan 1 *keyword* dari isi google drive mengenai cara kerja *fact checking* yang dilakukan oleh tim SI.

SEGMENT 1	SEGMENT 2
infrastruktur pemerintah tidak jwasentris	Komoditas indonesia bergantung pada pasar luar negeri
4 tahun wujudkan pemerataan pembangunan	Indonesia ekspor mentahan karet, kelapa sawit, komoditas perikanan, komoditas pertanian
	Anak-anak muda sudah membangun ekosistem online
	Pemerintah keluarkan kredit ultra mikro
	Pemerintah keluarkan bank wakaf mikro untuk pesantren
Blok Mahakam, Blok Rokan, Freeport kembali yang dikuasai negara	Nasabah Bank Wakaf 90 persen perempuan
	program mekaar jangkau 4,2 juta nasabah
	1,5 tahun 1,1 juta nasabah Ultra Mikro
	Mekaar meminjamkan 2 sampai 10 juta untuk pengusaha perempuan
	Peminjaman di atas 10 juta dialirkan ke Kredit Usaha Rakyat

Gambar 3.17. *Keyword* Debat Capres Ke-5

KLAIM	LINK BERITA
4 tahun wujudkan pemerataan pembangunan	<a href="https://nasional.kompas.com/read/2018/10/20/14144381/4-tahun-jokowi-jk-dan-cat">https://nasional.kompas.com/read/2018/10/20/14144381/4-tahun-jokowi-jk-dan-cat</a>
infrastruktur pemerintah tidak jwasentris	<a href="https://nasional.kompas.com/read/2019/04/01/05460961/andi-widjajanto-jokowi-leb">https://nasional.kompas.com/read/2019/04/01/05460961/andi-widjajanto-jokowi-leb</a>
	<a href="https://edukasi.kompas.com/read/2019/04/08/16285731/indonesia-startup-summit-">https://edukasi.kompas.com/read/2019/04/08/16285731/indonesia-startup-summit-</a>
	<a href="https://ekonomi.kompas.com/read/2018/08/01/225800626/ini-rencana-pertamina-sy">https://ekonomi.kompas.com/read/2018/08/01/225800626/ini-rencana-pertamina-sy</a>
Blok Mahakam, Blok Rokan, Freeport kembali yang dikuasai negara	<a href="https://ekonomi.kompas.com/read/2018/08/01/212800726/pertamina-siap-tampung">https://ekonomi.kompas.com/read/2018/08/01/212800726/pertamina-siap-tampung</a>
	<a href="https://ekonomi.kompas.com/read/2018/08/01/202400926/jonan-blok-rokan-diberi">https://ekonomi.kompas.com/read/2018/08/01/202400926/jonan-blok-rokan-diberi</a>
	<a href="https://ekonomi.kompas.com/read/2018/08/02/100937526/pengamat-finansial-pert-">https://ekonomi.kompas.com/read/2018/08/02/100937526/pengamat-finansial-pert-</a>
	<a href="https://ekonomi.kompas.com/read/2018/08/02/063000826/penyerahan-blok-rokan-">https://ekonomi.kompas.com/read/2018/08/02/063000826/penyerahan-blok-rokan-</a>
	<a href="https://ekonomi.kompas.com/read/2018/08/15/100200726/pengambilalihan-blok-ro">https://ekonomi.kompas.com/read/2018/08/15/100200726/pengambilalihan-blok-ro</a>
	<a href="https://money.kompas.com/read/2019/03/04/414700026/dibuat-tak-ada-pengurang">https://money.kompas.com/read/2019/03/04/414700026/dibuat-tak-ada-pengurang</a>
	<a href="https://ekonomi.kompas.com/read/2019/02/18/134903226/cek-fakta-jokowi-sebut-">https://ekonomi.kompas.com/read/2019/02/18/134903226/cek-fakta-jokowi-sebut-</a>
	<a href="https://ekonomi.kompas.com/read/2019/02/25/124100326/awal-2019-lifting-migas-">https://ekonomi.kompas.com/read/2019/02/25/124100326/awal-2019-lifting-migas-</a>
	<a href="https://nasional.kompas.com/read/2019/02/11/16083411/lmses-debat-kedua-jokow">https://nasional.kompas.com/read/2019/02/11/16083411/lmses-debat-kedua-jokow</a>
	<a href="https://nasional.kompas.com/read/2019/02/11/07380391/saat-jokowi-luapkan-keke-">https://nasional.kompas.com/read/2019/02/11/07380391/saat-jokowi-luapkan-keke-</a>
	<a href="https://nasional.kompas.com/read/2019/04/10/14511041/jokowi-saya-ditakut-takuti-">https://nasional.kompas.com/read/2019/04/10/14511041/jokowi-saya-ditakut-takuti-</a>
	<a href="https://ekonomi.kompas.com/read/2019/02/18/134903226/cek-fakta-jokowi-sebut-a">https://ekonomi.kompas.com/read/2019/02/18/134903226/cek-fakta-jokowi-sebut-a</a>
	<a href="https://nasional.kompas.com/read/2018/08/08/11481171/cerita-jokowi-ambil-alih-bl-">https://nasional.kompas.com/read/2018/08/08/11481171/cerita-jokowi-ambil-alih-bl-</a>
	<a href="https://nasional.komooas.com/read/2019/02/02/21330021/iokowi-minta-lulusan-pert-">https://nasional.komooas.com/read/2019/02/02/21330021/iokowi-minta-lulusan-pert-</a>

Gambar 3.18. Hasil *Link* Berita Debat Capres Ke-5

N U S A N T A R A

```
pc-bigdata-1@pc-bigdata-1:~/Desktop/Tugas Rapi$ python testlinkkeywordth1.py
Jumlah Data Berita yang tersedia: 31723
Masukkan Jumlah Keyword yang ingin di cari : 1
Masukkan Keyword1 : blok mahakam
1=>link: https://ekonomi.kompas.com/read/2018/03/26/120600026/dalam-tiga-bulan-h
arga-pertalite-naik-dua-kali
-->Date: 03/26/18 12:06 PM
==>Title: Dalam Tiga Bulan, Harga Pertalite Naik Dua Kali
2=>link: https://ekonomi.kompas.com/read/2018/03/26/070700226/esdm--jangan-ada-p
ernyataan-bahwa-pemerintah-membiarkan-pertamina-rugi
-->Date: 03/26/18 07:07 AM
==>Title: ESDM: Jangan Ada Pernyataan bahwa Pemerintah Membiarkan Pertamina R
ugi
3=>link: https://ekonomi.kompas.com/read/2018/08/01/212800726/pertamina-siap-tam
pung-karyawan-chevron-saat-kelola-blok-rokan
-->Date: 08/01/18 09:28 PM
==>Title: Pertamina Siap Tampung Karyawan Chevron Saat Kelola Blok Rokan - Ko
mpas.com
4=>link: https://ekonomi.kompas.com/read/2019/02/18/134903226/cek-fakta-jokowi-s
ebut-ambil-alih-blok-mahakam-rokan-kuasai-51-persen-saham
-->Date: 02/18/19 01:49 PM
==>Title: CEK FAKTA: Jokowi Sebut Ambil Alih Blok Mahakam-Rokan, Kuasai 51 Pe
rsen Saham Freeport - Kompas.com
```

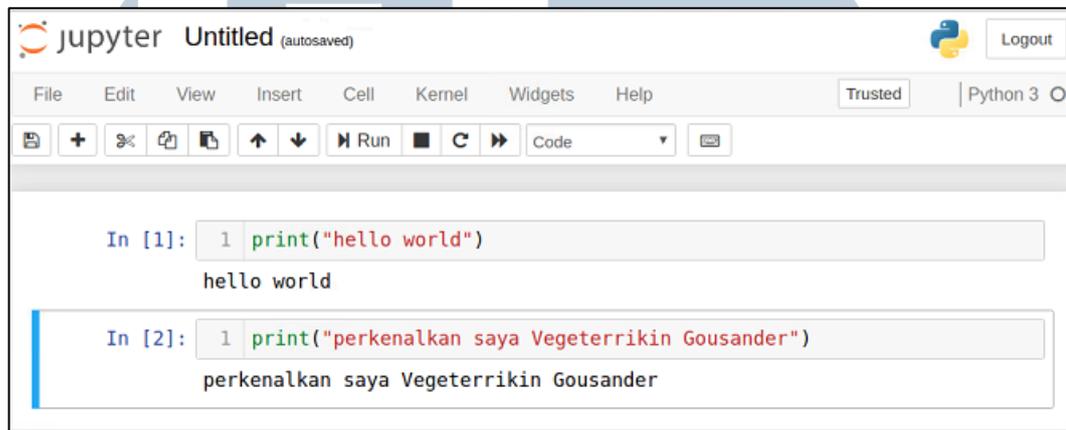
**Gambar 3.19. Pencarian *Link* Dengan *Keyword***

Dari hasil yang ditemukan oleh *script* tersebut tidak semuanya relevan sesuai isi berita yang ingin di cek kebenarannya. Oleh maka itu perlu dilakukan pemilihan lagi untuk di masukkan ke *online document* untuk di cek oleh tim Ilkom. Untuk melakukan hal tersebut tidak cukup hanya 1 *keyword* karena *keyword* yang disusun oleh tim Ilkom masih kurang (spesifik atau umum) sehingga berita yang dicari terkadang tidak menemukan hasil atau memang berita tersebut tidak ada di *publish* oleh Kompas.

