

**PENGEKSTRAKAN TEKS BERASASKAN  
PERATURAN BAGI PENAMBAHBAIKAN GRAF  
PENGETAHUAN UNTUK PEMAHAMAN IMEJ**

**IDZA AISARA BINTI NORABID**

**UNIVERSITI KEBANGSAAN MALAYSIA**

**PENGEKSTRAKAN TEKS BERASASKAN PERATURAN BAGI  
PENAMBAHBAIKAN GRAF PENGETAHUAN UNTUK PEMAHAMAN IMEJ**

**IDZA AISARA BINTI NORABID**

**DISERTASI YANG DIKEMUKAKAN UNTUK MEMENUHI SEBAHAGIAN  
DARIPADA SYARAT UNTUK MEMPEROLEH IJAZAH SARJANA SAINS  
KOMPUTER (KECERDASAN BUATAN)**

**FAKULTI TEKNOLOGI DAN SAINS MAKLUMAT  
UNIVERSITI KEBANGSAAN MALAYSIA  
BANGI**

**2022**

## PENGHARGAAN

Di kesempatan ini, saya ingin mendedikasikan ucapan penghargaan kepada semua pensyarah lebih lagi Dr Wan Fariza selaku penyelia saya atas tunjuk ajar serta bimbingan daripadanya bagi menyiapkan tesis ini dengan jayanya.

Saya juga ingin mengucapkan terima kasih yang tidak terhingga kepada ibu bapa saya yang memberi saya segala kemudahan dan sokongan moral yang tidak terhingga sepanjang tempoh menyiapkan tesis ini.

Ucapan penghargaan ini juga saya tujukan kepada rakan-rakan yang banyak memberi peringatan terhadap setiap apa yang saya telah alpa. Juga membantu saya dengan menjawab setiap persoalan saya.

Akhir kata, terima kasih khas ditujukan kepada pihak penaja (JPA) yang telah membantu saya dari segi kewangan sepanjang saya melanjutkan pengajian di UKM ini dan juga kepada semua yang terlibat secara langsung dan tidak langsung dalam memberikan sumbangan cadangan dan bantuan dalam menyiapkan tesis ini.

Semoga penyelidikan dan tesis ini dapat dijadikan wadah ilmu yang berguna untuk tatapan generasi akan datang.

Pusat Sumber  
FTSM

## **PENGAKUAN**

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang tiap-tiap satunya telah saya jelaskan sumbernya.

4 November 2022

IDZA AISARA BINTI NORABID  
P101080

Pusat Sumber  
FTSM

## PENGHARGAAN

Di kesempatan ini, saya ingin mendedikasikan ucapan penghargaan kepada semua pensyarah lebih lagi Dr Wan Fariza selaku penyelia saya atas tunjuk ajar serta bimbingan daripadanya bagi menyiapkan tesis ini dengan jayanya.

Saya juga ingin mengucapkan terima kasih yang tidak terhingga kepada ibu bapa saya yang memberi saya segala kemudahan dan sokongan moral yang tidak terhingga sepanjang tempoh menyiapkan tesis ini.

Ucapan penghargaan ini juga saya tujukan kepada rakan-rakan yang banyak memberi peringatan terhadap setiap apa yang saya telah alpa. Juga membantu saya dengan menjawab setiap persoalan saya.

Akhir kata, terima kasih khas ditujukan kepada pihak penaja (JPA) yang telah membantu saya dari segi kewangan sepanjang saya melanjutkan pengajian di UKM ini dan juga kepada semua yang terlibat secara langsung dan tidak langsung dalam memberikan sumbangan cadangan dan bantuan dalam menyiapkan tesis ini.

Semoga penyelidikan dan tesis ini dapat dijadikan wadah ilmu yang berguna untuk tatapan generasi akan datang.

Pusat Sumber  
FTSM

## ABSTRAK

Penyelidikan sebelum ini telah mengintegrasikan maklumat teks dengan tugas visual. Walau bagaimanapun, beberapa penambahbaikan masih boleh dibuat kerana masih terdapat penyelidikan yang mengabaikan pengetahuan latar belakang luas yang hanya boleh disokong oleh maklumat teks bagi imej. Graf pengetahuan (GP) adalah teknik yang sesuai yang boleh dilaksanakan untuk mengeskrak pengetahuan dan mengubah teks yang tidak berstruktur menjadi lebih berstruktur. Juga, mampu menggambarkan hubungan antara teks. Untuk turut sertakan hubungan antara imej dan teks, GP multimodal dicadangkan. Penyelidik lain yang telah membina GP multimodal dalam penyelidikan mereka hanya memberi tumpuan kepada proses visual lanjutan untuk mengekstrak entiti dan hubungan dari imej dan hanya menggunakan kaedah pemprosesan teks yang ringkas dan mudah. Contohnya, teknik tokenisasi dan penandaan *part-of-speech* (POS) yang hanya mengekstrak kata nama dalam teks. Secara tidak langsung mengabaikan maklumat lain yang kaya dalam teks itu sendiri. Oleh itu, pendekatan yang dicadangkan ini cuba menangani masalah tersebut dengan mengeskrak secara sepenuhnya maklumat sedia ada dalam membina GP multimodal. Seterusnya, GP ini dapat dijadikan sebagai latihan bagi tugas visual. Sebelum membina GP multimodal, rangkap tiga (yang terdiri daripada dua entiti dan satu hubungan yang menghubungkan kedua entiti tersebut) perlu dikenal pasti dan diekstrak terlebih dahulu. Kaedah ini dilakukan dengan menentukan peraturan pengekstrakan entiti dan hubungan berdasarkan penandaan POS dan kebergantungan tatabahasa. Pengekstrak rangkap tiga berdasarkan peraturan ini dibangunkan untuk mengekstrak rangkap tiga dari set data laman web berita dan mengisi GP multimodal yang dicadangkan. Kemudian, pencapaian pengekstrak rangkap tiga dinilai menggunakan formula ketepatan (*Precision*) dan perolehan kembali (*Recall*). Hasilnya, pendekatan yang dicadangkan ini dapat mengekstrak entiti dan hubungan yang wujud dalam set data dengan skor ketepatan 0.90 dan skor perolehan kembali 0.60. Namun, beberapa penambahbaikan masih dapat dilakukan terhadap peraturan pengekstrakan untuk menangkap semua pengetahuan.

## **RULE-BASED TEXT EXTRACTION TO ENHANCE KNOWLEDGE GRAPHS FOR IMAGE UNDERSTANDING**

### **ABSTRACT**

Previous research works have integrated textual information to visual tasks. However, some improvement still can be made as there are still work that ignore the wide background knowledge that text information can provide for the image. Knowledge graph (KG) is suited to capture the knowledge and transform the unstructured text into a structured one as well to show the relations between the texts. To preserve the relation between the image and text, a multimodal KG is proposed. Existing multimodal KG works have mainly focused on advanced visual processes for extracting entities and relations from images, and only employed standard text processing techniques such as tokenization and POS tagging which mostly captured the nouns in the text; indirectly neglecting other rich information in the text itself. Thus, the proposed approach attempts to address these issues by capturing to the fullest the existing information in building the multimodal KG which eventually can be used as a training seed for visual tasks. Before building the multimodal KG, triples need to be identified and extracted first. A triple consists of two entities and a relation which relates the two entities. Entity relation extraction rules are formulated based on part of speech (POS) tagging and grammar dependencies on a set of web news articles consisting of news images and their related text. A triple extractor incorporating the formulated rules, is developed to extract the triples from the news articles dataset and populate the proposed multimodal KG. The performance of the extractor is evaluated using the Precision and Recall metrics. The evaluation results show that the proposed approach is able to extract the entities and relations that exist in the dataset with the precision score of 0.90 and recall score of 0.60. While the results are promising, further improvement still can be made to the extraction rules in order to capture all the knowledge.

## KANDUNGAN

		<b>Halaman</b>
<b>PENGAKUAN</b>		<b>iii</b>
<b>PENGHARGAAN</b>		<b>iv</b>
<b>ABSTRAK</b>		<b>v</b>
<b>ABSTRACT</b>		<b>vi</b>
<b>KANDUNGAN</b>		<b>vii</b>
<b>SENARAI JADUAL</b>		<b>ix</b>
<b>SENARAI ILUSTRASI</b>		<b>x</b>
<b>BAB I</b>	<b>PENGENALAN</b>	
1.1	Pengenalan	1
1.2	Penyataan Masalah	6
1.3	Objektif kajian	8
1.4	Skop Kajian	8
1.5	Kesimpulan	9
<b>BAB II</b>	<b>KAJIAN SUSASTERAAN</b>	
2.1	Pengenalan	11
2.2	Graf pengetahuan (GP)	11
2.3	GP Multimodal - Gabungan Teks dan visual	15
2.4	Pengekstrakan hubungan	18
2.5	Metrik Pengukuran Pengekstrakan Teks	24
2.6	Hubungan Kebergantungan	26
2.7	Korpus	29
2.8	Kesimpulan	31
<b>BAB III</b>	<b>METODOLOGI</b>	
3.1	Pengenalan	33
3.2	Motivasi Kajian	33
3.3	Gambaran Rangka Kerja	34
	3.3.1 Fasa 1 (Pra proses dataset)	36
	3.3.2 Fasa 2 (Ekstrak entiti dan hubungan)	40



	3.3.3 Fasa 3 (Membina GP multimodal)	53
3.4	Kesimpulan	55
<b>BAB IV</b>	<b>DAPATAN KAJIAN</b>	
4.1	Pengenalan	56
4.2	Dataset	56
4.3	Pra-proses dataset	58
4.4	Penjanaan Rangkap Tiga	60
4.5	Perbandingan Kaedah Pengekstrakan Hubungan	64
4.6	Pembinaan GP multimodal	79
4.7	Kesimpulan	82
<b>BAB V</b>	<b>RUMUSAN DAN CADANGAN</b>	
5.1	Pengenalan	84
5.2	Kesimpulan Kajian	84
5.3	Sumbangan Kajian	85
5.4	Kekangan Kajian	87
5.5	Kajian Masa Depan	88
<b>RUJUKAN</b>		93
<b>LAMPIRAN</b>		
Lampiran A	Jadual penuh perbandingan antara dua kaedah lain	98
Lampiran B	Jadual ringkasan perbandingan antara dua kaedah lain (Dalam bentuk digit)	132
Lampiran C	Jadual Kekerapan Kata Ganti Nama ( <i>Pronoun</i> ) Dalam 109 Ayat Input	134

**SENARAI JADUAL**

<b>No. Jadual</b>		<b>Halaman</b>
Jadual 2.1	Contoh pengekstrakan maklumat bagi setiap jenis ayat	18
Jadual 2.2	Perbandingan teknik pengekstrakan berdasarkan peraturan dan berdasarkan pembelajaran mesin	21
Jadual 2.3	Kategori satu (Hubungan predikat klausa)	28
Jadual 2.4	Kategori empat (Hubungan pengubah kata nama)	28
Jadual 2.5	Jadual perbandingan kaedah-kaedah terhadap cara teks diproses	32
Jadual 3.1	Pelabelan untuk mengekstrak maklumat dari ayat mudah dan juga kompleks	42
Jadual 3.2	Hubungan kebergantungan dan keterangannya	45
Jadual 3.3	Senarai penandaan POS dan keterangannya	46
Jadual 3.4	Ringkasan komponen yang dikenal pasti bagi setiap peraturan	52
Jadual 3.5	Contoh pengekstrakan rangkap tiga mengikut jenis ayat	53
Jadual 4.1	Jadual perbandingan untuk beberapa artikel	67

## SENARAI ILUSTRASI

<b>No. Rajah</b>		<b>Halaman</b>
Rajah 1.1	Keratan akhbar dari News Straits Times (NST)	2
Rajah 1.2	Contoh beberapa topik yang dapat dicapai berkenaan dengan kata kunci <i>Stadium</i>	4
Rajah 1.3	Contoh untuk hubungan imej-teks berkenaan dengan kata kunci <i>Stadium</i>	5
Rajah 2.1	Contoh GP paling mudah	12
Rajah 2.2	Graf pengetahuan yang dihasilkan oleh P. Wu et al. (2018)	14
Rajah 2.3	Graf pengetahuan yang telah dicadangkan oleh Gong & Wang (2017)	16
Rajah 2.4	Teks yang diekstrak sebagai tag imej	17
Rajah 3.1	Rangka kerja bagi pembinaan GP multimodal	35
Rajah 3.2	Contoh artikel berita	37
Rajah 3.3	Sekumpulan perkataan yang harus dianggap sebagai token tunggal	38
Rajah 3.4	Perkataan yang telah dikumpulkan bersama	38
Rajah 3.5	Contoh perkataan <i>left</i> yang perlu dibuang	39
Rajah 3.6	Ayat mudah	41
Rajah 3.7	Ayat kompleks	41
Rajah 3.8	Maklumat yang harus diekstrak dari ayat mudah	43
Rajah 3.9	Maklumat yang harus diekstrak dari ayat kompleks	43
Rajah 3.10	Pokok kebergantungan yang dihasilkan oleh <i>SpaCy Dependency parser</i>	44
Rajah 3.11	Peraturan 1 (hubungan kata kerja)	49
Rajah 3.12	Peraturan 2 (hubungan kata kerja + preposisi)	50
Rajah 3.13	Peraturan 3 (hubungan <i>has-Attribute</i> )	51
Rajah 3.14	Peraturan 4 (hubungan <i>has-Image</i> )	51

Rajah 3.15	GP multimodal yang akan dihasilkan setelah kesemua rangkap tiga digabungkan	54
Rajah 3.16	Contoh GP multimodal yang mempunyai subgraf tergantung	55
Rajah 4.1	Contoh artikel berita yang lengkap dengan imej dan kapsyen	57
Rajah 4.2	Contoh data yang diekstrak dari fail JSON	58
Rajah 4.3	Kod bagi penyesuaian <i>tokenizer</i> dan <i>stopword removal</i>	59
Rajah 4.4	Pokok kebergantungan yang dihasilkan	60
Rajah 4.5	Kod bagi pengekstrakan <i>multiword entity</i>	61
Rajah 4.6	Kod bagi Peraturan 3 (mengekstrak entiti bagi hubungan <i>has-Attribute</i> )	61
Rajah 4.7	<i>Main entity</i> yang dikenal pasti	62
Rajah 4.8	Kod bagi Peraturan 4 (mengekstrak entiti bagi hubungan <i>has-Image</i> )	62
Rajah 4.9	Kod bagi Peraturan 1 (pengekstrakan hubungan kata kerja sahaja)	63
Rajah 4.10	Peraturan 2 (pengekstrakan gabungan hubungan kata kerja dan kata preposisi)	63
Rajah 4.11	Rangkap tiga diekstrak dari ayat input	64
Rajah 4.12	Rangkap tiga diekstrak dalam bentuk jadual	64
Rajah 4.13	Contoh artikel	75
Rajah 4.14	Contoh token yang dilabel tidak tepat	76
Rajah 4.15	Contoh token yang tidak sepatutnya dilabel <i>ROOT</i>	77
Rajah 4.16	Cebisan kod untuk visualisasikan GP	80
Rajah 4.17	GP setiap artikel	81
Rajah 4.18	Contoh GP yang diperbaiki	82
Rajah 5.1	Contoh graf tidak berarah dan imej ID yang mewakili imej	88
Rajah 5.2	Contoh ayat yang mempunyai hubungan kata preposisi	89
Rajah 5.3	Contoh GP multimodal yang mempunyai imej artikel	90
Rajah 5.4	Contoh graf yang berarah (directed)	91

Rajah 5.5	Contoh graf yang boleh dihasilkan dengan konsep imej yang sama	92
Rajah 5.6	Contoh graf yang seakan gaya <i>word cloud</i>	92

Pusat Sumber  
FTSM

## **BAB I**

### **PENGENALAN**

#### **1.1 PENGENALAN**

Pada masa kini, web telah menjadi salah satu sumber maklumat yang penting di dunia. Imej di web terdiri daripada maklumat tekstual yang bernilai yang boleh digunakan dalam aplikasi yang berbeza seperti anotasi imej, pengelompokan imej, pemahaman imej dan lain-lain. Maklumat ini dipanggil maklumat kontekstual yang merupakan teks yang berkaitan dengan imej (Tiwari 2014). Maklumat kontekstual kaya dengan konsep semantik peringkat tinggi dan mengandungi kedua-dua maklumat langsung dan tak langsung mengenai imej (Chan, Johar & Hong 2013).

Seseorang boleh mencari imej web dalam artikel berita, blog peribadi, laman web e-dagang dan lain-lain. Kebiasaannya, dalam artikel berita dan blog peribadi, seseorang boleh mencari halaman web dengan beberapa imej web diikuti dengan teks huraian untuk setiap imej. Teks inilah yang dikatakan sebagai maklumat kontekstual yang dinyatakan oleh Tiwari (2014).

Menurut Wang et al. (2016), terdapat jurang antara analisis tekstual dan visual untuk pemahaman imej. Sebagai contoh, algoritma sedia ada seperti rangkaian saraf yang mendalam, hanya tertumpu kepada perkara-perkara tertentu yang menggunakan ciri dalam imej itu sendiri, biasanya mengabaikan pengetahuan latar belakang yang luas tentang dunia nyata (Fang et al. 2017). Imej web seringkali dikaitkan dengan deskriptif tekstual yang kebiasaan berkait rapat dengan isi kandungan imej. Deskripsi tekstual banyak mengandungi pengetahuan latar belakang luas terutama untuk artikel berita dan blog.

Rajah 1.1 merupakan contoh keratan akhbar yang menunjukkan sebuah imej bersama huraianya. Huraian ini merupakan maklumat kontekstual bagi sebuah imej dan secara tidak langsung memberi pengetahuan latar belakang mengenai sesebuah imej. Contohnya, keterangan imej berikut menjelaskan dengan lebih jelas situasi yang telah berlaku di Larkin Sentral. Jika lihat pada imej sahaja, tidak menjelaskan keadaan secara mendalam. Sekiranya pengetahuan latar belakang yang luas ini iaitu dapat diperoleh dengan cara yang berstruktur, maka sebuah pangkalan pengetahuan dapat dihasilkan seterusnya boleh diterapkan dalam tugas pemahaman imej.

#### Larkin Sentral undergoes 9-hour sanitisation operation



Rajah 2.1 Keratan akhbar dari News Straits Times (NST)

Terdapat kajian mengenai analisis visual dan teks yang digabungkan bersama seperti (Vondrick et al. 2016; Wang et al. 2016). Wang et al. (2016) mencadangkan algoritma untuk mempelajari hubungan antara pemandangan, objek, dan teks dengan bantuan anotasi peringkat gambar. Kaedah yang dicadangkan adalah cekap dalam mewujudkan hubungan antara deskripsi teks dan konsep visual. Para penulis menunjukkan bahawa teks memberikan pelbagai konsep dunia nyata daripada objek kepentingan. Manakala Vondrick et al. (2016) memperkenalkan dataset baru orang yang sedang melakukan tindakan untuk meramalkan motivasi. Diberi imej orang menunggang basikal. Penulis mencadangkan model yang dapat meramalkan mengapa mereka menunggang basikal dengan memindahkan pengetahuan yang diperoleh dari

teks ke dalam visi komputer. Hasilnya disahkan bahawa ia dapat meramalkan mengapa seseorang melakukan tindakan yang lebih baik daripada pendekatan visi biasa sahaja.

Kedua-duanya menunjukkan bahawa teks memainkan peranan dalam menyokong bahagian visual kerana sifat teks tersebut yang mempunyai latar belakang yang luas. Walau bagaimanapun, Wang et al. (2016) dan Vondrick et al. (2016) lebih menumpukan pada bahagian visual dan bukannya teks kerana deskripsi yang digunakan merupakan teks yang diambil dari pakar.

Tambahan lagi, Chu et al. (2019) mencadangkan model dapatan semula multimodal untuk peroleh kembali imej (*image recall*) dalam gaya hidup teks-imej yang saling berkaitan. Apabila imej tidak dapat memberikan keseluruhan maklumat, maklumat teks adalah pelengkap untuk menggambarkan pengalaman hidup. Penulis menyatakan bahawa maklumat kontekstual berdasarkan teks seperti kapsyen dan keterangan yang berkaitan dengan gambar, menjadikannya iringan penting. Oleh itu, menjadikan teks-imej yang saling berkaitan menjadi maklumat yang lebih kaya untuk penarikan semula imej. Menurut Vondrick et al. (2016), hanya bergantung kepada imej sahaja (iaitu ciri visual) tidak mencukupi untuk mengekstrak makna tersirat imej.

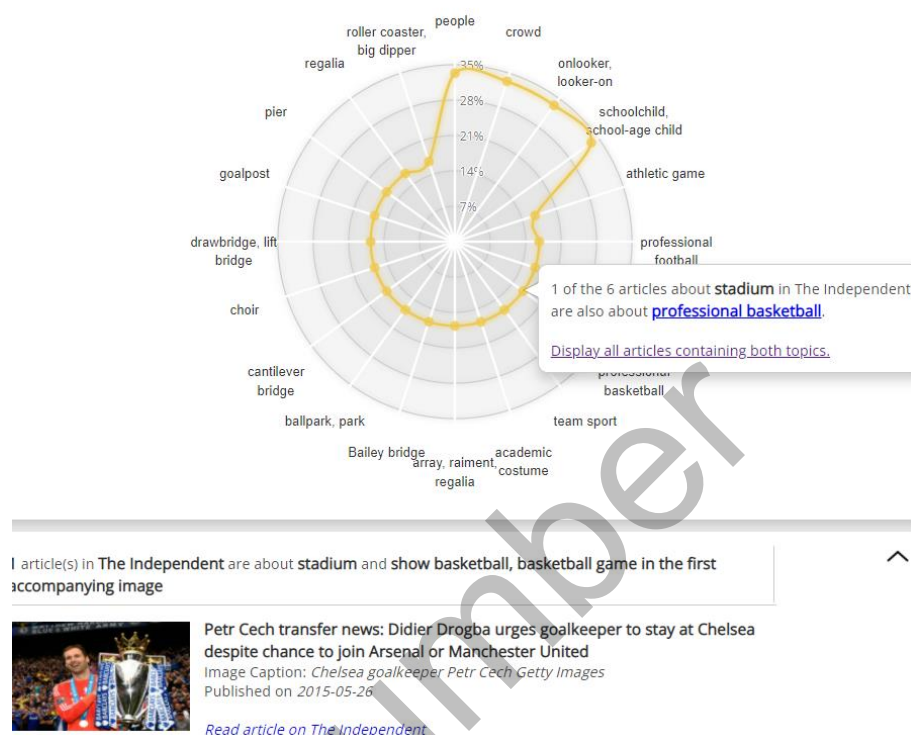
Seperti yang dinyatakan dalam kajian (Chu et al. 2019; Vondrick et al. 2016; Wang et al. 2016), analisis antara visual dan teks telah digabungkan bersama tetapi penulis-penulis ini lebih menumpukan pada bahagian visual dan bukannya teks kerana deskripsi yang digunakan diambil dari pakar. Walaupun dari segi pertalian imej-teks sangat tinggi, tapi anotasi dari pakar mempunyai kelemahan iaitu memakan masa dan juga mahal.

Karya Hollink et al. (2016) telah membuktikan kepentingan imej untuk berada dalam konteks tekstual semulajadi mereka, dimana imej diambil dari web berita dan kebiasaan akan diikuti oleh deskripsi mengenai sesebuah imej tanpa perlu diannotasi dari pakar dan juga kaya dengan makna, sama seperti yang diperhatikan pada keratan akhbar NST di Rajah 1.1. Hollink et al. juga menyatakan bahawa artikel berita mempunyai hubungan semulajadi antara teks dan imej.



Namun begitu, pengelas visual yang dipaparkan oleh Hollink et al. masih lagi terdapat batasan kes di mana imej-imej yang dijelaskan dalam output pengelas visual mereka masih lagi mempunyai ruang untuk diperbaiki. Kaedah yang digunakan oleh penulis untuk memproses teks adalah dengan menggunakan teknik *entity linking* untuk mendapatkan topik dan kategori yang dianotasi dimana beberapa topik berkait dapat dicapai berkenaan dengan sebuah kata kunci iaitu “Stadium” seperti di Rajah 1.2 berikut, tetapi ini hanya kukuh untuk hubungan teks-teks yang diekstrak daripada artikel berita. Dalam Rajah 1.3 merupakan contoh untuk hubungan imej-teks tetapi teks yang dipaparkan tidak berkait dengan imej yang dipaparkan.





Rajah 2.3 Contoh untuk hubungan imej-teks berkenaan dengan kata kunci *Stadium*

Berdasarkan Li et al. (2018), penulis menyatakan bahawa graf pengetahuan (GP) adalah teknik terbaik untuk mengekstrak pengetahuan dari sejumlah besar teks dan imej dan juga untuk menerangkan pelbagai entiti dan konsep yang wujud di dunia nyata, serta hubungan antara entiti dan konsep. Dalam erti kata lain, graf pengetahuan memudahkan penemuan maklumat dengan menganjurkannya kepada entiti dan menggambarkan hubungan antara entiti tersebut antara satu sama lain dan menubuhkan ontologi. Unsur penting graf pengetahuan adalah entiti dan hubungan (Auer & Mann 2019). Dua entiti dan hubungan antara entiti tersebut dikenali sebagai rangkap tiga (*triple*) yang merupakan binaan asas untuk GP.

Disimpulkan di sini, kajian-kajian sedia ada telah membuktikan bahawa pentingnya teks dalam tugas visual kerana ia menyediakan tentang pengetahuan latar belakang yang luas sebuah imej tidak dapat menyediakan. Juga, adalah lebih baik untuk teks itu diambil dari persekitaran semulajadi mereka seperti teks di dalam artikel berita. Pengetahuan yang sedia ada ini perlu diekstrak secara berstruktur dan GP merupakan struktur data yang mampu menangkap konsep dunia nyata dan hubungannya. Pangkalan pengetahuan berstruktur ini kemudian dapat digunakan

menjadi model latihan dalam tugas visual seperti pemahaman imej, pengesanan objek dan sebagainya.

## 1.2 PENYATAAN MASALAH

Dapat lihat bahawa walaupun adanya karya-karya yang melibatkan pemrosesan teks dimana proses hanya melibatkan anotasi terhadap bahagian teks sama ada didapati dari pakar ataupun dari teknik pemrosesan bahasa tabii (NLP) dalam analisis visual dan hasil eksperimen mereka berjaya tetapi bahagian ini masih diambil secara mudah. Disebabkan ketidakseimbangan dalam penggunaan teks dimana karya sedia ada lebih menumpukan pada bahagian visual dan kurangnya tumpuan dalam kaedah teks diproses, pembinaan GP dapat mengimbangi penggunaan kedua-dua elemen tersebut. Selain itu, dapat membantu dalam kajian yang melibatkan tekstual dan visual dengan meningkatkan lagi ketepatan hasil bagi tugas yang melibatkan dua elemen tersebut.