### **3.2.8. Data Preparation Twitter Kompas**

Penulis menggunakan *Jupyter Notebook* pada proses ini, agar setiap penulisan koding lebih rapi dan dapat dijalankan tiap *block code* untuk memudahkan pencarian jika terjadi *error*, serta ada beberapa petunjuk cara pengguna setiap fungsinya. *Jupyter Notebook* merupakan aplikasi web yang disediakan *open-source* yang berguna untuk membuat dan membagikan

dokumen koding secara *live*, perhitungan statistik, membuat visualisasi data, dan penulisan koding teks yang bersifat naratif.



**Gambar 3.20. Jupyter Notebook**

Penyajian data twitter kompas bertujuan untuk memudahkan penulis untuk menganalisa pola dan relasi data pada saat tahap *modelling data* pada tahap berikutnya. Proses ini mengambil *data-data* penting dari bentuk *file* JSON yang berisi berbagai macam informasi seperti isi Tweet, nama pengguna dan lain-lain, agar data tersebut lebih mudah dibaca diperlukan *script* untuk mengelola data tersebut. *Script* yang digunakan seperti di bawah ini:

Hal pertama yang dilakukan ialah *import* modul yang akan digunakan dan membuat *variable* baru berupa *list* kosong untuk menampung *data* seperti pada gambar di bawah ini.

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

```

1  #modul yang dipakai
2  import json
3  import pandas as pd
4
5  #membuat variabel untuk menampung data ke dalam list
6  tanggal_tweet = []
7  isi_tweet = []
8  hashtags = []
9  sumber_twitter = []
10 bahasa_twitter = []
11 nama_pengguna = []
12 lokasi_pengguna = []
13 jumlah_favorit = []
14 jumlah_status = []
15 jumlah_teman = []
16 jumlah_pengikut = []
17 pembuatan_userID = []

```

**Gambar 3.21. Import Modul dan List Kosong**

```

1  #membaca data json sebagai data_tweets
2  with open('../testing/kompasjson/kompas_all_170119.json') as data_tweets:
3      #membaca data per baris dengan looping menggunakan fungsi readlines
4      for row in data_tweets.readlines():
5          #try and except untuk menhandle error
6          try:
7              #variabel baru untuk menyimpan isi json
8              row_json = json.loads(row)
9              #mengambil isi data created_at tweet
10             created_at = row_json.get('created_at')
11             tanggal_tweet.append(created_at)
12             #mengambil isi data teks tweet
13             extended_tweet = row_json.get('extended_tweet', {})
14             isi = extended_tweet.get('full_text', '').lower()
15             isi_tweet.append(isi)
16             #mengambil isi data hashtags tweet
17             tags = extended_tweet.get('entities', {}).get('hashtags', '')
18             tags_list = []
19             for isi_tags in tags:
20                 tags_list.append(isi_tags.get('text'))
21             hashtags.append(','.join(tags_list))
22             #mengambil isi data sumber perangkat twitter
23             source = row_json.get('source')
24             sumber_twitter.append(source)
25             #mengambil isi data bahasa tweet
26             lang = row_json.get('lang')
27             bahasa_twitter.append(lang)

```

**Gambar 3.22. Pengambilan Data Twitter (1)**

M U L T I M E D I A  
N U S A N T A R A

```

28     #mengakses isi user
29     user = row_json.get('user', {})
30     #mengambil isi data nama pengguna
31     nama = user.get('screen_name')
32     nama_pengguna.append(nama)
33     #mengambil isi data lokasi pengguna
34     lokasi = user.get('location')
35     lokasi_pengguna.append(lokasi)
36     #mengambil isi data jumlah favorit tweet
37     fav = row_json.get('favorite_count')
38     jumlah_favorit.append(fav)
39     #mengambil isi data jumlah status yang dibuat oleh pengguna
40     status = user.get('statuses_count')
41     jumlah_status.append(status)
42     #mengambil isi data jumlah pertemana pengguna
43     teman = user.get('friends_count')
44     jumlah_teman.append(teman)
45     #mengambil isi data jumlah pengikut pengguna
46     foll = user.get('followers_count')
47     jumlah_pengikut.append(foll)
48     #mengambil isi data waktu pembuatan akun si user
49     user_created = user.get('created_at')
50     pembuatan_userID.append(user_created)
51
52     except:
53         #menampilkan isi data yang bermasalah
54         print('row')

```

**Gambar 3.23. Pengambilan Data Twitter (2)**

Langkah berikutnya membaca isi *file* JSON dengan modul fungsi yang telah disediakan, untuk mengambil isi-isi dalam JSON dengan menggunakan “*get*”. Hasil dari setiap fungsi “*get*” akan di tampung ke dalam *list* yang telah dibuat.

```

1  #mengubah stuktur data dari list menjadi pandas series
2  tanggal_tweets_series = pd.Series(tanggal_tweet, name='Tanggal_Tweet')
3  full_text_series = pd.Series(isi_tweet, name='Isi_Tweet')
4  hashtags_series = pd.Series(hashtags, name='Hashtags')
5  source_twitter_series = pd.Series(sumber_twitter, name='Sumber_Twitter')
6  lang_twitter_series = pd.Series(bahasa_twitter, name='Bahasa_Twitter')
7  screen_name_series = pd.Series(nama_pengguna, name='Nama_Pengguna')
8  locations_series = pd.Series(lokasi_pengguna, name='Lokasi_Pengguna')
9  favourites_count_series = pd.Series(jumlah_favorit, name='Jumlah_Favorit')
10 statuses_count_series = pd.Series(jumlah_status, name='Jumlah_Status')
11 friends_count_series = pd.Series(jumlah_teman, name='Jumlah_Teman')
12 follower_count_series = pd.Series(jumlah_pengikut, name='Jumlah_Pengikut')
13 create_at_userid_series = pd.Series(pembuatan_userID, name='Pembuatan_UserID')

```

**Gambar 3.24. Format Pandas Series**

Kemudian setelah semua data berhasil di ambil, berikutnya mengubah *data list* tersebut menjadi bentuk *pandas series* untuk menggunakan fungsi modul yang tersedia.

```

1 #membuat dataframe kosong
2 df = pd.DataFrame()
3 #menggabungkan seluruh data pandas series menjadi 1 dataframe
4 df = pd.concat([df, tanggal_tweets_series, full_text_series, hashtags_series,
5               source_twitter_series, lang_twitter_series, screen_name_series,
6               locations_series, favourites_count_series, statuses_count_series,
7               friends_count_series, follower_count_series, create_at_userid_series],
8               axis=1)
9 #mengeksport dataframe ke excel
10 df.to_excel('./testing/convertcsvuser/second/kompas_all_170119_test.xlsx',
11             index=False, na_rep="kosong")

```

**Gambar 3.25. Data Concat dan Output Excel**

Setelah langkah mengubah *format* berhasil, barulah data tersebut digabungkan dengan fungsi *concat* yang telah di sediakan *pandas series* ini. Lalu data-data tersebut dengan di ekspor ke bentuk *format* *xlsx*.

	Tanggal_Tweet	Isi_Tweet	Hashtags	Sumber_Twitter
0	Thu Jan 17 12:17:23 +0000 2019	visi misi yang sadar perihal ham #01Indonesiamaju	01Indonesiamaju,JokowiPeduliRakyat,jkwcerdas,j...	<a href="http://twitter.com" rel="nofollow">Tw...
1	Thu Jan 17 12:17:24 +0000 2019	jokowima"ruf enggak mau membongceng program mel...	SalamSatuJempol,01JokowiLagi,01Indonesiamaju,2...	<a href="https://chrome.google.com/extensions/...
2	Thu Jan 17 12:17:25 +0000 2019			<a href="https://mobile.twitter.com" rel="nofo...
3	Thu Jan 17 12:17:25 +0000 2019	sebenar lagi guys menuju #01Indonesiamaju sel...	01IndonesiaMaju,2019JokowiPresidenRI,Debat01KM...	<a href="http://twitter.com/download/android" ...
4	Thu Jan 17 12:17:27 +0000 2019			<a href="http://twitter.com/download/android" ...
	Thu Jan 17			<a href="http://twitter.com/download/android" ...