Beberapa karya yang ada seperti (Exner & Nugues 2012; Fang et al. 2017; P. Wu et al. 2018) melibatkan penggunaan GP dalam karya mereka sama ada graf yang dihasilkan sendiri atau juga graf yang sedia ada. Namun begitu, graf yang dicadangkan ini berasingan teks dan juga bahagian visual (imej).

Gong & Wang (2017) juga telah hasilkan GP dalam karya mereka di mana graf yang dibina merupakan graf pengetahuan multimodal iaitu graf pengetahuan yang digabung teks dan juga imej. Cara penulis ini memproses teks adalah dengan mengikuti langkah-langkah standard untuk pra-pemrosesan yang meliputi membersihkan teks artikel untuk mendapatkan teks biasa kemudian diikuti dengan token dan menerapkan penandaan POS pada teks. Kemudian, frasa nama dan kata nama sahaja diekstrak sebagai entiti dan berhenti begitu sahaja setelah objektif tercapai.

Kaedah yang digunakan ini masih lagi mempunyai ruang untuk penambahbaikan kerana bahagian pemrosesan teks masih lagi tidak memuaskan, dalam erti kata lain kaedah yang digunakan hanya menangkap kata nama atau frasa nama. Kekurangan kaedah ini menyebabkan pengabaian maklumat lain yang kaya seperti kata kerja atau kata sifat yang menerangkan tentang sesebuah imej, secara

tidak langsung, tidak menggunakan sepenuhnya pengetahuan latar belakang imej yang luas yang boleh diperoleh daripada teks tidak berstruktur.

Berdasarkan Rajah 1.1, teks yang terdapat dalam artikel berita ini boleh menjadi ayat pendek dan mudah atau ayat panjang dan kompleks. Ayat mudah ialah tajuk artikel berita, lazimnya ayat pendek yang hanya mempunyai satu kata kerja setiap ayat manakala ayat kompleks ialah kapsyen iaitu ayat panjang yang boleh terdiri daripada beberapa kata kerja dalam satu ayat. Bagi menunjukkan contoh mengikut kaedah yang digunakan Gong & Wang (2017), tajuk berita iaitu "*Larkin Sentral undergoes 9-hour sanitisation operation*" yang boleh dikategorikan sebagai ayat mudah, hanya frasa nama '*Larkin Sentral*' atau '*sanitisation operation*' sahaja yang akan diekstrak sebagai entiti.

Bagi membina GP, entiti dan hubungan adalah elemen penting yang diperlukan. Maka, teknik pengekstrakan maklumat perlu diterapkan. Seperti yang telah dinyatakan, karya sebelum ini hanya menggunakan teks yang mudah maka pendekatan teknik pengekstrakan hubungan boleh diaplikasikan untuk mengekstrak hubungan dan seterusnya entiti dari ayat yang lebih kompleks. Tugas pengekstrakan hubungan adalah pengesanan dan pengelasan hubungan semantik antara pasangan entiti (She et al. 2018).

Disimpulkan disini, teks penting dalam tugas visual kerana ia menyediakan tentang pengetahuan latar belakang yang luas sebuah imej tidak dapat sediakan. Beberapa kajian yang melibatkan tekstual dan visual hanya tumpu kepada bahagian visual. Maka penggunaan graf pengetahuan dapat seimbangkan tekstual dan visual, sekaligus menjadikan teks lebih berstruktur. Namun, GP yang sedia ada berasingan teks dan imej. GP multimodal pula menggabungkan antara dua elemen tersebut tetapi kajian sedia ada hanya menggunakan teks yang mudah untuk bahagian tekstual mereka.

Jelas bahawa masalah yang telah ditemui ialah pengetahuan latar belakang (teks) perlu disusun secara berstruktur sekaligus tidak menghiraukan hubungan imej- teks. Kedua, penggunaan teks yang mudah (kata nama) dalam pembinaan graf

pengetahuan multimodal menyebabkan pengabaian maklumat kaya lain yang boleh didapati dari teks..

Oleh itu, dengan menggunakan imej yang diiringi maklumat teks di dalam artikel berita bagi membina graf pengetahuan multimodal dan juga teknik pengekstrakan hubungan dapat mengatasi masalah yang telah dinyatakan.

### 1.3 OBJEKTIF KAJIAN

Matlamat keseluruhan adalah untuk menggambarkan pengetahuan latar belakang imej dengan cara yang lebih berstruktur daripada teks yang tidak berstruktur melalui pembinaan graf pengetahuan multimodal dengan mencapai objektif berikut iaitu:

1. Untuk menentukan peraturan pengekstrakan hubungan entiti berdasarkan kebergantungan tatabahasa untuk ayat yang mudah dan kompleks serta hubungan yang wujud antara imej dan teks. Entiti boleh menjadi entiti teks atau entiti imej.
2. Untuk membangunkan pengekstrak rangkap tiga bagi mengisi graf pengetahuan multimodal yang dicadangkan.
3. Untuk menilai hasil rangkap tiga yang diekstrak menggunakan formula perolehan kembali (*Recall*) dan ketepatan (*Precision*).

### 1.4 SKOP KAJIAN

Kajian ini secara khusus fokus dalam domain berita dan menekankan pada analisis teks berbanding bahagian visual. Domain berita menjadi fokus kerana kandungan teks yang mengelilingi imej dalam sebuah artikel berita mempunyai hubungan yang semulajadi antara imej dan teks.

Kajian ini juga menggunakan teks yang terdapat dalam artikel berita berbahasa Inggeris yang merangkumi ayat pendek dan mudah atau ayat panjang dan kompleks. Bagi skop pemilihan artikel, kajian ini mempertimbangkan imej dan teks, artikel yang tidak mempunyai imej akan digugurkan. Jika sesebuah artikel mempunyai imej tapi

tidak tersedia bersama kapsyen, artikel masih diterima kerana ada tajuk berita yang mewakili teks.

## 1.5 KESIMPULAN

Bab ini telah membincangkan serba sedikit mengenai kepentingan teks disamping imej dan juga kajian sedia ada yang melibatkan analisis tekstual dan juga visual. Pernyataan masalah mengenai bahagian analisis visual yang lebih difokuskan berbanding tekstual menyebabkan pengabaian latar belakang luas dari teks dan juga penggunaan teks yang mudah telah dikenal pasti di dalam bab ini. Objektif kajian juga dinyatakan bagi mencapai matlamat menggambarkan pengetahuan latar belakang dengan penggunaan graf pengetahuan multimodal. Skop kajian ini lebih menekankan pada analisis teks berbanding bahagian visual.

Bab 2 akan membincangkan mengenai GP yang sedia ada dan penggunaannya dalam dunia nyata. Selain itu, isu yang berkait dengan multimodal GP juga diketengahkan selain menjadi motivasi kepada kajian yang dijalankan ini. Kaedah yang akan digunakan dalam kajian ini iaitu teknik pengekstrakan hubungan dibincangkan dalam bab ini. Justifikasi terhadap kaedah yang telah dipilih juga diterangkan.

Bab 3 membincangkan dengan lebih mendalam mengenai cara GP multimodal dibangunkan. Rangka kerja ini mengandungi 3 fasa utama yang terdiri daripada Fasa 1: Pra proses dataset, Fasa 2: Ekstrak entiti dan hubungan, Fasa 3: Membina GP multimodal. Terdapat 3 sub-fasa dalam Fasa 2 iaitu Fasa2.1: Pra analisis data, Fasa 2.2: Pembangunan peraturan dan Fasa 2.3: Ekstrak berdasarkan peraturan. Maka akan terhasilnya sebuah GP multimodal yang menggambarkan hubungan antara teks-teks dan juga teks-imej.

Bab 4 menerangkan tentang hasil daripada eksperimen yang telah dijalankan. Dataset yang digunakan dalam eksperimen ini ialah dataset artikel berita dimana ia perlu menjalani pra-proses sebelum digunakan untuk mengekstrak entiti dan hubungan. Bagaimana rangkap tiga dan GP multimodal yang dihasilkan juga

dibincangkan. Setelah itu, perbandingan dilakukan terhadap dua kaedah dari karya lain seperti Gong & Wang (2017) dan Romadhony et al. (2018).

Bab 5 meringkaskan bab-bab yang sebelumnya dan juga sumbangan yang telah dilakukan oleh kajian ini. Beberapa kekangan yang telah dihadapi juga dibincangkan seperti batasan output dari penghurai (*parser*) yang digunakan, hasil GP multimodal hanya menggunakan imej ID dan bukan imej yang sebenar dan lain-lain. Perkara-perkara yang boleh diperbaiki di masa hadapan juga telah dinyatakan.

Pusat Sumber  
FTSM

## BAB II

### KAJIAN SUSASTERAAN

#### 2.1 PENGENALAN

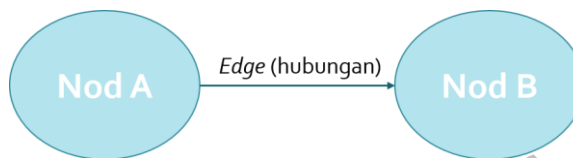
Bab ini membincangkan beberapa karya yang menggabungkan pengetahuan luar yang diwakili graf pengetahuan (GP) ke mesin telah membuktikan bahawa pentingnya gabungan ini untuk membantu mesin berfikir sama seperti manusia. Selain GP yang terhasil dari teks, terdapat juga GP multimodal iaitu gabungan dari beberapa elemen yang lain juga telah dibincangkan. Pendekatan teknik pengekstrakan hubungan juga akan diterangkan dengan lebih lanjut mengenai bagaimana entiti dan hubungan akan diekstrak yang akhirnya akan menghasilkan GP. Selain itu, topik kepentingan teks untuk berada dalam konteks tekstual semulajadi mereka juga akan disentuh.

#### 2.2 GRAF PENGETAHUAN (GP)

Secara ringkasnya, graf pengetahuan (GP) merupakan graf berlabel terarah yang terdiri daripada nod dan *edge* seperti di Rajah 2.1. Nod merupakan konsep dunia nyata dan selain daripada teks, imej juga boleh dijadikan sebagai nod. *Edge* pula menghubungkan sepasang nod dan menunjukkan hubungan antara nod (Chaudhri, 2021). Nod juga boleh dikenali sebagai entiti dan *edge* yang menghubungkan kedua-dua entiti ini dikenali sebagai hubungan. GP juga boleh mewakili koleksi huraian entiti yang saling berkait antara satu sama lain. Sebilangan besar GP sedia ada telah dibina seperti *Google Knowledge Graph*, *DBPedia*, *Wordnet*, *ConceptNet*. Graf sedia ada ini telah digunakan dalam banyak aplikasi dunia nyata dan secara ringkas, sebahagian daripadanya telah



digunakan dalam pengesanan objek, tugas menjawab soalan, klasifikasi imej dan banyak lagi.



Rajah 2.1 Contoh GP paling mudah

(Pan et al. (2019); Q. Wu et al., (2016)) telah membuktikan bahawa pentingnya untuk menggabungkan pengetahuan luar ke mesin untuk mengajuk bagaimana manusia menangani tugas jawapan pertanyaan (QA). Berdasarkan Vondrick et al. (2016), dengan menggabungkan pengetahuan bersama dari sumber lain secara tidak langsung dapat membantu mesin berfikir sama seperti manusia. Perwakilan pengetahuan berasaskan ontologi memberikan pengetahuan yang lebih berkembang tentang persekitaran untuk memproses tugas atau kaedah. Ontologi juga membantu dalam mewakili pengetahuan tentang persekitaran, peristiwa dan tindakan yang membantu dalam perancangan laluan dan membuat robot lebih autonomi (Gayathri & Uma 2018).

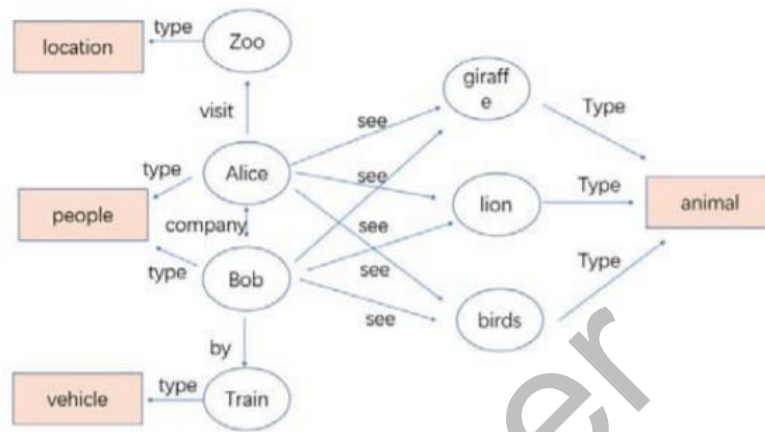
Penggunaan semula graf pengetahuan menandakan analisis yang masuk akal yang mirip dengan otak manusia, dengan kekuatan penalaran dan penghakiman logik. Oleh itu, penggunaan sepenuhnya graf pengetahuan dapat membantu mensimulasikan pemikiran manusia dengan lebih banyak (P. Wu et al. 2018).

Terdapat karya yang menggabungkan GP dalam kaedah yang dicadangkan untuk mencapai tujuan mereka seperti dalam ramalan berdasarkan teks (S. Jiang, Zhai & Mei 2018), dalam pengesanan objek (Fang et al. 2017) dan melombong konsep yang mendasari dalam video (Gao, Zhang & Xu 2019). Semua karya ini menunjukkan bahawa GP dapat membantu dalam mengoptimumkan kerjasama antara manusia dan sistem pembelajaran mesin dalam aplikasi dunia nyata. Walau bagaimanapun, GP yang digunakan adalah berdasarkan dari graf yang sedia ada dan tidak menumpukan pada

pemrosesan teks dalam membina GP. Graf yang dibina dari daripada teks tidak berstruktur itu sendiri akan mengandungi lebih banyak hubungan semula jadi.

Beberapa karya yang ada seperti (Exner & Nugues 2012; Kertkeidkachorn & Ichise 2017), kedua-duanya berjaya membina GP dari teks tidak berstruktur. Bagi membina graf ini, perkara asas yang diperlukan terlebih dahulu ialah elemen-elemen berikut iaitu entiti dan hubungan. Entiti adalah subjek atau objek dalam ayat manakala hubungan pula menghubungkan antara subjek dan objek. Elemen tersebut juga dikenali sebagai rangkap tiga (*triples*). Kaedah mereka dicadangkan dengan memproses terlebih dahulu teks tidak berstruktur dari artikel wikipedia. Langkah seterusnya ialah pemetaan entiti. Kemudian, hubungan dijanakan. Exner & Nugues (2012) telah menggunakan teknik pengekstrakan maklumat terbuka. Hasil dari proses di atas kemudian digunakan untuk memproses rangkap tiga.

Bagi hasil karya P. Wu et al. (2018), mereka menghasilkan GP dari teks dalam artikel berita seperti yang digambarkan di Rajah 2.2. Proses untuk membina graf adalah setelah memproses teks, mereka mengenal pasti entiti teks dengan mempertimbangkan model campuran pepadanan corak dan rangkaian neural. Bagi pengekstrakan hubungan pula dilakukan melalui Snorkel (alat sumber terbuka), TextCNN dan pohon kebergantungan sintaktik (*syntactic dependence tree*) dimana mereka rujuk kepada model pengekstrakan hubungan (Fan et al. 2014) dan pasangan hubungan dengan kebarangkalian tertinggi akan dipilih dengan cara pembelajaran melalui pengawasan jarak jauh (*distant supervision*). Akhirnya, entiti dan hubungan digunakan untuk menghasilkan subgraf graf pengetahuan.



Rajah 2.2 Graf pengetahuan yang dihasilkan oleh P. Wu et al. (2018)

Namun begitu, disebabkan karya ini menumpukan kepada output ringkasan maka graf yang dihasilkan merupakan GP ringkasan walaupun input asal merupakan ayat yang kompleks. Intipati teks telah diekstrak tetapi secara tidak langsung, ia mungkin tidak mengekstrak maklumat lain yang terdapat dalam ayat yang panjang seperti perkataan yang digunakan untuk menerangkan perkataan atau mana-mana perkataan yang boleh digabungkan bersama sebagai satu entiti. Contoh berikut :

*“Two more young black men join in the beating, which is caught on cameras.”*

*“Two men who are young black and join fight.”*

Ayat pertama merupakan ayat input panjang asal yang diringkaskan kepada ayat kedua yang lebih pendek. Rangkap tiga dari ayat yang diringkaskan ini digunakan untuk membina GP. Dapat dilihat bahawa output berikut tidak menangkap frasa *“caught on cameras”* yang juga boleh menjadi entiti dan hubungan bagi set rangkap tiga yang lain.

*“Alice and Bob took the train to visit the zoo. They saw a baby giraffe, a lion, and a flock of colorful tropical birds.”* merupakan contoh lain bagi input ayat asal yang dianggap panjang. Output dalam graf tidak mengambil *baby giraffe* sebagai entiti atau huraian tentang *tropical birds* yang berwarna-warni. Maklumat seperti ini boleh dianggap berguna terutamanya dalam tugas pengesanan objek walaupun ia hanya deskriptif bagi istilah tertentu.

Oleh demikian, berikut merupakan contoh kesukaran dalam mengekstrak maklumat dari ayat kompleks. Segelintir maklumat seperti frasa *caught on cameras* secara tidak langsung telah terabai. Terdapat juga perkataan yang tidak dikesan sebagai satu token. Seperti frasa *baby giraffe* yang tidak dilabel sebagai satu entiti yang sama. Hal ini berlaku kerana ayat kompleks mempunyai klausa pautan dimana penghurai akan mempunyai isu untuk mengesan subjek atau objek dalam klausa utama. Maka akan berlakunya isu segelintir maklumat penting yang tidak diekstrak.

### 2.3 GP MULTIMODAL - GABUNGAN TEKS DAN VISUAL

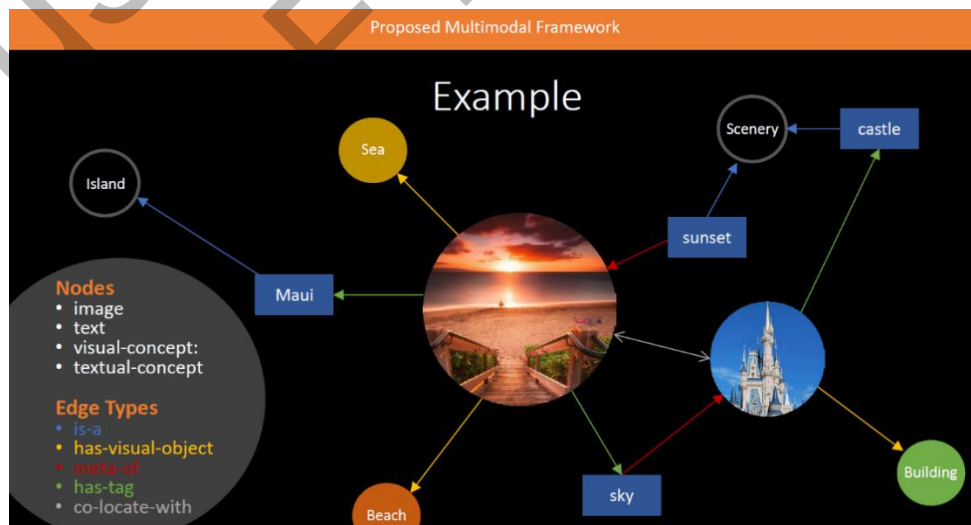
Seperti yang telah dinyatakan mengenai hasil karya Fang et al. (2017), mereka telah menggunakan graf pengetahuan dari sumber orang ramai iaitu MIT ConceptNet disebabkan lebih menumpukan bagaimana graf tersebut akan digunakan ke dalam visual pengesanan objek. Fang et al. telah menggabungkan graf dengan mengoptimumkan semula model pengesanan objek.

Selain itu, karya-karya berikut telah menghasilkan graf multimodal dimana entiti melibatkan bukan sahaja teks malah boleh merangkumi imej atau audio dan hubungan yang menghubungkan entiti sama ada dalam atau merentasi modaliti. Gong et al. (2017) telah mengemukakan pendekatan analisis multimodal untuk pengekstrakan maklumat teks. Kaedah yang dicadangkan mengkaji mengenai peraturan multimodal untuk meningkatkan kebolehpercayaan peraturan sintaktik. Gong et al. menyatakan bahawa maklumat multimodal seperti teks, gambar atau audio biasanya saling berkait dan melengkapi antara satu sama lain. Modaliti teks dan gambar difokuskan kerana kebolehsediaan maklumat yang tinggi.

Di samping itu, Gong & Wang (2017) mencadangkan algoritma pembelajaran multimodal untuk mengintegrasikan maklumat teks ke dalam pengekstrakan pengetahuan visual. Kaedah mereka menggunakan hubungan teks untuk meningkatkan ketepatan ramalan hubungan *has-visual-object* antara visual dan teks. Eksperimen penilaian menunjukkan bahawa jika dibandingkan dengan algoritma yang bergantung pada

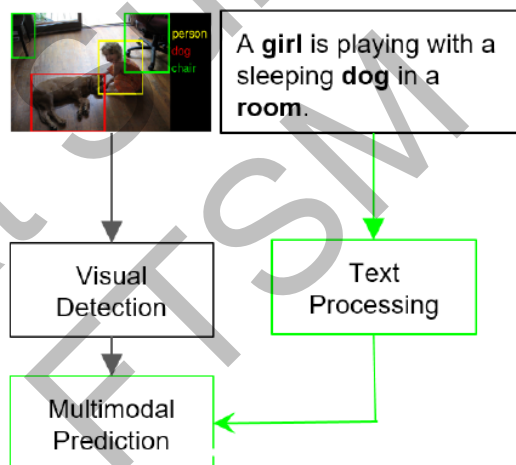
kandungan visual, algoritma multimodal meningkatkan ketepatan pengekstrakan pengetahuan visual dengan lebih ketara. Disini dapat lihat bahawa pertalian antara imej-teksts adalah sangat tinggi.

Namun begitu, kaedah Gong & Wang (2017) hanya ditumpukan pada cara teks diproses. Frasa teks dikenal pasti kerana ia berguna dalam dapatan semula imej (*image retrieval*). Kumpulan frasa teks adalah frasa nama yang diperolehi daripada teks yang sekeliling imej. Gong & Wang telah mengumpulkan laman web untuk mengekstrak teks dan imej. Setiap laman web akan melalui proses pembersihan untuk menghapus meta dan spam bagi mendapatkan hanya teks biasa sahaja. Langkah seterusnya adalah dengan tokenisasi dan menerapkan penandaan POS (*part-of-speech*) pada setiap teks. Secara ringkasnya, proses tokenisasi akan mengesan satu teks sebagai satu token dan penandaan POS pula menggambarkan struktur ciri istilah leksikal dalam ayat atau teks. Teks yang dilabel dengan kata nama kemudian diekstrak manakala untuk frasa nama pula, ia berdasarkan turutan kata nama. Akhirnya, teks ini akan dilabel sebagai tag untuk setiap imej. Seperti dalam Rajah 2.3, teks yang dilabel sebagai tag ini menjadi entiti kepada hubungan *has-tag* yang menghubungkan imej, secara tidak langsung terhasilnya graf pengetahuan multimodal dan mencapai objektif mereka.



Rajah 2.3 Graf pengetahuan yang telah dicadangkan oleh Gong & Wang (2017)

Artikel yang sedia ada ini telah kaitkan kedua bahagian visual dan teks dalam kaedah yang dicadangkan. Walaubagaimanapun, teks diproses secara mudah kerana ia sudah memadai untuk mencapai matlamat mereka yang membantu dalam meningkatkan keberkesanan pengetahuan visual yang diekstrak. Sebagaimana mereka hanya fokus pada hubungan *has-tag* yang hanya berkaitan dengan tag imej (kata nama atau frasa nama), secara tidak langsung mengabaikan maklumat kaya lain yang boleh didapati dari teks. Contohnya seperti dalam Rajah 2.4 , dalam ayat “A **girl** is playing with a sleeping **dog** in a **room**”, teks yang ditebalkan merupakan teks yang diekstrak sebagai tag imej. Maklumat seperti “*playing*” dan “*sleeping*” yang menjelaskan tentang situasi bagi kata nama “*girl*” dan “*dog*” tidak diekstrak. Jika teks tersebut diekstrak sekali, maka graf yang dihasilkan lebih lengkap dengan maklumat yang berkait dengan imej.



Rajah 2.4 Teks yang diekstrak sebagai tag imej

Atribut biasanya digunakan untuk menerangkan sesuatu objek. Menurut Krishna et al. (2017), atribut membolehkan untuk menerangkan, membandingkan dan mengkategorikan objek dengan mudah. Göring et al. (2014) telah membuktikan bahawa dengan penambahan atribut ini, terdapat peningkatan dalam tugas visual. Oleh itu, GP multimodal kini masih boleh dipertingkatkan lagi dengan mengisi lebih banyak maklumat yang boleh didapati dari teks ke dlm GP multimodal.

Jadual 2.1 berikut menunjukkan contoh pengekstrakan maklumat mengikut jenis ayat iaitu ayat mudah (mengandungi satu kata kerja) dan kompleks (mengandungi beberapa kata kerja). Contoh ayat mudah diambil dari karya (Gong & Wang, 2017) manakala ayat kompleks diambil dari (P. Wu et al., 2018). Maklumat yang diekstrak dalam ayat mudah hanya kata nama bagi dipasangkan kepada tag imej walaupun terdapat lagi maklumat latar belakang yang berkait dengan imej. Maklumat seperti kata kerja “*playing*” sepatutnya diekstrak dan juga atribut (*playing*) kepada kata nama “*dog*”, maka dapat digambarkan bahawa situasi yang sedang berlaku ialah seorang kanak perempuan sedang bermain dengan anjing yang sedang tidur (*girl -playing with - sleeping dog*). Bagi ayat kompleks pula dimana terdapat dua kata kerja dalam satu ayat, maklumat yang telah diekstrak merupakan ringkasan daripada ayat asal. Walaupun intipati ayat berikut telah diekstrak tetapi dapat dilihat bahawa output berikut tidak menangkap frasa “*caught on cameras*” yang juga boleh menjadi entiti dan hubungan bagi set rangkap tiga yang lain.