**Gambar 3.26. Menampilkan Hasil Gabungan Data**

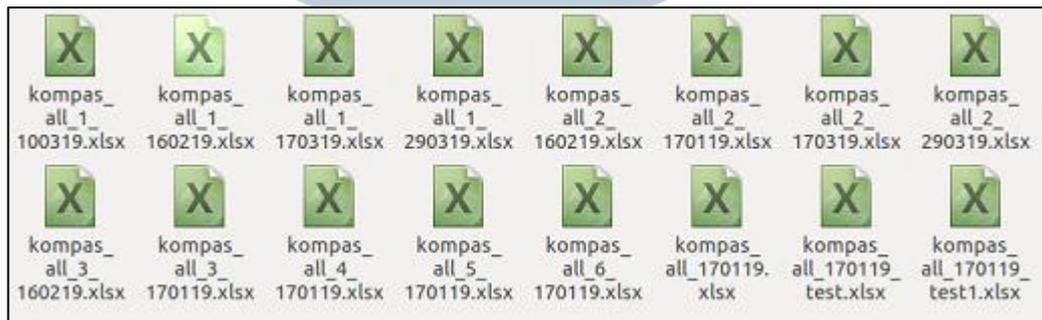
Jika ingin melihat isi struktur data yang telah berhasil digabungkan, dapat diketikkan fungsi *df* dan akan muncul seperti gambar 3.26 dan gambar 3.27.

In [5]: 1 df

ttter	Bahasa_Twitter	Nama_Pengguna	Lokasi_Pengguna	Jumlah_Favorit	Jumlah_Status	Jumlah_Teman	Jumlah_Pengikut	Pembuatan_UserID
dm"	in	KasepPurna	None	0	1013	117	74	Mon Oct 22 01:39:53 +0000 2018
<a	in	alfariziiziii	Bali, Indonesia	0	32	29	19	Tue Jun 26 09:41:36 +0000 2018
fo...	in	Irma_prana	Jawa timur, Indonesia	0	266	312	119	Mon Jan 07 16:00:53 +0000 2019
oid"	in	raisamarwan	Bandung, Indonesia	0	289	74	17	Mon Jan 07 11:42:58 +0000 2019
oid"	und	Jax76429941	None	0	2506	159	5	Wed Jul 11 01:29:37 +0000 2018
aird"								Sun Mar 16

**Gambar 3.27. Menampilkan Hasil Gabungan Data Lanjutan**

Hasil yang sudah berhasil dilakukan persiapan data dengan menggunakan *script* yang mengambil data agar lebih terstruktur dari bentuk JSON menjadi hasil bentuk *xlsx*.



**Gambar 3.28. Hasil Penyajian Data Twitter**

### 3.2.9. Data Modelling Twitter Kompas

#### A. Time Zone

Mengubah waktu *tweet* yang diperoleh dari JSON Twitter yang memiliki format *Time zone* UTC(*Coordinated Universal Time*)/GMT +0 menjadi waktu Indonesia dengan format *Time zone* GMT +7. Waktu

yang diubah ini berguna untuk mengetahui waktu posting yang sebenarnya dengan waktu yang di posting pengguna pada akun Twitter.

```
1 #modul yang dipakai
2 import pandas as pd
3 import pytz
4 from dateutil import parser
5 from datetime import time
6 import os
7 #mengset timezone gmt +7
8 local_tz = pytz.timezone('Asia/Jakarta')
```

**Gambar 3.29. Import Library dan Set Timezone**

Memasukkan modul yang digunakan untuk melakukan konversi waktu seperti pada gambar 3.29.

```
1 #membuat dataframe kosong
2 df = pd.DataFrame()
3 #mengset jalur letak direktori file
4 base_path= '/home/pc-bigdata-1/Desktop/testing/convertcsvuser/second/'
5 #membaca jalur direktori
6 for file in os.listdir(base_path):
7     #membuka data yang penamaan file mengandung kata test dan berakhir format .xlsx/excel
8     if file.endswith('test.xlsx'):
9         #melakukan print ke display bahwa file sedang dibaca
10        print('loading:', file)
11        #men append isi data 1 atau lebih
12        df = df.append(pd.read_excel(os.path.join(base_path, file)), ignore_index=True)
13 #mengprint notif apabila seluruh data berhasil dibaca
14 print("Done")
```

loading: kompas\_all\_170119\_test.xlsx  
Done

**Gambar 3.30. Membaca File XLSX**

Setelah modul berhasil di import, langkah berikutnya membaca *file* yang akan di kelola seperti pada gambar 3.30. Data yang digunakan dari hasil *script* persiapan *data* twitter.

```

1 #mengubah isi waktu tweet dari UTC ke GMT +7
2 try:
3     df['Tanggal_Tweet_Lokal'] = df['Tanggal_Tweet'].apply(
4         lambda x: parser.parse(x).astimezone(local_tz)).astype(str)
5 except:
6     pass
7 #mengubah isi waktu pembuatan akun dari UTC ke GMT +7
8 try:
9     df['Pembuatan_userID_Lokal'] = df['Pembuatan_UserID'].apply(
10        lambda x: parser.parse(x).astimezone(local_tz)).astype(str)
11 except:
12    pass

```

**Gambar 3.31. Waktu UTC Menjadi GMT +7**

Pada gambar 3.31, merupakan tahap mengubah *format* waktu terhadap tanggal tweet dan tanggal pembuatan *user ID*. Kemudian dilakukan ekspor data ke excel seperti gambar 3.32.

```

1 #mengeksport dataframe ke excel
2 df.to_excel(
3     '../testing/convertcsvuser/second/convertexceltolokaltimezone/kompas_all_170119_test.xlsx', index=None)

```

**Gambar 3.32. Ekspor File XLSX Timezone**

In [5]: 1 df

pengguna	Jumlah_Favorit	Jumlah_Status	Jumlah_Teman	Jumlah_Pengikut	Pembuatan_UserID	Tanggal_Tweet_Lokal	Pembuatan_userID_Lokal
kosong	0	1013	117	74	Mon Oct 22 01:39:53 +0000 2018	2019-01-17 19:17:23+07:00	2018-10-22 08:39:53+07:00
Indonesia	0	32	29	19	Tue Jun 26 09:41:36 +0000 2018	2019-01-17 19:17:24+07:00	2018-06-26 16:41:36+07:00
awa timur, Indonesia	0	266	312	119	Mon Jan 07 16:00:53 +0000 2019	2019-01-17 19:17:25+07:00	2019-01-07 23:00:53+07:00
Bandung, Indonesia	0	289	74	17	Mon Jan 07 11:42:58 +0000 2019	2019-01-17 19:17:25+07:00	2019-01-07 18:42:58+07:00
kosong	0	2506	159	5	Wed Jul 11 01:29:37 +0000 2018	2019-01-17 19:17:27+07:00	2018-07-11 08:29:37+07:00
amonnan					Sun Mar 16	2019-01-17	2014-03-17

**Gambar 3.33. Menampilkan Struktur Tabel Timezone**

Pada gambar 3.33, jika pengguna ingin melihat isi dari hasil perubahan yang terjadi setelah dilakukan konversi *format* waktu dapat menjalankan perintah `df` (dataframe).

## B. Twittbot.net

Pada tahap ini penulis ingin melihat siapa saja pengguna twittbot.net yang melakukan spam pada tweet mengenai debat capres 2019.

```
1 #modul yang dipakai
2 import os
3 import pandas as pd
4 from bs4 import BeautifulSoup
5 from collections import Counter, OrderedDict

1 #membuat dataframe kosong
2 df = pd.DataFrame()
3 #mengset jalur letak direktori file
4 base_path= '/home/pc-bigdata-1/Desktop/testing/convertcsvuser/second/convertexceltolokaltimezone/'
5 #membaca jalur direktori
6 for file in os.listdir(base_path):
7     #membuka data yang penamaan file mengandung kata test dan berakhir format .xlsx/excel
8     if file.endswith('test.xlsx'):
9         #melakukan print ke display bahwa file sedang dibaca
10        print('loading:', file)
11        #men append isi data 1 atau lebih
12        df = df.append(pd.read_excel(os.path.join(base_path, file)), ignore_index=True)
13 #mengprint notif apabila seluruh data berhasil dibaca
14 print("Done")
```

### Gambar 3.34. Modul Dan Membaca Data Analisa Twittbot.net

Pada gambar 3.34. merupakan langkah pertama melakukan *import* modul dan mulai membaca data yang akan digunakan untuk mengelola informasi ditahap berikutnya.

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

```

1 #membuat variabel list baru untuk menampung data sumber twitter
2 sumber_twitter_series = []
3 #membaca dataframe dengan looping
4 for row in df.iterrows():
5     #membuat variabel baru untuk menampung data sumber twitter yang memiliki format html dengan modul beautifulso
6     soup = BeautifulSoup(row[1]['Sumber_Twitter'], 'html.parser')
7     #kondisi mencari seluruh isi jika datanya ada dan =True
8     if soup.find('a'):
9         #maka mengambil isi teks sumber twitter
10        sumber_twitter_series.append(soup.find('a').get_text())
11    else:
12        #menambakan isi ke dalam Sumber_Twitter bahwa struktur data tidak terstruktur
13        sumber_twitter_series.append(row[1]['Sumber_Twitter'])
14        print('unrecognized:', row[1]['Sumber_Twitter'])

```

**Gambar 3.35. BeautifulSoup Sumber Twitter**

Gambar 3.35 merupakan *script* untuk mengambil isi teks tertentu dengan menggunakan *BeautifulSoup*. Kemudian, gambar 3.36 merupakan *script* untuk menampilkan isi tweet yang di post dan jumlah *like* yang disukai oleh pengguna tweet, hasil *filter* tersebut berdasarkan nama pengguna, pada contoh tersebut untuk *SilviaSandras1* yang merupakan pengguna *spam* *Twittbot.net*.

```

1 #membuat variabel list kosong
2 isi_tweet_pengguna = []
3 jumlah_favorit_tweet = []
4 #membaca dataframe nama pengguna yang sesuai dengan SilviaSandras1
5 for isi_tweet in df[df['Nama_Pengguna']=='SilviaSandras1'].iterrows():
6     #meng append data ke dalam list yang berisi Isi Tweet
7     isi_tweet_pengguna.append(isi_tweet[1]['Isi_Tweet'])
8     #meng append data ke dalam list yang berisi jumlah favorit tweet
9     jumlah_favorit_tweet.append(isi_tweet[1]['Jumlah_Favorit'])
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

**Gambar 3.36. Menampilkan Isi Konten User Twittbot**

Gambar 3.37 merupakan *script* untuk menampilkan jumlah banyak tweet yang dilakukan oleh *spam* *twittbot.net* berdasarkan isi.

```
In [9]: 1 Counter(isi_tweet_SilviaSandras1['Isi'].dropna())#.sort_values(ascending=False)[:15]

Out[9]: Counter({'belajar bukan hanya mengetahui apa yang harus dilakukan, tapi melakukan apa yang sudah kita ketahui #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'hanya satu motivasi yang ada, yaitu allah. adapun motivasi lainnya harus dalam rangka "karena dan/atau untuk" allah. #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'harapan itu seperti jalan di dalam hutan. di sana tak pernah ada jalan. tapi jika banyak orang menjalani jalan itu menjadi ada #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'kemerdekaan indonesia yang diraih adalah buah dari persatuan, perjuangan seluruh rakyat indonesia, dari berbagai latar belakang agama, suku, dan etnis menjadi kita ini sama #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'lakukanlah seluruh aktivitas kita dengan penuh cinta dan hanya mengharapkan keridhaan-nya sehingga akan bernilai ibadah. berlelah-lelahlah, manisnya hidup terasa setelah kita lelah berjuang. #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'manusia yang paling lemah ialah orang yang tidak mampu mencari teman. namun yang lebih lemah dari itu ialah orang yang mendapatkan banyak teman tetapi menyia - nyiakannya. #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'menghidupkan kembali agama berarti menghidupkan suatu bangsa. hidupnya agama berarti cahaya kehidupan #2019tetapjokowi #jokowi2periode #iwartaid': 1,
'niat adalah ukuran dalam menilai benarnya suatu perbuatan, oleh karenanya, ketika niatnya benar, maka perbuatan itu benar, dan jika niatnya buruk, maka perbuatan itu buruk #2019tetapjokowi #jokowi2periode #iwartaid': 1})
```

**Gambar 3.37. Isi Konten Twittbot SilviSandras1**

Gambar 3.38 merupakan *script* untuk menampilkan seluruh *filter* isi twittbot.net berdasarkan nama yang ingin dicari seperti pada gambar tersebut penulis ingin mencari pengguna twittbot.net dengan nama pengguna SilviaSandras1.

```
In [10]: 1 df[df['Nama_Pengguna']=='SilviaSandras1'].tail(50)

Out[10]:
```

	Tanggal_Tweet	Isi_Tweet	Hashtags	Sumber_Twitter	Bahasa_Twitter	Nama_Pengguna	Lokasi_Pengguna	Jurr
1893	Thu Jan 17 12:47:14 +0000 2019	kemerdekaan indonesia yang diraih adalah buah ...	2019TetapJokowi,Jokowi2Periode,iwartaid	href="http://twittbot.net/?rel=nofollow">...	in	SilviaSandras1	kosong	
4197	Thu Jan 17 13:17:14 +0000 2019	NaN	NaN	href="http://twittbot.net/?rel=nofollow">...	in	SilviaSandras1	kosong	
6989	Thu Jan 17 13:47:13 +0000 2019	NaN	NaN	href="http://twittbot.net/?rel=nofollow">...	in	SilviaSandras1	kosong	
12785	Thu Jan 17 14:17:13 +0000 2019	NaN	NaN	href="http://twittbot.net/?rel=nofollow">...	in	SilviaSandras1	kosong	
17254	Thu Jan 17 14:47:14 +0000 2019	hanya satu motivasi yang ada, yaitu allah. ada...	2019TetapJokowi,Jokowi2Periode,iwartaid	href="http://twittbot.net/?rel=nofollow">...	in	SilviaSandras1	kosong	

**Gambar 3.38. Menampilkan Semua Konten Nama SilviSandras1**

```
1 #modul Counter:berfungsi untuk menghitung jumlah kata yang sama seperti value.counts
2 from collections import Counter
3 #membuat variabel baru untuk mencounter.
4 nama_pengguna_counter = Counter()
5 #membaca seluruh dataframe src twitter yang mengandung twittbot.net
6 for row in df[df['src_twitter']=='twittbot.net'].iterrows():
7     #jika nama pengguna sesuai dengan sebelumnya maka nilai counter di tambah 1
8     nama_pengguna_counter.update([row[1]['Nama_Pengguna']])
9 #menampilkan nama dengan nilai count 15 terbesar
10 nama_pengguna_counter.most_common(15)
```

**Gambar 3.39. Filter Pengguna Twittbot.net**

Gambar 3.39 merupakan *script* untuk menghitung seberapa sering pengguna melakukan tweet yang menggunakan media twittbot.net sebagai tempat untuk melakukan *spam*.

### C. *Count* Perangkat Twitter, Bahasa Twitter, Kombinasi *Hashtag* dan *Active User*.

Pada *modelling data* berikutnya, penulis ingin mengetahui sumber perangkat yang paling banyak digunakan untuk melakukan Twitter berasal dari perangkat apa, bahasa yang paling sering digunakan dalam melakukan tweet pada debat capres 2019 ini menggunakan bahasa apa, mencari banyaknya kombinasi *hashtag* yang digunakan pengguna untuk melakukan sebuah *post* tweet seperti apakah pembicaraan tersebut bersifat *positif* atau *negatif* kepada kedua pasangan calon presiden 2019, dan ingin melihat siapa saja yang rajin melakukan tweet terhadap capres 2019 ini.

```
1 #modul yang di pakai
2 import os
3 import pandas as pd
4 from bs4 import BeautifulSoup
5 from collections import Counter, OrderedDict
```

**Gambar 3.40. *Import Library Count Twitter***

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA

```

1 #membuat dataframe kosong
2 df = pd.DataFrame()
3 #mengset jalur letak direktori file
4 base_path= '/home/pc-bigdata-1/Desktop/testing/convertcsvuser/second/convertexceltolokaltimezone/'
5 #membaca jalur direktori
6 for file in os.listdir(base_path):
7     #membuka data yang penamaan file mengandung kata test dan berakhir format .xlsx/excel
8     if file.endswith('test.xlsx'):
9         #melakukan print ke display bahwa file sedang dibaca
10        print('loading:', file)
11        #men append isi data 1 atau lebih
12        df = df.append(pd.read_excel(os.path.join(base_path, file)), ignore_index=True)
13 #mengprint notif apabila seluruh data berhasil dibaca
14 print("Done")

```

**Gambar 3.41. Membaca Data Excel**

Hal pertama yang harus dilakukan adalah import modul yang ingin dipakai pada *count* ini dapat dilihat seperti pada gambar 3.40 dan setelah modul berhasil maka mulai melakukan pembacaan data yang akan ditampung ke dalam *dataframe*, dimana fungsi *dataframe* ini untuk menampung data dengan 1 atau lebih dari 1 file *xlsx* sekaligus dapat dilihat seperti pada gambar 3.41.