Jadual 2.1 contoh pengekstrakan maklumat bagi setiap jenis ayat

Jenis ayat	Contoh	Maklumat yang diekstrak	Penambahbaikan yang boleh dibuat
Ayat mudah dari (Gong & Wang (2017))	<i>A girl is playing with a sleeping dog in a room</i>	▪ ‘girl’, ‘dog’, ‘room’	▪ <i>girl -playing with - sleeping dog</i>
Ayat kompleks dari (P. Wu et al. (2018))	<i>Two more young black men join in the beating, which is caught on cameras.</i>	▪ ‘Two men who are young black and join fight’	▪ <i>young black men -caught on -cameras</i>

## 2.4 PENGEKSTRAKAN HUBUNGAN

Untuk membina GP dari teks, amat penting untuk memahami teks sebelum mengekstrak hubungan. Oleh itu membawa kepada tugas pengekstrakan hubungan (*relation extraction (RE)*). RE adalah sub-tugas utama pengekstrakan maklumat (Miao et al. 2019) dan juga merupakan pengesanan dan pengelasan hubungan semantik antara pasangan entiti (She et al. 2018).

Menurut Zhou, Qian & Fan (2010), entiti adalah objek atau sekumpulan objek, sementara hubungan adalah hubungan yang dinyatakan secara eksplisit atau tersirat antara

dua entiti. Output untuk tugas biasanya disajikan dalam bentuk (e1, r, e2) atau lebih dikenali sebagai rangkap tiga, di mana r menunjukkan hubungan, dan e1 dan e2 menunjukkan entiti yang mempunyai hubungan r antara mereka.

Terdapat jenis hubungan yang wujud antara perkataan iaitu 1) hubungan semantik dan 2) hubungan sintaksis. Hubungan semantik ialah hubungan yang wujud antara makna perkataan. Terdapat beberapa kategori untuk hubungan seperti ini seperti sinonim, hiponimi dan banyak lagi. Secara ringkas, sinonim di mana dua perkataan mempunyai makna yang sama. Hiponimi bermaksud makna sesebuah itu termasuk dalam makna yang lain (Davies & Elder, 2004).

Hubungan yang lain pula merupakan hubungan sintaksis juga setara dengan hubungan tatabahasa, ini adalah hubungan tatabahasa yang wujud antara pasangan kata, tepatnya antara predikat dan frasa nama. Beberapa contoh hubungan adalah subjek, objek langsung, objek tidak langsung dan banyak lagi (Kibrik, 2001). Maka, kajian ini akan menggunakan hubungan sintaksis yang wujud secara semula jadi antara pasangan kata melalui penggunaan penghurai berdasarkan kebergantungan (*dependency-based parser*).

Pembangunan algoritma pengekstrakan hubungan ini mempunyai beberapa teknik. Antaranya ialah hubungan diekstrak berdasarkan peraturan dan diekstrak berdasarkan pembelajaran mesin (Li et al., 2018). Pengekstrakan hubungan berdasarkan peraturan adalah dengan menggunakan pengetahuan linguistik dan pengetahuan domain untuk membina kumpulan corak berdasarkan perkataan, bahagian pertuturan (*part-of-speech*) atau semantik berkolaborasi dengan pakar domain dan kemudian hubungan diekstrak oleh komputer mengikut peraturan yang telah ditetapkan. Beberapa karya seperti (Exner & Nugues 2012; Kertkeidkachorn & Ichise 2017; Romadhony et al. 2018) telah menggunakan pendekatan ini dalam mencapai objektif mereka.

Seterusnya, kaedah pengekstrakan hubungan berasaskan pembelajaran mesin menggunakan sejumlah besar data berlabel untuk latihan dan telah menunjukkan hasil yang baik dalam beberapa situasi, terutama jika dibandingkan dengan kaedah berdasarkan



peraturan. Terdapat juga beberapa sistem pengekstrakan maklumat terbuka yang menggunakan pendekatan ini seperti NELL (Carlson et al. 2010) dan OLLIE (Mausam et al. 2012). Karya (P. Wu et al., 2018) yang telah menggunakan pendekatan ini untuk mengekstrak hubungan yang akhirnya menghasilkan graf pengetahuan.

Pendekatan yang dinyatakan untuk mengekstrak hubungan dari ayat input mempunyai kebaikan dan keburukan mereka yang tersendiri. Sebagai contoh, pendekatan berdasarkan peraturan tidak memerlukan set data latihan terlebih dahulu. Karya Romadhony et al., (2018) memilih pendekatan ini kerana ia merupakan kelebihan utama kaedah tersebut. Manakala untuk pendekatan lain, umumnya memerlukan data latihan atau pangkalan pengetahuan yang sangat luas sebagai latihan.

Selain itu, manusia boleh mencipta corak peraturan yang cenderung untuk mempunyai ketepatan yang tinggi. Pendekatan berdasarkan peraturan digunakan oleh karya berikut (Jia et al. 2018; Zhao et al. 2017) atas sebab output yang dijanakan mempunyai ketepatan yang tinggi. Namun begitu, disebabkan sangat bergantung pada peraturan, prestasi generalisasi peraturan adalah tidak seberapa dan skor perolehan kembali (recall) menjadi rendah. Peraturan yang dirangka berkemungkinan terlalu sederhana atau terlalu rumit dalam aplikasi.

Pendekatan yang berasaskan pembelajaran mesin menggunakan sejumlah besar data latihan dan ujian berlabel. Kajian yang menggunakan pendekatan ini menunjukkan output yang baik dengan penghasilan perolehan kembali yang tinggi. Namun, masalah boleh timbul, seperti kekurangan data latihan dan prestasi generalisasi yang buruk. Set corak akan menjadi lebih terdedah kepada ralat dengan setiap lelaran.

(Chen et al. 2020) menyatakan bahawa apabila peraturan dijana secara automatik mengikut semantik mahupun diekstrak secara manual, satu set peraturan yang komprehensif dan tepat akan diperolehi. Dengan itu menjadikan ketepatan untuk mengisi graf adalah sangat tinggi. Oleh hal yang demikian, pengekstrakan hubungan berdasarkan peraturan dipilih untuk mengekstrak hubungan kerana tidak memerlukan data latihan

terlebih dahulu dan akan menghasilkan output yang lebih tepat walaupun terdapat kelemahan yang telah dinyatakan.

Jadual 2.2 Perbandingan teknik pengekstrakan berdasarkan peraturan dan pembelajaran mesin

Pengekstrakan berasaskan peraturan	Pengekstrakan berasaskan pembelajaran mesin
dengan menggunakan pengetahuan linguistik dan pengetahuan domain untuk membina kumpulan corak berdasarkan perkataan, bahagian pertuturan ( <i>part-of-speech</i> ) atau semantik berkolaborasi dengan pakar domain	dengan menggunakan sejumlah besar data berlabel untuk latihan bagi mengekstrak set hubungan yang baru
<b>Kelebihan</b>	
tidak memerlukan set data latihan terlebih dahulu	menunjukkan output yang baik dengan penghasilan perolehan kembali ( <i>recall</i> ) yang tinggi
corak peraturan yang dicipta cenderung untuk mempunyai ketepatan yang tinggi	
<b>Keburukan</b>	
akan mempunyai perolehan kembali ( <i>recall</i> ) yang rendah	memerlukan data latihan yang sangat luas sebagai latihan
prestasi generalisasi peraturan adalah tidak seberapa (mungkin terlalu sederhana atau rumit)	kekurangan data latihan akan menyebabkan lebih terdedah kepada ralat dengan setiap lelaran

Antara teknik dalam RE, Exner & Nugues (2012) adalah salah satu karya yang menggunakan pendekatan berasaskan peraturan. Pendekatan ini digunakan untuk memetakan predikat rangkap tiga kepada predikat yang sama dalam GP. Walau bagaimanapun, peraturan yang dijana tidak dapat merangkumi semua corak yang mungkin berada dalam domain terbuka. Serupa dengan Kertkeidkachorn & Ichise (2017) pendekatan berasaskan peraturan digunakan tetapi dengan penambahan pendekatan secara hibrid iaitu berasaskan vektor untuk mencapai objektif mereka. Di mana, peraturan yang terhasil dapat merangkumi semua corak yang mungkin yang menghasilkan rangkap tiga yang lebih lengkap.

Romadhony et al. (2018) telah menjalankan kajian untuk membina sistem pengekstrakan maklumat terbuka untuk bahasa Indonesia dengan kaedah berdasarkan peraturan. Prestasi yang dihasilkan Romadhony et al. telah membuktikan bahawa dengan

hanya menggunakan berasaskan peraturan masih boleh merumuskan peraturan umum yang boleh menangkap rangkap tiga dalam domain terbuka yang luas. Karya ini dirujuk dari segi kaedah yang digunakan untuk mengekstrak hubungan antara entiti.

Terdapat beberapa ciri yang sering digunakan dalam proses pengekstrakan hubungan seperti: 1) penandaan POS (*part-of-speech tag*), 2) jenis frasa, 3) jenis *named entity*, 4) jenis hubungan kebergantungan (*dependency relation*) dan 5) maklumat semantik. Ciri-ciri tersebut adalah faktor penting dalam kaedah pengekstrakan kerana ia menggambarkan kriteria sebuah perkataan atau frasa sebagai hubungan yang wajar. Karya ini menyatakan tiga ciri utama yang digunakan ialah penandaan POS, jenis frasa dan juga jenis hubungan kebergantungan (Romadhony et al., 2018).

Ciri gabungan dalam mengekstrak hubungan dan entiti adalah dengan menggunakan penandaan POS dan jenis hubungan kebergantungan. Penandaan POS digunakan untuk menangkap hubungan dengan mengenal kata kerja dalam ayat input. Daripada hanya mempertimbangkan kata kerja tunggal sebagai hubungan, kata kerja diikuti dengan kata keterangan (*adverb*) juga dipertimbangkan. Entiti dikenal pasti dengan menggunakan hubungan kebergantungan dari kata kerja yang telah diperoleh sebelumnya. Hubungan kebergantungan yang paling mudah adalah kebergantungan *subj* dan *obj* yang masing-masing merujuk kepada subjek dan objek.

Romadhony et al. merumuskan bahawa kaedah yang digunakan adalah dengan mengenal pasti hubungan berdasarkan penandaan POS *VERB* dan mendapatkan hubungan lengkap berdasarkan jenis hubungan kebergantungan *ADVMOD* (*adverbial modifier*). Diakhiri dengan mengekstrak entiti untuk kedua subjek dan objek. Maka akan terhasilnya rangkap tiga yang lengkap.

Langkah-langkah bagaimana penulis mendapatkan rangkap tiga dari teks dengan lebih lanjut telah dijelaskan. Pertama sekali, Penghurai berdasarkan kebergantungan (*dependency parser*) digunakan untuk mengenal pasti jenis hubungan kebergantungan antara perkataan. Kemudian set peraturan pengekstrakan digunakan untuk mengekstrak

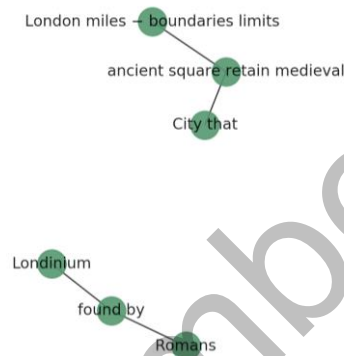
hubungan dan entiti. Peraturan ini mengesan berdasarkan penandaan POS dengan label kata kerja dan kemudian dengan penggunaan jenis hubungan kebergantungan *ADVMOD* untuk mendapatkan hubungan yang lebih lengkap. Pengekstrakan diselesaikan dengan pengekstrakan entiti, iaitu dengan mengesan berdasarkan hubungan kebergantungan yang menghubungkan klausa kepala argumen dan kata-kata lain. Oleh itu, pengekstrakan untuk mendapatkan rangkap tiga selesai.

Karya Romadhony et al. (2018) dirujuk berkenaan cara mereka menggunakan penandaan POS dan hubungan kebergantungan untuk mengekstrak hubungan. Namun begitu, mereka tidak mempertimbangkan hubungan sintaktik lain yang wujud antara teks iaitu teks yang menerangkan tentang sesuatu perkataan yang merupakan kata sifat kepada sesebuah kata nama. Seperti yang dinyatakan (Göring et al. 2014), terdapat peningkatan dalam tugas visual jika penambahan atribut dilakukan. Dengan mengekstrak hubungan jenis ini, kebanyakan hubungan yang wujud antara teks dapat diekstrak.

Terdapat juga beberapa sistem sedia ada (Joshi 2019; Borcan 2020) yang menggunakan hubungan kebergantungan untuk mengekstrak hubungan dari teks. Kaedah ini mengekalkan makna kaya setiap perkataan dalam sebuah ayat. Walau bagaimanapun, sistem ini dapat diperbaiki lagi kerana tumpuan hanya diberikan pada ayat yang ringkas. Dimana penggunaan ayat yang ringkas akan membuahkan hasil yang baik tetapi sebaliknya apabila ayat kompleks digunakan.

Secara tepatnya, salah satu sistem ini Joshi (2019) menggunakan ayat ringkas yang hanya mengandungi dua entiti dan satu hubungan. Pada hakikatnya dalam dunia sebenar, ayat yang digunakan kebanyakannya lebih kompleks. Borcan (2020) pula menggunakan ayat yang panjang tetapi, hanya ayat pendek sahaja yang menghasilkan graf pengetahuan yang lebih logik seperti yang tertera di Rajah 2.5. Sub graf yang dihasilkan di bahagian bawah merupakan graf yang terhasil dari ayat yang ringkas manakala sub graf di bahagian atas pula terhasil dari ayat yang kompleks iaitu ayat panjang yang boleh terdiri daripada beberapa kata kerja dalam satu ayat. Sebagai contoh, *The City of London, London's ancient core – an area of just 1.12 square miles (2.9 km<sup>2</sup>)*

*and colloquially known as the Square Mile – retains boundaries that follow closely its medieval limits.* Secara teori, beberapa rangkap tiga sepatutnya dihasilkan seperti berikut *City of London – retains – boundaries* dan *boundaries – follow – medieval limits.*



Rajah 2.5 Graf pengetahuan dari Borcan (2020)

Jadi, sistem yang dihasilkan oleh kedua penulis ini juga menjadi motivasi bagi kajian ini menghasilkan kaedah yang mampu mengekstrak rangkap tiga yang tepat malah mampu menangani ayat yang lebih kompleks.

## 2.5 Metrik Pengukuran Pengekstrakan Teks

(Mesquita, Schmidek & Barbosa 2013) mencadangkan pendekatan baharu dalam kaedah pengekstrakan hubungan terbuka (*Open RE*) Bagi menilai kaedah pengekstrakan hubungan ini, anotasi automatik mereka dibandingkan dengan anotasi manusia. Oleh itu, dua penyelidik dari kajian (Mesquita, Schmidek & Barbosa 2013) menganalisis set pembangunan hubungan dan menandakan setiap ayat sebagai betul atau salah.

Juga, dalam karya (Romadhony, Purwarianti & Widyantoro 2018), penilaian di peringkat ayat telah dijalankan, dengan membandingkan hasil pengekstrakan peraturan yang dicadangkan dan hasil daripada sistem penanda aras yang diekstrak secara manual. Skor ketepatan (*Precision*), perolehan kembali (*Recall*) dan F1 juga telah digunakan

sebagai metrik penilaian. Berikut merupakan formula bagi metrik pengukuran yang digunakan.

$$P = \frac{\text{bilangan pasangan entiti}}{\text{bilangan pasangan entiti yang diesktrak}} \quad \dots (2.1)$$

$$R = \frac{\text{bilangan pasangan entiti}}{\text{bilangan pasangan entiti yang tepat}} \quad \dots (2.2)$$

$$F1 = \frac{2PR}{P+R} \quad \dots (2.3)$$

Zhao et al. (2017) telah membina graf pengetahuan dalam domain kejuruteraan perisian. Bagi kaedah penilaian mereka, Zhao et al. memilih beberapa halaman web untuk pakar mengekstrak rangkap tiga daripada halaman web berikut. Oleh sebab kajian Zhao et al. mempunyai domain tertentu, iaitu dalam kejuruteraan perisian, pakar yang dipilih mempunyai latar belakang dalam kejuruteraan perisian. Rangkap tiga yang dihasilkan daripada pakar bertindak sebagai kebenaran asas (*ground truth*) untuk menilai graf pengetahuan mereka.

Chen et al. (2018) mencadangkan untuk membina GP untuk pendidikan. Bagi mendapatkan kebenaran asas bagi penilaian model tersebut, dua pakar domain yang terlibat dalam penggubalan standard pendidikan kurikulum dipilih untuk mengenalpasti perhubungan. Kemudian hasil daripada model akan dibandingkan bersama hasil daripada pakar.

Selain itu, Xu et al. (2021) mencadangkan pendekatan berasaskan peraturan untuk mengekstrak pengetahuan domain pengurusan keselamatan pembinaan. Dua pakar domain diminta memilih secara manual elemen domain. Oleh kerana penyelidik mengusaha dalam domain tertentu, pakar yang terlibat ialah seorang profesor universiti yang mempunyai pengetahuan teori yang kaya dan pengurus projek yang berpengalaman dalam pengurusan risiko keselamatan pembinaan. Keputusan pengestrakan dinilai dengan membandingkan senarai yang dijana oleh pakar domain dengan senarai yang dijana dari model daripada sampel ujian yang sama. Skor ketepatan (*Precision*) dan perolehan kembali (*Recall*) juga telah digunakan.

Dengan ini, jelas bahawa set input data perlu terlebih dahulu dianalisis secara manual dari pemikiran manusia untuk mendapatkan senarai rangkap tiga. Kemudian, senarai tersebut akan dijadikan sebagai kebenaran asas untuk dibandingkan dengan set rangkap tiga yang terhasil daripada model/sistem. Skor ketepatan (*Precision*) dan perolehan kembali (*Recall*) digunakan untuk menilai prestasi peraturan pengestrakan.

## 2.6 Hubungan Kebergantungan

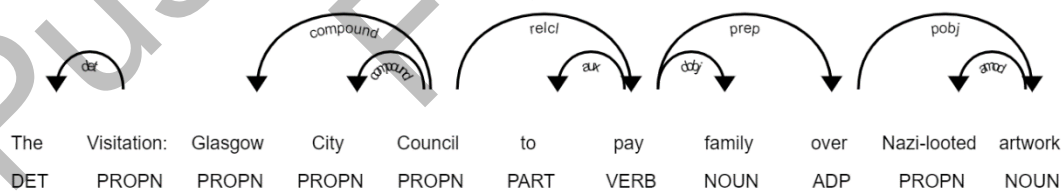
Penghuraian atau *parsing* merupakan istilah yang digunakan dalam menerangkan proses yang diperlukan dalam memecahkan ayat yang diberi untuk mencari komponen tatabahasa dari setiap perkataan dalam ayat tersebut. Penghuraian juga dapat didefinisikan sebagai penguraian input (ayat) ke dalam beberapa komponen yang diproses dengan lebih tepat. Dalam Bahasa tabii, penghuraian diklasifikasikan sebagai kaedah dalam menganalisis sebuah input ayat berdasarkan konstituen tatabahasa, hubungan sintaktik dan juga golongan kata. (Yusuf, Nwojo & Boukar 2018)

Penghurai berdasarkan kebergantungan (*dependency-based parser*) adalah untuk menyambungkan antara perkataan utama dengan perkataan yang bergantung dalam sesebuah ayat. Ia juga melabelkan hubungan yang bergantung pada perkataan utama untuk memberikan pendekatan terhadap hubungan antara predikat dan argumen (Cao et al. 2018).

Tugas penghurai kebergantungan adalah untuk mengambil teks input dan mengenakan ke atas teks itu set hubungan kebergantungan yang sesuai. Menghuraikan perkataan atau ayat melalui sebuah penghurai kebergantungan membantu menghasilkan pokok kebergantungan atau *dependency tree* iaitu sebuah model pokok yang berasaskan hubungan kebergantungan. (Yusuf et al., 2018).

Pengurai kebergantungan mempunyai kelebihan mengaburkan komponen dari realisasi permukaan teks ke tahap yang lebih besar, menghasilkan maklumat semantik yang lebih mudah diakses dalam perwakilan teks, yang bermanfaat untuk pengekstrakan maklumat (Yusuf et al., 2018).

Dalam pokok kebergantungan, nod dan *edge* mencerminkan sintaktik yang relevan antara perkataan dalam sebuah ayat seperti di Rajah 2.6. Kebergantungan adalah tanggapan bahawa unit linguistik saling berhubung dengan pautan yang berarah, dan pokok itu dapat menggambarkan pembinaan pautan tersebut (Y. Jiang, Hao & Zhu 2016). Dalam erti kata lain, sebuah pokok yang dihasilkan dari sebuah input ayat akan mewakili susunan secara grafik di mana penghasilan tatabahasa telah diterapkan semasa proses penghuraian.



Rajah 2.6 Contoh pokok kebergantungan yang dihasilkan oleh *Spacy dependency-based parser*

Pokok kebergantungan membantu model pengekstrakan hubungan menangkap hubungan jarak jauh antara perkataan. Kajian-kajian (Miwa & Bansal 2016; Xu et al. 2015) yang menggunakan model berdasarkan kebergantungan telah terbukti sangat berkesan dalam pengekstrakan hubungan. Hal ini kerana ia mampu menangkap hubungan sintaktik jarak jauh walaupun terdapat klausa yang panjang atau kompleks. Di samping



itu, pokok kebergantungan juga dapat menyampaikan maklumat struktur yang kaya dan juga telah terbukti berguna untuk menjalin hubungan antara entiti dalam teks (Guo, Zhang & Lu 2020).

Selain itu, berdasarkan kebergantungan universal (*Universal Dependencies(UD)*) (De Marneffe et al. 2014) , terdapat 42 hubungan yang boleh dikelompokkan dan ianya dapat menghasilkan sembilan kategori daripadanya. Antara contohnya adalah 1) predikat klausa, 2) kebergantungan bukan teras bagi predikat klausa, 3) klausa pautan, 4) kebergantungan kata nama, 5) koordinasi, 6) pengkompaunan dan tidak dianalisis, 7) penandaan huruf besar, preposisi, posesif, 8) hubungan *Loose joining* dan 9) lain-lain.

Tambahan pula, menurut kajian Cao et al. (2018), hubungan yang sering digunakan hanya tertumpu pada dua daripada sembilan kategori kebergantungan universal. Dua hubungan tersebut adalah daripada kategori satu dan kategori empat iaitu hubungan predikat klausa dan kebergantungan kata nama. Menurut Cao et al. juga dikenali sebagai hubungan pengubah kata nama. Jadual 2.3 dan Jadual 2.4 yang disertakan berikut merupakan antara contoh senarai hubungan kebergantungan universal bagi kategori satu dan empat.

Jadual 2.3 Kategori satu (Hubungan predikat klausa)

<b>Hubungan predikat klausa</b>	<b>Huraian</b>
<i>Nsubj</i>	<i>Nominal subject</i>
<i>Nsubjpass</i>	<i>Passive nominal</i>
<i>Csubj</i>	<i>Clausal subject</i>
<i>Csubjpass</i>	<i>Clausal passive subject</i>
<i>Dobj</i>	<i>Direct object</i>
<i>Iobj</i>	<i>Indirect object</i>
<i>Ccomp</i>	<i>Clausal complement</i>
<i>Xcomp</i>	<i>Open clausal complement</i>

Jadual 2.4 Kategori empat (Hubungan pengubah kata nama)

<b>Hubungan pengubah</b>	<b>Huraian</b>
<i>AMOD</i>	<i>Adjectival modifier</i>
<i>ADVMOD</i>	<i>Adverbial modifier</i>
<i>NMOD</i>	<i>Nominal modifier</i>
<i>NUMMOD</i>	<i>Numeric modifier</i>

bersambung...

...sambungan  
 APPOS  
 DET  
 COMPOUND

*Appositional modifier*  
*Determiner*  
*Compound*

---

## 2.7 KORPUS

Kebelakangan ini, terdapat pelbagai kenyataan (Hollink et al. 2016; Ramisa et al. 2018) yang mendakwa imej dan teks harus melengkap antara satu sama lain untuk menerima hasil yang diinginkan dalam hubungan imej-teks dan masing-masing telah mencadangkan dataset *Images in Online News (ION Corpus)* (Hollink et al. (2016) ) dan *Breaking News* (Ramisa et al. (2018) ) yang kedua-duanya mengandungi pelbagai artikel berita.

Hollink et al. (2016) menyatakan bahawa ia adalah satu keperluan untuk dataset imej web untuk berada dalam konteks tekstual semulajadi mereka. Kepentingan mengumpul dataset imej-teks adalah kerana hubungan yang lebih semula jadi antara teks dan imej. Walaupun terdapat dataset yang popular seperti MSCOCO yang boleh digunakan untuk melatih sistem untuk menghasilkan kapsyen seperti Flickr atau deskripsi orang ramai, Hollink et al. juga berpendapat bahawa adanya keperluan untuk banyakkkan lagi dataset yang menghasilkan kapsyen mengikut konteks. Rajah 2.7 berikut merupakan imej dari dataset MSCOCO dan deskripsi untuk imej tersebut. Berbanding dengan Rajah 2.8 iaitu imej yang diambil dari artikel berita, ia mempunyai deskripsi yang lebih abstrak dan semantik berperingkat tinggi yang sukar diperolehi melalui pemrosesan visual.

a black cat sitting on a sink in front of a mirror  
 black cat sitting on edge of sink next to granite wall.  
 a cat on a bath room sink in a reflection  
 a black cat sitting on sink next to a mirror.  
 a black cat sitting on top of a bathroom sink.



Rajah 2.7 Contoh imej dan kapsyen dari MSCOCO

### My Pet Story: The cat lover who used to dislike cats



Rajah 2.8 Contoh imej dan kapsyen dari artikel berita

Ramisa et al. (2018) juga menyatakan artikel berita bergambar tidak diterokai sepenuhnya. Domain artikel berita difokuskan kerana kandungan tekstual lebih menjelaskan hubungan konotatif dan samar-samar yang hanya dicadangkan tetapi tidak langsung disimpulkan dari imej (Ramisa et al. 2018).

## 2.8 KESIMPULAN

Kesimpulannya, kajian-kajian yang telah membina GP di dalam karya mereka masih lagi boleh diperbaiki. Karya P. Wu et al. (2018) membina GP dari teks masih lagi mempunyai batasan dari segi maklumat yang diekstrak tidak secara sepenuhnya. Manakala, karya Gong & Wang (2017) telah memberi motivasi untuk membina GP multimodal yang mempunyai hubungan antara teks dan visual. Namun, GP multimodal sedia ada ini boleh dipertingkatkan lagi.