```

1 #membuat variabel list baru untuk menampung data sumber twitter
2 sumber_twitter_series = []
3 #membaca dataframe dengan looping
4 for row in df.iterrows():
5     #membuat variabel baru untuk menampung data sumber twitter yang memiliki format html dengan modul beautifulso
6     soup = BeautifulSoup(row[1]['Sumber_Twitter'], 'html.parser')
7     #kondisi mencari seluruh isi jika datanya ada dan =True
8     if soup.find('a'):
9         #maka mengambil isi teks sumber twitter
10        sumber_twitter_series.append(soup.find('a').get_text())
11    else:
12        #menambakan isi ke dalam Sumber_Twitter bahwa struktur data tidak terstruktur
13        sumber_twitter_series.append(row[1]['Sumber_Twitter'])
14        print('unrecognized:', row[1]['Sumber_Twitter'])
15 #menampilan pesan Done jika sudah selesai kelola data
16 print("Done")

```

**Gambar 3.42. Filter Sumber Twitter**

Gambar 3.42 merupakan *script* untuk mengambil isi teks yang memiliki *format* html agar lebih rapi dan mudah dikelola dengan menambahkan fungsi *BeautifulSoup* ini kita dapat mengelola atau mengambil isi teks lebih mudah.

```

1 #menghitung jumlah sumber twitter dan di sorting Ascending
2 source_count_Series = pd.Series(Counter(sumber_twitter_series)).sort_values(ascending=False)[:15]
3 #menampilkan data sumber twitter
4 source_count_Series

```

Twitter for Android	56925
Twitter Web Client	9863
Twitter for iPhone	8348
Twitter for iPad	6492
Twitter Lite	3724
TweetDeck	2328
Twitter Web App	2037
twittbot.net	1529
IFTTT	1525
Mixero	699
Bad Kitty	672
Echofon	667
Tween	611
Seismic twirl	581
Krile2	576

dtype: int64

**Gambar 3.43. Count Sumber Twitter**

Pada gambar 3.43 merupakan *script* untuk menampilkan hasil perhitungan jumlah perangkat Twitter yang sering digunakan pengguna dalam melakukan tweet.

```

1 #membuat variabel baru untuk menampung dataframe bahasa twitter
2 bahasa_tweets = df['Bahasa_Twitter']
3 #menghitung jumlah bahasa tweet dan di sorting ascending
4 bahasa_tweet_Series = pd.Series(Counter(bahasa_tweets)).sort_values(ascending=False)[:15]
5 #menampilkan data bahasa tweet
6 bahasa_tweet_Series

```

in	87937
und	8220
en	1131
tl	438
eu	428
hu	263
es	146
da	133
ht	115
tr	55
et	48
hi	42
fi	38
nl	36
lt	29

dtype: int64

**Gambar 3.44. Count Bahasa Twitter**

Pada gambar 3.44 merupakan *script* untuk melakukan perhitungan jumlah bahasa yang paling sering digunakan untuk melakukan tweet

oleh pengguna twitter. Beberapa arti singkatan bahasa in = Indonesia, und = *undetermine*, dan en = *english*.

```

1 #menghitung jumlah kombinasi hashtags tweet
2 hastags_counts = pd.Series(df['Hashtags'].value_counts()).sort_values(ascending=False)
3 #menampilkan isi hashtags_counts
4 hastags_counts

```

PakWojumatanDimana, PasarkaretanJuaraATF2019, PrabowoSandiMenangDebat, PrabowoIndonesiaMenang, 2019GantiPresiden, 2019PrabowoPresidenRI, percayakumparan  
1187  
2019PilihJokowi, 01IndonesiaMaju, JokowiLagi, Debat01Jokowi, DebatKemenanganJokowi, JokowiAminMenangDebat  
680  
2019TetapJokowi, JokowiKerjanyaNyata, 01UnggulDebat, JumatBerkah, hoaxpendukungprabowo, hoaxpemilu  
557  
2019PrabowoSandi, 2019PrabowoPresiden, 2019GantiPresiden, 17AprilCoblosNomor2, 17AprilPrabowoPresiden, PrabowoIndonesiaMenang, 2019PrabowoPresidenRI, DebatPipres2019  
494  
2019PrabowoPresidenRI, PrabowoIndonesiaMenang, PrabowoSandiMenangDebat, RakyatSudahMuak, RakyatInginkanPerubahan, IndonesiaAdilMakmur  
355  
2019TetapJokowi, hoaxpendukungprabowo, hoaxpemilu, KamisManis  
336  
westpapua, freewestpapua, letwestpauavote, fwpc, papuawest, infowestpapua, westpauanid, 2019GantiPresiden  
333  
JokowiAminMenangDebat, 01IndonesiaMaju  
322  
westpaua, freewestpaua, letwestpauavote, fwpc, LawanGerakanSeparatis, 2019GantiPresiden

**Gambar 3.45. Kombinasi Hashtags**

Pada gambar 3.45 merupakan *script* untuk mencari jumlah kombinasi *hashtags* yang populer paling banyak digunakan *user* untuk melakukan tweet.

```

1 #membuat variabel list baru
2 isi_tweet_pengguna = []
3 usernya = []
4 #mengset keyword hashtags yang ingin dicari
5 keyword = 'JokowiAminMenangDebat,01IndonesiaMaju'
6 #mencari data dalam dataframe yang sesuai dengan keyword
7 for isi_tweet in df[df['Hashtags']== keyword].iterrows():
8     #meng append isi tweet
9     isi_tweet_pengguna.append(isi_tweet[1]['Isi_Tweet'])
10    #mengappend isi nama user
11    usernya.append(isi_tweet[1]['Nama_Pengguna'])
12
13 #mengubah format struktur data dari list ke pandas series
14 isi_tweet_pengguna_series = pd.Series(isi_tweet_pengguna, name='isi_tweet')
15 usernya_series = pd.Series(usernya, name='nama_user')
16 #menggabungkan 2 pandas series menjadi 1 dataframe
17 filter_by_hashtags = pd.concat([usernya_series, isi_tweet_pengguna_series], axis=1, ignore_index=False)
18 #menampilkan isi filter by hashtags yang sudah diconcat
19 filter_by_hashtags.head(20)

```

	nama_user	isi_tweet
0	Yohanabila01	memang tak mudah pemimpin yang baik adil dan j...
1	teenmoms2	setiap pemimpin negara kita mempunyai cara yan...
2	Yohanabila01	jangan ragukan dalam hal pembangunan, jangan t...
3	teenmoms2	begitu pula dengan pakde, pengentasan kemiskin...
4	VanessaAngela1	sebelum berangkat ke tempat debat capres perta...

**Gambar 3.46. Isi Tweet Berdasarkan Kombinasi Hashtags**

Pada gambar 3.46 merupakan *script* untuk melihat isi-isi tweet yang di *posting* oleh pengguna berdasarkan tingkat popularitas *hashtags* yang sering digunakan dari *keyword script* pada gambar 3.45. Kombinasi *Hashtag*.

```
1 #membuat variabel baru untuk menampung hasil count nama pengguna dan di sorting ascending
2 active_user = pd.Series(df['Nama_Pengguna'].value_counts()).sort_values(ascending=False)
3 # dict(active_user)
4 #menampilkan isi active_user 10 besar
5 active_user[:10]
```

khibar_p	457
avifrakhman	157
JokSusi2019	129
kabarkhusus	120
putra_bisma12	115
dadangpancong1	115
Barna_byo	107
makLambeTurah_	102
IslandBeby	99
BakriZulkarnain	96

Name: Nama\_Pengguna, dtype: int64

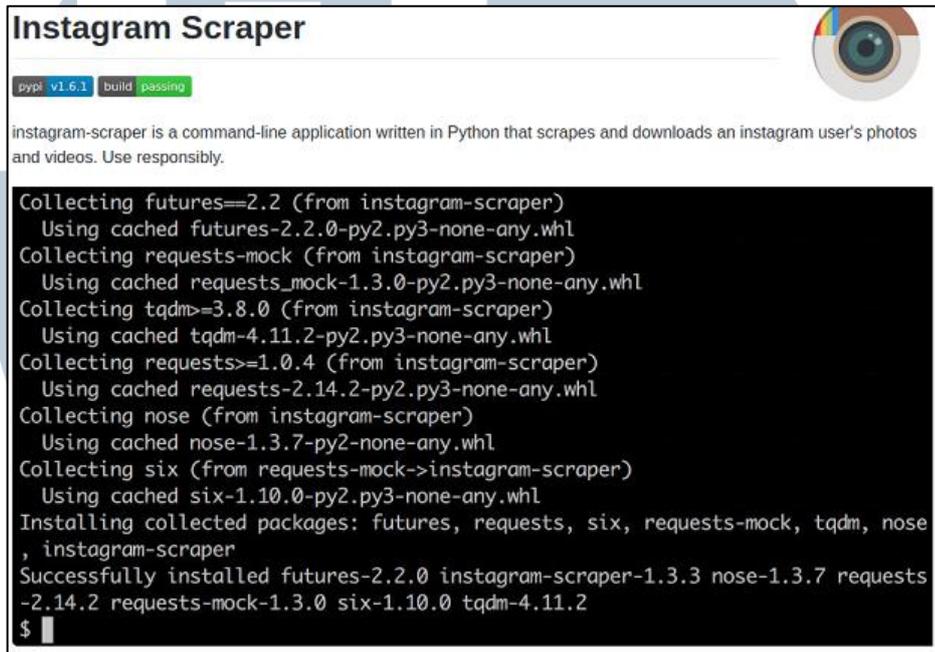
**Gambar 3.47. Active User**

Pada gambar 3.47 merupakan *script* untuk melihat pengguna aktif yang sering melakukan tweet pada akun Twitternya, terlihat bahwa *user* aktif tersebut bernama khibar\_p sebagai contohnya.

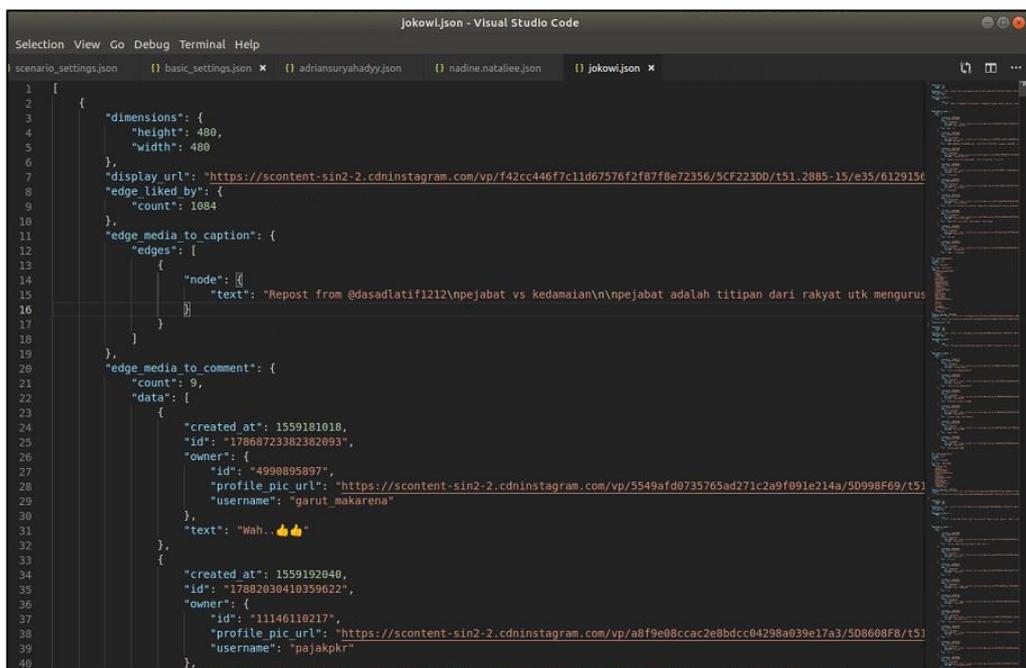
### 3.2.10. Instagram Scraping

Tugas tambahan yang di berikan oleh supervisor untuk melakukan Instagram *scraping* yang mengambil isi-isi data dengan menggunakan *keyword* tertentu. Penulis mencoba melakukan pencarian *internet* terkait Instagram *scraping* dan menemukan sebuah *program* yang dapat langsung digunakan dan sesuai kebutuhan. Instagram ini dapat melakukan *scraping* berdasarkan *keyword hashtags* dan nama pengguna Instagram. Isi data *scraping* berupa *file*

JSON yang berisi informasi Instagram seperti *link* foto, identitas pengguna, isi *posting*, komentar dan lainnya.

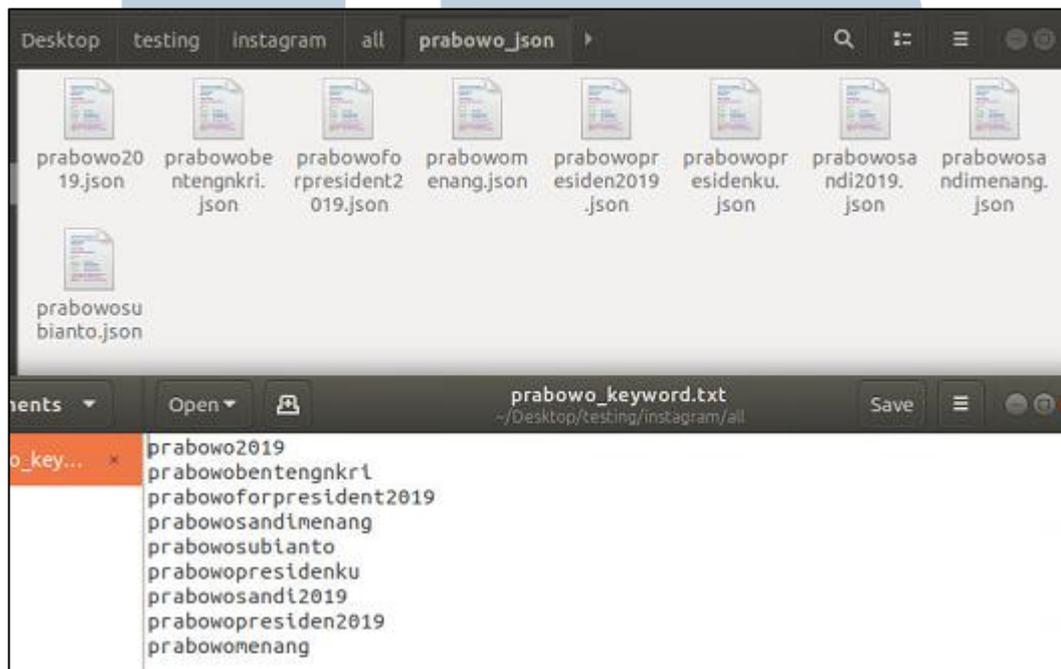


Gambar 3.48. Instagram Scarping



Gambar 3.49. Hasil JSON Instagram Scarping

Pada gambar diatas merupakan isi *file* JSON yang berhasil di *scraping* berdasarkan *keyword hashtags*, agar data pada gambar diatas lebih mudah di baca oleh penulis dilakukan tahap penyajian data Instagram yang akan di jelaskan di tugas berikutnya.



**Gambar 3.50. Scarping Instagram Dengan Keyword**

Pada gambar di atas merupakan hasil yang berhasil di *scarping* berdasarkan *keyword hashtags* mengenai prabowo.

### **3.2.11. Script Instagram Preparation Data**

Data dari Instagram *scarping* masih dalam bentuk JSON dan struktur data masih berantakan sehingga membuat penulis sulit untuk membaca isinya. Kemudian untuk mempermudah hal tersebut di buat sebuah *script* untuk *scarping* Instagram *preparation data* seperti gambar di bawah ini, merupakan *script* untuk mengambil data seperti jumlah komentar, *user ID*, isi *caption*,

jumlah *like*, lokasi pengguna, tanggal *posting*, *link* gambar, *link posting* Instagram dan *shortcode* sebagai keyword untuk nanti jika penulis tertarik melihat isi komentar-komentar berdasarkan posting yang diinginkan dan informasi lainnya.