Secara ringkasnya, bagi membina sebuah GP multimodal, dimulakan dengan memproses input data dari laman web untuk mengekstrak teks dan imej. Kemudian teknik NLP akan diterapkan pada setiap teks. Teks yang dilabel sebagai tag akan menjadi entiti kepada hubungan yang menghubungkan imej. Maka, terhasilnya hubungan antara visual-teks.

Bagi pembinaan hubungan antara teks-teks pula, teknik pengekstrakan hubungan digunakan iaitu pengekstrakan berasaskan peraturan (*rule-based RE*). Kajian ini merujuk kepada karya Romadhony et al. (2018) berkenaan cara mereka menggunakan penandaan POS dan hubungan kebergantungan untuk mengekstrak hubungan. Namun begitu, berbezanya kajian ini ialah selain dari pengekstrakan kata kerja tunggal, pengekstrakan kata kerja + preposisi juga diambil kira dan memanfaatkan hubungan kebergantungan *amod* dimana hubungan ini menerangkan sifat tentang sebuah entiti.

Karya Joshi (2019) hanya menggunakan ayat mudah dimana pada hakikatnya dalam dunia sebenar, ayat yang digunakan kebanyakannya lebih kompleks. Dalam karya berikut Borcan (2020), ayat lebih kompleks telah digunakan untuk mengekstrak set rangkap tiga tetapi rangkap tiga yang dihasilkan tidak logik dan tidak berapa tepat. Jadual 2.5 menunjukkan ringkasan perbandingan setiap kaedah terhadap cara penulis memproses teks.

Jadual 2.5 Jadual perbandingan kaedah-kaedah terhadap cara teks diproses

<b>Karya</b>	<b>Kaedah yg dicadangkan penulis</b>	<b>Batasan/had</b>
P. Wu et al. (2018)	GP yang dihasilkan ialah ringkasan daripada teks input panjang;	Output ringkasan tidak menangkap maklumat lain yang ditemui dalam ayat yang panjang
Gong & Wang (2017)	Mempunyai pertalian antara kedua visual dan teks	Hanya mengambil kata nama secara tidak langsung mengabaikan maklumat lain
Romadhony et al. (2018)	Menggunakan hubungan kebergantungan untuk mengekstrak hubungan daripada teks (ekstrak berdasarkan peraturan)	Tidak sertakan hubungan atribut
Joshi (2019)		Hanya menggunakan ayat mudah (Dua entiti dan satu hubungan (kata kerja))
Borcan (2020)		Ayat panjang/kompleks (terdapat beberapa kata kerja dalam satu ayat) digunakan tetapi menghasilkan rangkap tiga yang tidak logik

Pusat Sumber  
FTSM

## **BAB III**

### **METODOLOGI**

#### **3.1 PENGENALAN**

Analisis linguistik dilakukan pada setiap ayat untuk mendapatkan kebergantungan tatabahasa, dimana satu set hubungan kebergantungan akan dikenal pasti. Set ini akan digunakan untuk mengekstrak rangkap tiga dari teks. Akhirnya, rangkap tiga yang diekstrak digunakan untuk membina GP multimodal. Dalam bab ini, penjelasan terperinci mengenai rangka kerja untuk membina GP multimodal akan diterangkan

#### **3.2 MOTIVASI KAJIAN**

Kajian ini menyumbang kepada pengekstrakan hubungan dari ayat dunia nyata yang boleh menjadi ayat mudah pendek sehingga ayat yang panjang dan kompleks. Kajian ini bertujuan untuk mengekstrak hubungan antara dua entiti iaitu rangkap tiga dari teks. Teknik hubungan kebergantungan diterapkan kerana hubungan setiap perkataan tidak akan terputus.

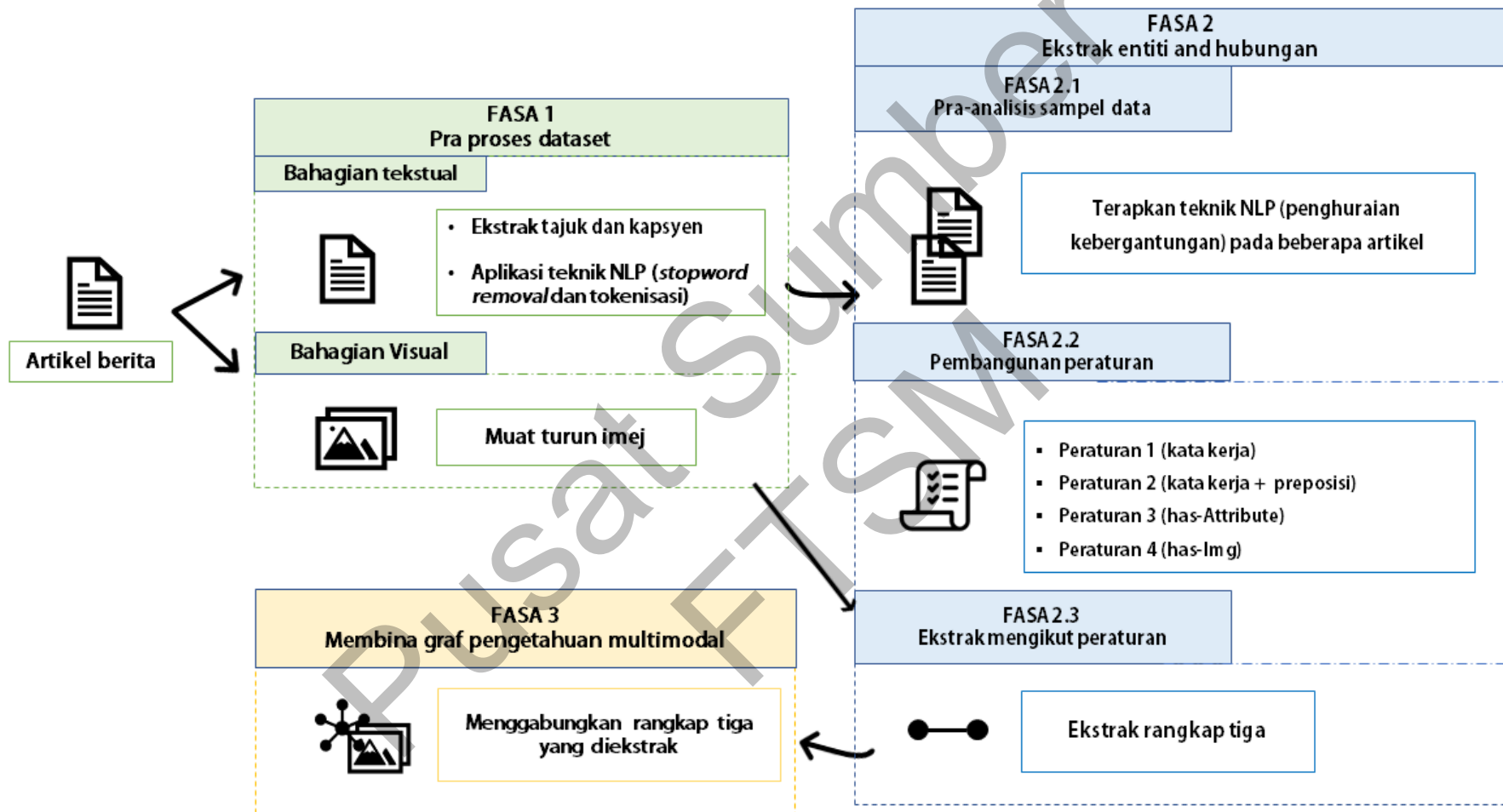
Motivasi utama kajian ini adalah beberapa graf pengetahuan yang sedia ada tidak mempertimbangkan hubungan yang wujud antara dua modaliti yang berbeza, iaitu teks dan visual. Maka, kajian ini mempertimbangkan hubungan ini dan menggunakan dataset imej dalam artikel berita yang disertakan dengan tajuk artikel dan kapsyen imej yang sangat berkait. Dataset yang digunakan meliputi pelbagai jenis topik seperti sukan, pendidikan, hiburan dan banyak lagi. Hal ini kerana pengetahuan yang wujud tidak terhad kepada satu topik sahaja malah sebaliknya. Tambahan lagi, terdapat penyelidik yang telah membina GP multimodal yang telah terbukti jauh lebih baik berbanding dengan GP unimodal.

Selain daripada menggunakan ayat mudah, kajian ini juga menerapkan penggunaan ayat kompleks (ayat yang boleh terdiri daripada beberapa kata kerja dalam satu ayat) kerana terdapat kajian yang menggunakan teks yang mudah sahaja. Juga dapat menangani ayat-ayat kompleks yang berkemampuan untuk menangkap beberapa rangkap tiga (*triples*) dari sebuah ayat.

### 3.3 GAMBARAN RANGKA KERJA

Rajah 3.1 menggambarkan rangka kerja bagi pembinaan GP multimodal. Rangka kerja ini mempunyai tiga fasa utama. Fasa 1: pra-proses dataset, Fasa 2: ekstrak entiti dan hubungan dan Fasa 3: membina GP multimodal. Setiap fasa akan dijelaskan secara terperinci dalam sub-bahagian berikut.

Pusat Sumber  
FTSM



Rajah 3.1 Rangka kerja bagi pembinaan GP multimodal



### 3.3.1 Fasa 1 (Pra proses dataset)

Bermula dengan fasa pertama, input Fasa 1 adalah sekumpulan artikel berita atas talian. Setiap artikel terdiri daripada imej dan artikel teks. Di sini, fokusnya adalah pada imej dan teks deskriptif yang mengiringi imej tersebut. Oleh itu, kajian ini hanya mempertimbangkan tajuk artikel, kapsyen dan imej. Sumber teks berasal dari tajuk dan kapsyen imej manakala sumber visual dari imej artikel.

Dataset yang digunakan merupakan artikel berita web yang terdiri daripada berita teks dan imej yang berkaitan dengan berita. Artikel berita perlu dibersihkan dan disaring kerana artikel berita mempunyai banyak maklumat yang dikumpulkan (tajuk, kapsyen, tarikh diterbitkan dan banyak lagi). Walau bagaimanapun, dalam kajian ini, hanya sebilangan kecil maklumat yang kaya ini akan digunakan. Hal ini kerana tajuk dan kapsyen yang dipilih ini sudah memadai untuk menerangkan tentang sesebuah imej tersebut.

Berdasarkan Rajah 3.1, bagi bahagian teks, kajian ini hanya mempertimbangkan tajuk dan kapsyen. Contohnya dalam Rajah 3.2, teks yang dibulatkan merupakan tajuk dan kapsyen yang digunakan dari setiap artikel. Bagi setiap artikel, tajuk dan kapsyen akan digunakan untuk mengenal pasti entiti dan hubungan yang akan diliputi pada fasa seterusnya.

## The Visitation: Glasgow City Council to pay family over Nazi-looted artwork

The tapestry fragment dates to 16th-century Switzerland

Kashmira Gander | @kashmiragander | Friday 14 August 2015 00:01



A 16th-century Swiss tapestry fragment entitled The Visitation, as a payment to be made to the heirs of a German-born Jewish woman after the artwork, looted by the Nazis, ended up in one of Scotland's best-known museums, the Burrell Collection in Glasg (D. Dosma/Glasgow Museums and Libraries Collections/PA Wire)

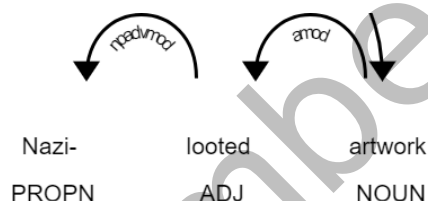
Rajah 3.2 Contoh artikel berita

Seterusnya, Teknik NLP seperti tokenisasi dan *stopword removal* akan diterapkan pada tajuk dan kapsyen. Kedua-dua teknik ini akan diubah supaya output yang dihasilkan sesuai dengan kajian. Sebab dan cara penyesuaian ini akan diterangkan di bawah.

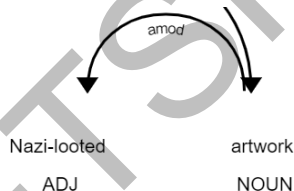
Bagi hampir semua tugas pemprosesan bahasa tabii (NLP), amat penting untuk memisahkan perkataan terlebih dahulu menjadikannya token, dan kemudian menggunakan pengumpulan token ini sebagai input untuk tugas pemprosesan teks lain. Secara umum, tokenisasi biasanya dilakukan dengan mengenali ruang sebagai pembatas sempadan (Panahandeh & Ghanbari 2019). Token boleh dibentuk dalam perkataan individual atau frasa tetapi kebiasaannya, satu perkataan dikesan sebagai satu token.

Walaupun begitu, dalam sesetengah kes, terdapat perkataan yang tidak boleh dianggap sebagai satu token. Sebagai contoh, dalam Rajah 3.3, terdapat sekumpulan perkataan yang dihubungkan oleh tanda hubung seperti *Nazi-looted* yang boleh dianggap

sebagai satu token tetapi disebabkan oleh keterbatasan penghurai (*parser*), ia mengesan secara berasingan. Oleh itu, perlu diubah suai untuk membuat penghurai ini mengesannya dengan betul. Tokenisasi disesuaikan dengan cara yang dapat menggabungkan beberapa perkataan sebagai token tunggal, bukan senario yang biasanya satu perkataan, satu token. Penyesuaian ini akan mengumpulkan semua perkataan yang mempunyai tanda sempang (-) dan menggabungkannya menjadi satu token seperti Rajah 3.4.



Rajah 3.3 Sekumpulan perkataan yang harus dianggap sebagai token tunggal



Rajah 3.4 Perkataan yang telah dikumpulkan bersama

Bagi langkah pra-pemprosesan yang seterusnya, *stopword removal* adalah proses menghilangkan perkataan umum yang sering muncul tetapi tidak mempunyai kepentingan dalam ayat. Perkataan yang paling umum dalam dokumen adalah artikel, preposisi, kata ganti nama dan lain-lain yang tidak memberikan makna (Imrona, Widyawan & Nugroho 2020). Senarai *stopword* yang asal tidak boleh digunakan dalam kajian ini. Perkataan seperti preposisi (*at*, *in*, *of* dan sebagainya) boleh dijadikan contoh mengapa senarai asal tidak boleh digunakan. Hal ini disebabkan perkataan yang berlabel kata preposisi akan diguna pakai kemudian untuk mengekstrak hubungan. Jadi, senarai *stopword* akan ditentukan sendiri bagi menyesuaikan dengan kajian ini. Secara khusus, perkataan seperti

*left*, *right*, *above*, *centre*, *below*. Perkataan ini umumnya untuk menggambarkan kedudukan objek tertentu dalam sesebuah imej. Perkataan ini dikeluarkan kerana ia mempengaruhi prestasi penghurai kebergantungan (*dependency-based parser*) yang menjadikannya kurang tepat.

Charles	Napier,	67,	left,
PROPN	PROPN	NUM	VERB

Rajah 3.5 Contoh perkataan *left* yang perlu dibuang

Sebagai contoh, dalam rajah 3.5, penghurai mengesan perkataan *left* sebagai kata kerja. Dalam konteks ayat ini, ia tidak betul kerana maksudnya berbeza dalam ayat tertentu. Perkataan *left* hanya untuk menggambarkan kedudukan orang dalam imej. Terdapat keperluan untuk senarai *stopword* diubah suai kerana ia mempengaruhi ketepatan output. Output penghurai perlu tepat yang mungkin supaya tidak mengalami masalah ketika mengekstrak entiti dan hubungan yang bergantung pada peraturan.

Bagi bahagian visual, setiap artikel berita biasanya mempunyai imej untuk menyokong berita tersebut. Setiap imej ini akan dimuat turun melalui URL. Kemudian ia disimpan untuk kegunaan fasa 2 untuk dipasangkan dengan hubungan *has-Image* yang akan dijelaskan dengan lebih lanjut dalam fasa 2.2. Ini untuk menunjukkan bahawa kapsyen berkait dengan imej.

Sebagai kesimpulan, output dari fasa ini adalah teks yang diekstrak merujuk kepada tajuk dan kapsyen, dan imej yang telah dimuat turun. Akhirnya, dataset,  $D$  akan terhasil.

$$D = \{T, V\} \quad \dots (3.1)$$

di mana  $T$  dan  $V$  masing-masing setara dengan bahagian teks dan visual. Bagi set,  $T$ , yang terdiri daripada  $h$  dan  $c$ .  $h$  mewakili tajuk sementara  $c$  mewakili kapsyen.

$$\{h, c\} \in T \quad \dots (3.2)$$

Bagi set  $V$ , yang hanya terdiri daripada satu elemen,  $img$ , iaitu imej yang dimuat turun.

$$\{img\} \in V \quad \dots (3.3)$$

### 3.3.2 Fasa 2 (Ekstrak entiti dan hubungan)

Fasa seterusnya merupakan lanjutan dari bahagian teks iaitu mengekstrak entiti dan hubungan. Ia akan dibahagikan menjadi tiga sub-fasa iaitu fasa 2.1: pra-analisis data, fasa 2.2: pembangunan peraturan dan fasa 2.3: ekstrak berdasarkan peraturan. Secara ringkasnya, pada fasa 2.1, beberapa artikel yang terdiri daripada ayat mudah dan kompleks dipilih dan dihuraikan melalui penghurai tatabahasa. Pada fasa 2.2, output dari penghurai tatabahasa yang merupakan pokok kebergantungan setiap artikel akan dianalisis. Ini adalah langkah awal untuk mengesan corak yang akan ditetapkan dengan kaedah pengekstrakan hubungan. Setelah peraturan telah ditentukan, pada fasa 2.3 rangkap tiga (entiti dan hubungan) akan diekstrak. Satu ayat terdiri daripada satu rangkap tiga. Namun, dalam beberapa kes, terdapat satu ayat yang terdiri daripada beberapa rangkap tiga.

#### a. Fasa 2.1 (Pra-analisis data)

Untuk sub-fasa pertama, Fasa 2.1, beberapa sampel data dipilih untuk menjalani analisis untuk menghasilkan peraturan bagi pengekstrakan rangkap tiga. Sampel data ini mengandungi teks yang diproses, iaitu tajuk dan kapsyen, dari beberapa artikel berita yang dihasilkan pada fasa sebelumnya, Fasa 1. Teks tersebut dipilih secara rawak dari beberapa artikel berita menjadikan keseluruhan 10 ayat yang terdiri daripada lima ayat mudah dan lima ayat kompleks seperti Rajah 3.6 dan Rajah 3.7. Ayat mudah dalam Rajah 3.6 kebanyakan diambil dari tajuk berita kerana ayat berikut biasa pendek dan dari segi tatabahasa, hanya mengandungi satu kata kerja, menepati definisi untuk ayat mudah.

Manakala, bagi ayat kompleks pula kebanyakan diambil dari kapsyen imej kerana mempunyai beberapa kata kerja dalam sebuah ayat.

- [1] Rihanna to appear as 'special adviser' on 'The Voice'
- [2] Stolen tortoise leads to arrest of sexual assault suspect
- [3] The Visitation: Glasgow City Council to pay family over Nazi-looted artwork
- [4] The 22-year-old recently won ATP Challenger tournament
- [5] Daniel Pena-Sanchez: Mother appeals for teenage son with heart problem to come home

#### Rajah 3.6 Ayat mudah

[1] A 16th-century Swiss tapestry fragment entitled *The Visitation*, as a payment is to be made to the heirs of a German-born Jewish woman after the artwork, looted by the Nazis, ended up in one of Scotland's best-known museums, the Burrell Collection in Glasgow. *Di Dosma/Glasgow Museums and Libraries Collections/PA Wire*

[2] Jailed: Former teacher Charles Napier, 67, left, who is related to Maldon MP John Whittingdale, right, conducted a 'campaign of abuse' at the school where he worked in the late 1960s and early 1970s

[3] The crew of an Emirates Airlines Boeing 777 prepares for passengers ahead of a demonstration flight in Dubai in 2007. Emirates announced plans Thursday to begin flying a Dubai to Panama City route on a 777, which will be the longest in the world

[4] Wolfgang Schuble, left, the German finance minister, and Yanis Varoufakis, the Greek finance minister, hold a joint news conference after a private meeting Thursday in Berlin

[5] Isaiah Ford, above, and Cam Phillips combined for 96 catches for 1,207 yards and nine touchdowns during their breakout freshmen seasons, but the Hokies other returning receivers had just 10 receptions between them. (Paul Vernon/Associated Press)

#### Rajah 3.7 Ayat kompleks

Ayat yang dipilih ini diteliti secara manual melalui perspektif manusia kerana manusia mempunyai pengetahuan untuk mengenal pasti teks yang saling berkait dan teks yang dapat menerangkan tentang sebuah imej juga. Jadual 3.1 menunjukkan panduan

untuk mengesktrak maklumat dari ayat input. Panduan ini digunakan oleh penganotasi untuk mengekstrak entiti dan hubungan bagi menghasilkan sebuah rangkap tiga yang lengkap. Penganotasi perlu mempunyai pengetahuan mengenai tatabahasa penandaan POS dan kebergantungan kerana panduan ini menggunakan pengetahuan tersebut untuk mengesan entiti dan hubungan. Sebagai asas, semua kata nama boleh dianggap sebagai entiti dan hubungan kebergantungan asas adalah kata kerja yang menghubungkan subjek dan objek dalam suatu ayat.

Jadual 3.1 Pelabelan untuk mengekstrak maklumat dari ayat mudah dan juga kompleks

Jenis entiti atau hubungan untuk diekstrak	Pelabelan untuk ekstrak maklumat	Contoh
Gabungan perkataan yang dianggap sebagai satu entiti	Perlu mencari gabungan kata nama; <i>NOUN + NOUN</i>	Perkataan ' <i>Glasgow city council</i> ' dianggap sebagai satu entiti
Entiti (teks) yang berkait dengan imej	Perkataan yang berkait dengan imej dipilih secara manual berdasarkan pemerhatian tetapi kebiasaan merupakan kata nama ( <i>NOUN</i> )	Seperti dalam Rajah 3.9, perkataan ' <i>Emirates Airline Boeing 777</i> ' ialah perkataan yang berkait dengan imej
Entiti yang menerangkan sesuatu sifat kata nama	Perlu mencari kata nama yang mempunyai kata adjektif; <i>ADJ + NOUN/PROP</i> N	' <i>Nazi-looted artwork</i> '; perkataan ' <i>artwork</i> ' mempunyai sifat ' <i>Nazi-looted</i> '
Hubungan berdasarkan kata kerja dan entity	Perlu mencari kata kerja; entiti ialah subjek dan objek untuk kata kerja tersebut; <i>NOUN-VERB-NOUN</i>	<i>Emirates – announced -plans</i>
Hubungan berdasarkan kata kerja bersama preposisi dan entiti	Perlu mencari kata kerja dan gabung bersama preposisi; entiti ialah subjek dan objek untuk kata kerja tersebut; <i>NOUN-VERB + PREP-NOUN</i>	<i>Emirates – flying to- Panama City</i>

Rajah 3.8 dan 3.9 menunjukkan contoh ayat yang mudah dan kompleks bersama imejnya. Dalam rajah di bawah, secara teori dengan hanya berpandukan penandaan POS, beberapa perkataan yang bertanda dianggap sebagai maklumat yang harus diekstrak.



The **Visitation: Glasgow City Council to pay family over Nazi-looted artwork**

Rajah 3.8 Maklumat yang harus diekstrak dari ayat mudah



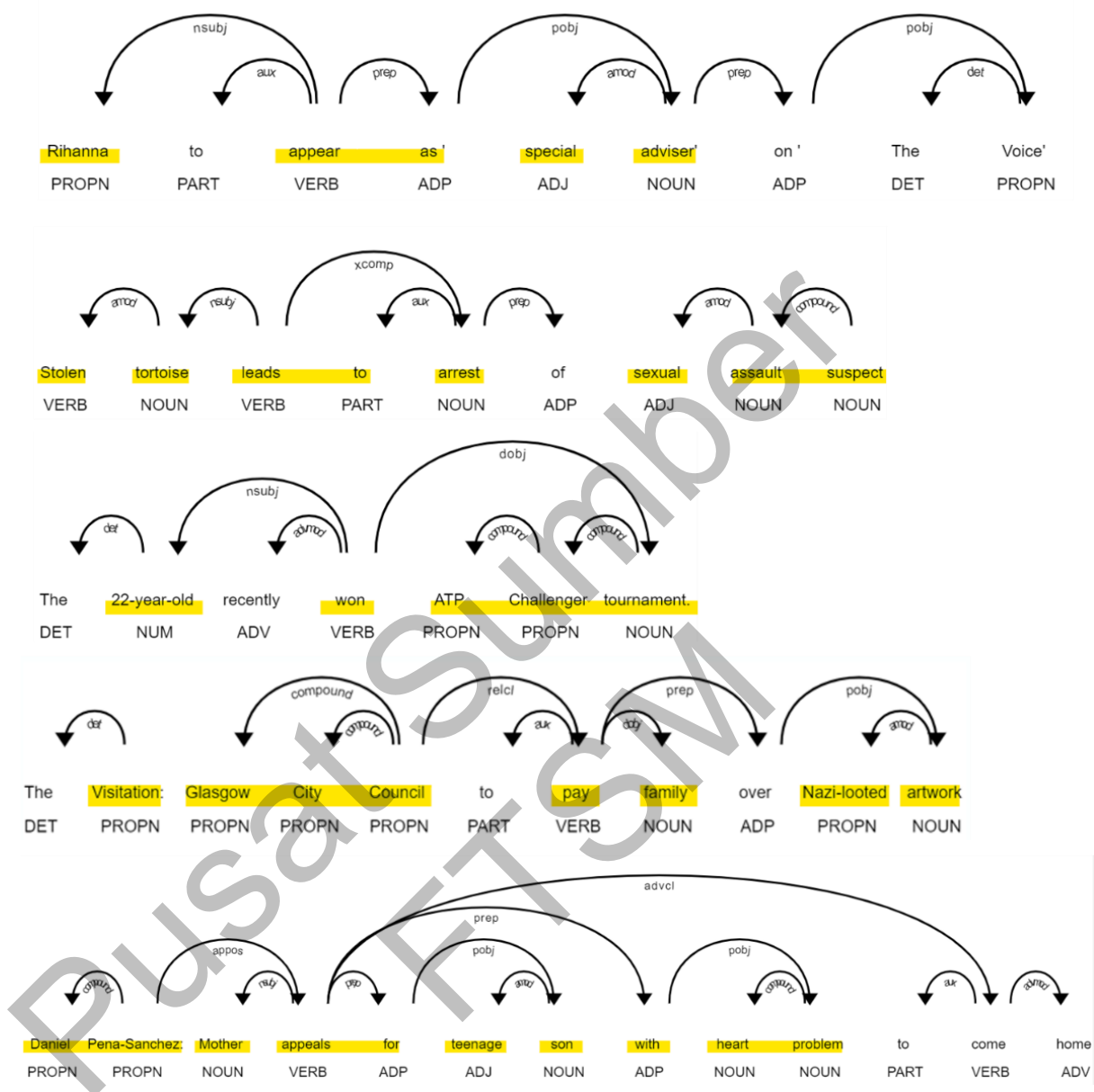
The **crew of an Emirates Airlines Boeing 777 prepares for passengers** ahead of a **demonstration flight** in Dubai in 2007. **Emirates announced plans** Thursday to begin **flying a Dubai to Panama City route** on a 777, which will be the longest in the world.

Rajah 3.9 Maklumat yang harus diekstrak dari ayat kompleks

Kemudian, untuk mengenal pasti kebergantungan yang penting dan penandaan POS, teknik NLP penghuraian kebergantungan (*dependency parsing*) dilaksanakan pada sampel data ini. Penghuraian ini menganalisis struktur tatabahasa ayat, menunjukkan hubungan yang ada antara perkataan utama dan perkataan yang mengubah inti tersebut. Penghurai jenis ini dipilih kerana output yang dihasilkan membantu dalam membina peraturan untuk mengekstrak entiti dan hubungan dalam fasa seterusnya. Seperti yang dinyatakan oleh (Guo et al. 2020), pohon kebergantungan menyampaikan maklumat struktur yang kaya yang telah terbukti berguna untuk membina hubungan antara entiti dalam teks. Oleh itu, teknik ini diterapkan kerana hubungan setiap perkataan tidak akan terputus.

Setiap input ayat dihurai melalui penghurai kebergantungan dan pokok kebergantungan (*dependency tree*) akan dihasilkan seperti rajah berikut. Rajah 3.10 merupakan output daripada beberapa ayat input kepada penghurai. Daripada output ini, corak konsisten yang ditunjukkan daripada teks yang terserlah diperhatikan.





Rajah 3.10 Pokok kebergantungan yang dihasilkan oleh *SpaCy Dependency parser*

Setiap perkataan yang telah dikenal pasti secara teori, dilihat pula struktur tatabahasanya iaitu hubungan kebergantungan dan penandaan POSnya seperti dalam contoh Rajah 3.10 yang menunjukkan struktur untuk beberapa ayat dari sampel data. Setelah diteliti akan setiap output tatabahasa yang dihasilkan, terdapat beberapa hubungan kebergantungan yang seringkali ada pada perkataan yang bertanda. Jadual 3.2 menunjukkan penerangan bagi hubungan kebergantungan (Choi, 2015; Altinok, 2021)

Jadual 3.2 Hubungan kebergantungan dan keterangannya

Jenis hubungan kebergantungan	Huraian
<i>ROOT</i>	topik utama (kata kerja) untuk ayat
<i>Compound</i>	kata nama mengubah suai kepala frasa nama
<i>Nsubj</i>	subjek nominal (frasa nama)
<i>Dobj</i>	objek langsung
<i>Pobj</i>	objek preposisi
<i>amod</i>	frasa adjektif yang mengubah makna frasa nama
<i>Prep</i>	frasa preposisi yang mengubah suai makna kepalanya.

Menurut (Cao et al. 2018), hubungan yang sering digunakan tertumpu pada dua kategori hubungan seperti yang telah dinyatakan dalam Jadual 2.2 (hubungan predikat klausa) dan 2.3 (hubungan pengubah) di dalam Bab 2.5, jika dibandingkan dengan Jadual 3.2, terdapat beberapa hubungan yang disenaraikan berada dalam kedua kategori hubungan tersebut. Hubungan tersebut ialah *Compound*, *Nsubj*, *Dobj* dan *amod*. Bagi teks seperti *appear as* dan *appeals for*, dapat dilihat bahawa hubungan *prep* perlu dipertimbangkan bagi menangkap secara keseluruhan kata kerja tersebut. *Pobj* pula menghubungkan entiti objek kepada hubungan *prep*.

Menurut Altinok (2021), *ROOT* merupakan label khas dalam pokok kebergantungan yang kebiasaan berada pada kata kerja utama sesebuah ayat. Bagi sesetengah kes, jika sebuah frasa itu diproses atau dengan kata lain bukan ayat penuh, *ROOT* diberikan kepada kata nama kepala frasa. Sepertimana dalam kajian ini, *ROOT* boleh menjadi kata kerja dan dalam sesetengah ayat menjadi kata nama. Dimana *ROOT* dalam kajian ini juga digunakan bagi mengekstrak entiti teks untuk berhubung dengan entiti imej selain untuk mengekstrak kata kerja utama.

Selain dari hubungan kebergantungan ialah penandaan POS. Ini adalah proses pelabelan tanda kata dalam bahasa berdasarkan klasifikasi kelas. Dengan kata lain, penandaan POS menunjukkan bahagian-bahagian pertuturan kepada setiap perkataan seperti kata nama, kata kerja, kata sifat dan banyak lagi. Jadual berikut adalah label yang sering terdapat pada perkataan yang telah dikenal pasti.

Jadual 3.3 Senarai penandaan POS dan keterangannya

Penandaan POS	Huraian
<i>ADP</i>	Kata depan
<i>ADJ</i>	Adjektif
<i>DET</i>	Penunjuk
<i>NOUN</i>	Kata nama
<i>PROPN</i>	Kata nama khas
<i>VERB</i>	Kata kerja

Setelah diperhatikan output hubungan kebergantungan dan penandaan POS bagi sampel data ini, Jadual 3.2 dan 3.3 merupakan corak konsisten yang dihasilkan oleh pengurai dan kesemua label ini digunakan sebagai asas kepada peraturan pengekstrakan entiti dan hubungan. Berdasarkan output dalam Rajah 3.10, secara asasnya bagi hubungan antara teks, teks yang akan diekstrak sebagai hubungan mempunyai penandaan POS berlabelkan *VERB* dan entiti pula mempunyai penandaan POS berlabelkan *NOUN* atau *PROPN*. Tetapi kaedah yang dicadangkan ini akan menggunakan juga maklumat hubungan kebergantungan yang telah dihasilkan oleh pengurai kebergantungan (*dependency parser*) bagi mengekstrak entiti yang lebih tepat. Corak bagi pengekstrakan entiti berdasarkan hubungan kebergantungan ialah teks yang berlabelkan *Nsubj*, *Dobj* atau *Pobj* akan diekstrak sebagai entiti. Hubungan kebergantungan yang diperoleh dari pokok kebergantungan ini juga menandakan sesuatu perkataan itu sama ada subjek atau objek.

Bagi mengekstrak secara keseluruhan hubungan kata kerja itu, teks yang mempunyai hubungan kebergantungan *prep* juga digabungkan bersama dengan kata kerja utama. Dengan ini, menjadikan hubungan yang diekstrak lebih sempurna. Tambahan lagi, teks yang menggambarkan tentang sesebuah entiti diekstrak berdasarkan hubungan kebergantungan *amod* menggunakan sepenuhnya semua hubungan yang wujud secara semula jadi antara teks. Lain pula dengan hubungan teks bersama imej, hubungan ini ditetapkan terlebih dahulu dan menggunakan label *ROOT* untuk mengekstrak entiti yang bersesuaian.

Jadi untuk mengimbas kembali, langkah awal untuk membuat peraturan untuk pengekstrakan hubungan berdasarkan peraturan adalah dengan mengenal pasti secara

manual teks yang harus diekstrak. Setelah itu, struktur tatabahasa bagi setiap perkataan yang dikenal pasti akan diteliti. Corak hubungan kebergantungan dan penandaan POS yang konsisten seperti dalam Jadual 3.2 dan 3.3 digunakan sebagai asas bagi peraturan untuk mengekstrak entiti dan hubungan.

Hubungan yang diekstrak dikategorikan kepada dua iaitu 1) teks-teks dan 2) teks-visual. Kategori 1 adalah hubungan yang diekstrak wujud secara semulajadi antara teks. Contohnya, hubungan berdasarkan kata kerja, hubungan berdasarkan kata kerja dan preposisi dan hubungan berdasarkan hubungan kebergantungan *amod* tetapi dinamakan semula sebagai *has-Attribute* sementara Kategori 2 pula hubungan yang ditetapkan secara manual bagi menghubungkan entiti yang berbeza iaitu teks dengan imej. Karya Romadhony et al. (2018) mengekstrak hubungan seperti dalam kategori 1 manakala karya Gong & Wang (2017) mengekstrak entiti bagi hubungan yang telah ditetapkan terlebih dahulu seperti dalam kategori 2. Kajian ini mempertimbangkan kedua kategori tersebut dalam pembinaan graf multimodal.

Sebelum menerangkan secara terperinci mengenai peraturan yang ditentukan, secara teorinya, ini adalah beberapa jenis entiti yang akan diekstrak dari ayat input. Selain daripada kata nama atau kata nama khas sebagai entiti yang merupakan peraturan pengekstrakan yang asas (Romadhony et al. (2018)) iaitu kata kerja sebagai hubungan dan diikuti dengan kata nama sebagai entiti.

1. *Multiword entity* merujuk kepada entiti yang mengandungi beberapa perkataan tetapi dianggap sebagai satu token atau entiti. Berdasarkan Gong & Wang (2017), entiti yang diekstrak adalah dari kata nama dan juga frasa nama. Namun, frasa nama yang diekstrak adalah berdasarkan turutan kata nama. Dalam kajian ini, entiti ini diekstrak dengan menggabungkan semua teks yang mempunyai hubungan kebergantungan *COMPOUND* yang dikesan oleh penghurai yang biasanya berada di sebelah satu sama lain seperti pokok ke-4 dalam Rajah 3.10. *Glasgow city council* akan digabungkan sebagai satu entiti dan bukan berasingan.

2. *Attribute entity* adalah entiti yang menghubungkan dengan hubungan *has-Attribute* yang kebanyakannya diekstrak dari penandaan POS yang biasanya dilabelkan sebagai ADJ (kata adjektif). Namun dalam sesetengah ayat, entiti ini juga boleh diekstrak dari perkataan berlabelkan PROP (kata nama khas). Atribut membolehkan untuk menerangkan, membandingkan dan mengkategorikan objek dengan mudah (Krishna et al. (2017)). Ia juga boleh menjadi sifat atau maklumat khas kepada *multiword entity*.
3. *Main entity* adalah entiti yang menghubungkan dengan hubungan *has-Image*. Merujuk kepada Gong & Wang (2017), teks (kata nama) diekstrak dan dilabel sebagai tag bagi setiap imej. Berbeza dengan kajian ini, perkataan yang berkait dengan imej sahaja yang asalnya dipilih secara manual oleh pemerhatian manusia mempunyai persamaan dari segi output penghurai iaitu kebanyakannya dilabel *ROOT* dan merupakan kata nama. Maka, kedua ciri tersebut digunakan sebagai asas untuk mengesktrak entiti bagi hubungan *has-Image*.

Setelah entiti dan hubungan berjaya diekstrak dengan lengkap, maka rangkap tiga, R yang selesai diekstrak akan terhasil dan dimana hubungan (h), subjek (s), dan objek (o) tersusun seperti susunan berikut iaitu

$$R = \{h,s,o\} \quad \dots (3.4)$$

Seperti yang dinyatakan, 10 ayat telah digunakan dalam analisis untuk menghasilkan peraturan bagi pengekstrakan rangkap tiga. Corak analisis seperti dalam Rajah 3.10 ini tidak menunjukkan sebarang perubahan setelah 10 ayat tersebut diproses. Oleh itu, bagi mengukuhkan lagi dapatan analisis tersebut, sebanyak lima ayat ditambah untuk memastikan tiada corak baharu terhasil melalui proses analisis tersebut. Setelah lima ayat tambahan diproses, corak analisis masih kekal sama. Justeru itu, dengan hanya menggunakan 10 ayat yang dipilih secara rawak bagi proses analisis ini dapat

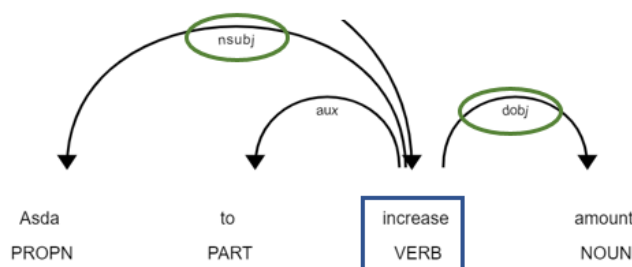
menghasilkan peraturan untuk mengekstrak entiti dan hubungan atau rangkap tiga. Subfasa seterusnya akan menerangkan lebih lanjut mengenai bagaimana peraturan untuk pengekstrakan hubungan ditentukan dengan menggunakan maklumat struktur tatabahasa yang telah dikenalpasti melalui teknik NLP iaitu penandaan POS dan hubungan kebergantungan semasa dalam proses analisis.

**b. Fasa 2.2 (Pembangunan peraturan)**

Seperti yang dinyatakan sebelumnya, kajian ini memberi tumpuan pada pengekstrakan hubungan berdasarkan peraturan dan tidak hanya merangkumi hubungan teks-teks tetapi juga hubungan teks-visual. Dalam subfasa ini, empat jenis peraturan ditetapkan untuk mengekstrak hubungan. Oleh itu, pembangunan peraturan dijelaskan dengan lebih terperinci. Setelah analisis dalam Fasa 2.1 dilakukan, berikut adalah peraturan yang ditentukan untuk mengekstrak hubungan:

1. Peraturan 1 (berlandaskan hubungan kata kerja)
2. Peraturan 2 (berlandaskan hubungan kata kerja + preposisi)
3. Peraturan 3 (berlandaskan hubungan *has-Attribute*)
4. Peraturan 4 (berlandaskan hubungan *has-Image*)

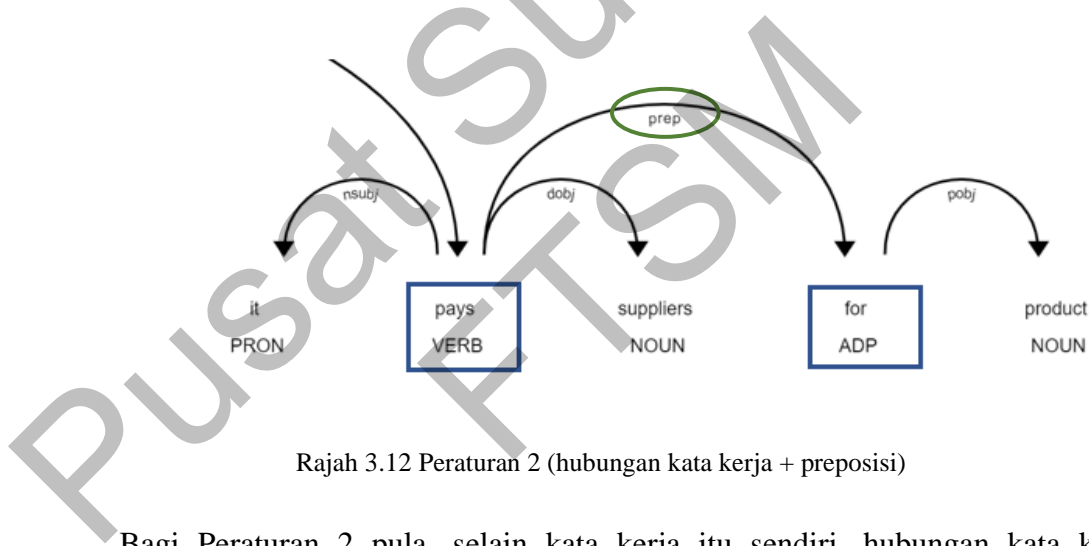
Peraturan 1 hingga 3 adalah peraturan untuk hubungan teks-teks (kategori 1) sementara peraturan 4 adalah untuk hubungan teks-visual (kategori 2).



Rajah 3.11 Peraturan 1 (hubungan kata kerja)

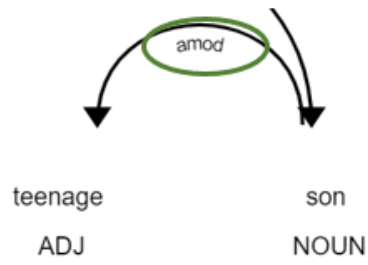
Bagi peraturan 1, hubungan diekstrak terlebih dahulu sebelum entiti, jadi langkah pertama adalah mengenal pasti kata kerja. Berdasarkan Rajah 3.11, ayat tersebut dihuraikan melalui penghurai dan token yang mempunyai label kebergantungan *ROOT* dan penandaan POS kata kerja (*VERB*) akan diekstrak sebagai hubungannya.

Proses seterusnya adalah mencari subjek dan objek sebagai entiti. Biasanya, bahagian kiri dan kanan kata kerja utama mengandungi kedua-dua entiti yang diperlukan. Ini adalah corak yang dapat dikenal pasti setelah menganalisis pokok kebergantungan dalam beberapa ayat. Setelah mengenal pasti hubungan kata kerja, maka subjek dikesan dengan mencari *subj* melalui anak panah dari pokok kebergantungan. Sama seperti subjek, objek tersebut dikesan menggunakan *obj*. Oleh itu, rangkap tiga yang selesai diekstrak akan menjadi  $R = \{increase, Asda, amount\}$  berdasarkan kepada Rajah 3.11.



Rajah 3.12 Peraturan 2 (hubungan kata kerja + preposisi)

Bagi Peraturan 2 pula, selain kata kerja itu sendiri, hubungan kata kerja + preposisi juga dipertimbangkan dalam kajian ini. Mirip dengan Peraturan 1, tetapi dengan penambahan hubungan kebergantungan *prep* seperti pada Rajah 3.12. Ayat akan melalui penghurai dan token dilabel dengan penandaan POS *VERB* akan disimpan sementara mencari token berlabelkan kebergantungan *prep*. Setelah token itu dikesan, kemudian digabungkan menjadi satu hubungan. Selanjutnya, untuk entiti, proses yang sama seperti dalam Peraturan 1 dilakukan untuk menangkap subjek dan objek. Contoh rangkap tiga yang akan terhasil berdasarkan ayat contoh di atas ialah  $R = \{pays\ for, it, product\}$ .



Rajah 3.13 Peraturan 3 (hubungan *has-Attribute*)

Seterusnya, peraturan 3 ini adalah berdasarkan hubungan kebergantungan *amod* yang dinamakan semula sebagai hubungan *has-Attribute* untuk menunjukkan ciri-ciri perkataan tertentu, yang kebanyakannya diambil dari kata adjektif. Lain dari peraturan sebelumnya, peraturan 3 ini akan mengenal pasti objek terlebih dahulu dengan mencari perkataan dengan kebergantungan *amod* seperti pada Rajah 3.13. Entiti ini juga boleh dikenalpasti sebagai *Attribute entity*. Kemudian, subjek iaitu kata nama yang berkait dengan objek akan dikenalpasti melalui anak panah di pohon kebergantungan. Oleh itu, rangkap tiga akan dihasilkan;  $R = \{has-Attribute, son, teenage\}$ .

```

det ( Visitation ||| The )
ROOT ( Visitation ||| Visitation )
punct ( Visitation ||| : )
    
```

Rajah 3.14 Peraturan 4 (hubungan *has-Image*)

Imej dan teks didalam sebuah artikel berita sangat berkait rapat. Peraturan 4 diwujudkan bagi mengekalkan hubungan yang wujud antara dua modaliti iaitu teks dan visual. Dalam erti kata lain, Peraturan 4 ini mewakili hubungan teks-visual. Berbeza dengan proses pengekstrakan sebelumnya, hubungan *has-Image* ditentukan secara manual sebelum entiti diekstrak. *Main entity* akan menjadi entiti yang mewakili teks, di mana ia juga merujuk sebagai subjek. Token yang mempunyai hubungan kebergantungan *ROOT* dan penandaan POS kata nama (*NOUN*) atau frasa nama (*PROPN*) akan dipilih sebagai subjek seperti pada Rajah 3.14 dan imej artikel akan menjadi objek untuk hubungan ini.



Jadual 3.4 di bawah merupakan ringkasan mengenai komponen yang dikenal pasti dalam setiap peraturan iaitu jenis hubungan kebergantungan dan penandaan POS. Peraturan 1 hingga 3 merupakan hubungan yang diekstrak dari teks manakala Peraturan 4 pula hubungan yang telah ditetapkan secara manual (*has-Img*). Dalam erti kata lain, Peraturan 1 hingga 3 adalah untuk mengekstrak entiti dan hubungan. Peraturan 4 adalah untuk mengesktrak entiti bagi hubungan yang menghubungkan imej dan entiti. Kesemua entiti dalam peraturan berikut mempunyai persamaan dalam penandaan POS kata nama, walaupun berbeza dari segi hubungan kebergantungannya. Entiti bagi hubungan *has-Attribute* salah satunya merupakan kata adjektif.

Jadual 3.4 Ringkasan komponen yang dikenal pasti bagi setiap peraturan

Peraturan	Entiti 1	Hubungan	Entiti 2
Peraturan 1 (berlandaskan hubungan kata kerja)	<i>subj</i> ( <i>NOUN</i> atau <i>PROPN</i> )	POS kata kerja ( <i>VERB</i> ) <i>ROOT</i>	<i>obj</i> ( <i>NOUN</i> atau <i>PROPN</i> )
Peraturan 2 (berlandaskan hubungan kata kerja + preposisi)	<i>subj</i> ( <i>NOUN</i> atau <i>PROPN</i> )	POS kata kerja ( <i>VERB</i> ) <i>prep</i>	<i>obj</i> ( <i>NOUN</i> atau <i>PROPN</i> )
Peraturan 3 (berlandaskan hubungan <i>has-Attribute</i> )	<i>NOUN</i>	<i>amod</i>	<i>ADJ</i>
Peraturan 4 (berlandaskan hubungan <i>has-Image</i> )	<i>ROOT</i> ( <i>NOUN</i> atau <i>PROPN</i> )	<i>has-Image</i>	Imej

### c. Fasa 2.3 (Ekstrak berdasarkan peraturan)

Setelah peraturan ditentukan, maka hubungan dan entiti akan diekstrak mengikut peraturan seperti yang diterangkan dalam fasa 2.2. Kesemua peraturan tersebut akan digunakan semasa pengekstrakan entiti dan hubungan. Data yang diproses akan melalui algoritma dan senarai rangkap tiga akan dihasilkan. Ia akan disusun mengikut artikel masing-masing. Dengan cara ini lebih jelas bahawa terdapat beberapa ayat yang mempunyai lebih daripada satu rangkap tiga terutamanya ayat-ayat yang kompleks.

Jadual 3.5 menunjukkan contoh rangkap tiga yang dihasilkan mengikut jenis ayat. Dapat dilihat bahawa peraturan-peraturan yang dihasilkan ternyata dapat mengendalikan kedua ayat mudah dan juga ayat kompleks. Rangkap tiga yang dihasilkan dari ayat kompleks mempunyai lebih bilangan berbanding daripada yang dihasilkan dari ayat mudah. Ternyata ayat kompleks dapat mengekstrak lebih banyak maklumat.

Jadual 3.5 Contoh pengekstrakan rangkap tiga mengikut jenis ayat

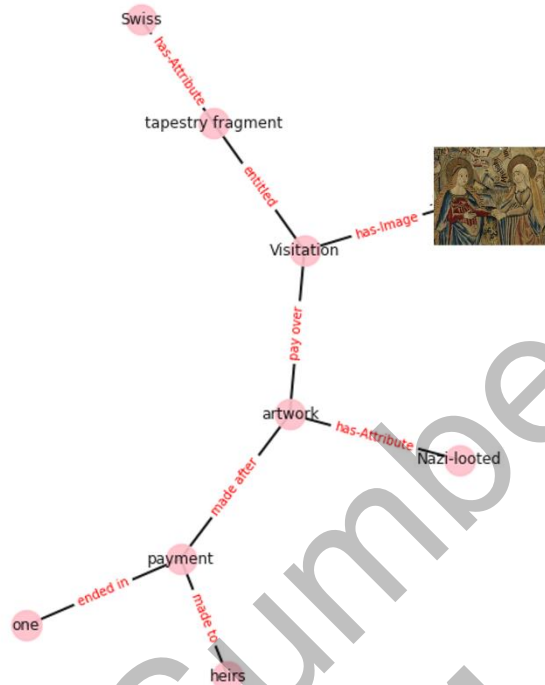
Jenis ayat	Contoh	Rangkap tiga yang diekstrak
Ayat mudah	<i>The Visitation: Glasgow City Council to pay family over Nazi-looted artwork</i>	<ul style="list-style-type: none"> <li>▪ (Visitation, has-Img, Image)</li> <li>▪ (Glasgow City Council, pay, family)</li> <li>▪ (Artwork, has-Attribute, Nazi-looted)</li> </ul>
Ayat kompleks dan panjang	<i>"The crew of an Emirates Airlines Boeing 777 prepares for passengers ahead of a demonstration flight in Dubai in 2007. Emirates announced plans Thursday to begin flying a Dubai to Panama City route on a 777, which will be the longest in the world."</i>	<ul style="list-style-type: none"> <li>▪ (Crew, prepares for, passengers)</li> <li>▪ (Flight, has-Attribute, demonstration)</li> <li>▪ (Emirates, announced, plans)</li> <li>▪ (Emirates, flying to, Panama City)</li> </ul>

### 3.3.3 Fasa 3 (Membina GP multimodal)

Pada fasa ini, hubungan teks-teks dan teks-visual akan digambarkan. Rangkap tiga yang diekstrak akan digabungkan dan menghasilkan GP yang boleh dirumuskan seperti berikut. G mewakili graf pengetahuan multimodal bagi sebuah artikel yang mengandungi gabungan beberapa subgraf atau juga rangkap tiga, R. n merupakan jumlah bilangan rangkap tiga yang diekstrak.

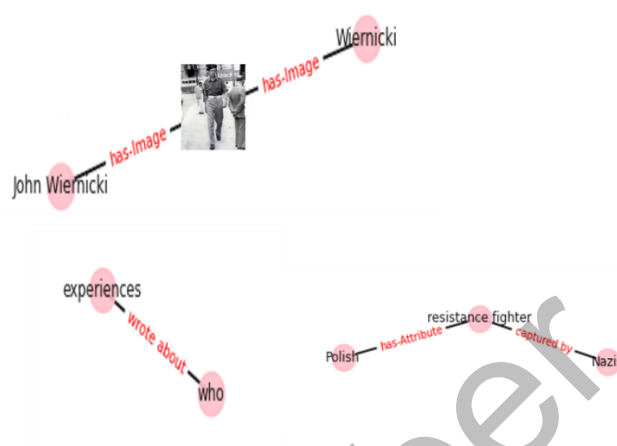
$$G = \{R_1, R_2, R_n \dots\} \quad \dots (3.5)$$

GP multimodal boleh divisualisasikan seperti dalam Rajah 3.15 dimana graf tersebut mempunyai dua jenis modaliti yang berbeza iaitu teks dan imej. Dari graf tersebut dapat dilihat terdapat subgraf yang lain. Satu subgraf yang dihasilkan mewakili satu rangkap tiga. Terdapat dua subgraf yang mempunyai entiti subjek yang sama, seperti token *firework*, masih mempunyai pertalian antara satu sama lain walaupun mereka mempunyai hubungan yang berbeza.



Rajah 3.15 GP multimodal yang akan dihasilkan setelah kesemua rangkap tiga digabungkan

Namun begitu, setelah eksperimen dilakukan didapati bahawa terdapat subgraf yang tergantung bersendirian seperti contoh Rajah 3.16 berikut. Jika berfikir seperti manusia, dapat dibuat kesimpulan bahawa subgraf yang tergantung itu masih mempunyai kaitan dengan imej. Hubungan ini boleh dilabelkan sebagai hubungan secara tidak langsung. Maka, pertambahan peraturan bagi hubungan secara tidak langsung akan dilaksanakan kepada subgraf yang tergantung itu. Hubungan ini akan dinamakan *has-Bg-Kg* dimana *Bg* mewakili *Background* manakala *Kg* mewakili *Knowledge*. Hubungan ini menunjukkan bahawa subgraf tersebut menerangkan pengetahuan latar belakang tentang sebuah imej selain mewujudkan pertalian yang lagi kukuh antara subgraf yang lain dan juga hubungan imej-teks tersebut.



Rajah 3.16 Contoh GP multimodal yang mempunyai subgraf tergantung

Peraturan 5 (hubungan *has-Bg-Kg*) ini terlebih dahulu perlu mengenal pasti entiti subjeknya. Oleh kerana *has-Bg-Kg* adalah hubungan yang menghubungkan antara imej dan teks maka imej nod akan dijadikan sebagai subjek. Bagi entiti objek pula, token yang bersesuaian untuk dijadikan sebagai objek merupakan subjek kepada subgraf yang tergantung itu. Kebiasaannya merupakan kata nama atau token yang mempunyai hubungan kebergantungan *nsubj*. Dengan adanya peraturan tambahan ini, maka masalah subgraf yang tergantung akan berkurangan juga.

### 3.4 KESIMPULAN

Bab ini menjelaskan mengenai rangka kerja bagi membina GP multimodal. Proses membina GP ini dimulakan dengan fasa 1 iaitu memproses dataset artikel berita yang terbahagi kepada dua bahagian iaitu tekstual dan visual. Hasil dari bahagian tekstual akan digunakan untuk mengekstrak entiti dan hubungan di fasa 2. Manakala, bahagian visual pula akan digunakan kemudian dalam fasa 2.3. Analisis perlu dilakukan terhadap beberapa artikel berita sebelum peraturan untuk mengekstrak entiti dan hubungan ditetapkan. Setelah itu, beberapa peraturan yang telah dikenal pasti seperti di fasa 2.2 akan digunakan untuk mengekstrak entiti dan hubungan dan akan terhasilnya rangkap tiga. Rangkap tiga ini kemudian akan digabungkan bagi membina GP multimodal dimana hubungan antara teks dapat dilihat dan juga hubungan imej-teks.

## BAB IV

### DAPATAN KAJIAN

#### 4.1 PENGENALAN

Bab ini membincangkan tentang dataset yang digunakan iaitu ION Corpus (Hollink et al. 2016) dan dataset ini ditapis dan diproses terlebih dahulu sebelum ia digunakan untuk pengesktrakan hubungan dan entiti. Juga menyatakan pustaka yang digunakan untuk mengaplikasikan teknik NLP (tokenisasi, *stopword removal*, *dependency parsing*) dan beberapa teknik NLP (tokenisasi, *stopword removal*) ini diubah suai mengikut keperluan kajian. Setelah rangkap tiga dihasilkan, perbandingan hasil antara kajian-kajian lain dan perkara yang boleh diperbaiki juga dibincangkan. Akhir sekali, pembinaan GP multimodal juga dinyatakan dalam bab ini.

#### 4.2 DATASET

Dalam kajian ini, dataset yang digunakan adalah dari *ION Corpus* (Hollink et al. 2016) yang terdiri daripada artikel berita yang diterbitkan di lima laman web surat khabar, yang dikumpulkan dari Ogos 2014 dan Ogos 2015. Terdapat sejumlah 319,475 artikel yang dikumpulkan. kebiasaannya, artikel berita mempunyai imej dan kapsyen disampingnya seperti yang ditunjukkan pada Rajah 4.1. Namun, ada kes di mana artikel tersebut mempunyai imej tetapi kapsyen tidak disediakan. Oleh kerana kajian ini mempertimbangkan imej dan teks, maka artikel yang tidak mempunyai imej akan digugurkan.

## The Visitation: Glasgow City Council to pay family over Nazi-looted artwork

The tapestry fragment dates to 16th-century Switzerland

Kashmira Gander | @kashmiragander | Friday 14 August 2015 00:01



A 16th-century Swiss tapestry fragment entitled *The Visitation* as a payment to be made to the heirs of a German-born Jewish woman after the artwork, looted by the Nazis, ended up in one of Scotland's best-known museums, the Burrell Collection in Glasg (Di Dosma/Glasgow Museums and Libraries Collections/PA Wire)

Rajah 4.1 Contoh artikel berita yang lengkap dengan imej dan kapsyen

Setelah semua artikel yang tidak mempunyai imej ditapis, jumlah artikel yang dikumpulkan menjadi 227,216 yang boleh dibahagikan seperti berikut:

- 21,462 (*Independent*)
- 82,424 (*Daily Mail*)
- 34,868 (*The Huffington Post*)
- 37,853 (*New York Times*)
- 50,609 (*The Washington Post*)

Daripada jumlah yang telah ditapis terlebih dahulu, kajian ini hanya menggunakan subset dari dataset ION Corpus. Subset dataset ini telah dilabel secara manual mengikut panduan yang diterangkan dalam Jadual 3.1 kerana tidak ada korpus artikel berita yang tersedia dengan label *triples*. Berdasarkan kajian Romadhony et al. (2018), mereka menggunakan 25 ayat dari dataset PANLOC bagi setiap 4 topik

menjadikan 100 ayat secara keseluruhan. Romadhony et al. juga memastikan ayat berikut merangkumi perbezaan dari segi tahap kesukaran (mudah dan sukar). Merujuk kepada Romadhony et al. (2018), eksperimen kajian ini pula menggunakan 60 artikel dari lima laman web surat khabar yang telah dinyatakan. 10 artikel pertama merupakan artikel yang dipilih secara rawak dan bagi selebih 50 artikel pula, setiap 10 diambil dari setiap laman web surat khabar. Secara keseluruhannya, 60 artikel merangkumi 109 ayat yang terdiri daripada ayat mudah dan juga kompleks.

### 4.3 PRA-PROSES DATASET

Artikel berita yang asal perlu disaring kerana mempunyai banyak maklumat yang dikumpulkan (tajuk, kapsyen, tarikh diterbitkan dan banyak lagi). Rajah 4.2 menunjukkan data yang telah diekstrak dari artikel berita yang akan digunakan terutamanya tajuk dan kapsyen.

A	B	C	D	E	F	G	H	I
0	ID	Headline	Caption	URL				
1	7d691fc97	The Visitation: Glasgow City Council to "A 16th-century Swiss tapestry fragment entitled Th		<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
2	fa2130bb2	Jack Warden: Admired political journa	Warden, centre, with Neil Kinnock and Margaret T	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
3	5beb2deb	John Wiernicki: Polish resistance fight	Wiernicki: escaped the Nazis on a march to a prisor	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
4	30a69ba0	Susan Sheridan: Voice actor behind a l	Sheridan, centre, with fellow â€”Hitchhikerâ€”s G	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
5	9057744b	Stolen tortoise leads to arrest of sexu	A one year old galapagos tortoise eats grass in Aust	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
6	5d4fd6d6c	Lewis Fogle: Prisoner freed after sper	Lewis Fogle with his wife, Deb, after his release fro	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
7	9de47518	Daniel Pena-Sanchez: Mother appeals	Daniel Pena-Sanchez as his mother has appealed fc	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
8	710aca6e	Milk farmer protests: Asda to increase	Asda is to increase the price it pays for milk MYCHE	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
9	d8241aa4	Rihanna to appear as 'special adviser'	Rihanna performs onstage at The Inaugural Diamoi	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
10	b6fdade6c	Burma's move towards democracy sul	The Union Solidarity and Development Party head	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
11	1ebf18af4	RNLI lifeboat crew rescue swimmer str	The RNLI rescue a swimmer in Gravesend RNLI	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				
12	1ebf18af4	RNLI lifeboat crew rescue swimmer str	The RNLI rescue a swimmer in Gravesend RNLI	<a href="http://static.independent.co.uk/s3fs-public/styles/s">http://static.independent.co.uk/s3fs-public/styles/s</a>				

Rajah 4.2 Contoh data yang diekstrak dari fail JSON

Namun, proses pembersihan dataset masih belum berakhir lagi. Teks ini perlu melalui proses tokenisasi dan *stopword removal*. Perpustakaan yang akan digunakan untuk teknik NLP yang dilaksanakan dalam eksperimen ini adalah *Spacy Library*. Kedua proses ini perlu diubah suai mengikut keperluan seperti yang diterangkan dalam Bab 3.2.1. Kod bagi *tokenizer* yang diubah suai dan senarai *stopword* yang digunakan seperti di Rajah 4.3. Secara umumnya, proses tokenisasi akan mengesan satu perkataan sebagai satu token tetapi terdapat kes dimana sesuatu perkataan itu perlu digabungkan menjadi satu token. Sebagai contoh, ayat yang mempunyai kata hubung ‘-’ dalam frasa ‘22-years-old’ akan dikesan sebagai beberapa token. Setelah diubah suai, frasa tersebut akan dikesan sebagai satu token sahaja.

```

def custom_tokenizer(nlp):
    infixes = (
        LIST_ELLIPSES
        + LIST_ICONS
        + [
            r"(?<=[0-9])[+\-\*^](?=[0-9])",
            r"(?<=[a-zA-Z0-9])\.(?=[a-zA-Z0-9])".format(
                al=ALPHA_LOWER, au=ALPHA_UPPER, q=CONCAT_QUOTES
            ),
            r"(?<=[a-zA-Z0-9])\.(?=[a-zA-Z0-9])".format(a=ALPHA),
            #r"(?<=[a-zA-Z0-9])\.(?=[a-zA-Z0-9])".format(a=ALPHA, h=HYPHENS),
            r"(?<=[a-zA-Z0-9])[:<=>|](?=[a-zA-Z0-9])".format(a=ALPHA),
        ]
    )

    infix_re = compile_infix_regex(infixes)

    return Tokenizer(nlp.vocab, prefix_search=nlp.tokenizer.prefix_search,
                    suffix_search=nlp.tokenizer.suffix_search,
                    infix_finder=infix_re.finditer,
                    token_match=nlp.tokenizer.token_match,
                    rules=nlp.Defaults.tokenizer_exceptions)

nlp = spacy.load("en_core_web_sm")
nlp.tokenizer = custom_tokenizer(nlp) #declare guna custom

stopwords = ['left', 'right', 'below', 'centre', 'above', 'left', 'right', 'below', 'centre', 'above']

```

Rajah 4.3 Kod bagi penyesuaian *tokenizer* dan *stopword removal*

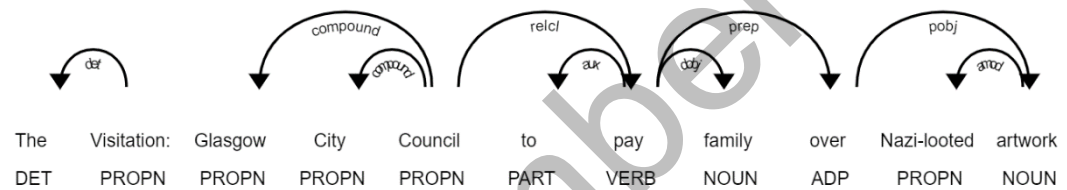
Bagi proses *stopword removal* pula, tertera di Rajah 4.3, beberapa perkataan telah dikenal pasti untuk dilabel sebagai *stopword*. Hal ini kerana perkataan-perkataan tersebut akan mempengaruhi prestasi penghurai ini. Output penghurai ini mesti tepat yang mungkin untuk mengelakkan ketidaktepatan dalam menangani ayat yang kompleks. Maka setelah teks-teks ini melalui proses pembersihan, seterusnya akan diproses berdasarkan 4 peraturan yang telah ditetapkan untuk menjadi pasangan entiti dan hubungan iaitu rangkap tiga bagi membina GP.

Seperti yang dinyatakan di rangka kerja fasa 1 (pra proses dataset), bagi bahagian visual, imej akan dimuat turun melalui senarai URL seperti di Rajah 4.2. Diberikan fail csv yang mengandungi senarai URL dan sistem akan memuat turun imej secara automatik. Dengan ini, terhasilnya dataset yang bersesuaian untuk kajian ini iaitu mengandungi kedua elemen (bahagian tekstual dan visual) yang akan digunakan bagi membina GP multimodal. Imej menjadi entiti yang mewakili bahagian visual yang akan berhubung dengan entiti teks. Namun begitu, semasa pembinaan GP multimodal, imej id digunakan bagi mewakili imej. Hal ini kerana terdapat cabaran dalam menulis kod untuk mengintegrasikan imej ke dalam nod.



#### 4.4 PENJANAAN RANGKAP TIGA

Dengan terhasilnya teks yang terhasil dari fasa pra proses dataset, teks ini akan melalui penghurai kebergantungan (*dependency parser*) dan akan mengeluarkan pokok kebergantungan (*dependency tree*) yang digambarkan seperti di rajah 4.4. Beberapa ayat akan digambarkan sebagai pokok kebergantungan sebagai langkah awal untuk membuat peraturan untuk mengekstrakan hubungan berdasarkan peraturan.



Rajah 4.4 Pokok kebergantungan yang dihasilkan

Seperti yang telah dinyatakan dalam bab 3, terdapat beberapa jenis entiti yang dikenal pasti. Sebagai contoh, *multiword entity*, *main entity* dan *attribute entity*. Bagi *multiword entity*, 'Council' mempunyai hubungan kebergantungan *COMPOUND* dengan dua perkataan yang lain maka frasa 'Glasgow City Council' perlu dianggap sebagai satu frasa. Jadi seperti Rajah 4.5, kod bagi menangkap beberapa token untuk dijadikan satu token ini adalah dengan melalui setiap token dan menangkap token yang mempunyai hubungan kebergantungan *COMPOUND*. Sistem hanya akan berhenti jika tiada lagi hubungan itu. Maka, terhasilnya satu entiti tapi mempunyai beberapa token di dalamnya.

```

### for multi-word entity ###
for token in doc:

    #extract compound nouns i.e. non-compositional phrases
    if(token.dep_=='compound'):
        if (not is_compound):
            is_compound = True
            start_i = token.i
        else:
            if (is_compound):
                #print('get end index\n')
                end_i = token.i
                compound_index_ls.append((start_i, end_i))
                #reset
                is_compound = False

    with doc.retokenize() as retokenizer:
        for indices in compound_index_ls:
            start, end = indices
            retokenizer.merge(doc[start:end+1])
            compound_ls.append(doc[start:end+1])

```

Rajah 4.5 Kod bagi pengekstrakan *multiword entity*

*Attribute entity* ialah dimana sebuah token akan menghuraikan ciri token yang lain. Dalam erti kata lain, berdasarkan Rajah 4.6, perkataan ‘*artwork*’ akan mempunyai ciri ‘*Nazi-looted*’. Juga menjadi entiti kepada hubungan *has-Attribute*. Jadi, kod untuk menangkap entiti jenis ini, seperti yang diterangkan dalam Peraturan 3 dalam Bab 3.2.2, adalah dengan mencari token yang mempunyai hubungan kebergantungan *AMOD* dan disimpan menjadi objek. Seterusnya, subjek akan dikenal pasti dan kebiasaannya merupakan token bersebelahan dengan *attribute entity* tersebut. Dengan ini, sebuah rangkap tiga yang mempunyai hubungan *has-Attribute* telah berjaya dikeluarkan.

```

### FOR AMOD(attribute) ENTITY ###
if(token.dep_=='amod'):
    subj_entity = token.nbor().text
    print( "amod", token.text , subj_entity)

    relation_attribute = "has-Attribute"
    triples.append(relation_attribute)

    amod_ent = token.text
    triples.append(subj_entity)
    triples.append(amod_ent)

```

Rajah 4.6 Kod bagi Peraturan 3 (mengekstrak entiti bagi hubungan *has-Attribute*)

*Main entity* merupakan entiti yang menghubungkan bahagian tekstual dan visual dengan perhubungan *has-Image*. Merujuk kepada Peraturan 4 dalam Bab 3.2.2, token yang dilabel *ROOT* akan dikenal pasti sebagai entiti tersebut seperti dalam Rajah 4.7.

```

det ( Visitation ||| The )
ROOT ( Visitation ||| Visitation )
punct ( Visitation ||| : )

```

Rajah 4.7 *Main entity* yang dikenal pasti

Bagi menangkap entiti jenis ini, berdasarkan Rajah 4.8, token yang mempunyai hubungan kebergantungan *ROOT* akan dikenal pasti dan token berkenaan juga mempunyai label penandaan POS kata nama (*NOUN*) atau frasa nama (*PROPN*). Kemudiannya akan disimpan menjadi objek. Bagi hubungan *has-Image* ini, subjeknya akan menjadi imej tetapi id imej akan mewakili imej yang sebenar.

```

### FOR STANDALONE ENTITY ###
#if(token.dep_=='ROOT' and token.pos_ == 'PROPN' or token.pos_ == 'NOUN'):
    if(token.dep_=='ROOT' and token.pos_ == 'PROPN'):
        #print("found root // propn token")
        #print(nexttoken.dep_ , nexttoken)
        try :
            nexttoken = token.nbor()

            if("punct" in nexttoken.dep_ and nexttoken.text == ':'):
                #print("next token is colon " , nexttoken)

                standalone_ent = token.text
                relation_standalone = "has-Image"

                triples.append(relation_standalone)
                triples.append(standalone_ent)
                triples.append(img_id)

```

Rajah 4.8 Kod bagi Peraturan 4 (mengeksktrak entiti bagi hubungan *has-Image*)

Seterusnya, Rajah 4.9 berikut merupakan Peraturan 1 yang ditetapkan untuk mengekstrak hubungan yang sedia ada dari ayat yang diberi. Pertama sekali, peraturan dimulakan dengan mengenal pasti token yang mempunyai kedua ciri tersebut iaitu hubungan kebergantungan *ROOT* dan penandaan POS *VERB*. Apabila hubungan tersebut telah dikenal pasti, fungsi pertalian dalam pokok kebergantungan dimanfaatkan. Jadi, subjek dan objek pula akan dikenal pasti yang kebiasaannya akan berada di bahagian kiri dan kanan dari token hubungan seperti yang dapat dilihat

dalam pokok kebergantungan di Rajah 4.4. Maka akan terhasilnya rangkap tiga yang diekstrak dari input ayat.

```

## rules-based statement for main verb ##

if(token.dep_=='ROOT' and token.pos_ == 'VERB'):

    #try:
    relation = token.text
    triples.append(relation)
    print("\n1r " + relation)
    print("head", token.head.text)

    for t in token.lefts:
        if 'subj' in t.dep_ :
            print(t.text)
            triples.append(t.text)
            #print("Left", token.Lefts)
    for t in token.rights:
        if 'obj' in t.dep_ :
            print(t.text)
            triples.append(t.text)

    print(token.rights)
    print("main verb ", triples)

```

Rajah 4.9 Kod bagi Peraturan 1 (pengekstrakan hubungan kata kerja sahaja)

Selain dari bergantung pada hubungan kata kerja sahaja, terdapat juga peraturan 2 iaitu untuk menggabungkan kata kerja dan kata preposisi seperti dalam Rajah 4.10. Pertama sekali, peraturan dimulakan dengan mengenal pasti token yang mempunyai ciri penandaan POS *VERB*. Kemudian, dari token tersebut, token lain yang mempunyai hubungan kebergantungan *prep* akan dikenal pasti untuk menggabungkan kedua token tersebut. Apabila hubungan tersebut telah diekstrak, subjek dan objek pula akan dikenal pasti sama kaedah seperti yang diterangkan sebelumnya.

```

### statement for combine relation (V + ADP) ###
elif(token.pos_=='VERB'):

    print("2r verb" , token.text)
    for t in token.rights:

        ##relation##
        if 'prep' in t.dep_ :
            print("combine ADP " + t.text)
            ADP = t.text
            gabung = token.text + " " + ADP
            triples.append(gabung)

        ##subject##
        subjL = [tsubj for tsubj in token.lefts if tsubj.dep_ == 'nsubj']
        print(subjL)
        if len(subjL) == 1:
            subj = subjL[0]
            triples.append(subj)

```

Rajah 4.10 Peraturan 2 (pengekstrakan gabungan hubungan kata kerja dan kata preposisi)

Setelah kesemua rangkap tiga telah dihasilkan, ia akan disimpan dalam sebuah senarai. Seperti yang digambarkan dalam Rajah 4.11, satu ayat dapat memiliki beberapa tiga rangkap jadinya terdapat senarai bersarang. Untuk menjelaskan secara ringkas, satu rangkap tiga disimpan dalam satu senarai dan jika dalam ayat yang sama, maka ia akan disimpan dalam senarai yang lain. Akan ada satu senarai besar yang menyimpan kesemua rangkap tiga yang telah diekstrak.

```

ALL
[[['has-Attribute', 'tapestry fragment', 'Swiss'], ['entitled', 'tapestry fragment', 'Visitation'], ['made to', 'payment', 'heirs'], ['made after', 'payment', 'artwork'], ['has-Attribute', 'Jewish', 'German-born'], ['has-Attribute', 'woman', 'Jewish'], ['looted by', 'Nazis'], ['ended up', 'payment'], ['ended in', 'payment', 'one'], ['has-Attribute', 'museums', 'best-known']], [['has-Attribute', 'lunch', 'centenary']], [['has-Image', 'Wiernicki', '5beb2deb2f802efd1697ac210d428bd21da8ae65'], ['escaped', 'Nazis']], [['has-Attribute', 'ãHitchhikerã99s', 'fellow'], ['cast', 'Shenidan', 'Guideã99', 'members', 'Simon Jones']], [['has-Attribute', 'galapagos tortoise', 'old'], ['eats', 'galapagos tortoise', 'grass'], ['has-Image', 'Getty Images News', '9057744b1920f4b7526bef7b0e5251f9a60aa747']], [], [['has-Attribute', 'teenage', 'unwell'], ['has-Attribute', 'son', 'teenage']], [['pays for', 'it', 'milk'], ['has-Image', 'Getty Images', '710aca6e60b69cf805ecda1fe2a8100f71b65d04']], [['performers', 'Rihanna'], ['presented by', 'Inaugural Diamond Ball', 'Rihanna'], ['presented at', 'Inaugural Diamond Ball', 'Vineyard'], ['presented on', 'Inaugural Diamond Ball', 'December']], [['remained', 'Party head office']], [['rescue', 'RNLI', 'swimmer']], [['rescue', 'RNLI', 'swimmer']], [['said'], ['has-Attribute', 'literary', 'Independent-sponsored'], ['has-Attribute', 'festivals', 'literary']], [['experienced', 'Xinjiang region', 'unrest'], ['has-Attribute', 'unrest', 'social']], [['shopping at', 'Holidaymakers', 'Heathrow'], ['has-Attribute', '5', 'terminal'], ['has-Image', 'Getty Images', '6bb9b07e60cd9582bde42b8df6cc5226a745edd8']], [['has-Attribute', 'Exeter University student accommodation blocks', 'many']], [], [], [['staged', 'Migrants', 'protest']], [['has-Image', 'CBI director general', '3f9a9e54cbb15fa3d3209c55e11449d53b522e44'], ['has-Attribute', 'flair', 'artisti c']]

```

Rajah 4.11 Rangkap tiga diekstrak dari ayat input

Rajah 4.12 menunjukkan transformasi rangkap tiga asalnya dalam bentuk senarai menjadi visualisasi yang lebih baik dan lebih kemas sehingga dapat difahami dengan mudah. Oleh kerana dataset mempunyai input ayat dari tajuk dan kapsyen, maka langkah-langkah ini perlu diulang kerana ia hanya mengekstrak satu demi satu. Ia hanya mengeluarkan rangkap tiga dari tajuk dan perlu diulang untuk bahagian kapsyen.

Headline	Caption	URL	Triples for headline
The Visitation: Glasgow City Council to pay fa...	"A 16th-century Swiss tapestry fragment entitl...	http://static.independent.co.uk/s3fs-public/st...	[[has-Image, Visitation, 7d691fc97d6123b2959fc...
Jack Warden: Admired political journalist whos...	"Warden, centre, with Neil Kinnock and Margare...	http://static.independent.co.uk/s3fs-public/st...	[[has-Image, Jack Warden, fa2130bbaf6a334c98cb...
John Wiernicki: Polish resistance fighter who ...	Wiernicki: escaped the Nazis on a march to a p...	http://static.independent.co.uk/s3fs-public/st...	[[has-Image, John Wiernicki, 5beb2deb2f802efd1...

Rajah 4.12 Rangkap tiga diekstrak dalam bentuk jadual

#### 4.5 PERBANDINGAN KAEDAH PENGEKSTRAKAN HUBUNGAN

Seksyen ini akan membincangkan mengenai perbandingan yang dilaksanakan antara dua kaedah dari karya lain seperti (Gong & Wang 2017; Romadhony et al. 2018).

Kaedah dua karya ini dibangunkan semula dan rangkap tiga yang terhasil akan dibandingkan dengan hasil rangkap tiga dari kaedah yang dicadangkan. Dalam Jadual 4.1, tertera imej bagi setiap artikel dan diiringi dua ayat iaitu tajuk dan kapsyen. Kaedah Gong & Wang (2017) akan dibandingkan dengan cara mereka mengekstrak hanya kata nama sahaja untuk menghubungkan antara teks dan juga visual. Kaedah Romadhony et al. (2018) pula mengekstrak hubungan yang boleh didapati dari teks itu sendiri. Peraturan yang mereka tetapkan ialah dari kata kerja dan juga hubungan kebergantungan *ADVMOD* dimana ia mengubah predikat atau kata kerja.

Jadual 4.1 berikut merupakan beberapa artikel terpilih daripada jumlah artikel iaitu 60 artikel yang digunakan untuk mengekstrak hubungan dan entiti bagi menghasilkan rangkap tiga. Pada ruangan di sebelah kiri sekali menunjukkan contoh setiap artikel yang mempunyai teks bersama imej nya, teks yang dilabel warna hijau merupakan entiti dan hubungan yang terlebih dahulu akan dianggap sebagai output rangkap tiga. Output ini dihasilkan secara manual oleh penganotasi berdasarkan panduan dalam Jadual 3.1 dan menjadi penanda aras bagi ketiga kaedah yang akan dibandingkan.

Terdapat juga teks yang dilabel warna kuning, ini merupakan rangkap tiga yang terhasil tetapi sama ada tidak lengkap ataupun terlebih. Sebuah rangkap tiga seharusnya mempunyai dua entiti dan satu hubungan sahaja. Rangkap tiga yang tidak menepati ciri ini akan diabaikan semasa proses pembinaan GP multimodal. Namun, hal seperti ini akan menyebabkan kualiti GP yang terhasil terjejas. Rangkap tiga yang diabaikan ini berkemungkinan mengandungi maklumat yang berkaitan disebabkan itu nilai perolehan kembali akan menjadi rendah. Faktor berlakunya pengekstrakan rangkap tiga yang tidak sempurna ini akan diberikan penjelasan.

Bagi teks yang berwarna merah pula merupakan rangkap tiga yang sempurna tetapi ia tidak sepatutnya diekstrak sebagai rangkap tiga kerana ia tidak berkait pada imej. Bagi rangkap tiga yang tidak sempurna dan tidak berkait ini akan diterangkan dengan lebih lanjut mengapa terjadi sedemikian. Teks berwarna biru menandakan teks yang masih boleh diterima sebagai entiti yang betul walaupun ianya kurang tepat. Teks berwarna oren menandakan perbezaan antara kaedah Romadhony et al. (2018) yang


mengekstrak kata kerja + *ADVMOD* dan kaedah yang dicadangkan iaitu kata kerja + preposisi.

Berdasarkan Jadual 4.1 berikut, dapat dilihat bahawa kaedah dicadangkan merupakan penambahbaikan daripada kedua hasil karya berikut. Hal ini kerana bagi kaedah Romadhony et al. (2018), penulis hanya bersasarkan dalam hubungan teks-teks sahaja. Bagi Gong & Wang (2017) pula, kaedah yang digunakan hanya fokus kepada hubungan imej-teks yang secara tidak langsung mengabaikan maklumat lain yang boleh didapati dari teks tersebut. Kaedah yang dicadangkan ini bukan sahaja berjaya mengekstrak hubungan teks-teks, malah hubungan imej-teks juga menjadikan ia sangat dipenuhi oleh maklumat untuk diisi di dalam GP multimodal yang akan dibina.

Selain itu, kaedah yang dicadangkan juga dapat menghasilkan beberapa rangkap tiga yang tepat dari sebuah ayat yang panjang dan kompleks. Hubungan *has-Attribute* menjadikan hubungan yang diekstrak lebih mencukupi dan kurangnya pengabaian maklumat dari teks. Hal ini kerana segala maklumat yang boleh diekstrak dari teks dapat digunakan dengan sepenuhnya. Bagi mengimbas kembali, berikut merupakan pelabelan berdasarkan warna yang berlainan dan maksudnya.

1. Label warna hijau merupakan rangkap tiga yang sepatutnya dihasilkan.
2. Label warna kuning ialah rangkap tiga yang terhasil tetapi sama ada tidak lengkap ataupun terlebih.
3. Label merah ialah rangkap tiga yang sempurna tetapi ia tidak sepatutnya diekstrak.
4. Label biru menandakan teks yang masih boleh diterima sebagai entiti yang betul walaupun ianya kurang tepat.
5. Label oren menandakan perbezaan antara kaedah Romadhony et al. (2018) yang mengekstrak kata kerja + *ADVMOD* dan kaedah yang dicadangkan iaitu kata kerja + preposisi.

Jadual 4.1 Jadual perbandingan untuk beberapa artikel

Artikel berita	Hasil rangkap tiga secara manual	Kaedah dicadangkan	Kaedah (Gong & Wang, 2017)	Kaedah (Romadhony et al., 2018)
<p>Artikel 3</p>  <p>[img03]</p> <p>John Wiernicki: Polish resistance fighter who was captured by the Nazis and later wrote about his experiences in Auschwitz</p> <p>Wiernicki: escaped the Nazis on a march to a prison camp</p>	<p>John Wiernicki – hasImg – img03, resistance fighter – hasAttr – Polish, resistance fighter – captured by – Nazis, resistance fighter- wrote – experiences</p> <p>Wiernicki – hasImg – img03, Wiernicki – escaped – Nazis</p>	<p>[[‘has-Image’, ‘John Wiernicki’, ‘img03’], [‘has-Attribute’, ‘resistance fighter’, ‘Polish’], [‘captured by’, ‘resistance fighter’, ‘Nazis’], [‘wrote about’, ‘who’, ‘experiences’]]</p> <p>[[‘has-Image’, ‘Wiernicki’, ‘img03’], [‘escaped’, ‘Nazis’]]</p>	<p>[[‘has-Image’, ‘John Wiernicki’, ‘img03’], [‘has-Image’, ‘resistance fighter’, ‘img03’], [‘has-Image’, ‘experiences’, ‘img03’]]</p> <p>[[‘has-Image’, ‘Wiernicki’, ‘img03’], [‘has-Image’, ‘march’, ‘img03’], [‘has-Image’, ‘prison camp’, ‘img03’]]</p>	<p>[[‘captured’, ‘resistance fighter’, ‘Nazis’], [‘later wrote’, ‘who’, ‘experiences’]]</p> <p>[‘escaped’, ‘Nazis’]</p> <p>bersambung...</p>
<p>Artikel 6</p>	<p>Lewis Fogle – hasImg – img06, prisoner – freed after – spending, prisoner – spending over – 30</p>	<p>[[‘spending in’, ‘after’, ‘prison’]]</p> <p>[‘freed’, ‘Prisoner’]</p>	<p>[[‘has-Image’, ‘Lewis Fogle’, img06], [‘has-Image’, ‘years’, img06], [‘has-Image’, ‘prison’,</p>	<p>[[‘spending’, ‘after’, years]]</p>



...sambungan



[img06]

"Lewis Fogle: Prisoner freed after spending over 30 years in prison, following DNA test"

"Lewis Fogle with his wife, Deb, after his release from the State Correctional Institution Bob Donaldson/Pittsburgh Post-Gazette via A

Artikel 9



years, prisoner – spending in – prison

Lewis Fogle – release from – State Correctional Institution Bob Donaldson

img06], ['has-Image', 'DNA test', img06]]

[[['has-Image', 'wife', img06], ['has-Image', 'release', img06]]

Rihanna – hasImg – img09, Rihanna – appear as – adviser, Rihanna – appear on – The Voice, adviser – hasAttr – special

[[['has-Attribute', 'adviser', 'special'

[[['has-Image', 'adviser', 'img09']]

Rihanna – performs – onstage, The Inaugural Diamond Ball – presented by – Rihanna, The Inaugural Diamond Ball – presented at – The Vineyard, The Inaugural Diamond Ball – presented

[[['presented by', 'Inaugural Diamond Ball', 'Rihanna', 'presented at', 'Inaugural Diamond Ball', 'Vineyard'], ['presented on', 'Inaugural Diamond Ball', 'December']] (3) ['appear', 'Rihanna']

[[['has-Image', 'Rihanna', 'img09'], ['Inaugural Diamond Ball', 'has-Image', 'img09']]

[[['perform onstage', 'Rihanna', 'Inaugural Diamond Ball']]

bersambung...

...sambungan

on – December

['performs', 'Rihanna']

[img09]

Rihanna to appear as special adviser on The Voice

Rihanna performs onstage at The Inaugural Diamond Ball presented by Rihanna and The Clara Lionel Foundation at The Vineyard on 11 December 2014.

Artikel 11



[img11]

RNLB lifeboat crew rescue swimmer stranded in water for four hours off Gravesend coast – video

The RNLB rescue a swimmer in Gravesend RNLB

RNLB lifeboat crew – hasImage – img11, RNLB lifeboat crew – rescue swimmer, swimmer – stranded in – water, swimmer – stranded for – four hours, swimmer – stranded off – Gravesend coast

RNLB – rescue – swimmer

[[ 'has-Image', 'RNLB lifeboat crew rescue swimmer', 'img11' ], [ 'stranded in', 'RNLB lifeboat crew rescue swimmer', 'water' ], [ 'stranded for', 'RNLB lifeboat crew rescue swimmer', 'hours' ]]

[[ 'rescue', 'RNLB', 'swimmer' ]], (1) [ 'stranded off', 'RNLB lifeboat crew rescue swimmer' ]

[[ 'has-Image', 'RNLB lifeboat crew rescue swimmer', 'img11' ], [ 'has-Image', 'water', 'img11' ], [ 'has-Image', 'hours', 'img11' ], [ 'has-Image', 'video', 'img11' ]],

[[ 'has-Image', 'swimmer', 'img11' ], [ 'has-Image', 'Gravesend RNLB', 'img11' ]]

[[ 'stranded', 'RNLB lifeboat crew rescue swimmer', 'water' ], [ 'stranded', 'RNLB lifeboat crew rescue swimmer', 'hours' ]]

[[ 'rescue', 'RNLB', 'swimmer' ]]

bersambung...

...sambungan

Artikel 17



[img17]

US marines who lowered flag at embassy in Havana fifty years ago set for Cuba return

"James Tracy, Larry Morris and Mike East US State Dept"

Artikel 21



[Img21]

E.T. drove home: Master mechanic makes space-age vehicles suitable for alien invaders from normal family cars

US marines – lowered – flag, US marines – lowered at – embassy, US marines – set for – Cuba return

James Tracy – hasImg – img17, Larry Morris – hasImg – img17, Mike East – hasImg – img17

[[ 'lowered at', 'who', 'embassy'], ['set for', 'US marines', 'Cuba'], ['return', 'US marines']]

[[ 'has-Image', 'US marines', 'img17'], ['has-Image', 'flag', 'img17'], ['has-Image', 'embassy', 'img17'], ['has-Image', 'years', 'img17']]

[[ 'lowered', 'who', 'flag']]

mechanic – hasImg – img21, mechanic – hasAttr – master, mechanic – makes – space-age vehicles, family cars – hasAttr – normal

mechanic – begins – process, mechanic – begins with – family cars, family cars – hasAttr – ordinary

[[ 'has-Attribute', 'invaders', 'alien'], ['has-Attribute', 'family cars', 'normal']], ['makes', 'Master mechanic']]

[[ 'begins', 'Master mechanic Michael Vetter', 'process'], ['has-Attribute', 'family cars', 'ordinary']]

[[ 'has-Image', 'Master mechanic', 'Img21'], ['has-Image', 'space-age vehicles', 'Img21'], ['has-Image', 'invaders', 'Img21'], ['has-Image', 'family cars', 'Img21']]

[[ 'has-Image', 'process', 'Img21'], ['has-Image', 'family cars', 'Img21'], ['has-Image', 'chassis', 'Img21']]

['begins', 'Master mechanic Michael Vetter', 'process']

bersambungan...

...sambungan

Master mechanic Michael Vetter begins the process with ordinary family cars before extending the chassis

Artikel 30



[Img30]

Widow sobs as inquest hears that Germanwings passenger's remains could only be identified by DNA samples after co-pilot deliberately crashed it into French Alps

"Martyn Matthews, 50, died in the Germanwings Airbus crash, pictured with his wife Sharon, 48, (centre right) and children, Jade, 20, and Nathan, 23"

Artikel 32

passenger's remains – identified by – DNA samples, co-pilot – crashed into – French Alps

[[ 'has-Attribute', 'hears', 'inquest', ['hears', 'Widow sobs'], ['crashed into', 'co-pilot', 'Alps'], ['has-Attribute', 'Alps', 'French'] ]]

[[ 'has-Image', 'Widow sobs', 'Img30'], ['has-Image', 'inquest', 'Img30'], ['has-Image', 'Germanwings passenger', 'Img30'], ['has-Image', 'remains', 'Img30'], ['has-Image', 'DNA samples', 'Img30'], ['has-Image', 'co-pilot', 'Img30'], ['has-Image', 'Alps', 'Img30'] ]]

['crashed', 'co-pilot', 'Alps'], ['hears', 'Widow sobs']

Martyn Matthews – hasImg – img30, Martyn Matthews – died in – Germanwings Airbus crash, Martyn Matthews – pictured with – his wife, Matthews – pictured with – children

[[ 'pictured with', 'Martyn Matthews', 'wife' ]], ['died', 'Martyn Matthews']

[[ 'has-Image', 'Germanwings Airbus crash', 'Img30'], ['has-Image', 'wife', 'Img30'], ['has-Image', 'centre', 'Img30'], ['has-Image', 'children', 'Img30'] ]]

[[ 'pictured 23', 'Martyn Matthews', 'wife' ]]

Police dog – hasImg – img32, Police dog – dies in – squad car

[[ 'Dies', 'Police Dog', 'Fails', 'Air Conditioner' ]]

[[ 'has-Image', 'Squad Car', 'Img32' ]]

-.

Wix – hasImg – img32, Wix – died in – Wisconsin, K-9 – hasAttr – 3-year-

[[ 'died in', 'who', 'Wisconsin'], ['being in', 'after', 'car'], ['has-

[[ 'has-Image', '3-year-old K-9', 'Img32'], ['has-Image', 'week', 'Img32'],

[[ 'died', 'who', 'Wisconsin'], ['being', 'after', 'car']] bersambung...

...sambungan



[Img32]

Police Dog Dies In Squad Car After Air Conditioner Fails

he 3-year-old K-9 who died in Wisconsin this week after being left in a squad car."

old, Wix – left in – squad car

Attribute', 'car', 'squad']  
['died after', 'who]

['has-Image', 'car', 'Img32']]

Artikel 54



[Img54]

Tiger Woods will play in Masters

Tiger Woods tees off in the first round of the Phoenix Open in January. (Rick Scuteri/AP)

Tiger Woods – hasImg – img54, Tiger Woods – play in – Masters

['play', 'Tiger Woods']]

['has-Image', 'Masters', 'Img54']]

Tiger Woods – tees off – round, round – hasAttr – first

['has-Image', 'Tiger Woods tees', 'Img54'],  
['has-Attribute', 'round', 'first']]

['has-Image', 'Tiger Woods tees', 'Img54'],  
['has-Image', 'round', 'Img54']]

Tambahan lagi, peraturan tambahan yang digunakan oleh Romadhony et al. (2018) ialah kata kerja dan juga hubungan kebergantungan *ADVMOD*. Namun begitu, dalam 60 artikel yang telah diproses terdapat beberapa sahaja rangkap tiga yang mempunyai hubungan sedemikian seperti dalam artikel 3, 9, 30 dan 48. Hal ini mungkin disebabkan oleh perbezaan keperluan dari segi bahasa yang digunakan. Romadhony et al. mengaplikasikan peraturan berdasarkan bahasa Indonesia manakala kajian ini pula berdasarkan bahasa Inggeris. Penggunaan hubungan kebergantungan ini adalah untuk mengekstrak kata kerja secara sepenuhnya seperti dalam artikel 3. Rangkap tiga yang dihasilkan ialah [*later wrote*, *who*, *experiences*] dimana kata kerja asalnya hanya *wrote* tapi dengan adanya penggunaan hubungan kebergantungan ini, dapat diketahui bahawa perbuatan ini telah dilakukan pada waktu bila.

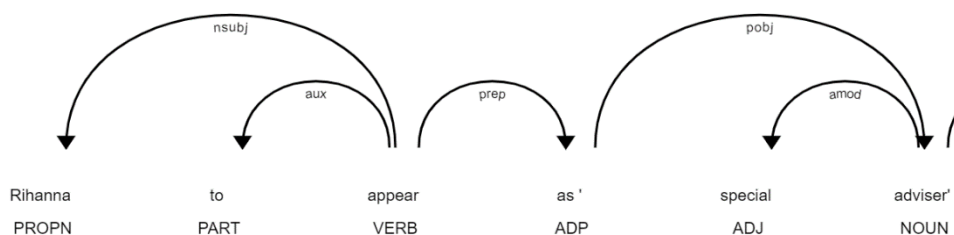
Namun, terdapat juga kes dimana rangkap tiga yang dihasilkan salah dan tersasar jauh dari makna asal ayat seperti dalam artikel 30, "*Martyn Matthews, 50, died in the Germanwings Airbus crash, pictured with his wife Sharon, 48, and children, Jade, 20, and Nathan, 23*". Rangkap tiga yang terhasil [*pictured 23*, *Martyn Matthews*, *wife*] memang jelas tidak tepat. Hal ini disebabkan oleh penghurai yang digunakan telah mengesan 23 itu mempunyai hubungan kebergantungan *ADVMOD*. Maka, kaedah yang dicadangkan iaitu berdasarkan Peraturan 2 (kata kerja + kata preposisi) adalah lebih sesuai digunakan untuk mengekstrak hubungan bagi dataset ini. Merujuk kepada artikel 30, berbeza dengan output dari kaedah Romadhony et al. (2018), kaedah yang dicadangkan telah berjaya mengekstrak rangkap tiga yang lebih tepat; [*pictured with*, *Martyn Matthews*, *wife*].

Hubungan *has-Attribute* yang diekstrak berdasarkan peraturan 3 menjadikan hubungan yang diekstrak lebih mencukupi. Hal ini kerana segala maklumat yang boleh diekstrak dari teks dapat digunakan dengan sepenuhnya. Sebagai contoh, bagi sebahagian kes seperti artikel 54 seharusnya mempunyai beberapa rangkap tiga yang terhasil berdasarkan hubungan kata kerja. Tetapi disebabkan kerumitan dari penghurai, maka hanya rangkap tiga bagi hubungan *'has-Image'* dan *'has-Attribute'* yang dapat diekstrak

yang memberi kelebihan dimana jika dibandingkan dengan output dari Romadhony et al. (2018), tiada rangkap yang dapat diekstrak.

Berdasarkan peraturan 4 yang telah ditetapkan untuk mengekstrak entiti subjek kepada hubungan *has-Image* ialah token yang mempunyai penandaan POS *NOUN* dan juga hubungan kebergantungan *ROOT*. Peraturan ini lebih sesuai untuk mengekstrak entiti bagi hubungan *has-Image* tanpa mengekstrak entiti yang tidak diperlukan. Sebagai contoh, kaedah (Gong & Wang, 2017) dapat mengekstrak banyak rangkap tiga untuk hubungan *has-Image* tetapi dapat dilihat dalam Jadual 4.1, hanya segelintir sahaja yang tepat untuk imej yang dipaparkan.

Kaedah yang dicadangkan ini masih lagi mempunyai ruang untuk diperbaiki. Antaranya ialah rangkap tiga yang diekstrak tidak lengkap seperti dalam artikel 3, 5, 6, 9 dan selebihnya berdasarkan di dalam Jadual 4.1 dimana subjek atau objek tidak diekstrak. Sebagai contoh, [*escaped*, *Nazis*], [*leads*, *tortoise*] dan juga [*appear*, *Rihanna*]. Berdasarkan Rajah 4.13, secara teori, penghurai mampu menangkap token yang mempunyai hubungan kebergantungan prep dan penandaan POS *VERB*. Hal ini kerana terdapat peraturan untuk menangkap kedua ciri tersebut untuk dijadikan sebagai satu hubungan [*appear as*] seperti dalam peraturan 2. Begitu juga dengan entiti objek [*adviser*] yang boleh dikenal pasti dengan hubungan kebergantungan *prep*. Akhirnya, akan menghasilkan rangkap tiga yang lengkap seperti [*appear as*, *Rihanna*, *adviser*]. Namun begitu, hasil daripada eksperimen ini mendapati bahawa hanya rangkap tiga [*appear*, *Rihanna*] mampu diekstrak oleh peraturan yang telah dibangunkan.



Rajah 4.13 Contoh artikel

Selain itu, prestasi penghurai kebergantungan menurun bagi ayat kompleks yang mengandungi klausa bergantung (*dependent clause*). Sebagai contoh, dalam artikel 17, rangkap tiga yang dihasilkan [*lowered at*, *who*, *embassy*] mempunyai subjek *who* dimana jika rujuk kepada ayat asalnya “*US marines who lowered flag at embassy in Havana fifty years ago set for Cuba return*”, triples yang harus diekstrak adalah [*set for*, *US marines*, *return*] dan [*lowered*, *US marines*, *flag*]. ‘*who*’ merupakan kata ganti relatif dalam klausa “*who lowered flag at embassy in Havana fifty years ago*” yang merujuk kepada subjek *US Marines* dalam ayat kompleks ini. Namun, ia disalah anggap sebagai entiti. Begitu juga dengan rangkap tiga yang terhasil dari artikel 32, [*died in*, *who*, *Wisconsin*] dimana “*who*” dalam klausa “*who died in Wisconsin this week*” itu merujuk kepada *Wix*, ayat asalnya “*Wix, the 3-year-old K-9 who died in Wisconsin this week after being left in a squad car.*””. Perkembangan dalam penyelidikan NLP akan meningkatkan lagi prestasi penghurai kebergantungan yang akan menambah baik hasil pengekstrakan teks. Teknik resolusi *co-reference* iaitu sesuatu tugas untuk mencari ungkapan yang merujuk kepada entiti yang sama dalam teks juga boleh dikaji bagi memperbaiki lagi hasil pengekstrakan.

Di samping itu, beberapa entiti yang diekstrak dari artikel 11, 18 dan 54 mempunyai hasil yang tidak berapa tepat. Sebagai contoh, rangkap tiga dari artikel 11, [*has-Image*, *RNLI lifeboat crew rescue swimmer*, *img11*], dimana entiti yang telah diekstrak adalah *RNLI lifeboat crew rescue swimmer*. Entiti ini sepatutnya menjadi satu set rangkap tiga yang lain iaitu (*rescue – RNLI lifeboat crew – swimmer*) dan bukannya diekstrak sebagai satu entiti. Entiti yang sepatutnya diekstrak adalah *RNLI lifeboat crew*



sahaja. Seterusnya, bagi artikel 18 dan 54 pula, [*'has-Attribute'*, *'peshmerga fighter fires'*, *'Kurdish'*] dan [*'has-Image'*, *'Tiger Woods tees'*, *'Img54'*], persamaan antara dua rangkap tiga ini ialah entiti yang seharusnya diekstrak adalah kata nama sahaja. Namun dalam kes ini, penghurai mengekstrak juga kata kerja yang seharusnya menjadi hubungan kepada rangkap tiga yang lain. Contoh dalam artikel 18, perkataan *fires* seharusnya diekstrak sebagai hubungan untuk menghasilkan rangkap tiga seperti (*fires on – fighter – Isis*). Kesemua entiti ini telah disalah ekstrak kerana penghurai telah melabel kata kerja berikut sebagai kata nama. Namun begitu, kajian ini masih mengambil kira rangkap tiga seperti berikut walaupun output yang dihasilkan tidak seperti yang diinginkan.

```
" PUNCT ROOT 0
Lewis Fogle PROPN dep 1
: PUNCT punct 2
```

Rajah 4.14 Contoh token yang dilabel tidak tepat

Tambahan lagi, beberapa artikel tidak mempunyai hubungan *has-Image* walaupun dataset yang digunakan telah disaring dimana artikel yang mempunyai imej sahaja akan digunakan. Salah satu contoh ialah pada hasil rangkap tiga bagi artikel 6 dalam Jadual 4.1. Hubungan *has-Image* tidak diekstrak walaupun artikel tersebut mempunyai imej. Bagi kajian ini, berdasarkan peraturan 4 yang ditetapkan untuk mengekstrak entiti subjek kepada hubungan *has-Image* ialah hanya token yang mempunyai penandaan POS *NOUN* dan juga hubungan kebergantungan *ROOT*. Disebabkan oleh output penghurai seperti Rajah 4.14, token *Lewis Fogle* yang seharusnya menjadi entiti subjek tidak dapat diekstrak kerana penghurai telah mengenal pasti symbol (“) sebagai token yang berlabel *ROOT* dan simbol berikut tidak mempunyai penandaan POS *NOUN*. Walaupun token *Lewis Fogle* tersebut mempunyai penandaan POS *NOUN* disebabkan ia tidak mempunyai hubungan kebergantungan *ROOT*, jadi ia tidak diekstrak sebagai entiti. Maka, tiada rangkap tiga yang dapat dihasilkan.

Manakala, rangkap tiga yang dihasilkan oleh kaedah Gong & Wang (2017), [*'has-Image'*, *'Lewis Fogle'*, *img06*], dapat menangkap entiti subjek bagi hubungan ini. Hal ini

kerana kaedah yang digunakan Gong & Wang hanya menangkap segala kata nama dan frasa nama yang boleh didapati dalam teks tersebut tanpa mengambil kira token yang menjadi topik utama (*ROOT*) untuk ayat tersebut. Berikut merupakan contoh antara rangkap tiga lain yang dihasilkan, ['has-Image', 'years', *img06*], ['has-Image', 'prison', *img06*], dimana imej bagi artikel 6 merupakan imej seorang lelaki yang bernama *Lewis Fogle* dan bukannya menggambarkan tahun atau penjara.

Stolen VERB amod 0	Milk farmer protests NOUN ROOT 0
tortoise NOUN nsubj 1	: PUNCT punct 1
leads VERB ROOT 2	Asda PROPN nsubj 2
to PART aux 3	to PART aux 3
arrest NOUN xcomp 4	increase VERB acl 4
of ADP prep 5	amount NOUN dobj 5
sexual ADJ amod 6	it PRON nsubj 6
assault suspect NOUN ROOT 7	pays VERB parataxis 7
	suppliers NOUN dobj 8
	for ADP prep 9
	product NOUN pobj 10

Rajah 4.15 Contoh token yang tidak sepatutnya dilabel *ROOT*

Di samping itu, merujuk kepada artikel 5 dan 8, imej yang dipaparkan merupakan imej seekor kura-kura dan juga segelas susu tetapi berikut merupakan rangkap tiga yang diekstrak ['has-Image', 'assault suspect', 'img05'] dan ['has-Image', 'Milk farmer protests', 'img08'] yang tidak mempunyai kaitan tentang imej yang dipaparkan. Sama seperti kes sebelumnya dimana output penghurai melabel token yang salah sebagai *ROOT* seperti dalam Rajah 4.16. Maka, rangkap tiga yang dihasilkan tidak berapa tepat.

Ringkasannya, dapat disimpulkan bahawa peraturan yang dibangunkan dapat mengekstrak rangkap tiga dari ayat mudah mahupun ayat kompleks. Peraturan 2 (kata kerja + preposisi) lebih sesuai digunakan untuk dataset Bahasa Inggeris berbanding dengan peraturan kata kerja + *ADVMOD*. Peraturan 3 (*has-Attribute*) dapat mengekstrak maklumat dengan lebih lengkap. Bagi Peraturan 4 (*has-Img*) hanya mengekstrak entiti penting yang berkait dgn imej sahaja. Tidak semua kata nama di dalam teks adalah entiti yang berkait dengan imej.

Eksperimen kajian ini menggunakan 60 artikel yang mempunyai jumlah 109 ayat dari tajuk dan kapsyen yang telah diteliti secara manual oleh penganotasi bagi menghasilkan senarai rangkap tiga. Berdasarkan Jadual 4.2, 319 rangkap tiga akan dijadikan sebagai penanda aras. Senarai ini kemudiannya akan dibandingkan dengan rangkap tiga yang dihasilkan oleh sistem. Prestasi sistem ini akan dinilai dengan formula *precision (P) and recall (R)*.

$$P = \frac{\text{jumlah rangkap tiga yang dihasilkan dengan tepat oleh algoritma}}{\text{jumlah rangkap tiga yang dihasilkan}} \quad \dots (4.1)$$

$$P = \frac{190}{209} = 0.90$$

$$R = \frac{\text{jumlah rangkap tiga yang dihasilkan dengan tepat oleh algoritma}}{\text{jumlah rangkap tiga yang sepatutnya diekstrak}} \quad \dots (4.2)$$

$$R = \frac{190}{319} = 0.60$$

Jadual 4.2 Ringkasan jumlah rangkap tiga yang terhasil

Jumlah // Kaedah		Kaedah yang dicadangkan	Kaedah (Gong & Wang, 2017)	Kaedah (Romadhony et al., 2018)
Rangkap tiga sepatutnya diekstrak (319)	Rangkap tiga yang diekstrak oleh algoritma	209	314	71

bersambung...

...sambungan				
	Rangkap tiga yang diekstrak dengan tepat	190	147	60
	Skor Ketepatan (Precision)	0.90	0.47	0.85
	Skor Perolehan Kembali (Recall)	0.60	0.46	0.20

Berdasarkan Jadual 4.2, disini dapat lihat bahawa  $P = 0.90$  dan  $R = 0.60$  bagi kajian ini merupakan yang tertinggi jika dibandingkan dengan dua lagi kaedah lain. Maka, dapat disimpulkan bahawa kaedah yang dicadangkan ini berjaya mengekstrak rangkap tiga yang lebih banyak dan yang penting rangkap tiga yang tepat. Nilai ketepatan (P) disini tinggi kerana antara 209 rangkap tiga yang dihasilkan, terdapat 190 rangkap tiga yang dihasilkan dengan betul. Namun begitu, nilai perolehan kembali (R) hanya mempunyai nilai 0.6 dimana hanya 190 rangkap tiga yang terhasil daripada 319 rangkap tiga yang secara teorinya mampu dihasilkan.

Kaedah yang dicadangkan ini sangat kepada bergantung prestasi penghurai kebergantungan. Jika prestasi penghurai ini rendah maka ia akan mempengaruhi hasilnya juga. Seperti yang telah diterangkan dalam perenggan sebelumnya, beberapa rangkap tiga yang terhasil tidak mencukupi entiti. Jika prestasi penghurai ini dapat ditingkatkan maka nilai R juga boleh ditingkatkan.

#### 4.6 PEMBINAAN GP MULTIMODAL

Seperti yang dinyatakan sebelum ini, entiti dan hubungan adalah elemen penting yang diperlukan untuk membina graf pengetahuan. Fasa sebelumnya telah pun berjaya mencapai dua elemen tersebut maka fasa seterusnya akan membina GP yang boleh dilihat secara keseluruhan iaitu gabungan kesemua artikel dan juga graf bagi setiap artikel berita. Nod akan mewakili entiti dan *edge* akan mewakili hubungan. Bagi membina graf ini, pustaka yang telah digunakan ialah pustaka *networkX*. Pustaka ini merupakan pakej *Python* untuk penciptaan, manipulasi, kajian struktur dinamik dan fungsi rangkaian yang

kompleks. Pakej *Matplotlib Python* juga digunakan bagi memvisualisasi kan graf yang telah dibina.

Berdasarkan Rajah 4.16, langkah pertama yang perlu dilakukan adalah dengan menggabungkan kedua bahagian, tajuk dan kapsyen. Kemudian, fungsi *Graph()* akan dipanggil tetapi ini hanyalah graf kosong sahaja tanpa maklumat di dalamnya. Jadi fungsi *add\_node* akan digunakan bagi melabelkan entiti kepada nod. Semasa proses mengekstrak rangkap tiga, terdapat sesetengah kes dimana rangkap tiga yang diekstrak tidak melengkap syarat utama nya iaitu perlu dua entiti dan satu hubungan. Jadi, sebarang rangkap tiga yang tidak lengkap akan dibuang terlebih dahulu sebelum menambahkan nod. Setelah itu, fungsi *add\_edge* akan digunakan bagi menghubungkan kedua nod yang telah dimasukkan.

```
import networkx as nx
import matplotlib.pyplot as plt
import matplotlib.image as mpimg

for document in dataset.iterrows():
    ## combine headline + caption per document(row) ##
    triples_ = dataset['Triples for headline'] + dataset['Triples for caption']

    G = nx.Graph()
    for triples in triples_: #[[,]]#
        for triple in triples:
            if len(triple) == 3:
                #print(triple)

                G.add_node(triple[1]) #ent1
                G.add_node(triple[2]) #ent2
                G.add_edge(triple[1], triple[2], label=triple[0])

    pos = nx.spring_layout(G)
    plt.rcParams['figure.figsize'] = [25,22]
    fig = plt.figure()

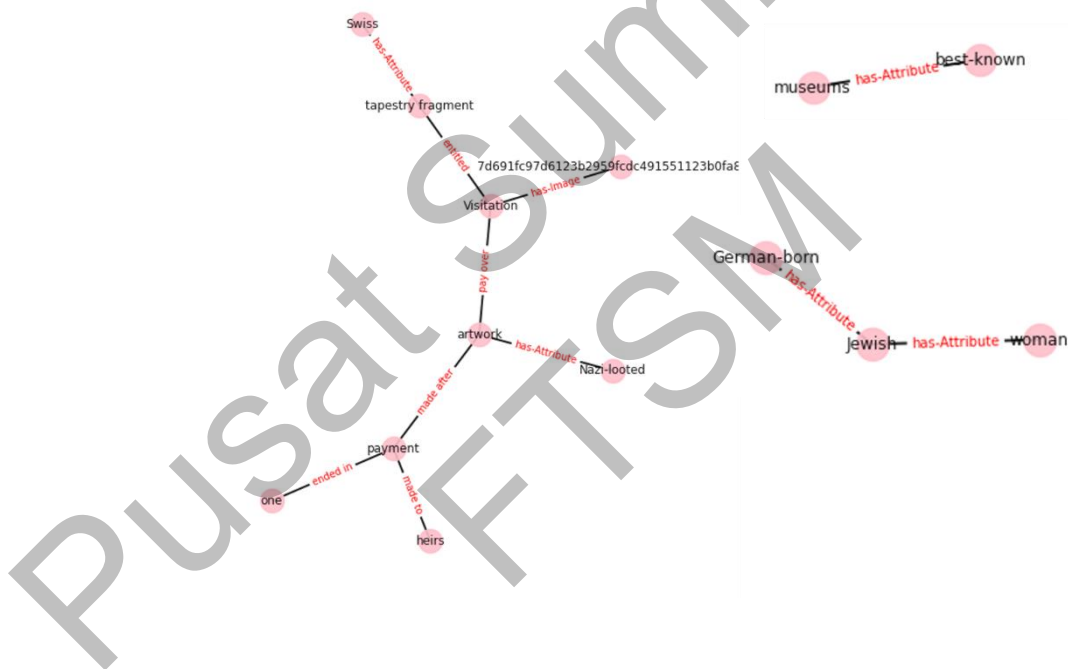
    ax = fig.add_subplot(111)
    ax.set_title('KG by article ', fontsize=18, fontweight='bold')
    nx.draw(G, pos, node_color = 'pink', width=2, linewidths=2,
            node_size=550, alpha=0.9,
            labels={node: node for node in G.nodes()})
```

Rajah 4.16 Cebisan kod untuk visualisasikan GP

Sesudah kesemua rangkap tiga dimasukkan ke nod dan *edge*, GP akan digambarkan secara keseluruhan dan juga satu persatu mengikut artikel. Dalam Rajah 4.17, disini dapat dilihat bahawa sebuah nod akan mempunyai sekurang-kurangnya satu

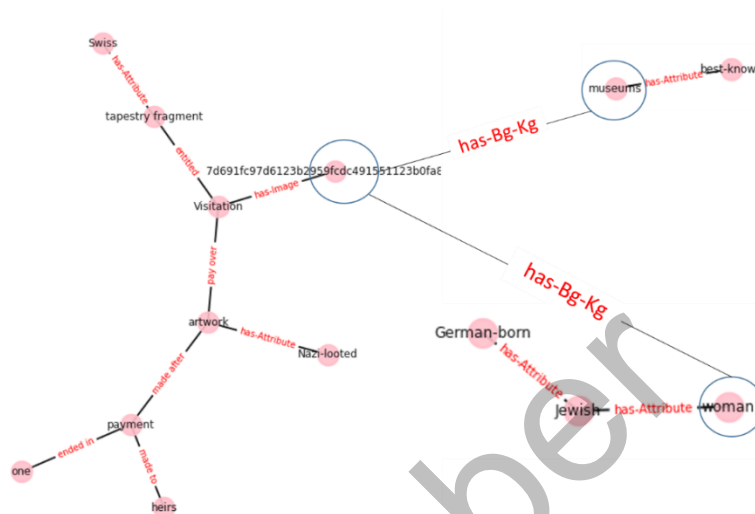
atau lebih hubungan dengan nod yang lain. Apabila rangkap tiga yang diekstrak dipaparkan sebagai graf, dapat dilihat dengan lebih jelas lagi bahawa sebuah rangkap tiga mempunyai kaitan dengan rangkap tiga yang lain.

Walaupun GP yang dihasilkan kelihatan mempunyai banyak perhubungan tetapi ada juga kes dimana subgraf yang dikeluarkan merupakan subgraf yang tergantung bersendirian seperti di Rajah 4.17. Perhubungan *has-Bg-Kg* akan diaplikasikan kepada subgraf yang tergantung ini bagi mewujudkan pertalian yang lagi kukuh antara subgraf yang lain dan juga hubungan imej-teks tersebut.



Rajah 4.17 GP setiap artikel

Subgraf yang tergantung akan dikenal pasti terlebih dahulu dengan melintasi semua hubungan graf dan kenal pasti graf yang terputus menggunakan fungsi *is\_connected*. Setelah itu, imej nod akan dilabel sebagai subjek dan nod yang mempunyai ciri subjek dalam subgraf tersebut akan menjadi objek bagi hubungan *has-Bg-Kg* dan maka akan terhasilnya GP yang lebih baik seperti contoh dalam Rajah 4.18.



Rajah 4.18 Contoh GP yang diperbaiki

Terdapat lagi ruang untuk baik pulih GP yang telah dibina ini. Antaranya ialah menggunakan imej yang sebenar daripada menggunakan imej ID supaya lebih jelas lagi bahawa GP yang dihasilkan ini mempunyai kedua hubungan teks-teks dan imej-teks. Juga, selain imej, konsep imej boleh dimasukkan di dalam GP multimodal ini. Konsep imej merupakan sesuatu konsep yang menggambarkan tentang sesebuah imej. Imej akan dijadikan sebagai subjek dan konsep imej sebagai objek yang berhubung dengan hubungan *has-Img-Concept*. Maka, akan wujudnya hubungan antara visual-visual. Di samping itu, graf yang dihasilkan merupakan graf yang tidak berarah (*undirected*) yang mungkin akan menimbulkan kekeliruan untuk membaca graf. Jadi lebih sesuai untuk menjadikannya graf yang berarah (*directed*) dimana nod yang mempunyai anak panah yang keluar darinya akan dikenali sebagai nod subjek dan nod objek pula dimana anak panah itu berpandukan.

#### 4.7 KESIMPULAN

Bab ini telah membincangkan secara teknikal bagaimana artikel berita diproses sehingga menjadi GP multimodal. Bermula dengan artikel berita ditapis dan dibersihkan, dan seterusnya teks yang dihasilkan akan digunakan untuk mengekstrak entiti dan hubungan

berdasarkan peraturan-peraturan yang telah dibangun. Maka akan terhasil rangkap tiga (entiti dan hubungan) yang kemudian akan digabungkan untuk menjadi GP multimodal. Berikut merupakan beberapa peraturan yang telah dibangun :

1. Peraturan 1 (hubungan kata kerja)
2. Peraturan 2 (hubungan kata kerja + preposisi)
3. Peraturan 3 (hubungan *has-Attribute*)
4. Peraturan 4 (hubungan *has-Image*)
5. Peraturan 5 (hubungan *has-Bg-Kg*)

Kaedah yang telah dicadangkan juga telah dibandingkan dan dinilai ketepatan melalui kiraan formula *Precision & Recall*. Beberapa cadangan untuk perbaiki lagi peraturan mengekstrak hubungan dan entiti dan juga pembinaan GP multimodal juga telah dinyatakan.



## **BAB V**

### **RUMUSAN DAN CADANGAN**

#### **5.1 PENGENALAN**

Bab ini akan merumuskan mengenai hasil kajian dan objektif yang telah berjaya dicapai. Selain membincangkan mengenai kekangan yang dihadapi semasa membuat kajian ini, beberapa cadangan untuk penambahbaikan dalam mengekstrak maklumat dengan lebih berkesan dan membina GP multimodal yang lebih jelas telah dinyatakan. Cadangan-cadangan ini boleh dijadikan sebagai tonggak untuk kajian masa depan.

#### **5.2 KESIMPULAN KAJIAN**

Orang ramai cenderung untuk terlepas fakta bahawa teks masih mendominasi dalam pertukaran maklumat atas talian. Kebanyakan aktiviti yang dilakukan secara atas talian masih berdasarkan teks bacaan, contohnya media sosial, berita, pencarian, e-mel, ulasan produk dan banyak lagi. Namun begitu, sebagaimana perkataan tidak boleh berubah menjadi imej, imej pula tidak dapat menggantikan perkataan dari segi keupayaan mereka untuk menyampaikan maklumat yang jelas dan komprehensif. (Gaudiano, 2014)

Secara rumusnya, kajian ini bermatlamat untuk menggambarkan pengetahuan latar belakang imej dengan cara yang lebih berstruktur daripada teks yang tidak berstruktur melalui pembinaan GP multimodal dimana pembinaan ini dapat menggabungkan dua modaliti yang berbeza iaitu bahagian tekstual dan visual. Maka, peraturan pengekstrakan hubungan entiti berdasarkan kebergantungan tatabahasa ditentukan serta hubungan yang wujud antara imej dan teks. Seterusnya, pengekstrak rangkap tiga juga dibangunkan bagi mengisi GP multimodal yang dicadangkan.

Terdapat beberapa kajian yang telah menggabungkan tekstual dan visual untuk mencapai objektif masing-masing. Namun begitu, masih lagi terdapat kekurangan dalam konteks analisis teks dimana kajian sedia ada tidak menumpukan terhadap analisis teks. Oleh itu, kajian ini memberi penekanan terhadap kaedah memproses teks dengan menyesuaikan penggunaan GP multimodal. Graf pengetahuan yang terhasil ini boleh digunakan untuk menyelesaikan tugas visual.

Motivasi utama untuk kajian ini adalah terdapat beberapa GP yang sedia ada tidak mempertimbangkan hubungan yang wujud antara dua modaliti yang berbeza, seperti tekstual dan visual. Walau bagaimanapun, terdapat juga penyelidik yang telah membina GP multimodal yang telah terbukti jauh lebih baik berbanding dengan GP unimodal.

Setelah menjalankan eksperimen untuk mengekstrak rangkap tiga dan telah menggabungkan rangkap tiga yang dihasilkan untuk dijadikan sebagai GP multimodal, maka matlamat untuk menggambarkan pengetahuan latar belakang dengan cara yang lebih berstruktur telah pun tercapai. Selain itu, hubungan teks-teks dan imej-teks juga dapat dibina dengan menggunakan kaedah yang dicadangkan iaitu peraturan pengestrakan yang telah ditetapkan. Hasil rangkap tiga ini juga dinilai menggunakan formula perolehan kembali (Recall) dan ketepatan (Precision) di mana masing-masing keputusannya adalah 0.60 dan 0.90.

Seperti yang dinyatakan sebelum ini, rangkap tiga (entiti dan hubungan) adalah elemen penting yang diperlukan untuk membina graf pengetahuan. Kriteria bagi menghasilkan GP yang berkualiti adalah dengan melupuskan rangkap tiga yang tidak sempurna iaitu rangkap tiga yang tidak lengkap formatnya (hubungan – subjek – objek). Maka hanya rangkap tiga yang lengkap sahaja akan digunakan untuk menghasilkan GP. Juga, dengan memastikan hampir kesemua graf bersambungan dan tiada subgraf yang tergantung.

### **5.3 SUMBANGAN KAJIAN**

Kaedah yang dicadangkan ialah menetapkan peraturan bagi mengekstrak hubungan dari ayat input. Teknik hubungan kebergantungan juga diterapkan bagi mengekalkan

hubungan antara setiap perkataan supaya tidak terputus. Jadi, penghurai berdasarkan kebergantungan digunakan dalam kajian ini atas kekuatan penghurai untuk mengekalkan hubungan yang wujud antara teks. Hasil kajian ini telah terbukti bahawa kaedah ini dapat mengekstrak hampir kesemua maklumat kaya yang boleh didapati dari ayat input tersebut. Tambahan lagi, kaedah yang dicadangkan ini berjaya menghasilkan rangkap tiga yang mempunyai nilai ketepatan yang tinggi (0.90). Selain daripada itu, kajian ini menggunakan ayat kompleks juga kerana didapati kajian yang lain menggunakan teks mudah dan bukannya ayat kompleks. Juga dapat menangani ayat-ayat kompleks yang berkemampuan menangkap beberapa rangkap tiga dalam satu ayat.

Seterusnya, peraturan yang dibangunkan mampu mengekstrak rangkap tiga dari ayat mudah mahupun ayat kompleks. Keunikan bagi peraturan yang dibangunkan adalah seperti berikut. Contoh, Peraturan 2 (kata kerja + preposisi) adalah sama seperti Peraturan 1 (kata kerja), tetapi dibangunkan bagi mengekstrak ayat dengan lebih lengkap dan juga lebih sesuai digunakan untuk dataset Bahasa Inggeris berbanding dengan peraturan (kata kerja + *ADVMOD*) yang digunakan oleh Romadhony et al. (2018). Peraturan 3 (*has-Attribute*) dapat memberi maklumat tambahan (menerangkan tentang suatu sifat) bagi satu kata nama. Bagi Peraturan 4 (*has-Img*) hanya mengekstrak entiti penting yang berkait dengan imej sahaja. Tidak semua kata nama di dalam teks adalah entiti yang berkait dengan imej. Pertambahan Peraturan 5 (hubungan *has-Bg-Kg*) pula dilaksanakan pada hubungan secara tidak langsung kepada subgraf yang tergantung di dalam GP.

Di samping itu, rangkap tiga yang terhasil ini digambarkan menjadi GP multimodal untuk menunjukkan dengan lebih jelas lagi pertalian antara teks-teks dan juga imej-teks. GP multimodal yang dihasilkan ini boleh dijadikan sebagai model latihan bagi tugas-tugas visual yang memerlukan asas pengetahuan untuk mencapai objektif mereka.

Konklusinya, kaedah dicadangkan ini merupakan penambahbaikan daripada kedua hasil karya berikut (Romadhony et al. (2018); Gong & Wang (2017)). Hal ini kerana kaedah Romadhony et al. (2018) yang hanya bersasarkan dalam hubungan

teks-teks sahaja. Bagi Gong & Wang (2017) pula, kaedah yang digunakan hanya fokus kepada hubungan imej-teks yang secara tidak langsung mengabaikan maklumat lain yang boleh didapati dari teks tersebut. Kaedah yang dicadangkan ini bukan sahaja berjaya mengekstrak hubungan teks-teks, malah hubungan imej-teks juga menjadikan ia sangat dipenuhi oleh maklumat untuk diisi di dalam GP multimodal yang akan dibina.

#### 5.4 KEKANGAN KAJIAN

Terdapat beberapa batasan yang dapat disenaraikan daripada hasil kajian ini. Berikut ialah antara kekangan yang telah dihadapi.

Batasan output dari penghuraian kebergantungan (*dependency parser*) yang digunakan. Hal ini telah diketengahkan dalam Bab 4 dimana beberapa rangkap tiga yang dihasilkan sama ada tidak melengkap atau entiti yang diekstrak tidak menggambarkan dengan imej yang tersedia. Salah satu contohnya, hasil daripada eksperimen ini mendapati bahawa hanya rangkap tiga [*'appear', 'Rihanna'*] mampu diekstrak oleh peraturan yang telah dibangunkan. Secara teori, penghurai mampu menangkap token yang mempunyai hubungan kebergantungan prep dan penandaan POS kata kerja. Hal ini kerana terdapat peraturan untuk menangkap kedua ciri tersebut untuk dijadikan sebagai satu hubungan [*'appear as'*]. Begitu juga dengan entiti objek [*'adviser'*] yang boleh dikenal pasti dengan hubungan kebergantungan *pobj*. Akhirnya, akan menghasilkan rangkap tiga yang lengkap seperti [*'appear as', 'Rihanna', 'adviser'*].

Kedua, rangkap tiga yang dihasilkan [*'lowered at', who, 'embassy'*] mempunyai subjek *who* dimana jika rujuk kepada ayat asalnya “*US marines who lowered flag at embassy in Havana fifty years ago set for Cuba return*”, didapati bahawa *who* sebenarnya merujuk kepada subjek *US Marines*. Hal ini kerana kajian ini terhad kepada teknik resolusi *co-reference* iaitu sesuatu tugas untuk mencari ungkapan yang merujuk kepada entiti yang sama dalam teks. Kajian ini tidak mempertimbangkan teknik ini atas sebab dataset yang digunakan mempunyai kadar peratusan kata ganti nama yang rendah (3.29%) dalam keseluruhan ayat input yang

merangkumi tajuk dan kapsyen imej. Rujuk Lampiran C bagi pengiraan kadar peratusan kata ganti nama.

Selain itu, dari segi visualisasi, GP multimodal yang dihasilkan hanya menggunakan imej ID sahaja yang dibulatkan seperti dalam Rajah 5.1 bagi mewakili imej artikel yang sebenar. Hal ini menjadikan graf kurang jelas terhadap wujudnya dua modaliti yang berbeza iaitu terdapatnya tekstual dan visual. Juga, GP yang dihasilkan merupakan graf yang tidak berarah (*undirected*) yang mungkin akan menimbulkan kekeliruan untuk membaca graf. Graf yang berarah (*directed*) merupakan graf yang mempunyai anak panah yang menunjukkan cara bacaan entiti yang wujud di dalam graf pengetahuan.

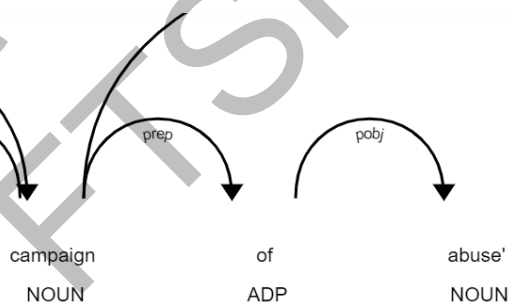


Rajah 5.1 Contoh graf tidak berarah dan imej ID yang mewakili imej

## 5.5 KAJIAN MASA DEPAN

Masih terdapat lagi ruang yang boleh ditambah baik dari segi pembangunan peraturan untuk mengekstrak hubungan dan juga pembinaan GP multimodal. Antaranya ialah :

1. Bagi menyelesaikan masalah yang mendapati entiti *who* sebenarnya merujuk kepada subjek yang betul, contoh *US Marines* atau *Wix*, teknik resolusi *co-reference* iaitu sesuatu tugas untuk mencari ungkapan yang merujuk kepada entiti yang sama dalam teks seperti yang telah dilakukan dalam kajian (Exner & Nugues 2012; Kertkeidkachorn & Ichise 2017). Kedua penulis berikut menggunakan teknik resolusi *co-reference* bagi mengesan rantai entiti dalam teks tidak berstruktur.
2. Bagi memperhalusi lagi peraturan sedia ada, penambahan peraturan seperti hanya kata preposisi boleh ditambah lagi. Berdasarkan Rajah 5.2 berikut, peraturan akan kenal pasti token yang mempunyai penandaan POS *ADP* sebagai hubungan dan subjek pula ialah token yang mempunyai hubungan kebergantungan *prep* terhadap token hubungan tersebut. Objek pula dikenal pasti dengan hubungan kebergantungan *pobj*.



Rajah 5.2 Contoh ayat yang mempunyai hubungan kata preposisi

3. Menggunakan imej yang sebenar daripada menggunakan imej ID supaya lebih jelas lagi bahawa GP multimodal yang dihasilkan ini mempunyai kedua hubungan teks-teks dan imej-teks. Rajah 5.3 menunjukkan imej sebenar bagi artikel tersebut telah menjadi nod kepada hubungan *has-Image* yang berhubung dengan entiti (*Visitation*). Disini dapat lihat dengan lebih nyata bahawa secara tidak langsungnya imej yang dipaparkan mempunyai kaitan dengan entiti