```
1 df = pd.DataFrame() #tampung bacaan file json ke dalam dataframe
2 df1 = pd.DataFrame() #tampung data hasil gabungan concat
3
4 base_path= '/home/pc-bigdata-1/Desktop/testing/nadine/'
5 save_path = 'adrian.csv'
6
7 for file in os.listdir(base_path):
8     if file.endswith('nadine.nataliee.json'):
9         print('loading:', file)
10        df = df.append(pd.read_json(os.path.join(base_path, file)))
11
12 print("Data Siap")
13
14 jumlah_comment = []
15 for user_komen in df['edge_media_to_comment']:
16     jumlah_comment.append(user_komen['count'])
17 print("Jumlah_Comment Selesai")
18
19 id_user = []
20 for userid in df['owner']:
21     id_user.append(userid['id'])
22 print("ID_User Selesai")
23
24 # jumlah like = []
25 # for like in df['edge_liked_by']:
26 #     jumlah_like.append(like['count'])
27 # print("Jumlah_Like Selesai")
28
29 isi_caption = []
30 for texts in df['edge_media_to_caption']:
31     try:
32         isi_caption.append(texts['edges'][0]['node']['text'].replace('\n', ' ')) #untuk mengambil text caption in
33     except:
34         pass
35 print("Isi_Caption Selesai")
36
37 #mengambil lokasi
38 daftar_alamat = []
39 daftar_kota = []
40 for alamat in df['location']:
41     if alamat:
42         daftar_alamat.append(alamat['address_json'])
43     else:
44         pass
45 daftar_alamat.append({})
```

**Gambar 3.51. Script Penyajian Data Instagram (1)**

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

```

46
47 for lok in daftar_alamat:
48     if lok:
49         kota = json.loads(lok)
50         print(nama_kota.get('city_name'))
51         nama_kota = kota.get('city_name')
52         daftar_kota.append(nama_kota)
53     else:
54         daftar_kota.append(0)
55 print("Lokasi Selesai")
56
57 shortcode = []
58 link_posting_instagram = []
59 for posting in df['shortcode']:
60     link_posting_instagram.append('https://www.instagram.com/p/'+posting)
61     shortcode.append(posting)
62 print("Link Posting Selesai")
63
64 url_gambar = []
65 for img_vid in df['display_url']:
66     url_gambar.append(img_vid)
67 print("Url Gambar Selesai")
68
69 date_time = df['taken_at_timestamp'].apply(lambda x: datetime.fromtimestamp(x))
70 print("Tanggal Posting Selesai")
71
72 # jumlah like series = pd.Series(jumlah_like, name='Jumlah Like Post')
73 id_user_series = pd.Series(id_user, name='User ID')
74 jumlah_comment_series = pd.Series(jumlah_comment, name='Jumlah Komen')
75 isi_caption_series = pd.Series(isi_caption, name='Isi Posting')
76 daftar_kota_series = pd.Series(daftar_kota, name='Lokasi')
77 link_posting_instagram_series = pd.Series(link_posting_instagram, name='Link Instagram')
78 url_gambar_series = pd.Series(url_gambar, name='Link Gambar')
79 key_shortcode_series = pd.Series(shortcode, name='Keyword Shortcode')
80 print("Convert list to pandas.series")
81
82 df1 = pd.concat([date_time, id_user_series, jumlah_comment_series, isi_caption_series, daftar_kota_series, url_gambar_series])
83 print("Data Concat Selesai")
84
85 df1.rename(columns={'taken_at_timestamp': 'Tanggal Posting'}, inplace=True)
86 print("rename Selesai")
87
88 df1.to_csv(os.path.join(base_path, save_path), sep='-', na_rep="Kosong", index=None)
89 print("Eksport CSV Selesai")

```

**Gambar 3.52 Script Penyajian Data Instagram (2)**

```

87
88 df1.to_csv(os.path.join(base_path, save_path), sep='-', na_rep="Kosong", index=None)
89 print("Eksport CSV Selesai")

```

Loading: nadine.nataliee.json  
Data Siap  
Jumlah Comment Selesai  
ID User Selesai  
Isi Caption Selesai  
Lokasi Selesai  
Link Posting Selesai  
Url Gambar Selesai  
Tanggal Posting Selesai  
Convert list to pandas.series  
Data Concat Selesai  
rename Selesai  
Eksport CSV Selesai

**Gambar 3.53. Penyajian Data Instagram Eksport CSV**

Pada gambar di atas apabila *script* sudah berhasil mengambil data-datanya kemudian dilakukan ekspor untuk menyimpan hasilnya yang disimpan dalam bentuk CSV.

Gambar berikutnya di bawah ini merupakan hasil yang didapatkan mengenai informasi yang dibutuhkan.

	A	B	C	D	E	F	G	H	I	
	Tanggal Posting	User ID	Jumlah Like	Jumlah Post	Jumlah Komen	Isi Posting	Lokasi	Link_Gambar	Link_Instagram	Keyword_ShortCode
1	05/30/19 08:45 AM	6281944016	1084	9	Repost from @da			0 https://scontent-sin2-2.https://www.instagram.com/p/ByEYuz2nW/ByEYuz2nW		
2	05/30/19 10:45 PM	2175610388	185	6	Kesalamu'alaikum			0 https://scontent-sin2-2.https://www.instagram.com/p/ByF5kAFD3N/ByF5kAFD3N		
3	05/30/19 08:21 PM	11591774092	133	9	👏👏👏 bagi 10-10			0 https://scontent-sin2-2.https://www.instagram.com/p/ByFqPpGjIN/ByFqPpGjIN		
4	05/29/19 11:59 AM	5882100072	1199	23	BRABOWO GILA			0 https://scontent-sin2-2.https://www.instagram.com/p/ByCKv2nkhU/ByCKv2nkhU		
5	05/30/19 04:08 PM	13614595624	1008	17	So Sweet Sekali			0 https://scontent-sin2-2.https://www.instagram.com/p/ByFMNNOBw/ByFMNNOBw		
6	05/30/19 07:18 PM	5719709738	645	129	Kalo benar alasan			0 https://scontent-sin2-2.https://www.instagram.com/p/ByFh814JiH/ByFh814JiH		
7	05/30/19 11:41 AM	5719709738	715	37	Gobloknya gak ae			0 https://scontent-sin2-2.https://www.instagram.com/p/ByEthlpPFm/ByEthlpPFm		
8	05/30/19 03:52 AM	5895177921	2227	29	#fahrhmanzah #gg			0 https://scontent-sin2-2.https://www.instagram.com/p/ByD38FvBPw/ByD38FvBPw		
9	05/29/19 05:47 AM	5895177921	3023	45	#fahrhmanzah #gg			0 https://scontent-sin2-2.https://www.instagram.com/p/ByBgQMizy/ByBgQMizy		
10	05/30/19 11:42 PM	7340239598	0	0				0 https://scontent-sin2-2.https://www.instagram.com/p/ByFzAVn1e/ByFzAVn1e		
11	05/30/19 11:42 PM	6312672796	0	0	"Di sini kita berbic			0 https://scontent-sin2-2.https://www.instagram.com/p/ByGAFrJuC/ByGAFrJuC		
12	05/30/19 11:41 PM	3677148196	0	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByGACjkACl/ByGACjkACl		
13	05/30/19 11:41 PM	3677148196	0	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByGAB7sACl/ByGAB7sACl		
14	05/30/19 11:41 PM	3677148196	1	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByGABL7gP/ByGABL7gP		
15	05/30/19 11:41 PM	3677148196	1	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByGAAfagXZ/ByGAAfagXZ		
16	05/30/19 11:41 PM	3677148196	1	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByFvlgInW/ByFvlgInW		
17	05/30/19 11:41 PM	3677148196	0	0	Semua salah pent			0 https://scontent-sin2-2.https://www.instagram.com/p/ByFu4JcgW/ByFu4JcgW		
18	05/30/19 11:41 PM	12750979030	0	0	#jagapersatuand			0 https://scontent-sin2-2.https://www.instagram.com/p/ByF6_oACx/ByF6_oACx		
19	05/30/19 11:41 PM	3677148196	2	0	#01jokowilagi #20			0 https://scontent-sin2-2.https://www.instagram.com/p/ByF8_YBgK/ByF8_YBgK		
20	05/30/19 11:40 PM	9895458547	1	1	TARIF			0 https://scontent-sin2-2.https://www.instagram.com/p/ByF8_MnV/ByF8_MnV		

Gambar 3.54. Output CSV Penyajian Data Instagram

```

1 import json
2 import csv
3 from datetime import datetime
4 import pandas as pd
5 import os.path

1 df = pd.DataFrame()
2
3 base_path= '/home/pc-bigdata-1/Desktop/Tugas Rapi/joko/'
4 save_path = 'komen.csv'
5
6 for file in os.listdir(base_path):
7     if file.endswith('.json'):
8         print('loading:', file)
9         df = df.append(pd.read_json(os.path.join(base_path, file)))
10
11 print("Data Siap")

```

Gambar 3.55. Import Module dan Read Data untuk script Instagram

Pada gambar di atas merupakan *script* yang nantinya untuk menghasilkan isi komentar-komentar tiap *caption posting*, hal yang dilakukan adalah *import* modul-modul yang akan digunakan. Setelah itu membaca letak *file* yang akan dikelola.

```

1 komen_list = []
2
3 isi_komen = []
4 tanggal_komen = []
5 nama_pengguna = []
6 gambar_profil = []
7
8 keyword_shortcode = input("Masukkan keyword Shortcode: ")
9
10 for filter_shortcode in df[df['shortcode'] == keyword_shortcode].iterrows():
11     komen_list.append(filter_shortcode[1]['edge_media_to_comment']['data'])
12     print("Jumlah_Komen: ", filter_shortcode[1]['edge_media_to_comment']['count'], "Reply_Posting")
13
14
15 komen_pandas_dataframe = pd.DataFrame(komen_list)
16 reply_list = []
17
18 for komen_iterrows in komen_pandas_dataframe.iterrows():
19     reply_list.append(komen_iterrows[1])
20
21 identitas_list = []
22 for kumpulan_reply in reply_list:
23     for identitas_comment in kumpulan_reply:
24         tanggal_komen.append(identitas_comment['created_at'])
25         isi_komen.append(identitas_comment['text'])
26         identitas_list.append(identitas_comment['owner'])
27
28 for detail_identitas in identitas_list:
29     nama_pengguna.append(detail_identitas['username'])
30     gambar_profil.append(detail_identitas['profile_pic_url'])
31
32 isi_komen_series = pd.Series(isi_komen, name="Posting_Reply")
33 waktu_komen_series = pd.Series(tanggal_komen).apply(lambda x: datetime.fromtimestamp(x))
34 nama_pengguna_series = pd.Series(nama_pengguna, name='Nama')
35 gambar_profil_series = pd.Series(gambar_profil, name='Gambar_Profile')
36
37 df1 = pd.concat([waktu_komen_series, nama_pengguna_series, gambar_profil_series, isi_komen_series], axis=1)
38 df1.rename(columns={0: "Waktu_Reply"}, inplace=True)
39
40 df1.to_csv(os.path.join(base_path, save_path), sep='-', na_rep="Kosong", index=None)

```

**Gambar 3.56. Penyajian Data Isi Komentar Instagram Tiap Posting**

Pada gambar di atas pertama membuat sebuah *variable* baru *list* kosong untuk menampung isi data yang akan dijalankan di koding berikutnya. Pada *script* ini data-data yang diambil berupa semua isi komentar orang yang terdiri dari tanggal komentar, gambar profile, *User ID* komentar, dan isi komentarnya. Setelah selesai diambil datanya maka akan di ekspor ke bentuk CSV.

U N I V E R S I T A S  
M U L T I M E D I A  
N U S A N T A R A

Waktu_Reply	Nama	Gambar_Profile	Posting_Reply
05/30/19 09:44 PM	genius.property		https://scontent-sin2-2@nyatakan_bising https://www.instagram.com/p/Bv7KmQpJ4xt/?igshid=ujj6axbJep
05/30/19 09:45 PM	alexandrovictoria		https://scontent-sin2-2@ya benar...jaman jokowi asn sangat sulit pungli...walaupun msh ada tp sulit dan asn tdk bs santai spt jaman seblm jokowi...dulu a
05/30/19 09:46 PM	nurtazizabsari		https://scontent-lax3-2@MANTUL..
05/30/19 09:46 PM	aguschan1		https://scontent-sin2-2@nyatakan_bising mau aseng mau apa kita Bhineka tunggal ika mari kita jaga Persatuan dan Kesatuan
05/30/19 09:48 PM	alexandrovictoria		https://scontent-sin2-2@alingbulin yg mlh dr kalangan wong cilik
05/30/19 09:51 PM	hendri1360		https://scontent-sin2-2@Tidak semua PNS...
05/30/19 09:51 PM	putricobain		https://scontent-sin2-2@Benar benar benar
05/30/19 10:00 PM	anjelle18		https://scontent-sin2-2@nyatakan_bising mending gitu...drpada bilang antek Aseng.tp sendirinya bergerombol sama Aseng??havoo...@
05/30/19 10:01 PM	veinugui		https://scontent-sin2-2@sejuju
05/30/19 10:02 PM	najah9362		https://scontent-sin2-2@Boor masih banyak pungli bro
05/30/19 10:04 PM	elva_cake		https://scontent-sin2-2@Cocokkkkhh
05/30/19 10:09 PM	erwinyudas		https://scontent-sin2-2@nyatakan_bising akun bodong...ibumu jancuk
05/30/19 10:09 PM	popi.yusmira		https://scontent-sin2-2@hendri1360 80% ASN/BUMN pasti gitu
05/30/19 10:11 PM	popi.yusmira		https://scontent-sin2-2@alexandrovictoria bener bgt. 80 % pelayanan ASN kepublik itu omong kosong.semua pungen gaji buta
05/30/19 10:26 PM	indralerianto		https://scontent-sin2-2@alingbulin yang mlh saya, bukan ASN atau pegawai BUMN, cuma pegawai biasa
05/30/19 10:28 PM	arishanyant		https://scontent-sin2-2@bs jadi kam masih bnyak orang bodoh dan tidak waras, demi pungli, anjrit
05/30/19 10:30 PM	mir_idax707172		https://scontent-sin2-2@nyatakan_bising @jacklyn_choppers ini ada yg mau ngeboom nih bang...
05/30/19 10:32 PM	kapt_yudi		https://scontent-sin2-2@Memang bener.karena kran pundiz oleh Jokowi di tutup, jadi mereka mlh pak wowo...siapa tahu kran pundiz terbuka kembali....
05/30/19 10:36 PM	hadl_zoax		https://scontent-sin2-2@Tp sayang d daerah masih ada pungli pengurusan ktp dan kk
05/30/19 10:37 PM	hari.anto.391		https://scontent-sin2-2@alingbulin Mang rakyat berapa persen yg asn
05/30/19 10:38 PM	awabiru6356		https://scontent-sin2-2@alingbulin memang karena punya otak..yg kalah ngaku menang itu baru gk punya otak@@@
05/30/19 10:38 PM	shabina.intan		https://scontent-sin2-2@Terbukti integritas lingai
05/30/19 10:39 PM	lesdewi07		https://scontent-sin2-2@
05/30/19 10:40 PM	jancokmerdekarebom		https://scontent-sin2-2@hadl_zoax foto bukti kita bantu share...@
05/30/19 10:40 PM	jancokmerdekarebom		https://scontent-sin2-2@kapt_yudi @
05/30/19 10:41 PM	jancokmerdekarebom		https://scontent-sin2-2@hendri1360 Memang
05/30/19 10:41 PM	hadl_zoax		https://scontent-sin2-2@jancokmerdekarebom ada d hp bini om, bsok qw dm deh
05/30/19 10:41 PM	jancokmerdekarebom		https://scontent-sin2-2@lesdewi07 @
05/30/19 10:41 PM	krisna.alfath		https://scontent-sin2-2@Emang fakta tmt sy ASN bnyk yg pilih Prabowo
05/30/19 10:42 PM	jancokmerdekarebom		https://scontent-sin2-2@alexandrovictoria
05/30/19 10:42 PM	jancokmerdekarebom		https://scontent-sin2-2@yuni_darwati
05/30/19 10:43 PM	jancokmerdekarebom		https://scontent-sin2-2@sarinahusein

**Gambar 3.57. Hasil CSV Isi Komentar Instagram**

Pada gambar diatas merupakan hasil CSV yang di dapatkan seperti waktu\_reply, Nama, Gambar profile, dan posting\_reply.

### 3.3. Kendala

Selama melakukan kegiatan kerja magang di Universitas Multimedia Nusantara tepatnya di *Lab Big Data* sebagai *Text Mining*, terdapat beberapa kendala yang dialami diantaranya:

1. Masih pemula dalam bahasa pemrograman Python dan penggunaan sistem operasi Linux sehingga memperlambat pengerjaan karena harus belajar terlebih dahulu.
2. Pemberian *keyword* oleh tim Ilkom terkadang tidak mendapatkan hasil karena *keyword* yang diberikan antara tidak spesifik atau sangat spesifik, namun jika diambil beberapa saja *keyword* artikel malah menjadi tidak relevan dan ketersediaan database untuk melakukan fact checking debat capres 2019 terbatas.

### 3.4. Solusi

Solusi dari kendala yang dialami ketika melakukan kegiatan kerja magang di Universitas Multimedia Nusantara di *Lab Big Data* sebagai *Text Mining*, yaitu:

1. Menanyakan yang tidak dimengerti kepada Pak Yustinus Eko selaku pembimbing kerja magang yang lebih mengetahui akan materi Python, dan sistem operasi Linux, serta melakukan pembelajaran otodidak melalui *internet*.
2. Memperbanyak jumlah *data* dengan melakukan *scraping website* Kompas.com untuk mendapatkan *data* terbaru.

UMMN

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA