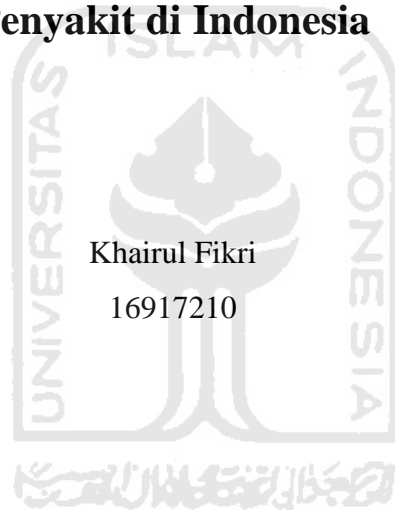




## **Analisis Korelasi antara *Google Trends* dengan Pengawasan Penyakit di Indonesia**



*Tesis diajukan sebagai syarat untuk meraih gelar Magister Komputer*

*Konsentrasi Sistem Informasi Enterprise*

*Program Studi Informatika Program Magister*

*Fakultas Teknologi Industri*

*Universitas Islam Indonesia*

2020

**Lembar Pengesahan Pembimbing**

**Analisis Korelasi antara *Google Trends* dengan Pengawasan Penyakit di Indonesia**

Khairul Fikri

16917210



الجامعة الإسلامية  
الابستاد الباندو

Pembimbing

A handwritten signature in blue ink, which appears to be 'DThomas Hatta Fudholi'. The signature is fluid and stylized, with a large initial 'D' and 'F'.

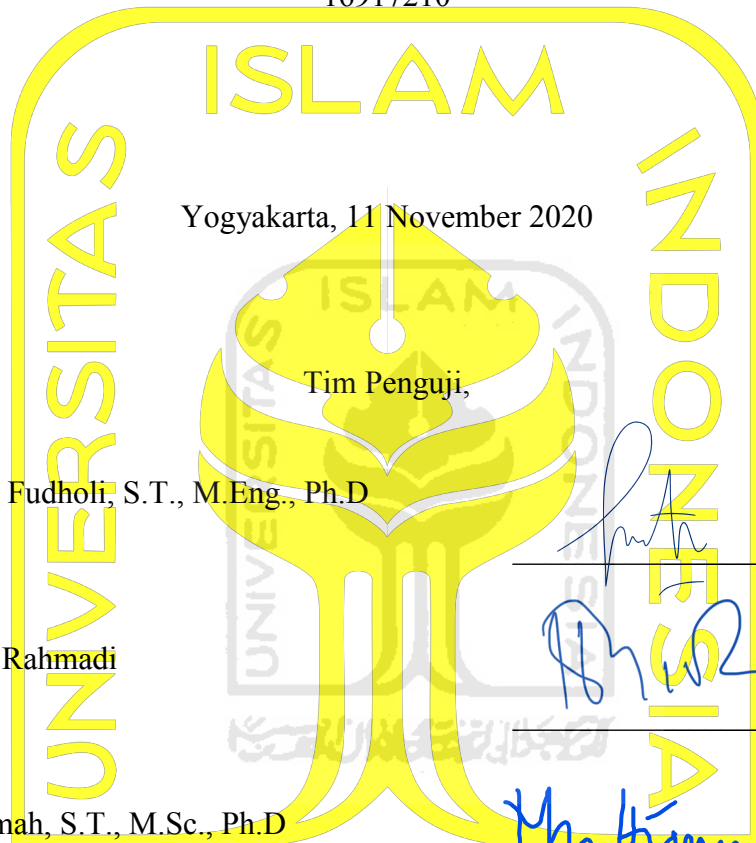
DThomas Hatta Fudholi, S.T., M.Eng., Ph.D

## Lembar Pengesahan Penguji

### Analisis Korelasi antara *Google Trends* dengan Pengawasan Penyakit di Indonesia

Khairul Fikri

16917210



Yogyakarta, 11 November 2020

Tim Penguji,

Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D

Ketua

Dr. Ing Ridho Rahmadi

Anggota I

Izzati Muhimmah, S.T., M.Sc., Ph.D

Anggota II

Mengetahui,

Ketua Program Studi Informatika Program Magister

Universitas Islam Indonesia



Izzati Muhimmah, S.T., M.Sc., Ph.D.

## Abstrak

### **Analisis Korelasi antara *Google Trends* dengan Pengawasan Penyakit di Indonesia**

Latar Belakang: Aktivitas pengguna internet meninggalkan jejak digital yang dapat dianalisis. *Google Trends* adalah salah satu kolektor jejak digital ini. Sejak 2016, *Google Trends* telah diuji oleh para peneliti seperti di bidang kesehatan. Pemantauan penyebaran penyakit merupakan topik yang paling banyak dibicarakan di luar Indonesia. Tujuan: Sangat penting untuk menganalisis korelasi antara tren *Google Trends* dan pemantauan penyakit di Indonesia. Palsunya, *Google Trends* telah diteliti di luar Indonesia dan pola penyebaran penyakit di Indonesia berbeda dengan negara lain. Tujuan penelitian adalah untuk mengevaluasi korelasi *Google Trends* dengan pengawasan penyakit dengan 1) menguji istilah penelusuran sebelumnya, 2) menemukan sumber untuk istilah penelusuran baru, 3) membuat istilah penelusuran baru yang relevan, dan 4) membuat templat istilah penelusuran praktik terbaik. Desain: peneliti mengambil laporan *surveillance* penyakit konvensional dari Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia. Tren penyakit dari *Google Trends* diambil untuk melihat derajat korelasi dengan laporan pengawasan penyakit konvensional. Buku Kapita Selekta Kedokteran di Indonesia dan *Related Query* di *Google Trends* merupakan sumber baru dalam membuat istilah pencarian. Perhitungan korelasi menggunakan *Pearson Correlation* dengan pengujian lebih lanjut berupa uji konsistensi istilah pencarian memiliki tingkat korelasi yang tinggi ( $\geq 0,80$ ). Hasil: Sedikitnya terdapat 32 istilah pencarian dengan tingkat korelasi tertinggi yang dapat digunakan untuk memantau Tuberkulosis di Indonesia. Sedangkan istilah "adalah" dan "ciri-ciri" dapat digunakan untuk penyakit selain Tuberkulosis dengan aturan penggunaan yang telah ditetapkan. Faktor berita, bahasa, populasi, dan akses internet menjadi perhatian penting dalam pengendalian penyakit melalui *Google Trends*. Masyarakat di Indonesia cenderung mendiagnosis diri sendiri daripada mencari definisi penyakit tertentu. Hal tersebut ditunjukkan dengan banyaknya kecenderungan yang berkaitan dengan diagnosis dibandingkan definisi, epidemiologi, faktor risiko dan tata laksana. Area dengan akses internet minimal dan populasi rendah menjadi sulit untuk dipantau melalui *Google Trends*. Meskipun pengawasan dilakukan sewaktu-waktu, namun penelitian lebih lanjut tetap diperlukan mengingat belum ditemukan waktu / musim yang efektif untuk melakukan pengawasan atau peringatan dini bagi pemerintah.

### **Kata kunci**

*Google trends*, Tuberkulosis, istilah pencarian, *Pearson correlation*, korelasi.

## Abstract

### *Correlation Analysis between Google Trends and Disease Monitoring in Indonesia*

*Background: Internet user activity leaves digital traces that can be analyzed. Google Trends is one such digital trail collector. Since 2016, Google Trends has been tested by researchers such as those in the health sector. Monitoring the spread of disease is the most discussed topic outside Indonesia. Purpose: It is imperative to analyze the correlation between Google trends and disease monitoring in Indonesia. The reason is, Google Trends has been studied outside Indonesia and the pattern of disease spread in Indonesia is different from other countries. The aim of the study was to evaluate Google Trends' correlation with disease surveillance by 1) testing previous search terms, 2) finding sources for new search terms, 3) creating new relevant search terms, and 4) creating best practice search terms templates. Design: researchers took conventional disease surveillance reports from the Directorate General of P2P Ministry of Health of the Republic of Indonesia. Disease trends from Google Trends are taken to see the degree of correlation with conventional disease surveillance reports. Kaipita Selektta Kedoteran book in Indonesia and Related Query on Google Trends are new sources for creating search terms. Correlation calculation using Pearson Correlation with further testing in the form of consistency test of search terms has a high correlation level ( $\geq 0.80$ ). Results: There are at least 32 search terms with the highest level of correlation that can be used to monitor Tuberculosis in Indonesia. Meanwhile, the terms "are" and "characteristics" can be used for diseases other than tuberculosis with predetermined rules of use. News, language, population, and internet access factors are important concerns in disease control through Google Trends. People in Indonesia tend to self-diagnose rather than looking for definitions of certain diseases. This is indicated by the many trends associated with diagnosis versus definition, epidemiology, risk factors and management. Areas with minimal internet access and low population can be difficult to monitor through Google Trends. Even though supervision is carried out from time to time, further research is still needed considering that there has not been found an effective time / season to carry out supervision or an early warning for the government.*

#### **Keywords**

*google trends, tuberculosis, search terms, pearson correlation, correlation.*

## **Pernyataan Keaslian Tulisan**

Dengan ini saya menyatakan bahwa tesis ini merupakan tulisan asli dari penulis, dan tidak berisi material yang telah diterbitkan sebelumnya atau tulisan dari penulis lain terkecuali referensi atas material tersebut telah disebutkan dalam tesis. Apabila ada kontribusi dari penulis lain dalam tesis ini, maka penulis lain tersebut secara eksplisit telah disebutkan dalam tesis ini.

Dengan ini saya juga menyatakan bahwa segala kontribusi dari pihak lain terhadap tesis ini, termasuk bantuan analisis statistik, desain survei, analisis data, prosedur teknis yang bersifat signifikan, dan segala bentuk aktivitas penelitian yang dipergunakan atau dilaporkan dalam tesis ini telah secara eksplisit disebutkan dalam tesis ini.

Segala bentuk hak cipta yang terdapat dalam material dokumen tesis ini berada dalam kepemilikan pemilik hak cipta masing-masing. Apabila dibutuhkan, penulis juga telah mendapatkan izin dari pemilik hak cipta untuk menggunakan ulang materialnya dalam tesis ini.

Yogyakarta, 11 November 2020



Khairul Fikri, S.Kom

## Daftar Publikasi

- Fikri, K. (2018). Analisis Pengaruh Penggunaan Teks di Gambar Iklan terhadap Jangkauan dan Interaksi Konsumen. *Profetik: Jurnal Komunikasi*, 11(2), 46. <https://doi.org/10.14421/pjk.v11i2.1444>
- Fikri, K. (2020). Konsep Pembaharuan Sistem Katalog Perpustakaan (Studi Kasus: Perpustakaan Universitas Di D.I Yogyakarta). *Jurnal Nasional Teknologi Dan Sistem Informasi*, 6(3).
- Fudholi, D. H., & Fikri, K. (2020). Towards an Effective Tuberculosis Surveillance in Indonesia through Google Trends. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 5(4).

### **Publikasi yang menjadi bagian dari tesis**

Publikasi berikut menjadi bagian dari Bab 3 dan Bab 4:

*Sitasi publikasi 3*

Kontributor	Jenis Kontribusi
Author Dthomas Hatta Fudholi	Mendesain eksperimen (30%) Menulis dan mengedit <i>paper</i> (45%)
Author Khairul Fikri	Mendesain eksperimen (70%) Menulis dan mengedit <i>paper</i> (55%)

## Halaman Kontribusi

Terimakasih untuk bapak Fathul Wahid atas ide awal penelitian yaitu tentang *Google Trends*. Meskipun tidak membahas tentang *e-government*, ide tersebut dilanjutkan dengan pembahasan di bidang medis atau kesehatan. Selanjutnya, terimakasih untuk bapak DThomas Hatta Fudholi sebagai pembimbing penelitian yang melanjutkan ide *Google Trends* mengenai pengawasan penyakit di Indonesia. Tidak lupa pula, terimakasih untuk kakak Husnayain sebagai penyumbang konsep penelitian yang diadopsi pada penelitian tesis ini. Terimakasih juga untuk KASUBID Tuberkulosis, Direktorat Jendral Pencegahan dan Penyendalian Penyakit (P2P) Kementerian Kesehatan Republik Indonesia yang berkontribusi dalam *dataset* jumlah kasus penyakit Tuberkulosis. Terakhir, terimakasih untuk para penguji yang telah memberikan saran dalam analisis, konsep, pengujian tambahan, ide baru, dan pengkoreksian hasil penelitian tesis ini.





## Halaman Persembahan

Alhamdulillah, peneliti ucapkan atas rahmat dan hidayah dari Allah subhanahu wa ta'ala sehingga penelitian tesis ini bisa selesai. Terimakasih juga buat keluarga besar yang telah mendo'akan, membiayai dan menyemangati peneliti untuk menyelesaikan studi magister ini. Tidak lupa juga teman-teman seperjuangan yang selalu saling mendukung sehingga studi ini selesai dengan perasaan senang. Peneliti ucapkan juga untuk Magister Informatika yang juga menyumbangkan dana dalam menyelesaikan penelitian tesis ini. Terakhir, terimakasih buat Direktorat Jendral Pencegahan dan Penyendalian Penyakit (P2P) Kementerian Kesehatan Republik Indonesia dalam kerjasamanya. Terkhusus KASUBID Tuberkulosis yang telah memberikan data untuk keperluan penelitian tesis ini.



## Kata Pengantar

Syukur allhamdulillah penulis ucapkan atas rahmat dan nikmat dari Allah SWT, sehingga studi magister ini bisa diselesaikan dengan baik. Selain itu, penulis juga bersyukur atas selesainya pengerjaan tesis dengan judul “Analisis Korelasi antara *Google Trends* dengan Pengawasan Penyakit di Indonesia”. Sholawat beriringkan salam, penulis curahkan buat Nabi Muhammad SAW. Allahumma sholli ala sayyidina Muhammad wa ala ali sayyidina muhammad. Semoga kita semua mendapatkan syaf’at beliau di hari akhirat nantinya, aamiin. Tesis ini dibuat untuk persyaratan kelulusan program Magister Informatika di Fakultas Teknologi Industri di Universitas Islam Indonesia.

Penulisan tesis ini berhasil berkat adanya bimbingan dan masukan dari berbagai pihak. Penulis mengucapkan terimakasih kepada pihak-pihak yang telah membantu dalam penyelesaian tesis ini, yaitu:

1. Bapak Prof Fathul Wahid, ST, M,Sc., Ph.D, selaku Rektor Universitas Islam Indonesia.
2. Bapak Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D, selaku pembimbing penulisan tesis.
3. Bapak dan ibu dewan penguji sidang proposal, progres dan pendataan tesis. Beserta pihak Universitas yang membantu dalam kegiatan terkait tesis ini.
4. Bapak dan ibu Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia.
5. Seluruh keluarga besar yang selalu men-*support* penulis.
6. Seluruh sahabat dan rekan seperjuangan yang selalu bersama-sama dalam menyelesaikan studi Magister Informatika ini.

Penulis berharap semoga Allah SWT membalas semua kebaikan dan kerjasama bapak/ibu. Selanjutnya, penulis juga berharap saran dan kritikan dari pembaca tesis ini untuk kesempurnaan penulisan berikutnya.

Yogyakarta, 11 November 2020

Penulis,



Khairul Fikri, S.Kom

## Daftar Isi

Lembar Pengesahan Pembimbing .....	i
Lembar Pengesahan Penguji.....	ii
Abstrak .....	iii
Abstract.....	iv
Pernyataan Keaslian Tulisan .....	v
Daftar Publikasi .....	vi
Halaman Kontribusi.....	vii
Halaman Persembahan .....	viii
Kata Pengantar.....	ix
Daftar Isi .....	x
Daftar Tabel.....	xiii
Daftar Gambar .....	xiv
Glosarium .....	xv
<b>BAB 1 Pendahuluan .....</b>	<b>1</b>
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan dan Manfaat .....	4
<b>BAB 2 Tinjauan Pustaka .....</b>	<b>5</b>
2.1 Landasan Teori .....	5
2.1.1 Penetrasi Internet Di Indonesia.....	5
2.1.2 Kenapa Harus <i>Google Trends</i> ?.....	5
2.1.3 <i>Google Trends</i> (GT) .....	5
2.1.4 Kata Kunci.....	8
2.1.5 Pengawasan Berbasis Internet .....	8

2.1.6	Normalisasi Data Ke Dalam Rentang Nilai 0 – 100.....	8
2.1.7	Korelasi.....	9
2.1.8	Pearson Correlation .....	9
2.1.9	Signifikansi Koefisien Korelasi.....	9
2.1.10	Penafsiran Koefisien Korelasi .....	10
2.1.11	Tingkat Variabel X Mewakili Keadaan Variabel Y .....	10
2.2	Penggunaan <i>Google Trends</i> Sebelumnya .....	11
2.3	Penelitian Sebelumnya dalam Bidang Kesehatan.....	14
BAB 3 Metodologi .....		15
3.1	Data .....	15
3.2	Lokasi dan Waktu Penelitian .....	16
3.3	Langkah-langkah Penelitian .....	16
3.3.1	Pencarian Tema Penelitian .....	16
3.3.2	Pendalaman Tema Penelitian.....	16
3.3.3	Penentuan Fokus Penelitian.....	16
3.3.4	Tahapan Inti Penelitian.....	17
3.3.5	Tahapan Pembuatan Laporan Hasil Akhir.....	20
BAB 4 Hasil dan Pembahasan.....		21
4.1	Hasil.....	21
4.1.1	<i>Disease Cases</i> .....	21
4.1.2	<i>General Category</i> .....	22
4.1.3	<i>Generate New Terms</i> .....	24
4.1.4	Disease Trend in Google in Google Trends .....	24
4.1.5	<i>Correlation Calculation</i> .....	25
4.1.6	<i>Analysis and Evaluation</i> .....	25
4.2	Pembahasan .....	27
4.2.1	Sumber Istilah Pencarian Baru .....	27

4.2.2	Uji Konsistensi <i>Template</i> Istilah Pencarian Usulan .....	32
4.2.3	Faktor yang Mempengaruhi Istilah Pecarian.....	34
4.2.4	Penafsiran Terhadap Ketegori Utama ( <i>General Category</i> ).....	35
4.2.5	Waktu yang Bagus untuk Pengawasan Penyakit dengan GT .....	36
BAB 5 Kesimpulan dan Saran.....		37
5.1	Kesimpulan .....	37
5.2	Saran .....	38
Daftar Pustaka .....		39
LAMPIRAN A .....		45
LAMPIRAN B.....		59
LAMPIRAN C.....		62



## Daftar Tabel

Tabel 4.1 Template istilah kombinasi untuk domain penyakit menular.....	32
Tabel 4.2 Praktik terbaik <i>template</i> istilah pencarian untuk domain penyakit umum. ....	33
Tabel 4.3 Praktik terbaik <i>template</i> istilah pencarian untuk domain penyakit Tuberkulosis. .....	33



## Daftar Gambar

Gambar 3.1 Metodologi Penelitian (Secara Umum).....	17
Gambar 4.1 Kasus penyakit Tuberkulosis di Indonesia tahun 2014-2018.....	21
Gambar 4.2 Normalisasi kasus penyakit Tuberkulosis di Indonesia dalam range nilai 0-100.....	22
Gambar 4.3 Rancangan kategori utama istilah pencarian.....	23
Gambar 4.4 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'sakit tbc' (level negara).....	26
Gambar 4.5 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'gejala tb' (level negara).....	27
Gambar 4.6 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'obat tb' (level negara).....	30
Gambar 4.7 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'kode diagnosa tb paru' (level negara).....	30
Gambar 4.8 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'icd 10 limfadenitis tb' (level negara).....	31
Gambar 4.9 Perbandingan jumlah istilah pencarian berdasarkan Kategori Utama.....	36

## Glosarium

GT	- <i>Google Trends</i>
RSV	- <i>Relative Search Volume</i>
DITJEN P2P KEMKES RI	- Direktorat Jendral Pencegahan dan Pengendalian Penyakit Kementerian Kesehatan Republik Indonesia





# BAB 1

## Pendahuluan

### 1.1 Latar Belakang

Google Trends (GT) pada 11 Mei 2006 dikenalkan oleh perusahaan *Google Inc.* dan dapat diakses secara gratis dan umum. GT ditujukan sebagai sebuah *tool* yang menampilkan kata kunci atau sebuah kata/beberapa kata yang populer digunakan oleh penggiat internet ketika *searching* (mencari informasi) di *website* *www.google.com* (mesin pencari milik *Google Inc.*). Kata kunci juga disebut dengan sebutan “istilah pencarian (*search term*)” yang berartikan istilah pencarian. Konsep data tren pencarian pengguna ini sudah dibahas pada tahun sebelumnya dan semakin banyak penelitian dilakukan setelah itu (Jun, Yoo, & Choi, 2018).

Empat tahun setelah dirilis GT, peneliti (Carneiro & Mylonakis, 2009a; Olson et al., 2009; Pelat, Turbelin, Bar-Hen, Flahault, & Valleron, 2009) telah menggunakan GT dalam penelitiannya yang membahas tentang pengawasan wabah penyakit secara *online*. Hingga tahun 2019 GT telah dibahas pada bidang kesehatan, bisnis, *research* (akademik), dan hukum. Beberapa kasus penelitian menghasilkan suatu kemajuan atau keberhasilan, namun dibebberapa kasus mengalami kemunduran atau kegagalan. Seperti penelitian (Carneiro & Mylonakis, 2009a) memberikan sebuah contoh kasus “flu burung” yang mengalami kelonjakan (*volume*) tren dalam GT pada daerah tertentu, akan tetapi kenyataannya pada daerah tersebut tidak ada kasus penyakit yang sebenarnya. Para peneliti mencoba menguji korelasi tren GT dengan *dataset* kasus yang memang pernah terjadi disebuah daerah. Terdapat kurang dari 50% (dari total kasus) kasus tren GT yang memiliki korelasi/kecocokan dengan kasus kenyataan di lapangan (Cho et al., 2013; Kang, Zhong, He, Rutherford, & Yang, 2013; Seifter, Schwarzwald, Geis, & Aucott, 2010).

Memang diakui beberapa peneliti mengusulkan *dataset* GT digunakan untuk peringatan dini, monitoring, pengecekan, prediksi/ramalan, dan pengawasan. Namun, perlu diingat bahwa beberapa peneliti juga menemukan bias/kekurangan pada GT untuk kasus tertentu.

Beberapa masalah muncul seperti a) Istilah pencarian yang digunakan ambigu atau bermakna banyak (Kristoufek, 2013), b) kesalahan membuat istilah pencarian dalam melihat tren (Dilmaghani, 2019), dan c) belum ada standardisasi atau pedoman dalam menggunakan GT. Hal itu ditandainya dengan hasil penelitian sebelumnya menyarankan penggunaan GT harus mempertimbangkan populasi kota yang besar, bahasa yang didukung, perilaku masyarakat dalam menggunakan internet, dan latar pendidikan masyarakat (Husnayain, Fuad, & Lazuardi, 2019; Kang et al., 2013).

Apabila tren GT tidak valid, maka akan menyebabkan alokasi bantuan salah target serta penumpukan obat-obatan disatu daerah. Hal ini bisa terjadi, mengingat kasus pada (Cho et al., 2013) menyebutkan bahwa pemberitaan di media masa (televisi) bisa menyebabkan tren GT naik. Misalnya media masa memberitakan penyakit flu di daerah A, akan tetapi yang lebih banyak melakukan *searching* ialah daerah B. Selain itu, kesalahan hasil studi bagi akademis karena *dataset* yang digunakan dalam penelitian bermasalah, dan kerugian finansial bagi pengusaha/pebisnis yang menggunakan GT untuk prediksi saham ataupun strategi bisnis/pemasaran. Saat ini, teknik GT dalam mengkalkulasikan *dataset* pencarian para pengguna internet sehingga muncul tren, belum diketahui secara pasti. Sehingga diperlukannya uji dan diskusi yang serius lagi tentang istilah pencarian yang berkorelasi dengan kasus penyakit tertentu (yang diamati), perlunya transparansi dari pihak GT untuk meningkatkan kepercayaan publik, dan penyampaian algoritma dalam penentuan RSV yang digunakan pada (Carneiro & Mylonakis, 2009a; Husnayain et al., 2019; Nuti et al., 2014a).

Ada 24 literatur yang peneliti rujuk tentang penggunaan GT dan enam literatur fokus pada bidang kesehatan. Penelitian ini juga berfokus pada bidang kesehatan. Penelitian (Husnayain et al., 2019) telah mendapatkan hasil bahwa secara statistik GT memiliki korelasi dengan laporan resmi kasus Demam Berdarah tahunan di Indonesia. Tahun 2017 Indonesia menjadi nomor ketiga tertinggi di dunia untuk kasus Tuberkulosis (TBC) (Kusnandar, 2019). TBC juga menjadi salah satu penyebab kematian terbesar sepanjang tahun 2007-2017 (Pranita, 2020). WHO juga melaporkan TBC termasuk salah satu dari 10 penyebab kematian terbesar tahun 2012 di Indonesia (WHO, 2015). Tahun 2014 penyakit TBC menyumbang 5,7% sebagai penyakit terbesar diderita masyarakat di Indonesia (Rokom, 2017). Diketahui juga bahwa 75 persen dari pasien TBC tergolong kelompok produktif yaitu berusia 15-55 tahun (Ihsanuddin, 2020). TBC sebagai penyakit menular mendapatkan perhatian khusus dari Ikatan Dokter Indonesia (IDI) dan mendesak pemerintah

untuk menarapkan protokol Kejadian Luar Biasa (KLB) dalam penanganannya (Tamba, 2018). Oleh sebab itu, peneliti menjadikan TBC sebagai objek penelitian ini.

Berdasarkan temuan di atas, dapat disimpulkan bahwa pengawasan penyebaran penyakit TBC sangat perlu dilakukan. Saat ini, pelaporan kasus TBC dari daerah ke tingkat provinsi masih membutuhkan waktu lama dan pelaporan dari provinsi dikeluarkan setiap awal tahun (misalnya tahun 2014 akan dikeluarkan awal tahun 2015). GT dengan laporan tren *realtime* menjadi salah satu alternatif untuk mengawasi penyebaran penyakit secara cepat (*real-time*).

Penelitian ini akan melalui lima tahapan umum yaitu *diases cases, general category, generate new terms, disease trends in GT, correlation calculation, analysis dan evaluation*. Perhitungan korelasi menggunakan *Correlation Pearson* dan sumber istilah pencarian dari buku Kapita Selekta Kedokteran dan *Related Query on GT*. Penelitian ini diharapkan menghasilkan standarisasi berupa *template* istilah pencarian yang siap digunakan untuk melakukan pengawasan penyakit melalui GT. Template diharapkan mampu mengatasi permasalahan istilah pencarian yang ambigu atau kesalahan dalam pembuatan istilah pencarian seperti yang muncul pada penelitian-penelitian sebelumnya. Laporan penelitian ini dibuat dengan lima bagian yaitu Pendahuluan, Tinjauan Pustaka, Metodologi, Hasil & Pembahasan dan Kesimpulan & Saran.

## **1.2 Rumusan Masalah**

Berdasarkan permasalahan dan peluang yang terdapat pada bagian latar belakang serta kondisi kasus penyakit di Indonesia, maka peneliti merumuskan beberapa masalah dalam bentuk pertanyaan yaitu:

1. Bagaimana *template* istilah pencarian yang tepat untuk melakukan pengawasan penyakit melalui GT ?.
2. Bagaimana potensi penggunaan GT untuk pengawasan penyakit Tuberkulosis di Indonesia ?.

## **1.3 Batasan Masalah**

Agar penelitian ini memiliki arah yang jelas, maka peneliti membuat beberapa batasan dalam melakukan penelitian yaitu:

1. Topik dan bidang yang dibahas ialah tentang kesehatan, sub-bidangnya yaitu penyebaran/kasus penyakit. Kasus penyakit yang dibahas adalah penyakit Tuberkulosis.

2. Pengujian dan eksperimen menggunakan *dataset* kasus penyakit dari lembaga pengawasan / sejenisnya tentang kasus penyakit. Data kedua ialah *dataset* tren GT yang nantinya akan dilihat korelasi dengan data kasus penyakit.
3. Pembahasan dilakukan pada lingkup negara Indonesia dengan bahasanya bahasa Indonesia.
4. Studi ini semata-mata melihat korelasi dan asosiasi secara matematika.
5. Tidak membandingkan antar grup dan tidak membahas Simpsons's Paradox.

#### **1.4 Tujuan dan Manfaat**

Tujuan penelitian ini dilakukan ialah untuk menghasilkan *template* istilah pencarian khusus pada kasus penyakit dan validasi korelasi kasus penyakit Tuberkulosis. DITJEN P2P KEMKES RI merupakan sumber data kasus penyakit tersebut yang nantinya di korelasikan dengan tren GT.

Adapun manfaat yang diharapkan dari penelitian ini, bisa dilihat dari beberapa sudut pandang yaitu:

1. Peneliti: dapat menambah wawasan dan ilmu pengetahuan serta melatih dalam melakukan penelitian (termasuk dalam penulisan karya ilmiah).
2. Pemerintah: mendapatkan *best practice* khususnya dalam pengawasan oleh lembaga tertentu, dapat mengetahui kondisi atau masalah terkini di dalam masyarakat melalui pemanfaatan GT, dan bisa memberikan saran atau bahan pertimbangan dalam melakukan pemberian bantuan bencana. *Best practice* yang dimaksud ialah istilah pencarian yang siap digunakan dan nantinya bisa menghasilkan data tren yang valid (berdasarkan hasil uji penelitian ini).
3. Universitas dan para Peneliti: dapat menjadi pintu gerbang untuk penelitian lanjutan atau baru terkait tentang GT.

## **BAB 2**

### **Tinjauan Pustaka**

#### **2.1 Landasan Teori**

##### **2.1.1 Penetrasi Internet Di Indonesia**

Merujuk pada (KKBI, 2016e) istilah penetrasi berartikan penerobosan seperti budaya luar memengaruhi budaya suatu daerah tertentu. Apabila kata penetrasi disandingkan dengan kata internet maka memiliki makna persentase penggunaan internet. Sehingga kalimat “penetrasi internet di Indonesia” bermakna persentase penggunaan internet oleh masyarakat Indonesia.

Bulan Januari 2019 telah terbit buletin dari (APJII, 2019) yang mengatakan survei terakhir tentang penetrasi internet ialah tahun 2017. Penetrasi internet di Indonesia sudah mencapai 50% dari populasi (masyarakat yang terhubung ke jaringan). Jumlahnya yaitu sekitar 143,26 juta jiwa pengguna internet di Indonesia. Pada tanggal 30 Januari 2019, dua perusahaan media sosial yaitu *We Are Social* (berasal dari Inggris) bersama *platform Hoot Suite* (berasal dari Kanada) melalui laporan “Digital 2019” melaporkan bahwa penetrasi internet di Indonesia sudah mencapai 56% dengan jumlahnya sekitar 150 juta pengguna seperti pada Lampiran A.1 (Orenzi, 2019).

##### **2.1.2 Kenapa Harus Google Trends?**

Mesin pencari memang tidak hanya *Google.com*, ada juga *Yahoo.com* dan *Bing.com*. Akan tetapi berdasarkan laporan *Hoot Suite* dan *We Are Social*, mesin pencari *Google.com* menduduki posisi pertama sebagai *website* paling banyak dikunjungi perbulannya oleh masyarakat Indonesia seperti pada Lampiran A.2. Sedangkan *Yahoo.com* berada pada urutan ke delapan dan *Bing.com* tidak termasuk dalam sepuluh *top website* yang paling banyak dikunjungi. Sehingga *Google Trends* menjadi pilihan utama yang tepat untuk diteliti.

##### **2.1.3 Google Trends (GT)**

Perilaku para pengguna mesin pencari *Google* disimpan di *database Google*, hal yang disimpan ialah kegiatan pengguna mencari informasi menggunakan istilah pencarian di mesin pencari (Husnayain et al., 2019). Selanjutnya pada tahun 2006 dirilis lah produk baru yaitu *Google Trends* (GT) (Jun et al., 2018). GT menampilkan istilah pencarian yang populer dicari atau dimasukkan oleh pengguna internet ke mesin pencari.

Geografis dari istilah pencarian juga dimunculkan sehingga terlihat jelas asal tren tersebut seperti Lampiran A.3. Peneliti memberikan tiga contoh istilah pencarian yaitu tuberkulosis, tbc dan tb dengan area pencarian yaitu negara Indonesia untuk rentang waktu lima tahun (untuk semua kategori dan berdasarkan penelusuran *web*).

GT sendiri memiliki 3 pola layanan yang disediakan yaitu:

#### 1. Pelajari

Layanan “Pelajari” ini memungkinkan untuk melakukan eksperimen untuk melihat tren dari suatu topik atau istilah pencarian berdasarkan lokasi, waktu, kategori, jenis penelusuran. Alamat *website* nya yaitu <https://trends.google.com/trends/explore> dan ini peneliti contohkan seperti pada Lampiran A.3. Kategori disini merupakan kategori yang dibuat dari permintaan pencarian di mesin pencari yang kemudian di klasifikasi secara otomatis. Ada sebanyak 25 kategori induk dan termasuk kategori kesehatan dan 300 sub-kategori yang disediakan di GT (Arora, McKee, & Stuckler, 2019). Sedangkan jenis penelusuran ialah jenis media yang digunakan oleh pengguna internet dalam mencari informasi. Jenisnya yaitu:

##### a. Penelusuran *Web*

Pencarian informasi yang dilakukan secara umum yang menampilkan informasi berupa teks dan alamat *website* pemilik teks tersebut. Pengguna cukup mengunjungi alamat [www.google.com](http://www.google.com).

##### b. Penelusuran Gambar

Ketika pengguna ingin mencari informasi berupa gambar, hal ini bisa memilih menu “Gambar” pada pencarian [google.com](http://google.com) atau mengunjungi alamat: <https://www.google.com/imghp>.

##### c. Penelusuran Berita

Sama seperti penelusuran gambar, untuk berita bisa memilih menu “Berita” ketika ingin mencari informasi atau mengunjungi alamat pencarian berita langsung yaitu <https://www.google.com/webhp?tbm=nws>.

##### d. *Google Shopping*

Jenis ini sama seperti penelusuran berita dan gambar, sehingga pengguna yang mencari informasi melalui menu “*Shopping/Belanja*” atau alamat *link* <https://www.google.com/webhp?tbm=shop> akan disimpan dan ditampilkan di *Google Trends* (GT).

##### e. Penelusuran YouTube

Jenis ini dikhususkan tren dari pencarian informasi melalui *Youtube.com* yang merupakan salah satu produk dari perusahaan *Google Inc.* juga.

Berdasarkan ketiga pola layanan di atas, peneliti menggunakan pola layanan “Pelajari”, karena hanya pola ini yang memungkinkan untuk bereksperimen atau melihat tren suatu topik yang bisa ditentukan sendiri. Sedangkan yang dua lainnya, topik ditentukan oleh sistem dan tidak dapat dirubah. Data tren dari GT disesuaikan oleh sistem dengan maksud mempermudah dalam perbandingan topik atau istilah pencarian. Pihak GT juga menyebutkan bahwa hasil penelusuran yang ada bersifat proporsional terhadap waktu dan lokasi *query* berdasarkan proses berikut ini (Trends, 2019).

1. Setiap poin data yang ada akan dibagi. Pembagian tersebut didasarkan menurut penelusuran total terkait geografis dan rentang waktu yang diwakili data tersebut. Hal ini dilakukan untuk perbandingan popularitas relatif dan mencegah *volume* penelusuran terbanyak selalu berada *diranking* tertinggi.
2. Hasil dari proses pertama akan dirubah dalam bentuk skala yaitu rentang 0 sampai 100 berdasarkan proporsi topik untuk semua penelusuran yang ada pada semua topik.
3. Harus diingat bahwa suatu wilayah yang minat penelusuran istilah yang sama dengan wilayah lain, tidak akan selalu sama *volume* total penelusuran.

Pengumpulan data tren menggunakan persentase penelusuran dan pihak GT menyebutkan bahwa data GT merupakan sampel daya penelusuran mesin *Google* yang tidak dibiaskan. Data tren GT memiliki dua cara pembentukan tren atau bisa disebut juga ada dua jenis data tren GT yaitu:

1. Data *real-time*

Data ini diciptakan dengan cara mengambil sampel acak penelusuran dalam rentang waktu 7 hari terakhir.

2. Data *non-real-time*

Data yang kedua ini diciptakan dari pengambilan sampel acak yang dimulai dari tahun 2004 hingga 36 jam sebelum pengguna melihat tren.

Data yang sudah dikumpulkan, akan dikelompokkan dan dihubungkan dengan topik serta melakukan penghapusan informasi pribadi pengguna yang melakukan penelusuran. Ada beberapa data yang dibuang atau tidak dimasukkan untuk dilihat trennya yaitu penelusuran yang dilakukan sedikit pengguna, penelusuran duplikat (penelusuran sama atau diulangi yang dilakukan oleh pengguna yang sama dengan jarak waktu pengulangan yang dekat), dan karakter khusus seperti kasus *apostrof* yaitu menghilangkan bagian kata dan

diganti dengan simbol. Contohnya yaitu kata “Akan” menjadi “ ’kan ”. Ada tanda petik satu sebagai simbol menyatakan huruf “A” (Trends, 2019).

#### 2.1.4 Kata Kunci

Kata kunci menurut (KKBI, 2016c) ialah sebuah kata atau beberapa kata yang mewakili suatu gagasan/konsep. Ketika melakukan pencarian informasi di internet (misalnya *website www.google.com*) maka para pengguna memasukkan/mengetikan kata yang mewakili apa yang ingin dicari. Contohnya, ketika ingin mencari tentang penyakit Tuberkolosis maka pengguna bisa mengetik tuberkolosis atau tbc. Kata “tuberkolosis” atau “tbc” itulah yang dimaksud dengan kata kunci yang selanjutnya disebut dengan istilah pencarian.

#### 2.1.5 Pengawasan Berbasis Internet

Pengawasan dalam bahasa Inggris disebut juga dengan *surveillance* dan saat ini pengawasan berbasis internet mulai di minati dan diteliti. Kerangka dari pengawasan ini ialah mengumpulkan data *real-time* atau data terkini tentang kejadian wabah penyakit maupun sugesti/dugaan sedang terjadi. Proses kegiatan tersebut juga disebut dengan istilah *Syndromic surveillance* dan bahkan dijadikan sebagai sebuah sistem seperti yang dilakukan oleh lembaga dari departemen kesehatan (Thomas, Yoon, Collins, Davidson, & Mac Kenzie, 2018).

#### 2.1.6 Normalisasi Data Ke Dalam Rentang Nilai 0 – 100

Normalisasi data dilakukan untuk menyamakan rentang nilai dari suatu *dataset* yang selanjutnya akan dikorelasikan. Penelitian ini memiliki dua *dataset* yaitu tren GT yang memiliki rentang nilai (skala) dari nol (0) hingga seratus (100), sedangkan *dataset* kedua yaitu kasus penyakit dengan rentang nilai nol hingga puluhan ribu (10000-an). Karena rentang nilai GT adalah paling rendah, sehingga membuat *dataset* kasus penyakit mengikuti rentang nilai GT.

Rumus yang digunakan untuk normalisasi *dataset* ialah rumus (2.1) dan disesuaikan menjadi rumus (2.2) atau (2.3) sesuai kondisi dataset dan diterapkan saat proses normalisasi sebelum *dataset* tersebut digunakan untuk di korelasi dengan *dataset* tren GT.

$$Data_i = \frac{\text{nilai maksimal normalisasi data}}{\text{nilai maksimal dari semua data}} * Data_i \quad (\text{Husnayain et al., 2019})(2.1)$$

Laporan data kasus penyakit berupa laporan perbulan, maka rumus menjadi:

$$Bulan_i = \frac{100}{\text{nilai maksimal dari semua data}} * \text{data bulan}_i \quad (2.2)$$

Sedangkan jika format data berada dalam kurun waktu triwulan, maka rumusnya yaitu:

$$Triwulan_i = \frac{100}{\text{nilai maksimal dari semua data}} * \text{data triwulan}_i \quad (2.3)$$



### 2.1.7 Korelasi

Korelasi dalam (KKBI, 2016d) diartikan sebagai hubungan sebab-akibat dari beberapa variabel/hal/sesuatu. Hubungan yang dimaksud juga termasuk hubungan timbal balik seperti variabel iklim dengan variabel dunia tumbuh-tumbuhan. Korelasi sendiri merupakan suatu teknik untuk pengujian suatu kasus/pendapat/kesimpulan. Salah satu teknik atau metode korelasi yaitu *Pearson correlation*. *Pearson correlation* digunakan untuk menguji hubungan antara dua variabel secara *linear* dengan asumsi nilai level atau ratio sudah ditentukan. Data yang digunakan ialah data yang memiliki skala berupa skala interval/rasio, sedangkan bagi data ordinal menggunakan korelasi *Spearman* (Mirabel, Palit, & Handojo, 2017).

### 2.1.8 Pearson Correlation

*Pearson correlation* merupakan metode untuk menguji hubungan dua variabel (dua *set* data) menggunakan rumus. Hasil dari perhitungan (berupa nilai tunggal) akan menunjukkan nilai keeratan korelasi/hubungan. Nilai tersebut disebut dengan koefisien *Pearson* atau koefisien korelasi (Emerson, 2015; Teng, Cheng, & Zhao, 2017). Selain itu, korelasi *Pearson* dikenal juga dengan nama korelasi *Product Moment*. Berikut rumus dari *Pearson correlation*:

$$\rho_{X,Y} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{N \sum X^2 - (\sum x)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}} \quad (\text{Teng et al., 2017})(2.4)$$

$$r_{x,y} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{(n \sum x_i^2 - (\sum x_i)^2) (n \sum y_i^2 - (\sum y_i)^2)}} \quad (\text{Sugiyono, 2007})(2.5)$$

Apabila nilai koefisiennya 1, maka itu menandakan bahwa variabel benar-benar positif atau hubungannya erat sekali (sempurna) (Teng et al., 2017). Sedangkan jika 0, maka hubungannya tidak ada atau variabel tidak saling terkait dan jika -1 maka menunjukkan korelasinya negatif secara sempurna. Korelasi yang bernilai positif, berartikan bahwa ketika variabel X naik maka variabel Y juga ikut naik. Sedangkan korelasi bernilai negatif, memiliki arti bahwa ketika variabel X naik maka variabel Y turun (Emerson, 2015). Perlu diketahui juga bahwa simbol r digunakan untuk sampel, simbol R (huruf kapital) digunakan untuk korelasi ganda, dan simbol  $\rho$  (rho) digunakan untuk populasi (Sugiyono, 2007).

### 2.1.9 Signifikansi Koefisien Korelasi

Pengecekan signifikansi pada koefisien hasil korelasi dilakukan untuk melihat apakah variabel atau korelasi tersebut dapat digeneralisasikan didalam populasi atau untuk melihat apakah variabel x dapat mencerminkan variabel y.

Pengecekan koefisien dapat dilakukan dengan dua cara yaitu (Sugiyono, 2007):

- Uji Tabel

Pengujian dilakukan dengan cara melihat nilai  $r$  pada sampel dan membandingkannya dengan nilai  $r$  yang terdapat pada tabel nilai-nilai  $r$  *Product Moment*. Tabel tersebut dapat melihat Lampiran A.4. Peneliti harus menetapkan berapa tingkat kepercayaan/keyakinan akan hasil penelitian. Misalnya, peneliti yakin 95% hasil penelitian yang dilakukan ialah benar sehingga hanya 5% saja kesalahan (*error*) pada hasil penelitian tersebut. Maka pada tabel nilai  $r$ , yang dilihat ialah taraf signifikan sebesar 5% dengan nilai  $n$  sesuai jumlah *dataset*/sampel yang diuji korelasinya. Sebagai contoh,  $n = 5$ , maka nilai  $r$  tabel yaitu 0,878. Sehingga, jika nilai koefisien korelasi ( $r$  sampel) lebih besar atau sama dengan nilai  $r$  tabel maka koefisien korelasi sampel signifikan.

- Uji  $t$

Pengujian yang kedua ini menggunakan rumus  $t$  (2.6) untuk mendapatkan nilai  $t$  sampel. Selanjutnya nilai  $t$  sampel akan dibandingkan dengan harga  $t$  tabel. Selain itu, peneliti harus menetapkan tingkat kepercayaan, nilai  $dk$ , dan uji dua pihak atau satu pihak. Nilai  $dk$  didapat dari perhitungan data dikurang 2 ( $dk = n - 2$ ). Apabila nilai  $t$  sampel lebih besar atau sama dengan harga  $t$  tabel, maka koefisien korelasi sampel signifikan.

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

(Sugiyono, 2007)(2.6)

### 2.1.10 Penafsiran Koefisien Korelasi

Penafsiran dilakukan untuk melakukan interpretasi atau melihat apakah hubungan korelasi itu kuat/lemah atau besar/kecil. Tolak ukuran kuat/lemah atau besar/kecil terhadap interpretasi koefisien korelasi, dapat melihat tabel tingkat hubungan pada Lampiran A.5. Misalnya, nilai koefisien korelasi sampel sebesar 0,45. Maka sesuai tabel tingkat hubungan, sampel tersebut memiliki tingkat hubungan yang “Sedang”.

### 2.1.11 Tingkat Variabel X Mewakili Keadaan Variabel Y

Melihat tingkat pengaruh variabel luar terhadap tingkat hubungan sampel, dapat menggunakan Determinasi. Nilai koefisien Determinasi dari  $r$  sampel dihitung menggunakan rumus (2.7).

$$\text{Determinasi} = r^2$$

(Sugiyono, 2007)(2.7)

Nilai koefisien Determinasi menjadi penentu representasi variabel x terhadap y dan variabel luar yang juga merepresentasikan variabel y. Sebagai contoh, x merupakan jumlah kasus yang pernah terjadi di suatu wilayah dan variabel y ialah tren dari istilah pencarian ‘penyakit liver’. Peneliti ingin melihat korelasi kedua variabel tersebut. Setelah dilakukan perhitungan, maka dapat lah nilai r dari sampel tersebut sebesar 0,58. Sehingga nilai determinasi yaitu:

$$\text{Determinasi} = r^2 = 0,58^2 = 0,3364$$

Nilai determinasi diubah menjadi bilangan persen sehingga 0,3364 menjadi 33,64%. Kesimpulannya, tren istilah pencarian hanya mampu menggambarkan atau merepresentasikan atau memprediksi kasus penyakit sedang terjadi di wilayah sebesar 33,64%. Artinya 66,36% keadaan di lapangan atau kondisi wilayah tersebut diwakili oleh hal lain.

## 2.2 Penggunaan Google Trends Sebelumnya

Peneliti telah melakukan studi pustaka terhadap literatur (jurnal, *report*, prosiding, artikel) yaitu 13 pembahasan GT yang dilakukan di Indonesia dan 12 pembahasan di luar Indonesia. Secara garis besar, penerapan GT sudah dilakukan di beberapa bidang seperti kesehatan (paling banyak diterapkan), bisnis, hukum/tindak kriminal dengan tujuan untuk pengawasan, perolehan strategi, pengecekan dan pembuktian kasus. Hal tersebut merupakan analisa awal terhadap 24 literatur tersebut.

Penelitian tentang GT dimulai di sektor kesehatan oleh (Carneiro & Mylonakis, 2009a; Olson et al., 2009; Pelat et al., 2009) dengan diskusi tentang penyebaran kasus wabah, dapat dilihat pada Lampiran A.7. Ketiga studi dilakukan di luar Indonesia. Dua tahun kemudian, pada 2011, penelitian Google Trends mulai dilakukan oleh (Falgenti, 2011; Rahmanto, Wiyadi, & Isa, 2011) dalam bisnis di Indonesia, sementara (Vosen & Schmidt, 2011) meneliti konsumsi konsumen di *United States of America* (USA). Pada 2015, penelitian (Nijman, 2015) mencoba menggunakan GT di bidang hukum, yaitu mengangkat masalah kemungkinan pelanggaran peraturan dan sebagian memantau ekspor barang.

Pemanfaatan GT sebagai dukungan untuk studi dimulai pada 2015 serta para peneliti (Sulistiyanto, Nugraha, Sari, Karima, & Asrori, 2015) yang menggunakan GT sebagai penguat dalam pemilihan *embedded* sistem (*Arduino*, *Atmega*, X86, dan *Raspberry Pi*), ini diilustrasikan dalam Lampiran A.8. Dari empat pilihan tersebut, para peneliti memilih

*Arduino* dengan pernyataan yang menguatkan bahwa *Arduino* adalah *platform* dengan permintaan terbanyak berdasarkan tren GT pada 2005-2015.

Penggunaan GT didasarkan pada 24 literatur yang dibahas para peneliti, ada enam pola penggunaan teknis seperti di Lampiran A.8. GT digunakan untuk mengoptimalkan situs web sehingga nilai optimasi mesin pencari (SEO) meningkat. Implementasi di Indonesia juga menggunakan GT untuk membuktikan pernyataan peneliti tentang sesuatu. Kemudian peneliti (Laila Ratna Swari & Lakoro, 2016) menggunakan GT untuk membantu dalam menentukan tema, judul, dan membuat latar belakang penelitiannya. Selain itu, di Indonesia GT digunakan untuk membantu pemasaran produk, iklan, dan promosi.

Tidak seperti di Indonesia, luar negeri memiliki dua pola, yaitu pengecekan dan prediksi. Untuk pengecekan, dimaksudkan bahwa GT pada awalnya digunakan untuk memeriksa tren kasus dan pada akhirnya disimpulkan kemungkinan dampak / manfaat jika menggunakan GT seperti pemantauan, peringatan dini, dan pemantauan reaksi masyarakat. Sementara pola prediksi, peneliti (Choi & Varian, 2012a; Vosen & Schmidt, 2011) segera mengusulkan GT untuk memprediksi sesuatu dan dibuktikan dengan membandingkan model prediksi dari konvensional dengan GT.

Beberapa penelitian mengusulkan dan memberikan saran dalam mengambil data tren GT sehingga informasi yang diperoleh sesuai dengan apa yang dimaksudkan atau kemungkinan biasanya kecil. Seperti dalam Lampiran A.9, peneliti (Carneiro & Mylonakis, 2009a) mengatakan bahwa lokasi yang akan dilihat sebagai tren harus memiliki populasi yang besar, negara-negara maju karena penggunaan internet lebih merata dan area yang tersedia di GT karena beberapa daerah tidak memiliki tren. Selain itu, juga harus mempertimbangkan bahasa istilah pencarian yang dianalisis di GT. *Dataset* konvensional seperti survei atau laporan pemantauan, diperlukan untuk memvalidasi data dari tren GT karena sampai sekarang belum ada pendapat yang mengatakan data GT dapat digunakan sebagai data *primer*. Untuk rentang waktu *dataset* harus disinkronkan jika melakukan analisis atau uji kompatibilitas dengan *dataset* lain (Cho et al., 2013; Olson et al., 2009).

Data GT perlu diuji validitasnya sehingga di masa depan tren GT dapat digunakan untuk berbagai keperluan publik. Ada empat peneliti yang telah memvalidasi atau memeriksa data GT seperti yang ditunjukkan pada Lampiran A.10. Semua peneliti diuji dengan teknik yang sama yaitu korelasi, korelasi dilihat dengan *dataset* konvensional (survei / laporan pengawasan) sebagai perbandingan. Hasil tes menetapkan bahwa GT harus digunakan sebagai data tambahan atau data sekunder bukan sebagai data *primer*.

Ada empat *domain* utama atau bidang dalam penggunaan GT dan dari domain utama ini muncul *sub-domain* yang dibahas oleh peneliti sebelumnya. Lampiran A.11 menunjukkan bahwa sektor kesehatan sendiri telah membahas tujuh masalah, yaitu penyakit menular, kesehatan mental dan penggunaan narkoba, penyakit tidak menular lainnya, populasi perilaku umum, pemantauan reaksi masyarakat, peringatan dini dan pengawasan.

Ada hal yang menarik yaitu penelitian (Nijman, 2015), peneliti memasukkannya ke dalam bidang hukum yaitu kejahatan dengan *sub-domain* analisis kejahatan. Studi ini (Nijman, 2015) mengangkat isu larangan perdagangan belut untuk negara-negara anggota Uni Eropa (UE) dengan meningkatkan ekspor belut di Indonesia. Larangan ekspor belut *Anguilla Eropa* diputuskan oleh *the Convention on International Trade in Endangered Species of Wild Fauna and Flora* (CITES) pada tahun 2007 dan pada tahun 2010 juga mengeluarkan larangan terhadap semua jenis belut. Dua puluh delapan anggota Uni Eropa yaitu Swedia, Austria, Slovakia, Irlandia, Belanda, Spanyol, Siprus, Italia, Belgia, Portugal, Kroasia, Bulgaria, Prancis, Lithuania, Republik Ceko, Luksemburg, Yunani, Malta, Denmark, Latvia, Polandia, Estonia, Jerman, Rumania, Finlandia, Inggris, Slovenia, dan Hongaria (EEAS, 2017). Sehingga negara Indonesia tidak termasuk dalam larangan tersebut.

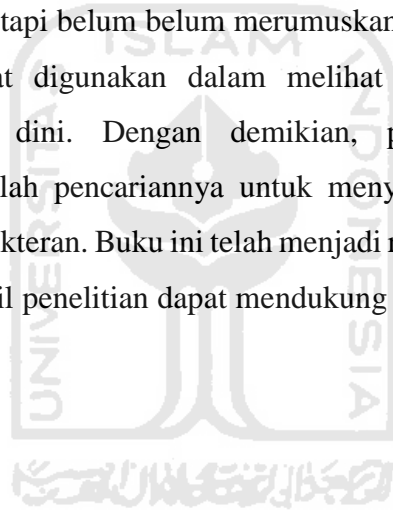
Indonesia juga membuat peraturan yang membatasi ekspor belut pada tahun 1974, tetapi kegiatan ekspor semakin besar setiap tahun. Sejak 2010, ekspor di Indonesia telah tinggi pada 4,0 juta kg / tahun pada 2010-2012 dan pada 2013 menjadi 6,1 juta kg. Nijman memeriksa tren GT, tampaknya sejak 2010 pencarian ikan "belut" telah meningkat seiring dengan peningkatan data ekspor Indonesia. Ini membuat kesimpulan bahwa GT dapat digunakan untuk peringatan dini atas tindakan kriminal atau sejenisnya. Nijman sendiri belum mengkonfirmasi bahwa informasi dari media dan dokumen pemerintah Indonesia sudah benar. Jika benar, maka tentu saja pemerintah harus memperhatikan agar perdagangan internasional tidak menghambat konservasi belut (belut tropis). Sedangkan bidang penelitian menggunakan GT sebagai alat bantu dan bidang periklanan / bisnis menggunakan GT untuk berbagai tujuan seperti yang ditunjukkan pada Lampiran A.11.

Ada dua belas model penggunaan GT bersama dengan rekomendasi untuk penggunaan model GT dari masing-masing peneliti, dapat dilihat pada Lampiran A.12. Sedangkan metode atau teknik yang digunakan dalam penelitian ini, ada sembilan metode / pendekatan. Usulan Penelitian ini akan fokus pada delapan model penggunaan yaitu pelacakan, pemantauan reaksi publik, verifikasi, pemantauan, peringatan dini, prediksi, dan penelitian. Sedangkan metode yang digunakan sama dengan metode yang digunakan oleh

peneliti (Husnayain et al., 2019). Perbaharuan yang ingin dicapai dari penelitian ini adalah ketersediaan daftar templat istilah pencarian yang kemudian menghasilkan istilah pencarian untuk melihat tren penyakit di GT untuk wilayah Indonesia.

### **2.3 Penelitian Sebelumnya dalam Bidang Kesehatan**

Ada dua sisi perbedaan yang ingin difokuskan dalam penelitian ini, yaitu domain penelitian dan evaluasi penelitian. Penelitian tren GT dalam kasus penyakit di Indonesia masih membahas satu penyakit menular, yaitu Demam Berdarah *Dengue* (DBD), sementara peneliti akan memeriksa penyakit menular lainnya (Tuberkulosis / TB). Lampiran A.13 menunjukkan bahwa penelitian sebelumnya (Carneiro & Mylonakis, 2009a; Nuti et al., 2014a; Olson et al., 2009) tidak membahas secara khusus bagaimana menentukan istilah Pencarian yang akan dilihat tren-nya. Sementara peneliti (Cho et al., 2013; Husnayain et al., 2019; Kang et al., 2013; Pelat et al., 2009; Seifter et al., 2010) telah memberikan asal-usul menentukan istilah pencarian, tetapi belum belum merumuskan dengan jelas tentang istilah pencarian yang nantinya dapat digunakan dalam melihat tren di GT untuk tujuan pengawasan atau peringatan dini. Dengan demikian, peneliti akan memperluas (mengembangkan) kategori istilah pencariannya untuk menyesuaikan dengan informasi dalam buku Kapita Selekta Kedokteran. Buku ini telah menjadi referensi di dunia kedokteran / kesehatan. Di masa depan, hasil penelitian dapat mendukung realisasi sistem pengawasan wabah penyakit berbasis *online*.



## **BAB 3**

### **Metodologi**

#### **3.1 Data**

Penelitian ini sebenarnya menggunakan dua data inti yang menjadi fokus penelitian, yaitu data dari tren *Google Trends* dan data dari lembaga pengawasan penyakit di Indonesia. Badan tersebut adalah Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia yang memiliki dua direktorat lagi di bawah pengawasannya, yaitu direktorat P2 Penyakit Tidak Menular dan direktorat P2 Penyakit Menular Langsung (untuk Tuberkulosis). Namun selain itu ada juga yang mengambil dari sumber lain seperti buku Kapita Selekta Kedokteran.

Data diambil dari beberapa sumber sehingga, jika dikelompokkan berdasarkan sumber, itu menjadi sebagai berikut:

##### 1. Data Primer

Peneliti melakukan pengamatan langsung ke situs <http://trends.google.com> untuk mendapatkan data tren tentang Tuberkulosis.

##### 2. Data Sekunder

a. Melihat dan membaca literatur dilakukan untuk mengumpulkan data suatu kasus penyakit di Indonesia. Data ini diambil dari Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia dalam bentuk file dokumen yang berisi jumlah kasus TBC yang pernah terjadi di Indonesia.

##### b. Istilah pencarian

Istilah pencarian akan merujuk pada literatur (Cho et al., 2013; Husnayain et al., 2019; Kang et al., 2013) dan buku Kapita Selekta Kedokteran.

Penelitian ini menggunakan data statistik berupa jumlah kasus penyakit (angka) dan jumlah tren GT (angka), sehingga menjadi penelitian kuantitatif. *Google Trends* data tren dan data pelaporan kasus penyakit memiliki tipe data deret waktu, karena data dikumpulkan secara berkala dari waktu ke waktu, yaitu 2014 hingga 2018. Penelitian ini akan mengeksplorasi sebuah fenomena, yaitu kasus penyakit di Indonesia dengan tren dalam mencari informasi tentang penyakit di internet. Jadi, penelitian deskriptif juga digunakan dalam penelitian ini. Data yang digunakan adalah data sekunder, terutama untuk data kasus penyakit di Indonesia.

Peneliti mengumpulkan data dalam tiga cara, yaitu:

#### 1. Tinjauan Literatur

Studi penelitian sebelumnya dilakukan untuk mengetahui bagaimana peneliti sebelumnya melakukan analisis dan penentuan istilah pencarian. Peneliti juga mengambil metode yang sama dan menambahkan sumber istilah pencarian yang baru.

#### 2. Pengamatan

Pengamatan yang dimaksud di sini adalah peneliti secara langsung mengambil *dataset* tren dari situs *web Google Trends* resmi. Sedangkan dataset kasus penyakit diambil langsung dengan mengunjungi kantor resmi Kementerian Kesehatan Indonesia.

### 3.2 Lokasi dan Waktu Penelitian

Peneliti melakukan selama sepuluh bulan mulai dari studi literatur hingga membuat laporan penelitian. Rincian jadwal untuk melakukan penelitian dapat dilihat pada Lampiran B.1. Penelitian ini dilakukan di Indonesia dengan tingkat analisis di tingkat Negara dengan data yang digunakan berasal dari Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia.

### 3.3 Langkah-langkah Penelitian

Secara garis besar, langkah-langkah yang dilalui dalam penelitian ini adalah lima langkah yang dimulai dengan pencarian tema penelitian hingga membuat laporan hasil penelitian.

#### 3.3.1 Pencarian Tema Penelitian

Peneliti membaca satu hasil penelitian (Riyanto, 2014) yang membahas penggunaan GT dalam menentukan istilah pencarian suatu produk. Selain itu, peneliti tertarik untuk mengeksplorasi topik GT, karena diharapkan akan menjadi populer di masa depan (mengingat peluang pemanfaatannya yang besar).

#### 3.3.2 Pendalaman Tema Penelitian

Peneliti mencari beberapa hasil penelitian atau aplikasi sederhana (tinjauan literatur) terkait dengan GT yang diterapkan di Indonesia dan di luar Indonesia. Ada 24 literatur yang diperoleh untuk mengeksplorasi tema GT.

#### 3.3.3 Penentuan Fokus Penelitian

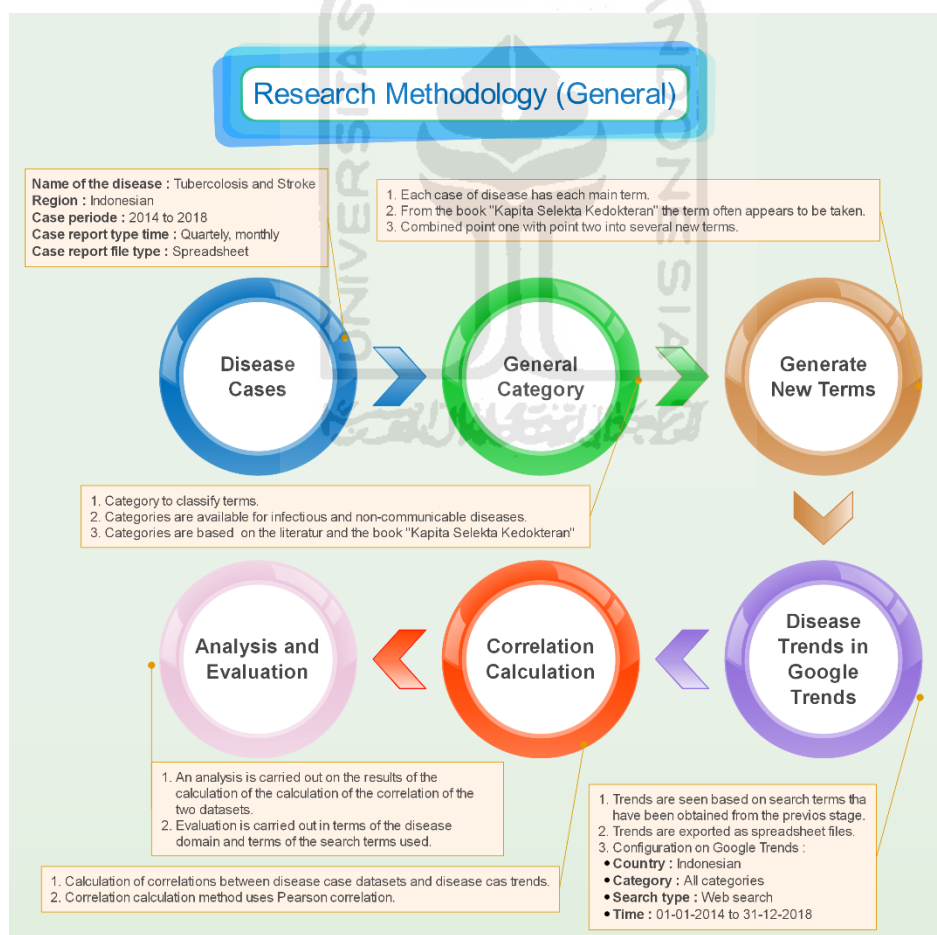
Penelitian ini berfokus pada membuat praktik terbaik dalam penggunaan data GT dengan menggunakan studi kasus di sektor kesehatan. Sektor kesehatan diambil karena efek yang ditimbulkan secara luas yang menyangkut masyarakat (publik). Pada tahun 2017 Indonesia menjadi nomor ketiga tertinggi didunia untuk kasus Tuberkulosis (TBC) (Kusnandar, 2019). TBC juga menjadi salah satu penyebab kematian terbesar sepanjang tahun 2007-2017



(Pranita, 2020). WHO juga melaporkan TBC termasuk salah satu dari 10 penyebab kematian terbesar tahun 2012 di Indonesia (WHO, 2015). Tahun 2014 penyakit TBC menyumbang 5,7% sebagai penyakit terbesar diderita masyarakat di Indonesia (Rokom, 2017). Diketahui juga bahwa 75 persen dari pasien TBC tergolong kelompok produktif yaitu berusia 15-55 tahun (Ihsanuddin, 2020). TBC sebagai penyakit menular mendapatkan perhatian khusus dari Ikatan Dokter Indonesia (IDI) dan mendesak pemerintah untuk menarapkan protikol Kejadian Luar Biasa (KLB) dalam penanganannya (Tamba, 2018). Melihat gejolak kasus kematian dan kewaspadaan tinggi dari berbagai lembaga terhadap penyakit Tuberkulosis maka, peneliti memilih untuk mengangkat penyakit Tuberkulosis (TBC) sebagai kasus penyakit yang diteliti.

### 3.3.4 Tahapan Inti Penelitian

Secara umum, langkah keempat ini memiliki enam kegiatan / proses yang dilakukan dalam penelitian ini, seperti yang diilustrasikan dalam Gambar 3.1.



Gambar 3.1 Metodologi Penelitian (Secara Umum).

Keenam kegiatan ini dimulai dengan mengumpulkan data tentang kasus Tuberkulosis. Data kasus penyakit akan disesuaikan sehingga dapat dikorelasikan dengan

data kedua, yaitu tren GT. Rincian kegiatan ini dapat dilihat pada Lampiran B.1 yang secara teknis menggambarkan kegiatan analisis dalam penelitian ini. Kegiatan-kegiatan berikut dilakukan pada tahapan keempat:

1. *Disease Cases*

Data kasus penyakit yang telah terjadi di Indonesia akan dikumpulkan dalam kurun waktu lima tahun, yaitu 2014, 2015, 2016, 2017 dan 2018. Data diambil dari laporan di Direktorat Jenderal P2P Kementerian Kesehatan Republik Indonesia, selanjutnya disebut sebagai *dataset* kasus penyakit. Laporan tersebut kemudian diproses dengan melihat kelengkapan, struktur, atau keseragaman struktur (variabel) sehingga nantinya dapat dihitung korelasi *Pearson*. Data kasus penyakit akan dirubah rentang nilainya menjadi rentang nilai seperti tren GT, yaitu rentang nilai 0-100. Perubahan nilai menggunakan rumus (2.3).

2. *General Category*

Membuat kategori umum dimaksudkan untuk mengklasifikasikan istilah pencarian dan juga untuk memunculkan kemungkinan istilah pencarian baru. Kategori umum diperoleh dari literatur dan *sub-bab* yang membahas beberapa penyakit dalam buku *Kapita Selekta Kedokteran*.

3. *Generate New Term*

Istilah pencarian akan dibuat dari beberapa aspek dan menjadi aturan untuk menghasilkan daftar istilah pencarian. Ini disebut generator istilah pencarian. Aspek yang menghasilkan generator istilah pencarian adalah **cara membuat istilah pencarian dari buku Kapita Selekta Kedokteran** yang merupakan buku referensi yang paling banyak digunakan di dunia medis, terutama di Indonesia. Istilah pencarian baru dibentuk dari **dua aspek, yaitu istilah utama** (*main term*) dan **istilah kombinasi** (*combination term*). Bentuk gandingan dari dua aspek ini mengikuti persamaan (3.1) dan hanya digunakan untuk **pembuatan istilah pencarian dari sumber buku Kapita Selekta Kedokteran**.

$$\text{Istilah Pencarian} = \text{istilah utama} + \text{istilah kombinasi} \quad (3.1)$$

Istilah utama adalah istilah yang ditetapkan dari definisi umum objek penelitian, yaitu Tuberkulosis. Sedangkan istilah kombinasi diperoleh dari proses pencarian frekuensi kata dari masing-masing *sub bab* dari pembahasan penyakit dalam buku *Kapita Selekta*

Kedokteran. Jika *sub bab* diskusi (kategori umum) memiliki lima, maka istilah kombinasi akan menjadi lima sumber yang berbeda. Pada saat **memilih istilah kombinasi juga melalui beberapa kegiatan yaitu istilah peringkat yang paling sering muncul**. Dua puluh istilah teratas diambil untuk difilter lagi. Filter selanjutnya berupa: a) istilah yang tidak memiliki dua makna; b) istilah yang bukan angka / simbol; c) bukan konjungsi; d) dan bukan istilah yang memiliki makna jauh yang berkaitan dengan penyakit (misalnya nama penyakit lain).

#### 4. *Disease Trends in Google Trends*

Setelah kegiatan pertama (1) dan ketiga (3) selesai, data tren GT akan dikumpulkan dengan istilah Pencarian yang telah dibuat dalam kegiatan ketiga (3). Aturan dalam melihat tren adalah untuk Indonesia, kategorinya adalah "Semua kategori", jenis pencariannya adalah "Pencarian Web" dan rentang waktunya dari 1 Januari 2014 hingga 31 Desember 2018. Kegiatan ini akan menghasilkan beberapa *spreadsheet* per istilah pencarian yang berformat laporan mingguan. Format laporan akan diubah menjadi laporan triwulan sesuai format *dataset* kasus penyakit. Perubahan format laporan menggunakan fungsi 'AVERAGE' yang terdapat pada aplikasi *Miscrosoft Office Excel*.

#### 5. *Correlation Calculation*

Ada tiga data yang dibahas dalam proses ini, yaitu daftar istilah pencarian, dataset tren kasus penyakit, dan dataset kasus penyakit. Seperti dalam Lampiran B.2 tampak bahwa data baru (hasil korelasi) akan muncul ketika dataset tren GT dikorelasikan dengan dataset kasus penyakit. Korelasi menggunakan metode korelasi Pearson untuk melihat tingkat hubungan antara dua set data.

#### 6. *Analysis and Evaluation*

Analisis dilakukan dari hasil perhitungan korelasi yang diperoleh pada tahap sebelumnya. Penelitian ini menetapkan tingkat kepercayaan 95% dalam hasil, sehingga ada kesalahan 5% atau tingkat signifikansi. Setiap tahun memiliki empat data (triwulanan) dan penelitian ini memakan waktu lima tahun sehingga  $n = 20$ . Mengacu pada Lampiran A.4, penelitian ini memiliki koefisien korelasi yang signifikan pada  $r \geq 0,444$ . Dengan demikian, istilah pencarian yang diklasifikasikan sebagai signifikan (koefisien korelasi  $\geq 0,444$ ) akan diambil untuk analisis lebih lanjut.

Penelitian ini memiliki keinginan yang kuat untuk mendapatkan templat yang baik, sehingga istilah pencarian yang melewati filter signifikansi akan difilter lagi. Filter kedua ini mengacu pada Lampiran A.5. Tingkat hubungan terendah disebut 'Sangat Rendah'

dengan interval koefisien korelasi 0,00-0,199 dan tertinggi disebut "Sangat Kuat" dengan interval koefisien korelasi 0,80-1,000 (Sugiyono, 2007). **Penelitian ini mengambil tingkat hubungan tertinggi, sehingga istilah pencarian yang akan dianalisis lebih lanjut memiliki  $r \geq 0,80$ .** Pada tahap ini, penelitian sudah memiliki *template* yang baik untuk membangun istilah pencarian. **Nilai korelasi positif tinggi menjadi patokan bahwa istilah pencarian dapat menampilkan tren yang mampu menggambarkan keadaan kasus penyakit di area yang ditampilkan.** Istilah pencarian dievaluasi untuk membuat templat istilah pencarian. Analisis faktor penetrasi internet di satu wilayah terhadap tren GT akan terlihat menggunakan data tambahan dari Asosiasi Penyedia Layanan Internet Indonesia atau disingkat APJII.

Evaluasi dilakukan dalam beberapa hal yaitu: a) Segi domain untuk melihat korelasi domain penyakit dengan kasus penyakit yang telah terjadi sebelumnya; b) Melihat tingkat pengaruh apakah itu istilah utama atau istilah kombinasi dalam membuat istilah pencarian dengan korelasi positif yang tinggi; c) Segi istilah pencarian untuk melihat pengaruh setiap kategori istilah pencarian terhadap tingkat korelasi yang positif (bagus); d) Segi tren GT untuk melihat faktor yang mempengaruhi tinggi rendahnya tren; e) Segi tren tren GT untuk melihat faktor penetrasi internet dalam kesuksesan penggunaan GT sebagai alat peringatan dini atau pengawasan.

### **3.3.5 Tahapan Pembuatan Laporan Hasil Akhir**

Langkah terakhir yang dilakukan ialah pembuatan laporan serta penyampaian hasil penelitian yang telah dilakukan.

## BAB 4

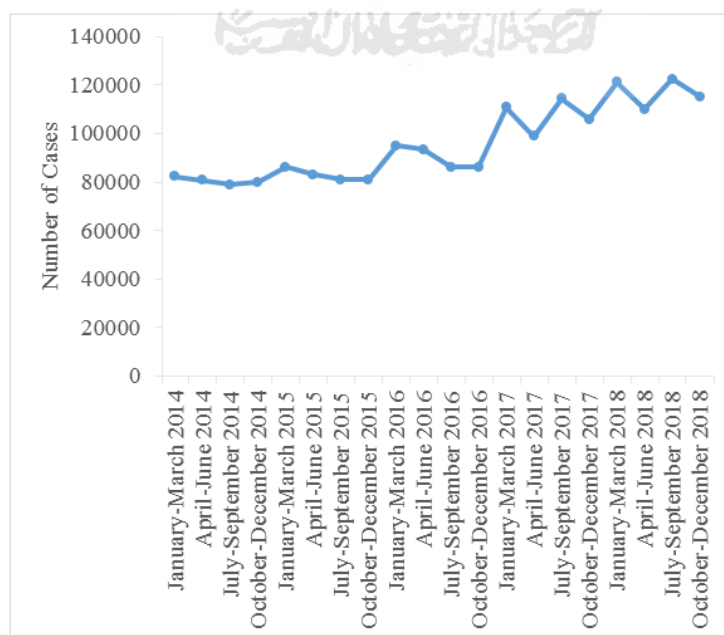
### Hasil dan Pembahasan

#### 4.1 Hasil

Berdasarkan metodologi penelitian yang dibahas sebelumnya, langkah yang diambil adalah pengumpulan dan persiapan *dataset* seperti kasus penyakit, istilah pencarian, tren istilah pencarian yang ada dan proses korelasi data kasus penyakit dengan tren istilah pencarian. Sementara analisis dibahas dalam *sub bab* diskusi, baik analisis dalam hal korelasi, domain dan istilah pencarian.

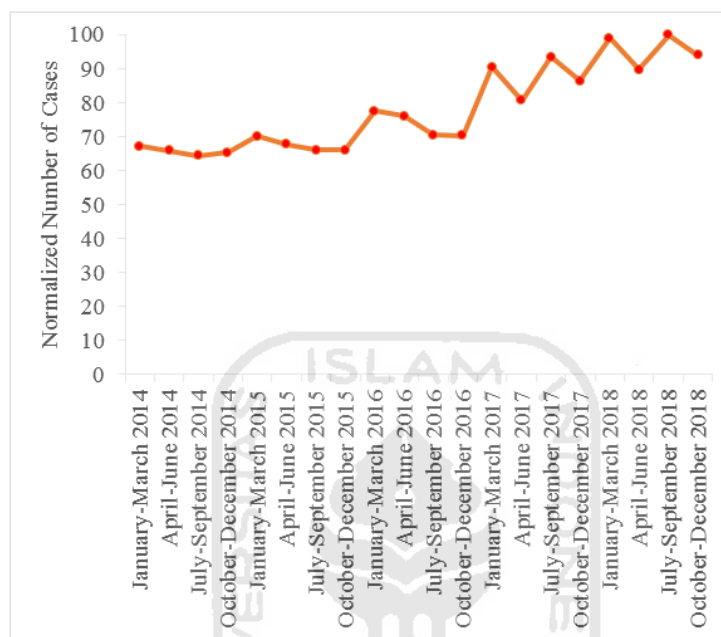
##### 4.1.1 Disease Cases

Kasus penyakit yang diambil ialah Tuberkulosis yang bersumber dari Ditjen P2P Kemkes RI. Ditjen tersebut yaitu Direktorat Jendral Pencegahan dan Pengendalian Penyakit Menular Langsung (P2PML). Pada Ditjen P2PML memiliki Kepala *Sub* Direktorat (Kasubdit) untuk beberapa penyakit yaitu Kasubdit Tuberkulosis, Kasubdit Infeksi Saluran Pernapasan Akut, dan Kasubdit Penyakit Tropis Menular Langsung. *Dataset* kasus penyakit Tuberkulosis diperoleh dari Kasubdit Tuberkulosis dengan data rentang waktu tahun 2014 hingga 2018. *Dataset* berupa *files spreadsheet* dengan laporan triwulan untuk setiap kabupaten di seluruh Indonesia. Pada Gambar 4.1 terlihat bahwa kasus penyakit Tuberkulosis cenderung naik tiap tahunnya.



Gambar 4.1 Kasus penyakit Tuberkulosis di Indonesia tahun 2014-2018.

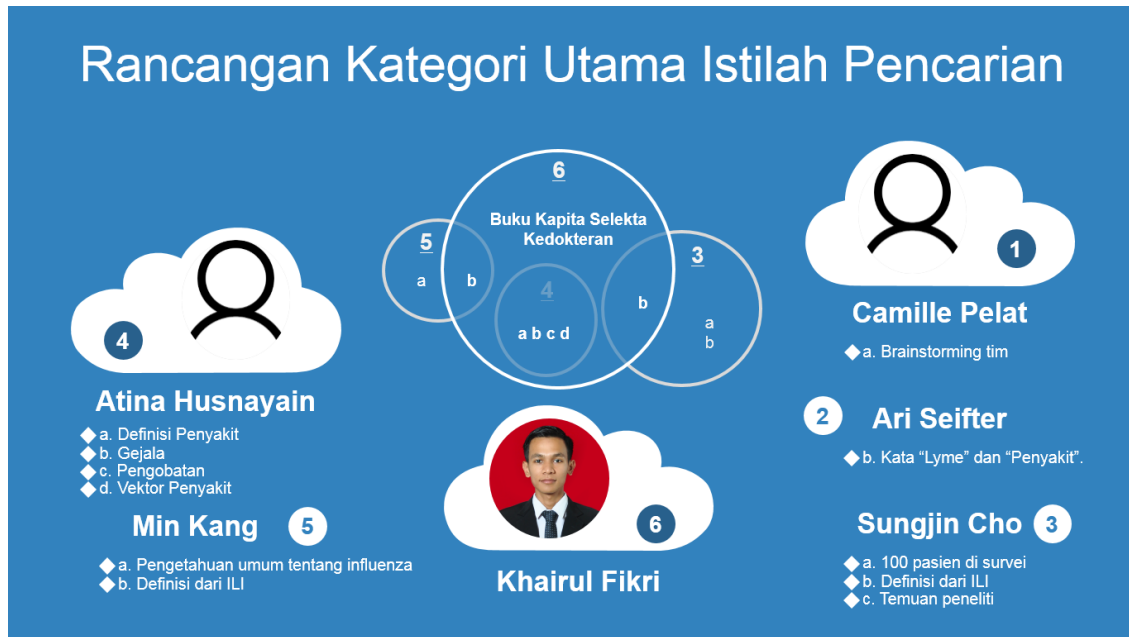
*Dataset* kasus penyakit akan dilakukan normalisasi kedalam interval nilai *range* 0-100, mengikuti interval nilai dari tren GT. Sehingga saat korelasi, kedua variabel ( $X = \text{dataset kasus penyakit}$ ,  $Y = \text{dataset tren GT}$ ) memiliki interval nilai yang sama. Pada Gambar 4.2 merupakan grafik normalisasi kasus penyakit Tuberkulosis di Indonesia dalam range nilai 0-100.



Gambar 4.2 Normalisasi kasus penyakit Tuberkulosis di Indonesia dalam range nilai 0-100.

#### 4.1.2 General Category

Istilah pencarian menjadi kunci untuk melihat terlihat tren di GT. Peneliti mengumpulkan lima cara untuk menentukan istilah pencarian yang telah digunakan di sektor kesehatan, terutama kasus penyebaran penyakit. Seperti pada Gambar 4.3, peneliti (Pelat et al., 2009) melakukan Tim *Brainstorming* untuk memunculkan istilah pencarian yang akan digunakan. Sedangkan (Seifter et al., 2010) menambahkan kata "penyakit" untuk membuat istilah pencarian. Berbeda dengan penelitian (Cho et al., 2013) yang melakukan survei terhadap 100 pasien untuk mendapatkan istilah pencarian dan definisi penyakit yang diteliti. Peneliti (Husnayain et al., 2019) membuat kategorisasi untuk memunculkan istilah pencarian, yaitu definisi penyakit, gejala, pengobatan, dan vektor penyakit. Peneliti (Kang et al., 2013) kembali menggunakan definisi penyakit untuk memunculkan istilah pencarian dan dilengkapi dengan pengetahuan umum tentang penyakit yang diteliti.



Gambar 4.3 Rancangan kategori utama istilah pencarian.

Studi ini mengambil tiga studi sebelumnya untuk menentukan kategori yaitu a) “definisi” diambil dari (Choi & Varian, 2012a; Husnayain et al., 2019; Kang et al., 2013) b) “gejala”, “pengobatan” dan “vektor penyakit” dari penelitian (Husnayain et al., 2019). Jadi, ada empat kategori yang terbentuk dari literatur yaitu Definisi, Gejala, Pengobatan, dan Vektor Penyakit yang dapat dilihat pada Lampiran C.1.

Peneliti menambahkan **pengkategorian baru yang berasal dari buku Kapita Selekt Kedokteran**. Seperti pembahasan Bab 2 (dua) tentang fokus penelitian, peneliti mengambil sepuluh penyakit utama dari dua puluh penyebab kematian terbesar di Indonesia. Sepuluh penyakit tersebut akan dilihat pembahasannya didalam buku Kapita Selekt Kedokteran. Setiap pembahasan untuk satu penyakit pada buku Kapita Selekt Kedokteran, memiliki beberapa *sub-pembahasan* seperti *sub bab* definisi, *sub bab* epidemiologi, dan lainnya seperti yang terlihat pada Lampiran C.2. Ada tiga ketentuan yang ditetapkan yaitu a) penyakit yang dibahas akan dikelompokkan menjadi penyakit menular dan tidak menular; b) kategori utama di buat berdasarkan nama *sub bab*; c) syarat *sub bab* yang menjadi kategori utama yaitu *sub bab* yang tersedia diseluruh penyakit yang dibahas. Akhirnya, ada lima kategori utama untuk penyakit menular yaitu **Definisi, Faktor Risiko, Epidemiologi, Diagnosis, dan Tata Laksana**. Sedangkan untuk penyakit tidak menular ada tiga yaitu **Definisi, Diagnosis, dan Tata Laksana**.

Kategori dari literatur yaitu Definisi digabungkan ke kategori utama yang juga ada Definisi. Kategori Gejala digabungkan ke kategori utama yaitu Diagnosis karena didalam

mendiagnosis juga membahas gejala-gejala. Sedangkan kategori Pengobatan digabungkan ke kategori utama yaitu kategori Tata Laksana. Kategori literatur terakhir yaitu Vektor Penyakit digabungkan ke kategori Epidemiologi. **Sehingga kategori utama (*General Category*) yang digunakan tetap ada lima untuk penyakit menular.**

#### **4.1.3 Generate New Terms**

*Main term* atau istilah utama / kata kunci utama ditentukan untuk setiap domain penyakit. Penyakit Tuberkulosis ditetapkan ada tiga istilah utama yaitu ‘tuberkulosis’, ‘tb’, dan ‘tbc’. Pemunculan *term* baru, dibutuhkan istilah kombinasi (*combination term*) yang didapatkan dari proses perangkian frekuensi kata (sering muncul) dalam pembahasan di setiap penyakit untuk masing-masing *sub bab*. Sehingga nanti *term* baru akan muncul dari lima *sub bab* (yang juga menjadi *General Category*) untuk penyakit menular.

Proses perangkian frekuensi kata diawali dengan penyimpulan teks pembahasan masing-masing *sub bab* penyakit di dalam buku KSK, nomor halaman buku setiap penyakit yang di bahas dapat dilihat pada Lampiran C.2. Setelah semua teks untuk masing-masing *sub bab* didapat, teks tersebut diolah melalui *website voyant-tools.org*. Hasil dari olahan tersebut akan diambil dan dilihat untuk dua puluh *term* teratas. Sebagai contoh dapat dilihat pada Lampiran C.3, dua puluh *term* teratas dipilih dan difilter lagi seperti apakah *term* tersebut berupa angka (*number*), nama penyakit (*disease name*), atau lainnya yang termasuk jenis *combination term* yang ditolak seperti pada Lampiran C.4. Selanjutnya, semua *term* yang lolos filter, akan dikombinasikan dengan *main term* sehingga satu *combination term* memunculkan tiga *term* baru. Seperti Lampiran C.5, sebagai contoh *combination term* ‘pemeriksaan’ apabila dikombinasikan dengan *main term* maka membentuk tiga *term* baru yaitu ‘pemeriksaan tuberkulosis’, ‘pemeriksaan tb’, dan ‘pemeriksaan tbc’. Ada 129 istilah pencarian baru (unik) yang terbuat untuk domain Tuberkulosis.

#### **4.1.4 Disease Trend in Google in Google Trends**

Istilah pencarian baru yang didapat dari tahapan sebelumnya akan dilihat tren pada GT, hal ini akan mewakili tren kasus setiap penyakit versi GT. Ada beberapa istilah pencarian yang tidak memiliki tren seperti pada Lampiran C.6, sehingga nantinya nilai korelasi akan diset menjadi nol (0) dan sekaligus menandakan bahwa istilah pencarian tersebut secara mutlak tidak bisa menggambarkan keadaan kasus penyakit disuatu wilayah.

Setingan pada GT untuk melihat tren tetap sama untuk semua istilah pencarian dengan menyesuaikan seperti pola *dataset* kasus penyakit (*disease case*) yaitu negara Indonesia, kategory nya ‘*All Category*’, jenis penelusurannya ‘*Web Search*’ dan rentang



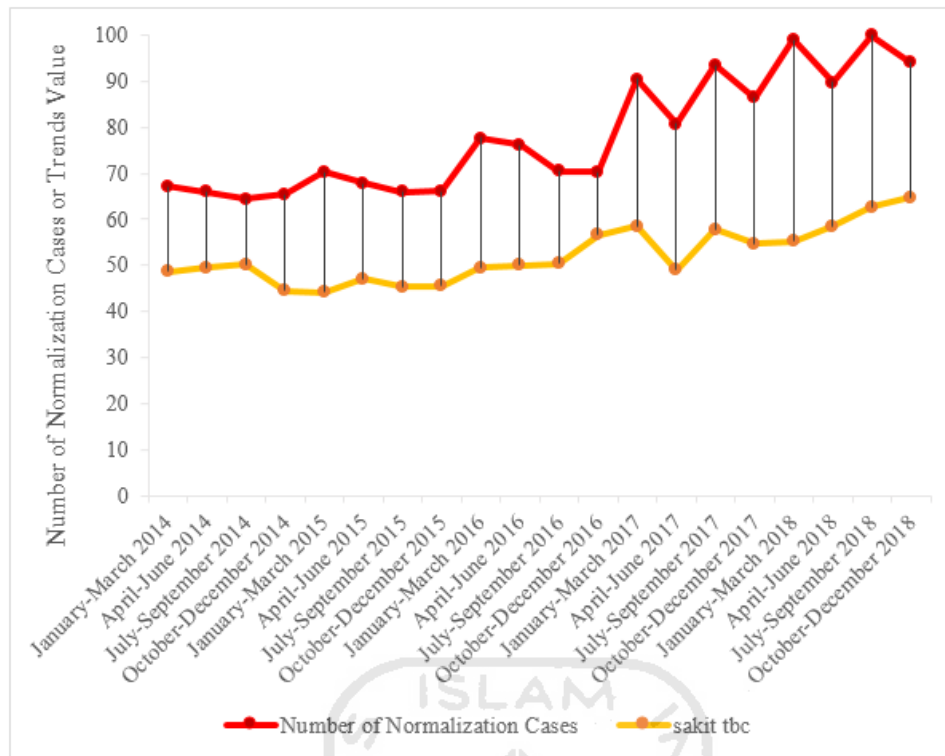
waktu '01 Januari 2014 hingga 31 Desember 2018. Hasil tren setiap istilah pencarian berupa *file* berekstensi \*.xls dengan format laporan di dalam *file* yaitu format laporan mingguan. Hal tersebut membuat perlunya normalisasi untuk menyamakan dengan kondisi *dataset* kasus penyakit yang berformat laporan triwulan. Proses normalisasi merubah dari mingguan ke triwulan menggunakan fungsi 'AVERAGE' yang disediakan pada aplikasi *Microsoft Office Excel*, seperti pada Lampiran C.7.

#### **4.1.5 Correlation Calculation**

Pada tahap ini, sudah ada dua dataset yaitu *disease case* dan *disease trends*. Kedua dataset ini akan dilihat korelasinya menggunakan *correlation pearson*. Pada korelasi *Pearson* membutuhkan setidaknya dua variabel *set data*, misalnya X dan Y. Sehingga ditetapkan bahwa variabel X adalah *disease case* dan variabel Y adalah *disease trends*. Pada Lampiran C.8 dapat dilihat contoh perhitungan korelasi yang dilakukan, perhitungan ini memanfaatkan fungsi korelasi 'Pearson' yang terdapat pada aplikasi *Microsoft Office Excel*. Perhitungan dilakukan untuk setiap kelompok istilah pencarian yaitu Definisi, Diagnosis, Epidemiologi, Faktor Risiko, dan Tata Laksana untuk domain penyakit Tuberkulosis (penyakit menular). Semua hasil perhitungan korelasi dapat dilihat pada Lampiran C.9 dan ini sudah difilter bahwa istilah pencarian yang diambil tahap pertama yaitu istilah pencarian dengan koefisien korelasi besar sama dengan ( $\geq$ ) 0,444.

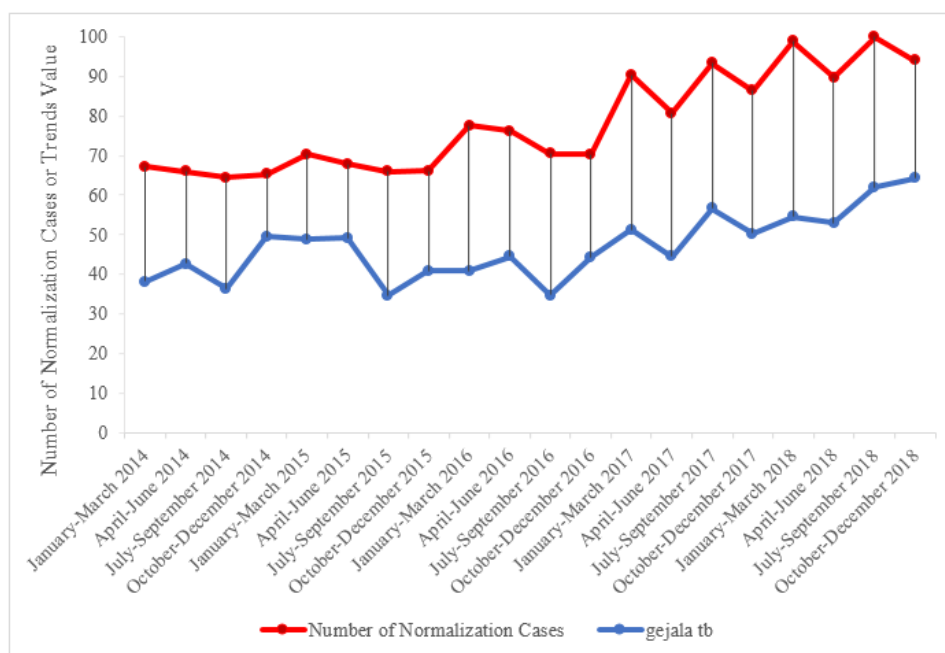
#### **4.1.6 Analysis and Evaluation**

Ada lima belas istilah kombinasi unik dan dua istilah utama yang didapatkan dari sembilan belas istilah pencarian unik pada filter tahap pertama ( $r \geq 0,444$ ). Segi istilah utama, hanya istilah 'tbc' dan 'tb' yang muncul dan ini memberikan tanda bahwa singkatan dari nama penyakit merupakan istilah pencarian yang sering digunakan oleh pengguna internet dalam mencari informasi tentang penyakit. Ada dua istilah pencarian yang lolos filter tahap kedua ( $r \geq 0,80$ ) yaitu 'sakit tbc' dan 'gejala tb'. Pada Gambar 4.4 dapat dilihat bahwa pola grafik tren *term* 'sakit tbc' hampir menyerupai pola grafik kasus penyakit.



Gambar 4.4 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'sakit tbc' (level negara).

Istilah pencarian 'sakit tbc' dari buku KSK merupakan gabungan dari istilah utama 'tbc' dan istilah kombinasi 'sakit'. Istilah pencarian 'sakit tbc' dari buku KSK menggambarkan bahwa dari segi domain penyakit, pengguna lebih cenderung menggunakan singkatan (jika ada) dari nama penyakit untuk mencari informasi di internet. Sedangkan pada Gambar 4.5, pola tren istilah pencarian 'gejala tb' juga hampir menyerupai pola tren kasus penyakit. Meskipun, ada lima titik yang polanya tidak sama yaitu April-June 2014, April-June 2015, April-June 2016, April-Juni 2018 dan October-December 2018 untuk istilah pencarian 'sakit tbc'. Berdasarkan hasil untuk buku KSK, maka hanya dua *template* istilah pencarian yang didapat yaitu 'sakit' dan 'gejala'.



Gambar 4.5 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'gejala tb' (level negara).

## 4.2 Pembahasan

Penelitian ini menambah satu sumber istilah pencarian baru selain buku KSK yaitu kueri terkait pada GT. Sumber ini bertujuan untuk memunculkan istilah pencarian baru yang nantinya akan dianalisis dan evaluasi lagi. Saat melihat tren pada GT menggunakan satu istilah pencarian, akan ada bagian khusus dari GT yang menampilkan kueri terkait istilah pencarian tersebut. **Ini lah yang akan diambil sebagai sumber istilah pencarian baru.**

### 4.2.1 Sumber Istilah Pencarian Baru

**Tahap pertama**, menetapkan istilah utama yaitu 'tb', 'tbc', dan 'tuberkulosis'. Kegunaan dari istilah utama pada saat ini ialah sebagai istilah pencarian yang akan dilihat kueri terkait terhadap istilah pencarian tersebut. Settingan pada GT sama seperti sebelumnya yaitu wilayah Indonesia, *All Category*, jenis penelusurannya "*Web search*", dan rentang waktunya "01 Januari 2019 hingga 31 Desember 2018". Untuk istilah utama 'tb' memiliki lima puluh kueri terkait (total dari *rissing* dan *top*) dan setelah difilter duplikat, ada satu yang duplikat sehingga untuk istilah 'tb' memunculkan 49 istilah pencairan baru. Istilah utama kedua yaitu 'tbc', memiliki lima puluh kueri terkait yang unik. Sedangkan istilah 'tuberkulosis' memiliki 35 kueri terkait dan hanya dua puluh istilah saja yang unik. Total awal ada 119 kueri terkait dari ketiga istilah utama. **Namun, ada satu kueri terkait yang duplikat sehingga hanya ada 118 istilah baru dari kueri terkait pada GT.**

**Tahap kedua**, menghitung korelasi antara 118 istilah baru dengan kasus penyakit Tuberkulosis. Ada 50 istilah pencarian yang memiliki koefisien korelasi  $r \geq 0,444$ . Penelitian ini menoleh kemajuan besar karena ada tiga istilah pencarian yang mencapai nilai korelasi diatas 0,9. **Tahap ketiga**, mengelompokkan kueri terkait kedalam Kategori Utama. Namun, ada beberapa istilah yang tidak dimasukkan ke Kategori Utama. Hal itu karena beda makna atau beda topik seperti ‘amel alvi’ yang merupakan nama seseorang. Sehingga, ditetapkan lah kategori baru yaitu a) “Other 1” berisi term berupa nama/inisial/informasi terkait orang atau kelompok tertentu; b) “Other 2” berisi nama merek suatu produk; c) “Other 3” berisi nama penyakit yang berbeda dengan yang diteliti; d) “Other 4” berisi nama website tertentu; dan e) “Other 5” berisi nama tempat seperti perkantoran;. **Tahap keempat**, memfilter 50 istilah pencarian untuk kedua kalinya dengan syarat  $r \geq 0,80$ . Ada 21 istilah pencarian yang lolos seperti pada Lampiran C.10.

**Tahap terakhir**, analisa dan evaluasi terhadap 21 istilah pencarian yang lolos. Ada empat istilah yang melenceng dari domain penelitian atau ditolak yaitu ‘bronkitis’, ‘google scholar’, ‘plaza oleos tb simatupang’, ‘siloam tb simatupang’, ‘mdr’, ‘tcm’, dan ‘toss’. Istilah ‘mdr’ merupakan singkatan dari *Multi-drug Resistant* yang berkaitan tentang kondisi pasien yang kebal terhadap obat Tuberkulosis. Beberapa penyakit tidak memiliki kondisi ini sehingga istilah ini ditolak. Istilah ‘tcm’ atau kepanjangan dari Tes Cepat Molekul merupakan metode yang digunakan dalam pendiagnosaan penyakit Tuberkulosis. Selain itu TCM juga memiliki makna berbeda yaitu singkatan dari Traditional Chinese Medice. Hal ini tentu membuat term ‘tcm’ tidak bagus digunakan sebagai Istilah Kombinasi Utama.

Ada satu istilah kombinasi dengan nilai 0,945384891 yang tidak dijadikan sebagai Istilah Kombinasi Utama yaitu ‘toss’. Hal ini disebabkan karena istilah ‘toss’ memiliki banyak makna seperti kegiatan tepuk tangan, singkatan dari gerakan pencarian obat penyakit, sistem rumah sakit, dan selain itu tidak terdapat seragaman untuk semua penyakit. Misalnya untuk penyakit Tuberkulosis memiliki TOSS TB (sebuah gerakan mencari obat sampai sembuh), namun untuk penyakit seperti liver tidak memiliki toss liver.

Akhirnya, ada tujuh istilah yang tergolong Istilah Kombinasi Utama. Maksudnya, istilah tersebut memiliki kemungkinan bisa digunakan untuk penyakit menular lainnya. Misalnya, term ‘obat’ memiliki peluang untuk semua penyakit sehingga bisa menjadi istilah pencarian ‘obat liver’, ‘obat hipertensi’, atau lainnya. Namun ada satu istilah kombinasi yang sendiri muncul yaitu ‘icd’ dengan penggunaan yang beragam seperti ‘icd 10 tb kelenjar’, ‘kode icd tb paru’, ‘kode icd 10 tb’, ‘icd 10 tb’, ‘icd tb paru’, dan ‘icd 10 limfadenitis’.

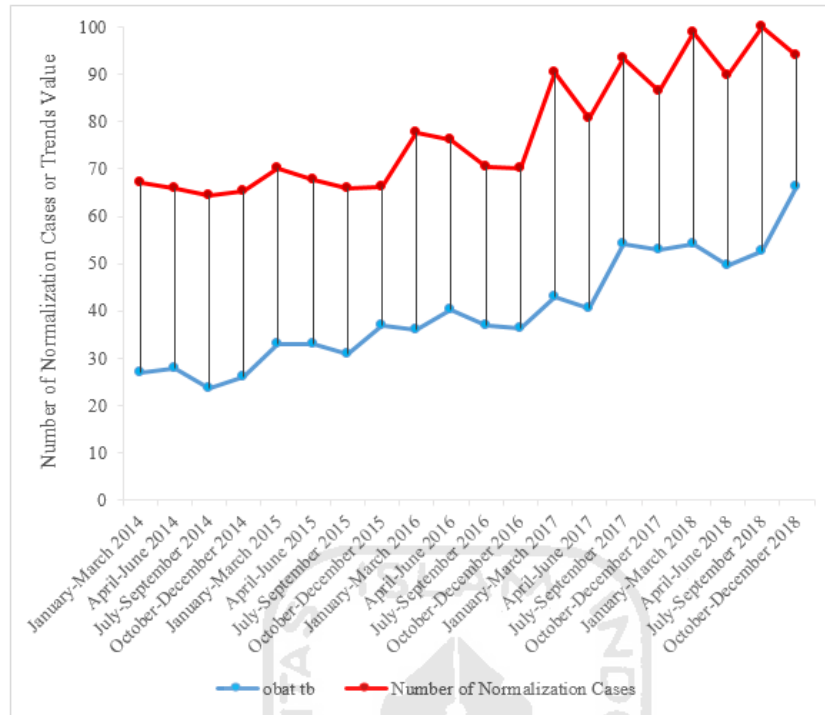
Melihat frekuensi yang banyak ini, maka penelitian ini mencoba membuat istilah perwakilan yaitu ‘kode icd’ dan diharapkan bisa digunakan untuk semua penyakit. **Sehingga ada delapan istilah kombinasi pilihan yang bisa digunakan untuk semua domain penyakit khususnya penyakit menular, seperti pada Lampiran C.11.**

Istilah 'obat tb', ‘kode diagnosa tb paru’, dan ‘icd 10 limfadenitis tb’ cenderung memiliki pola grafik yang sama dengan pola grafik kasus penyakit Tuberkulosis di Indonesia. Hal ini bisa dilihat pada Gambar 4.6, Gambar 4.7, dan Gambar 4.8 yang membuktikan bahwa tren dari sebuah istilah pencarian bisa menggambarkan kejadian kasus penyakit di lapangan (lingkungan masyarakat) meski tidak seratus persen faktual. Tentunya interpretasi causal yang peneliti gunakan ialah secara matematika atau statistik. Setiap pola tren atau tren yang dihasilkan GT, membutuhkan variabel pendukung lain untuk menguatkan interpretasi tersebut.

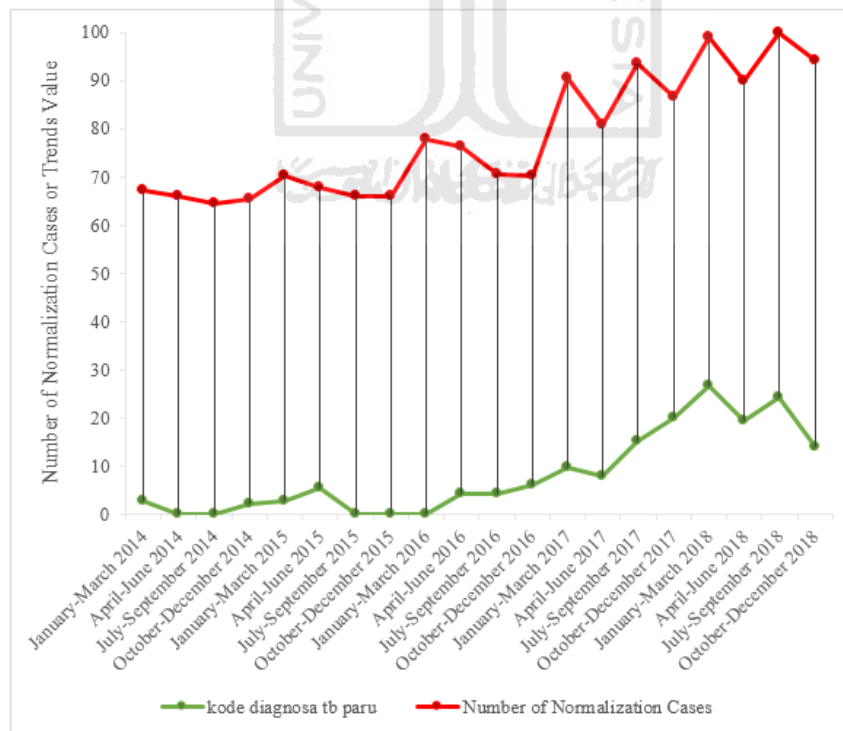
Perhitungan saat ini masih di level negara (kelompok utama), peneliti tidak menutup kemungkinan bahwa jika perhitungan di sub-kelompok akan menghasilkan tingkat korelasi yang berbeda. Seperti penelitian (Kievit, Frankenhuis, Waldorp, & Borsboom, 2013) menyebutkan bahwa pemberian tingkat obat yang tinggi dapat memberikan kesembuhan yang tinggi juga. Hal itu dimungkinkan berlaku di tingkat umum atau kelompok utama. Namun kenyataannya di sub-kelompok seperti pria dan wanita memiliki perbedaan hasil. Fenomena ini dikenal dengan sebutan Simpson’s Paradox (Selvitella, 2017). Sama halnya dengan penelitian ini, beberapa istilah pencarian yang berkorelasi tinggi di level negara, dimungkinkan akan berkorelasi rendah atau tetap tinggi di level kota/perdesaan. Hal itu dikarenakan asumsi bahwa masyarakat perkotaan lebih cenderung banyak penggunaan layanan internet dibanding masyarakat perdesaan.

Variabel pengganggu tren GT juga tidak peneliti abaikan, karena variabel tersebut bisa memiliki pengaruh terhadap hasil tren GT. Hal sama biasa terjadi di berbagai penelitian yang memiliki variabel pengganggu sehingga muncul Paradoks Simpson (Ameringer, Serlin, & Ward, 2009). Misalnya, populasi kota yang besar, bahasa yang didukung, perilaku masyarakat dalam menggunakan internet, dan latar pendidikan masyarakat (Husnayain et al., 2019; Kang et al., 2013). Pemberitaan media menjadi variabel pengganggu terbesar bagi tren GT sehingga membuat tren penyakit tinggi di daerah yang memberitakan penyakit tersebut dibanding daerah yang sedang terjangkit penyakit tersebut namun sedikit pemberitaan media masa (Carneiro & Mylonakis, 2009a; Kang et al., 2013). Peneliti

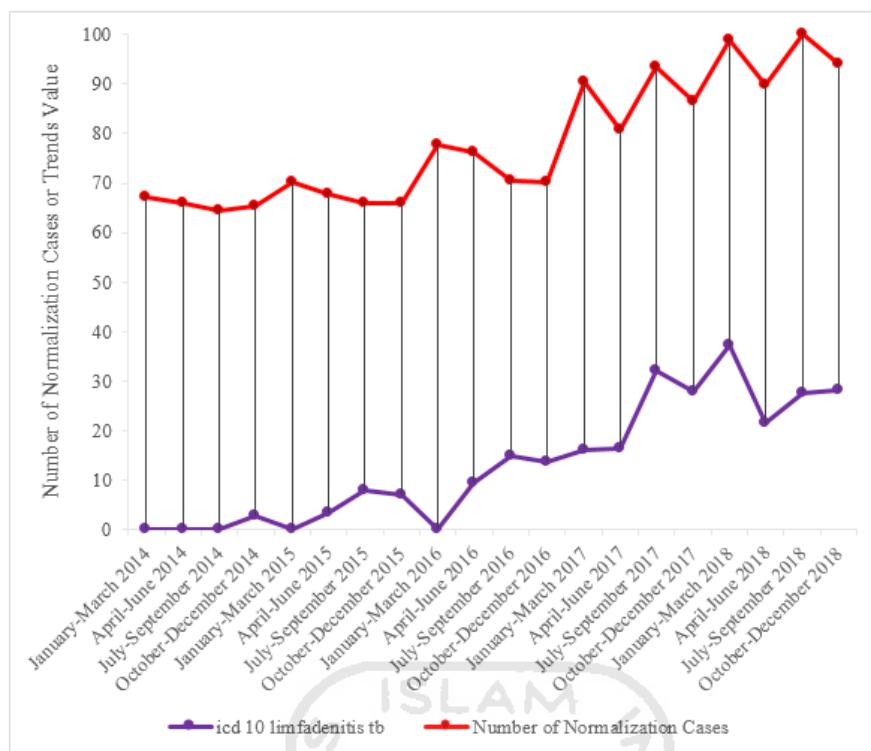
mencoba menghindari interpretasi berlebihan terhadap korelasi yang didapat, sehingga meminimalisir muncul paradoks pada hasil penelitian.



Gambar 4.6 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'obat tb' (level negara).



Gambar 4.7 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'kode diagnosa tb paru' (level negara).



Gambar 4.8 Perbandingan pola grafik antara kasus penyakit Tuberkulosis dengan tren istilah pencarian 'icd 10 limfadenitis tb' (level negara).

Apabila dievaluasi terhadap istilah 'sakit' dari segi kondisi yang digambarkannya, maka kemungkinan pengguna yang menggunakan istilah ini sedang dalam kondisi sakit. Sedangkan istilah 'gejala' yang kemungkinan besar pengguna sedang merasa sakit atau terjadi tanda-tanda aneh pada anggota tubuh pengguna. Misalnya, istilah 'gejala liver' maka dimungkinkan pengguna sedang ada kendali atau sedikit rasa sakit pada bagian organ tubuhnya terutama hati sehingga dia mencari informasi terkait gejala penyakit yang berhubungan dengan hati dan dalam hal ini ialah liver. **Hal ini sejalan dengan penemuan penelitian (Husnayain et al., 2019) yang juga mendapati istilah 'gejala' pada istilah pencarian 'gejala demam berdarah' memiliki nilai korelasi tertinggi yaitu sebesar 0,937.**

Akhirnya, seperti pada Tabel 4.1 penelitian ini menghasilkan suatu *template* istilah pencarian umum (sementara) yang bisa digunakan untuk melihat tren pada GT, dengan tujuan untuk pengawasan kasus penyakit menular di Indonesia. Tingkat korelasi terendah pada istilah pencarian ialah sebesar 0,80776689 yaitu istilah 'tb adalah'. Sedangkan tingkat korelasi tertinggi dimiliki istilah pencarian 'obat tb' sebesar 0,906818458 yang berasal dari kueri terkait pada GT. Penelitian selanjutnya memiliki tantangan, bagaimana membuat atau memunculkan *template* istilah pencarian untuk kategori Epidemiologi dan Faktor Risiko.

Selain itu juga mungkin menemukan template istilah pencarian yang lain, baik untuk kategori Definisi, Diagnosis, maupun Tata Laksana.

Tabel 4.1 Template istilah kombinasi untuk domain penyakit menular.

<i>Number</i>	<i>General Category</i>	<i>Source Term</i>	<i>Combinate Term</i>
1	Definisi	Buku KSK, Related Query on GT	'sakit'
2		Related Query on GT	'adalah'
3	Diagnosis	Related Query on GT	'ciri ciri'
4		Related Query on GT	'ciri'
5		Buku KSK, Related Query on GT	'gejala'
6		Related Query on GT	'kode icd'
7		Related Query on GT	'kode diagnosa'
8	Epidemiologi	-	-
9	Faktor Risiko	-	-
10	Tata Laksana	Related Query on GT	'obat'

#### 4.2.2 Uji Konsistensi *Template* Istilah Pencarian Usulan

Penelitian ini mencoba menggali lebih dalam tentang GT sebagai media pengawasan penyakit dimasa depan. *Template* istilah pencarian yang diusulkan sebelumnya akan diuji kembali. Faktor-faktor yang mempengaruhi hasil tren dari *template* istilah pencarian, perlu untuk dicek lagi. Selain itu, penggunaan *template* istilah pencarian juga akan ditinjau dari segi waktu. Hal itu semua dilakukan agar benar-benar dimasa depan akan terwujud pengawasan penyakit melalui GT di Indonesia. Tentunya, penelitian tidak menutup kemungkinan akan membahas bagian lain yang bisa diteliti dari GT untuk kasus pengawasan penyakit. Uji konsistensi yang dimaksud pada penelitian ini yaitu pengecekan terhadap *template* istilah pencarian (istilah kombinasi). Istilah kombinasi dicek ketika digabungkan dengan istilah utama, apakah selalu konsisten menghasilkan korelasi positif tinggi walau posisi penggabungan diubah. Rumus istilah pencairan awal yaitu rumus (3.1), dimodifikasi menjadi rumus (4.1). Sehingga ada dua rumus atau dua pola posisi kombinasi istilah dan ini digunakan dalam uji konsistensi.

$$\text{Istilah Pencarian} = \text{istilah kombinasi} + \text{istilah utama} \quad (4.1)$$

Uji konsistensi menggunakan dua istilah utama yaitu 'tb' dan 'tbc'. Hal itu dikarenakan keduanya merupakan istilah utama yang terkandung dalam istilah pencairan dengan koefisien korelasi positif tinggi. Sedangkan istilah kombinasi diambil dari Tabel 4.1, sehingga ada 32 istilah pencarian (2 istilah utama \* 8 istilah kombinasi \* 2 kondisi gabungan) yang diuji. **Tolak ukur bagi *template* (istilah kombinasi) dinyatakan konsisten yaitu**



ketika *template* digunakan saat kedua posisi penggabungan mendapatkan koefisien korelasi diatas sama dengan ( $\geq$ ) 0,8. Apabila tolak ukur tersebut tidak terpenuhi, maka *template* dinyatakan tidak konsisten.

Hasil uji konsistensi mendapati bahwa hanya ada dua *template* yang 100% konsisten memiliki koefisien korelasi diatas sama dengan ( $\geq$ ) 0,8. Dua *template* tersebut yaitu istilah kombinasi ‘adalah’ dan ‘ciri ciri’. Pada Lampiran C.12, dapat diketahui juga bahwa lima istilah kombinasi hanya konsisten jika digabungkan dengan istilah utama ‘tb’. Lima istilah tersebut yaitu ‘ciri’, ‘gejala’, ‘kode icd’, ‘kode diagnosa’, dan ‘obat’. Hal ini membuat kelima istilah hanya bisa digunakan untuk kasus penyakit Tuberkulosis dengan syarat penggunaan istilah utama yaitu ‘tb’ saat pengawasan penyakit dengan GT. Berdasarkan 32 istilah pencarian pada uji konsistensi ini (secara kuantitas dan kualitas), istilah utama seimbang jika diletakkan di depan maupun di belakang istilah pencarian. Selain itu, istilah utama ‘tbc’ terlihat jelas memiliki tingkat penggunaan yang buruk baik itu diletakkan diawal maupun di akhir istilah pencarian. Akhirnya pada Tabel 4.2, **peneliti menemukan dua istilah kombinasi** yang berkemungkinan besar bisa **digunakan untuk umum (segala kasus penyakit)**.

Tabel 4.2 Praktik terbaik *template* istilah pencarian untuk domain penyakit umum.

No	Kategori Umum	Istilah Kombinasi	Cara Penggunaan
1.	Definisi	‘adalah’	‘istilah kombinasi’ + penyakit atau
2.	Diagnosis	‘ciri ciri’	penyakit + ‘istilah kombinasi’

*Template* istilah pencarian untuk umum dapat digunakan dengan cara menambahkan istilah dari nama penyakit yang sedang diawasi. Misalnya, jika ingin mengawasi penyakit Diabetes maka istilah pencarian menjadi ‘diabetes adalah’, ‘adalah diabetes’, ‘ciri ciri diabetes’ ataupun ‘diabetes ciri ciri’. Sedangkan untuk istilah pencarian pada Tabel 4.3 bisa digunakan langsung tanpa menambah apapun dan hanya digunakan untuk mengawasi pengakit Tuberkulosis saja.

Tabel 4.3 Praktik terbaik *template* istilah pencarian untuk domain penyakit Tuberkulosis.

Number	General Category	Search Term	How to Use
1	Definisi	sakit tbc	Best practice search term untuk pengawasan penyakit
2	Diagnosis	gejala tb	
3	Diagnosis	ciri ciri tbc	
4	Diagnosis	ciri tb	
5	Diagnosis	icd 10 limfadenitis tb	

<i>Number</i>	<i>General Category</i>	<i>Search Term</i>	<i>How to Use</i>
6	Diagnosis	icd 10 tb	Tuberkulosis di Indonesia
7	Diagnosis	icd 10 tb kelenjar	
8	Diagnosis	icd tb paru	
9	Definisi	kelenjar tb	
10	Diagnosis	kode diagnosa tb paru	
11	Diagnosis	kode icd 10 tb	
12	Diagnosis	kode icd tb paru	
13	Tata Laksana	obat tb	
14	Tata Laksana	pro tb 4	
15	Definisi	tb adalah	
16	Definisi	tb mdr	
17	Definisi	tbc adalah	
18	Tata Laksana	tcm tb	
19	Tata Laksana	toss tb	
20	Definisi	adalah tb	
21	Diagnosis	ciri ciri tb	
22	Diagnosis	kode icd tb	
23	Diagnosis	kode diagnosa tb	
24	Definisi	adalah tbc	
25	Diagnosis	tb ciri ciri	
26	Diagnosis	tb ciri	
27	Diagnosis	tb gejala	
28	Diagnosis	tb kode icd	
29	Diagnosis	tb kode diagnosa	
30	Tata Laksana	tb obat	
31	Diagnosis	tbc ciri ciri	
32	Diagnosis	tbc ciri	

#### 4.2.3 Faktor yang Mempengaruhi Istilah Pecarian

Peneliti melihat **faktor unsur dari sebuah istilah pencarian** menjadi paling utama berefek pada tren GT. Penggunaan istilah ‘tuberkulosis’ dengan ‘tb’ memiliki jarak perbedaan tren yang sangat jauh. Hal itu bisa dilihat dari hasil penelitian bahwa istilah pencarian yang memiliki unsur kata ‘tuberkulosis’ tidak ada yang memiliki koefisien korelasi diatas sama dengan ( $\geq$ ) 0,8. Selain itu, koefisien korelasi tertinggi yang dimiliki istilah pencarian dengan unsur kata ‘tuberkulosis adalah sebesar 0,70514106 (istilah pencarian ‘ciri ciri tuberkulosis’). Sedangkan koefisien korelasi terendah-nya yaitu sebesar -0,741220611 (istilah pencarian ‘tuberkulosis paru’).

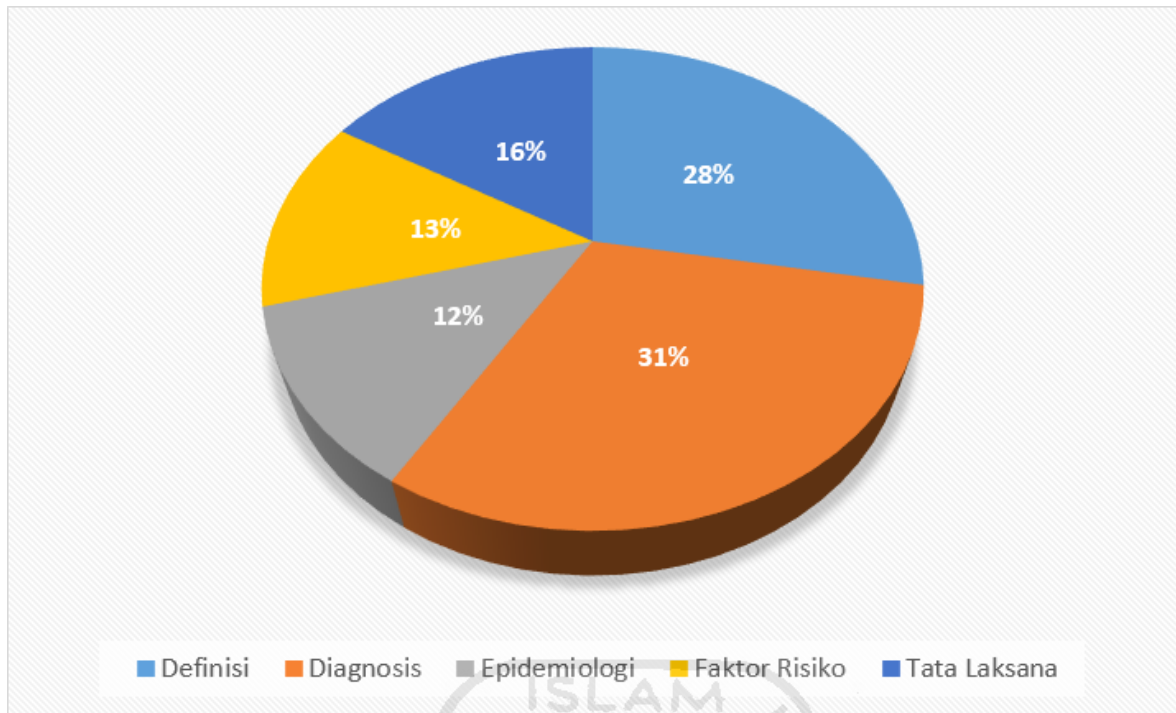
**Faktor bahasa istilah** yang digunakan juga memiliki pengaruh. Seperti, istilah ‘*tuberculosis*’ mendapatkan koefisien korelasi sebesar -0.250105084. Hal ini menandakan penyesuaian bahasa istilah perlu dilakukan. Penelitian ini telah membatasi jenis bahasa isitilah yang digunakan yaitu Bahasa Indonesia. Namun, istilah pencarian ‘*tuberculosis*’

dibahas sebagai contoh untuk memaparkan pengaruh dari segi bahasa. Tentunya, semua istilah pencarian yang diteliti untuk menjadi *template* akhir berbahasa Indonesia.

**Faktor jumlah kata** dalam sebuah istilah pencarian memiliki pengaruh yang sedang atau tidak terlalu berpengaruh. Ada lima istilah pencarian yang terdiri dari satu suku kata (tunggal) yaitu a) 'tb' = 0,400320344; b) 'tbc' = 0,719201143; c) 'tuberkulosis' = -0,186542877; d) 'tuberkolosis' = -0,451810629; dan e) '*tuberculosis*' = -0,250105084. Ketika kelima istilah tunggal ini disandingkan dengan istilah lain maka hanya sedikit yang menghasilkan koefisien korelasi tinggi positif. Berdasarkan penelitian ini, kata tunggal 'tb' dan 'tbc' yang mampu menghasilkan koefisien korelasi tinggi positif yaitu diatas sama dengan ( $\geq$ ) 0,8. Kata tunggal 'tb' lebih dominan menghasilkan koefisien korelasi tinggi dibandingkan dengan kata tunggal 'tbc'.

#### **4.2.4 Penafsiran Terhadap Kategori Utama (*General Category*)**

Ada 215 istilah pencarian unik yang diteliti didalam penelitian ini (selain istilah pencarian yang ditolak karena tidak berhubungan dengan Tuberkulosis). Semua istilah pencarian sudah dikelompokkan kedalam kategori utama, secara terurut yaitu : a) kategori Definisi sebanyak 60 istilah pencarian; b) kategori Diagnosis sebanyak 66 istilah pencarian; c) kategori Epidemiologi sebanyak 26 istilah pencarian; d) kategori Faktor Risiko sebanyak 28 istilah pencarian; dan e) kategori Tata Laksana sebanyak 35 istilah pencarian. Perbandingan persentasi jumlah istilah pencarian dapat dilihat pada Gambar 4.9. Berdasarkan data tersebut, maka di Indonesia dapat disimpulkan bahwa masyarakat lebih membutuhkan informasi tentang cara mendiagnosa suatu penyakit atau mengenali gejala-gejalanya dibanding informasi tentang definisi maupun obat-obatan. Hal tersebut menjadi peluang bagi para akademis maupun lembaga terkait untuk menyediakan informasi lebih banyak tentang mendiagnosa penyakit pada laman *website* atau internet. Informasi dapat disampaikan melalui laman resmi lembaga ataupun dalam bentuk karya ilmiah.



Gambar 4.9 Perbandingan jumlah istilah pencarian berdasarkan Kategori Utama.

Saat melakukan pengawasan, jika suatu tren naik dari istilah pencarian tertentu maka perlu dilihat juga dari sisi kategori umum istilah pencarian tersebut. Misalnya, istilah pencarian ‘ciri ciri tbc’ memiliki kategori Diagnosis sehingga dugaan yang muncul ketika tren istilah pencarian ini naik adalah sedang terjadi kasus Tuberkulosis. Asumsi yang muncul terhadap masyarakat daerah tren yaitu ada beberapa orang yang sedang terkena penyakit Tuberkulosis dan masyarakat lainnya menjadi waspada hingga mencari informasi tentang penyakit tersebut (salah satunya melalui pencarian di internet). Bagi lembaga terkait, bisa memanfaatkan dugaan awal ini untuk cepat kelapangan dan sekaligus melakukan sosialisasi pencegahan. Pemberitaan media masa juga bisa mendukung untuk klarifikasi atas tren dari GT. Selain itu, pemberitaan media masa juga mempengaruhi turun naiknya tren pada GT. Sehingga ketika tren naik, maka pemberitaan media masa akan mulai ramai tentang penyakit tersebut.

#### 4.2.5 Waktu yang Bagus untuk Pengawasan Penyakit dengan GT

Peneliti ingin mengetahui mencoloknya suatu kasus penyakit terjadi pada suatu waktu tertentu. Berdasarkan data kasus dari P2P KEMKES RI, seperti Lampiran C.13 tidak ada waktu tertentu yang mencolok. Hal tersebut dikarenakan setiap triwulan disetiap tahun selalu mengalami kenaikan tanpa penurunan.

## BAB 5

### Kesimpulan dan Saran

#### 5.1 Kesimpulan

*Google Trends* terbukti mampu mewakili keadaan sebenarnya di suatu wilayah tertentu. Hal itu tentunya secara statistik korelasi. Sensitifitas dari sebuah istilah pencarian perlu disorot dalam pemanfaatan *Google Trends*. Pengujian istilah pencairan harus dilakukan sebelum istilah pencarian digunakan untuk pengawasan dimasa sekarang atau masa depan. Ketika istilah pencairan sudah minimal terhadap *error*/bias maka keakuratan pengawasan semakin tinggi.

Penelitian ini menghasilkan tiga *template* istilah pencarian yang tepat dalam melakukan pengawasan penyakit melalui GT di Indonesia. *Template* pertama dikhususkan untuk penyakit Tuberkulosis, ada 32 istilah pencarian yang siap digunakan langsung untuk pengawasan. *Template* kedua dikhususkan untuk penyakit menular saja, ada tujuh istilah kombinasi (*Combinate Term*) yang siap digunakan dengan aturan penggunaannya harus di kombinasi dengan istilah utama (*main term*). Sedangkan *template* ketiga bisa digunakan untuk semua jenis pengakit dan harus mengikuti aturan penggunaan istilah pencairan dari penelitian ini.

Peneliti melihat potensi besar GT untuk pengawasan penyakit Tuberklosis di Indonesia. Hal itu melihat dari tingkat korelasi mencapai 0,94 untuk istilah pencarian ‘toss tb’ dan sebesar 0.93 untuk istilah pencarian ‘tb obat’. Tentunya tren GT membutuhkan variabel atau komponen lainnya untuk menjadi bahan dalam penentuan kebijakan saat pengawasan penyebaran penyakit. Tren GT *second opinion* untuk pengawasan penyebaran penyakit dan *first opinion* sebaiknya mengambil dari variabel lain yang bukan hasil perhitungan statistik.

Pengawasan tidak hanya sebatas mengetahui tinggi atau rendah kasus penyakit di suatu daerah. Namun, prilaku masyarakat juga dapat dilihat dari kemunculan istilah pencarian yang sering muncul. Gambaran prilaku masyarakat bisa menjadi opini dalam pembuatan kebijakan atau bantuan oleh pihak terkait. Faktor waktu/musim suatu penyakit juga memberi pengaruh pada tren *Google Trends*.

## 5.2 Saran

Peneliti berkeyakinan bahwa *Google Trends* memiliki peluang besar untuk membantu lembaga kesehatan di suatu negara. Tentunya, kedepan diharapkan terciptanya sistem berupa aplikasi perangkat lunak yang bisa mengawasi penyebaran penyakit di suatu negara. Pengujian terhadap penyakit lain (selain Tuberkulosis), diharapkan membuat kevalidatan analisis penelitian ini semakin lengkap. Kolaborasi dengan pihak yang kompeten terhadap cuaca atau iklim suatu negara dan pakar kesehatan, diharapkan mampu menemukan waktu pengawasan untuk setiap penyakit dengan kevalidatan yang tinggi. Sedangkan ketidakmeratanya akses internet membuat ketidak seimbangan dalam tren antar daerah. Daerah yang memiliki kasus tapi tidak memiliki akses internet, berkemungkinan besar tidak terdeteksi oleh tren *Google Trends*.



## Daftar Pustaka

- Abiyyu, J. P., Andreswari, R., & Hasibuan, M. A. (2018). Implementasi Aplikasi Mobile Modul Penyelenggara Dan Konsumen Kegiatan Di Kota Bandung Menggunakan Metode Iterative Incremental Untuk Meningkatkan Minat Terhadap Kegiatan Di Kota Bandung. *E-Proceeding of Engineering*, 5(1), 1381–1391.
- Ameringer, S., Serlin, R. C., & Ward, S. (2009). Simpson's Paradox and Experimental Research Suzanne. *Nursing Research*, 58(2), 123–127.  
<https://doi.org/10.1097/NNR.0b013e318199b517.Simpson>
- APJII. (2019). Mengawali Integritas Era Digital 2019. In *Buletin APJII*. Retrieved from <https://apjii.or.id/downfile/file/BULETINAPJIIEDISI33Januari2019.pdf>
- Arora, V. S., McKee, M., & Stuckler, D. (2019). Google Trends: Opportunities and limitations in health and health policy research. *Health Policy*, 123(3), 338–341.  
<https://doi.org/10.1016/j.healthpol.2019.01.001>
- Carneiro, H. A., & Mylonakis, E. (2009a). Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases*, 49(10), 1557–1564. <https://doi.org/10.1086/630200>
- Carneiro, H. A., & Mylonakis, E. (2009b). Google Trends: A Web-Based Tool for Real-Time Surveillance of Disease Outbreaks. *Clinical Infectious Diseases*, 49(10), 1557–1564. <https://doi.org/10.1086/630200>
- Cho, S., Sohn, C. H., Jo, M. W., Shin, S. Y., Lee, J. H., Ryoo, S. M., ... Seo, D. W. (2013). Correlation between national influenza surveillance data and Google Trends in South Korea. *PLoS ONE*, 8(12). <https://doi.org/10.1371/journal.pone.0081422>
- Choi, H., & Varian, H. (2012a). Predicting the Present with Google Trends. *Economic Record*, 88(SUPPL.1), 2–9. <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
- Choi, H., & Varian, H. (2012b). Predicting the Present with Google Trends. *Economic Record*, 88(SUPPL.1), 2–9. <https://doi.org/10.1111/j.1475-4932.2012.00809.x>
- Dilmaghani, M. (2019). Workopolis or The Pirate Bay: what does Google Trends say about the unemployment rate? *Journal of Economic Studies*, 46(2), 422–445.  
<https://doi.org/10.1108/JES-11-2017-0346>
- EEAS, D. of the E. U. to I. and B. D. (2017). Uni Eropa berikan Beasiswa Erasmus Plus kepada 225 Mahasiswa Indonesia untuk menempuh pendidikan di Eropa. Retrieved

- February 19, 2020, from Delegation of the European Union to Indonesia and Brunei Darussalam website: [https://eeas.europa.eu/delegations/indonesia\\_hi/31212/Uni-Eropa-berikan-Beasiswa-Erasmus-Plus-kepada-225-Mahasiswa-Indonesia-untuk-menempuh-pendidikan-di-Eropa](https://eeas.europa.eu/delegations/indonesia_hi/31212/Uni-Eropa-berikan-Beasiswa-Erasmus-Plus-kepada-225-Mahasiswa-Indonesia-untuk-menempuh-pendidikan-di-Eropa)
- Emerson, R. W. (2015). Causation and Pearson's correlation coefficient. *Journal of Visual Impairment and Blindness*, 109(3), 242–244.  
<https://doi.org/10.1177/0145482x1510900311>
- Falgenti, K. (2011). Transformasi UKM ke Bisnis Online dengan Internet Marketing Tools. *Ilmiah Faktor Exacta*, 4(1), 62–73.
- Husnayain, A., Fuad, A., & Lazuardi, L. (2019). Correlation between Google Trends on dengue fever and national surveillance report in Indonesia. *Global Health Action*, 12(1). <https://doi.org/10.1080/16549716.2018.1552652>
- Ihsanuddin. (2020). Jokowi: Indonesia Peringkat Ketiga Penderita TBC Tertinggi di Dunia. Retrieved November 11, 2020, from Kompas website: <https://nasional.kompas.com/read/2020/07/21/12031081/jokowi-indonesia-peringkat-ketiga-penderita-tbc-tertinggi-di-dunia>
- Jun, S. P., Yoo, H. S., & Choi, S. (2018). Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological Forecasting and Social Change*, 130(November 2017), 69–87.  
<https://doi.org/10.1016/j.techfore.2017.11.009>
- Kang, M., Zhong, H., He, J., Rutherford, S., & Yang, F. (2013). Using Google Trends for Influenza Surveillance in South China. *PLoS ONE*, 8(1), 2009–2014.  
<https://doi.org/10.1371/journal.pone.0055205>
- Kievit, R. A., Frankenhuys, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, 4(AUG), 1–14. <https://doi.org/10.3389/fpsyg.2013.00513>
- KKBI. (2016a). de.fi.ni.si. Retrieved February 28, 2020, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website: <https://kbbi.kemdikbud.go.id/entri/Definisi>
- KKBI. (2016b). ge.ja.la. Retrieved February 28, 2020, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website: <https://kbbi.kemdikbud.go.id/entri/Gejala>
- KKBI. (2016c). kata kunci. Retrieved December 6, 2019, from KBBI Daring Kementerian



- Pendidikan dan Kebudayaan Republik Indonesia website:  
[https://kbbi.kemdikbud.go.id/entri/Kata kunci](https://kbbi.kemdikbud.go.id/entri/Kata%20kunci)
- KKBI. (2016d). ko.re.la.si. Retrieved December 6, 2019, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website:  
<https://kbbi.kemdikbud.go.id/entri/Korelasi>
- KKBI. (2016e). pe.net.ra.si. Retrieved December 6, 2019, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website:  
<https://kbbi.kemdikbud.go.id/entri/Penetrasi>
- KKBI. (2016f). peng.o.bat.an. Retrieved February 28, 2020, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website:  
<https://kbbi.kemdikbud.go.id/entri/Pengobatan>
- KKBI. (2016g). vek.tor. Retrieved February 28, 2020, from KBBI Daring Kementerian Pendidikan dan Kebudayaan Republik Indonesia website:  
<https://kbbi.kemdikbud.go.id/entri/Vektor>
- Kristoufek, L. (2013). Can google trends search queries contribute to risk diversification? *Scientific Reports*, 3, 1–5. <https://doi.org/10.1038/srep02713>
- Kusnandar, V. B. (2019). WHO: Kasus TBC Indonesia 2017 Terbesar Ketiga Dunia. Retrieved November 11, 2020, from Databoks website:  
<https://databoks.katadata.co.id/datapublish/2019/10/09/who-kasus-tbc-indonesia-2017-terbesar-ketiga-dunia>
- Mirabel, M. E., Palit, H. N., & Handojo, A. (2017). Analisis Korelasi Pada Data Yahoo ! Properties dan Instant Messaging dengan Menggunakan Hadoop. *Jurnal Infra*, 5(1). Retrieved from <http://publication.petra.ac.id/index.php/teknik-informatika/article/view/5154/4744>
- Nijman, V. (2015). CITES-listings, EU eel trade bans and the increase of export of tropical eels out of Indonesia. *Marine Policy*, 58, 36–41.  
<https://doi.org/10.1016/j.marpol.2015.04.006>
- Nuruddin, M., & Udin, A. R. A. (2017). Penerapan Internet Sehat Dan Produktif ( Insap ) Bagi Kelompok Remaja Di Lingkungan Sumber Ketangi Kelurahan Wirolegi Kecamatan Sumbersari Kabupaten Jember. *Seminal Nasional Hasil Pengabdian Kepada Masyarakat*, 247–250.
- Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., & Murugiah, K. (2014a). The use of google trends in health care research: A systematic review.

- PLoS ONE*, 9(10). <https://doi.org/10.1371/journal.pone.0109583>
- Nuti, S. V., Wayda, B., Ranasinghe, I., Wang, S., Dreyer, R. P., Chen, S. I., & Murugiah, K. (2014b). The Use of Google Trends in Health Care Research: A Systematic Review. *PLoS ONE*, 9(10), 1–49. <https://doi.org/10.1371/journal.pone.0109583>
- Olson, D. R., Baer, A., Coletta, M. A., Deyneka, L., Gentry, R., Ising, A., ... Mostashari, F. (2009). Searching for better flu surveillance? A brief communication arising from Ginsberg et al. *Nature* 457, 1012-1014 (2009). *Nature Precedings*, 1014(August), 1012–1014. <https://doi.org/10.1038/npre.2009.3493>
- Orenzi, B. (2019). Statistik Pengguna Digital Dan Internet Indonesia 2019. Retrieved December 3, 2019, from BOC Indonesia website: <http://www.boc.web.id/statistik-pengguna-digital-dan-internet-indonesia-2019/>
- Pelat, C., Turbelin, C., Bar-Hen, A., Flahault, A., & Valleron, A.-J. (2009). More Diseases Tracked by Using Google Trends. *Clinical Infectious Diseases*, 15(8), 1327–1328. <https://doi.org/10.3201/eid1508.090299>
- Pranita, E. (2020). Indonesia Disebut Endemis TBC, Jumlah Kasus Tertinggi Ketiga di Dunia. Retrieved November 11, 2020, from Kompas website: <https://www.kompas.com/sains/read/2020/03/23/120300923/indonesia-disebut-endemis-tbc-jumlah-kasus-tertinggi-ketiga-di-dunia?page=all>
- Preis, T., Moat, H. S., & Eugene Stanley, H. (2013). Quantifying trading behavior in financial markets using google trends. *Scientific Reports*, 3, 1–6. <https://doi.org/10.1038/srep01684>
- Rahmanto, D., Wiyadi, & Isa, M. (2011). Analisis Permintaan Pasar Online Produk Batik Di Indonesia (Universitas Muhammadiyah Surakarta). Retrieved from <http://eprints.ums.ac.id/60940/>
- Riyanto, A. D. (2014). Pemanfaatan Google Trends Dalam Penentuan Kata Kunci Sebuah Produk Untuk Meningkatkan Daya Saing Pelaku Bisnis Di Dunia Internet. *Seminar Nasional Informatika*, 52–59.
- Rokom. (2017). Tekan Angka Kematian Melalui Program Indonesia Sehat dengan Pendekatan Keluarga. Retrieved November 11, 2020, from Kementerian Kesehatan RI website: <http://sehatnegeriku.kemkes.go.id/baca/rilis-media/20170614/2921397/tekan-angka-kematian-melalui-program-indonesia-sehat-pendekatan-keluarga/>
- Satibi, A. F., Suharyono, & Abdillah, Y. (2017). Analisis Pemanfaatan Search Engine

- Optimization Dalam Meningkatkan Penjualan Produk UKM Di Pasar Internasional (Studi Kasus Pada CV . Ayung Sportindo). *Jurnal Administrasi Bisnis (JAB)*, 50(6), 96–105.
- Seifter, A., Schwarzwald, A., Geis, K., & Aucott, J. (2010). The utility of “Google Trends” for epidemiological research: Lyme disease as an example. *Geospatial Health*, 4(2), 135–137. <https://doi.org/10.4081/gh.2010.195>
- Selvitella, A. (2017). The ubiquity of the Simpson’s Paradox. *Journal of Statistical Distributions and Applications*, 4(1). <https://doi.org/10.1186/s40488-017-0056-5>
- Siswanto, M., & Fahriannur, A. (2016). Google Trend untuk Analisa Pasar Bisnis Online & Pemilihan Keywords pada E-Commerce Web. *Seminar Hasil Penelitian Dan Pengabdian Pada Masyarakat Dana BOPTN*, 272–277.
- Sugiyono. (2007). *Statistika untuk Penelitian*. Bandung: Alfabeta.
- Suharyadi, I. (2016). Peran Penting Asia Africa Smart City Summit (AASCS) 2015 Terhadap Perkembangan Paradiplomasi Kota Bandung. *GLOBAL*, 18(1), 95–107.
- Sulistiyanto, M. P. T., & Nugraha, D. A. (2015). Implementasi IoT ( Internet of Things ) dalam pembelajaran di Universitas Kanjuruhan Malang. *SMARTICS Journal*, 1(1), 20–23.
- Sulistiyanto, M. P. T., Nugraha, D. A., Sari, N., Karima, N., & Asrori, W. (2015). Implementasi IoT (Internet of Things) dalam pembelajaran di Universitas Kanjuruhan Malang. *SMARTICS Journal*, 1(1), 20–23. Retrieved from <http://ejournal.unikama.ac.id/index.php/jst/article/view/842>
- Susanto, C. O. N. (2016). *Internet Sehat*. Retrieved from <http://repository.umy.ac.id/bitstream/handle/123456789/1820/Internet>
- Swari, Laila R., & Lakoro, R. (2016). Perancangan Video Promosi “ Jelajah Pantai Tulungagung" untuk Menunjang Potensi Wisata Pantai di Kabupaten Tulungagung. *Sains Dan Seni ITS*, 5(2).
- Swari, Laila Ratna, & Lakoro, R. (2016). Perancangan Video Promosi “Jelajah Pantai Tulungagung” untuk Menunjang Potensi Wisata Pantai di Kabupaten Tulungagung. *Jurnal Sains Dan Seni ITS*, 5(2). <https://doi.org/10.12962/j23373520.v5i2.19934>
- Tamba, A. (2018, March 23). Protokol KLB Perlu Jadi Acuan Penanganan Tuberkulosis. *Harian Nasional*. Retrieved from <http://harnas.co/2018/03/22/protokol-klb-perlu-jadi-acuan-penanganan-tuberkulosis>
- Teng, W., Cheng, L. P., & Zhao, K. J. (2017). Application of kernel principal component

and Pearson correlation coefficient in prediction of mine pressure failure. *Proceedings - 2017 Chinese Automation Congress, CAC 2017, 2017-Janua*, 5704–5708.

<https://doi.org/10.1109/CAC.2017.8243801>

Thomas, M. J., Yoon, P. W., Collins, J. M., Davidson, A. J., & Mac Kenzie, W. R. (2018). Evaluation of Syndromic Surveillance Systems in 6 US State and Local Health Departments. *Journal of Public Health Management and Practice*, 24(3), 235–240.

<https://doi.org/10.1097/PHH.0000000000000679>

Trends, G. (2019). Trends 2019. Retrieved December 6, 2019, from Google Trends website:

[https://support.google.com/trends/answer/4365533?hl=id&ref\\_topic=6248052](https://support.google.com/trends/answer/4365533?hl=id&ref_topic=6248052)

Vosen, S., & Schmidt, T. (2011). Forecasting private consumption: Survey-based indicators vs. Google trends. *Journal of Forecasting*, 30(6), 565–578.

<https://doi.org/10.1002/for.1213>

WHO. (2015). *Indonesia: WHO statistical profile Basic*. Retrieved from

[http://www.who.int/Gho/Countries/Idn/Country\\_Profiles/En/](http://www.who.int/Gho/Countries/Idn/Country_Profiles/En/)



# LAMPIRAN A

## 1. Penetrasi internet di Indonesia



Sumber : (Orenzi, 2019).

## 2. Website paling banyak dikunjungi masyarakat Indonesia

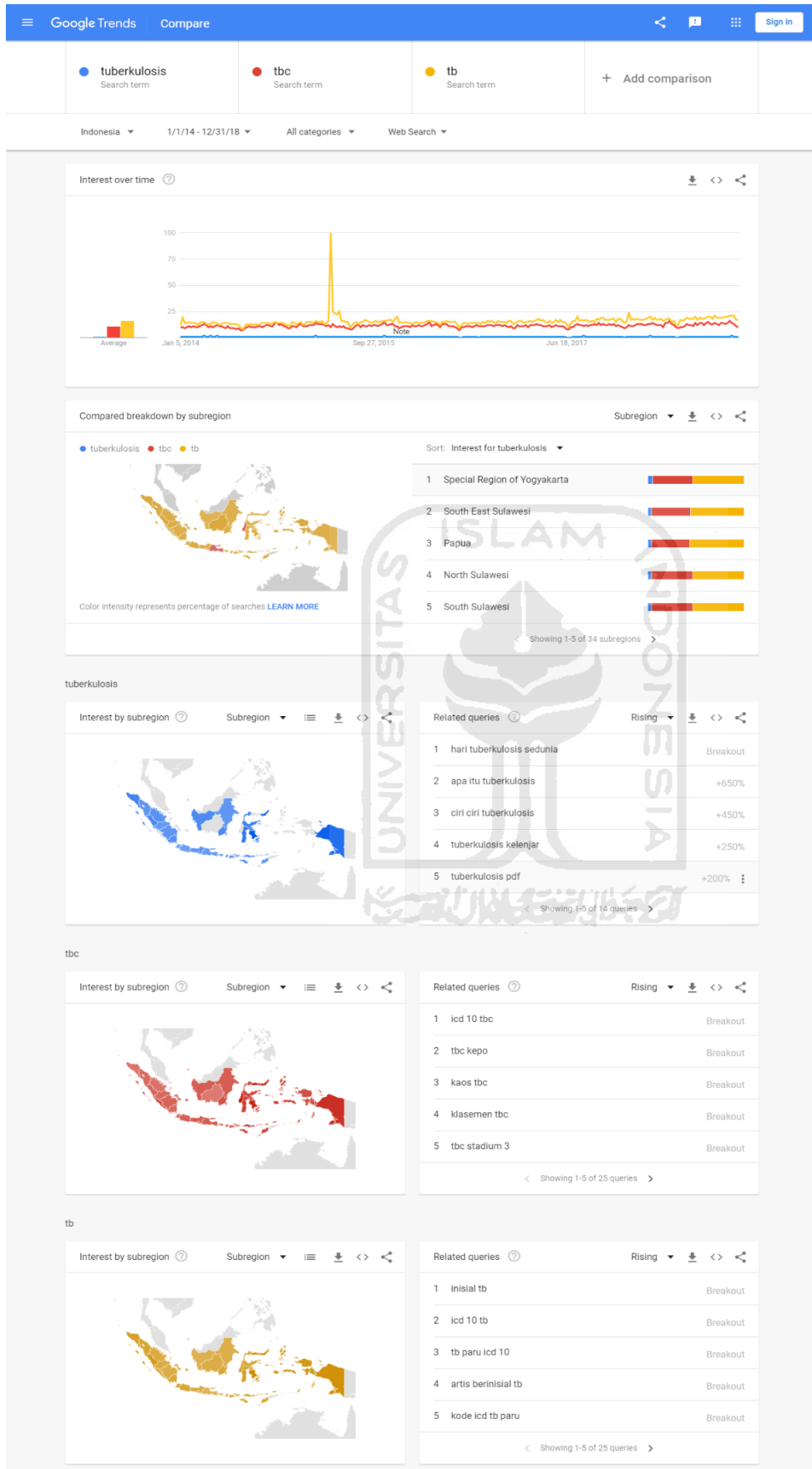
**JAN 2019** **SIMILARWEB'S TOP WEBSITES**  
RANKING OF WEBSITES BY AVERAGE MONTHLY TRAFFIC

#	WEBSITE	CATEGORY	MONTHLY TRAFFIC	TIME PER VISIT	PAGES PER VISIT
01	GOOGLE.COM	SEARCH	1,028,900,000	08M 06S	7.7
02	GOOGLE.CO.ID	SEARCH	702,300,000	07M 16S	6.6
03	FACEBOOK.COM	SOCIAL	605,200,000	10M 13S	10.3
04	YOUTUBE.COM	TV & VIDEO	504,400,000	26M 07S	11.2
05	TRIBUNNEWS.COM	NEWS	215,000,000	05M 42S	4.1
06	DETIK.COM	NEWS	137,300,000	06M 50S	3.8
07	TOKOPEDIA.COM	SHOPPING	124,500,000	04M 26S	4.3
08	YAHOO.COM	NEWS	89,800,000	07M 18S	6.6
09	WHATSAPP.COM	COMMUNICATION	88,500,000	02M 29S	2.0
10	BUKALAPAK.COM	SHOPPING	87,200,000	04M 41S	4.1

SOURCE: SIMILARWEB (JANUARY 2019). FIGURES BASED ON MONTHLY AVERAGES FOR Q4 2018. NOTES: 'MONTHLY TRAFFIC' DOES NOT REPRESENT UNIQUE VISITORS. 'TIME PER VISIT' FIGURES REPRESENT THE AVERAGE DURATION OF USERS' VISITS, MEASURED IN MINUTES AND SECONDS. ADVISORY: SOME WEBSITES FEATURED IN THIS RANKING MAY CONTAIN ADULT CONTENT. PLEASE USE CAUTION WHEN VISITING UNKNOWN WEBSITES.

Sumber : (Orenzi, 2019).

### 3. Tampilan halaman *website* GT



Alamat URL: <https://trends.google.com/trends/explore?date=2014-01-01%202018-12-31&geo=ID&q=tuberkulosis,tbc,tb>

Waktu akses: Selasa, 18 Februari 2020 Pukul 14:14 Wib.

4. Nilai-Nilai  $r$  Product Moment

**TABEL III**  
**NILAI-NILAI  $r$  PRODUCT MOMENT**

N	Taraf Signifikan		N	Taraf Signifikan		N	Taraf Signifikan	
	5%	1%		5%	1%		5%	1%
3	0,997	0,999	27	0,381	0,487	55	0,266	0,345
4	0,950	0,990	28	0,374	0,478	60	0,254	0,330
5	0,878	0,959	29	0,367	0,470	65	0,244	0,317
6	0,811	0,917	30	0,361	0,463	70	0,235	0,306
7	0,754	0,874	31	0,355	0,456	75	0,227	0,296
8	0,707	0,834	32	0,349	0,449	80	0,220	0,286
9	0,666	0,798	33	0,344	0,442	85	0,213	0,278
10	0,632	0,765	34	0,339	0,436	90	0,207	0,270
11	0,602	0,735	35	0,334	0,430	95	0,202	0,263
12	0,576	0,708	36	0,329	0,424	100	0,195	0,256
13	0,553	0,684	37	0,325	0,418	125	0,176	0,230
14	0,532	0,661	38	0,320	0,413	150	0,159	0,210
15	0,514	0,641	39	0,316	0,408	175	0,148	0,194
16	0,497	0,623	40	0,312	0,403	200	0,138	0,181
17	0,482	0,606	41	0,308	0,398	300	0,113	0,148
18	0,468	0,590	42	0,304	0,393	400	0,098	0,128
19	0,456	0,575	43	0,301	0,389	500	0,088	0,115
20	0,444	0,561	44	0,297	0,384	600	0,080	0,105
21	0,433	0,549	45	0,294	0,380	700	0,074	0,097
22	0,423	0,537	46	0,291	0,376	800	0,070	0,091
23	0,413	0,526	47	0,288	0,372	900	0,065	0,086
24	0,404	0,515	48	0,284	0,368	1000	0,062	0,081
25	0,396	0,505	49	0,281	0,364			
26	0,388	0,496	50	0,279	0,361			

Sumber: (Sugiyono, 2007).

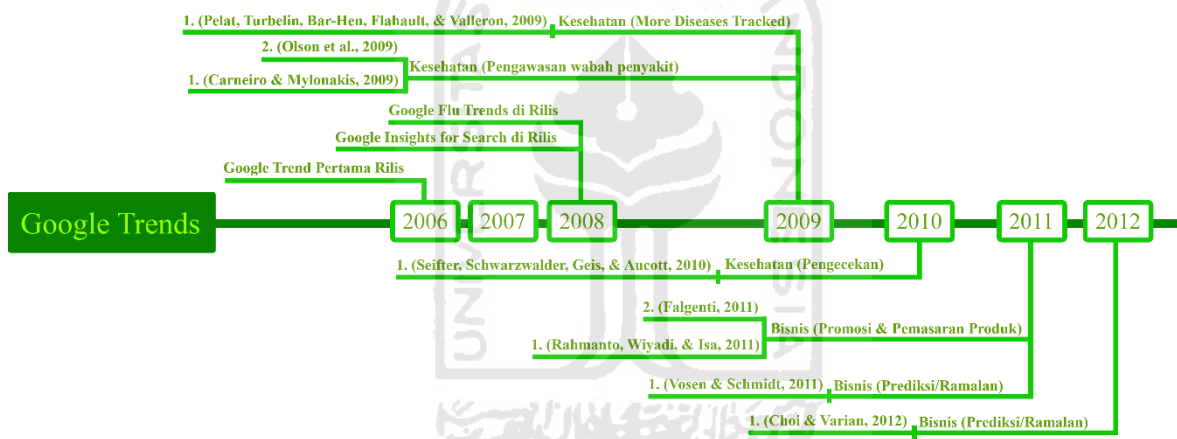
5. Tolak ukur pemberian penafsiran terhadap koefisien korelasi

**TABEL 7.3**  
**PEDOMAN UNTUK MEMBERIKAN INTERPRESTASI**  
**TERHADAP KOEFISIEN KORELASI**

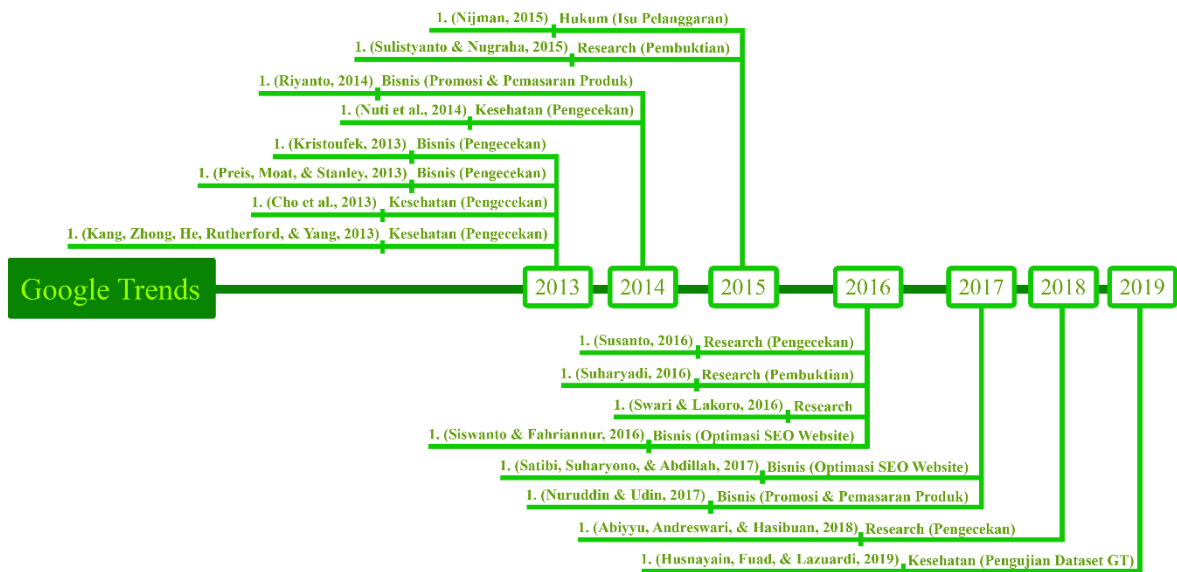
Interval Koefisien	Tingkat Hubungan
0,00 – 0,199	Sangat rendah
0,20 – 0,399	Rendah
0,40 – 0,599	Sedang
0,60 – 0,799	Kuat
0,80 – 1,000	Sangat kuat

Sumber: (Sugiyono, 2007).

6. Timeline penggunaan GT pada tahun 2006 sampai tahun 2012

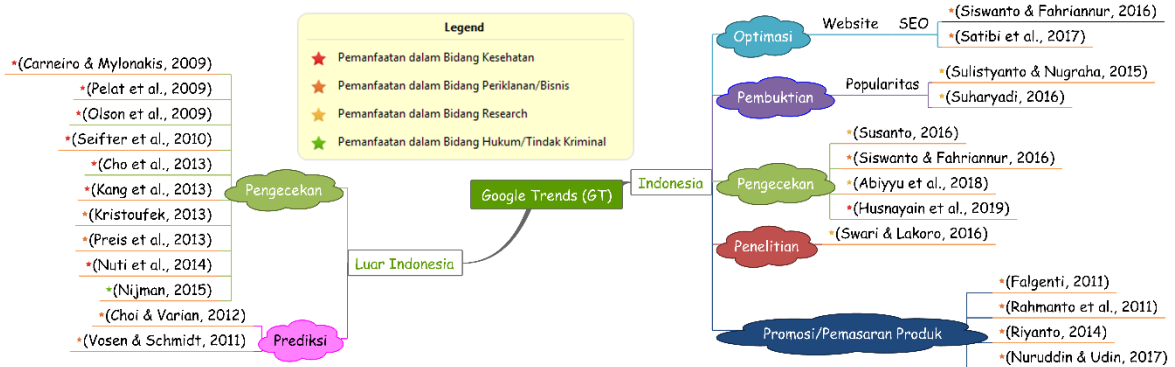


7. Timeline penggunaan GT pada tahun 2013 sampai tahun 2019

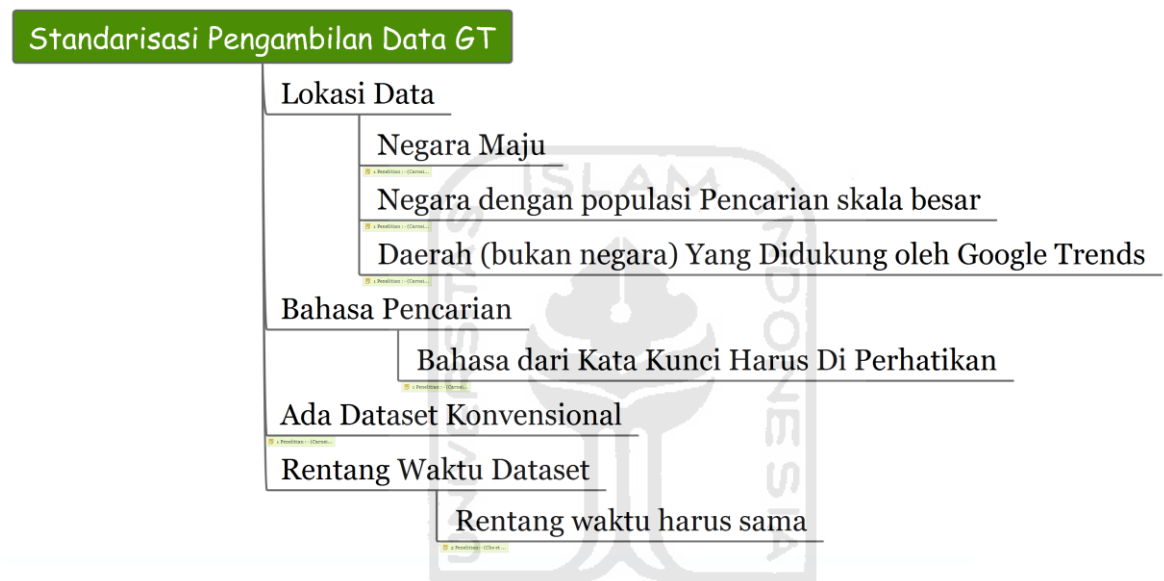




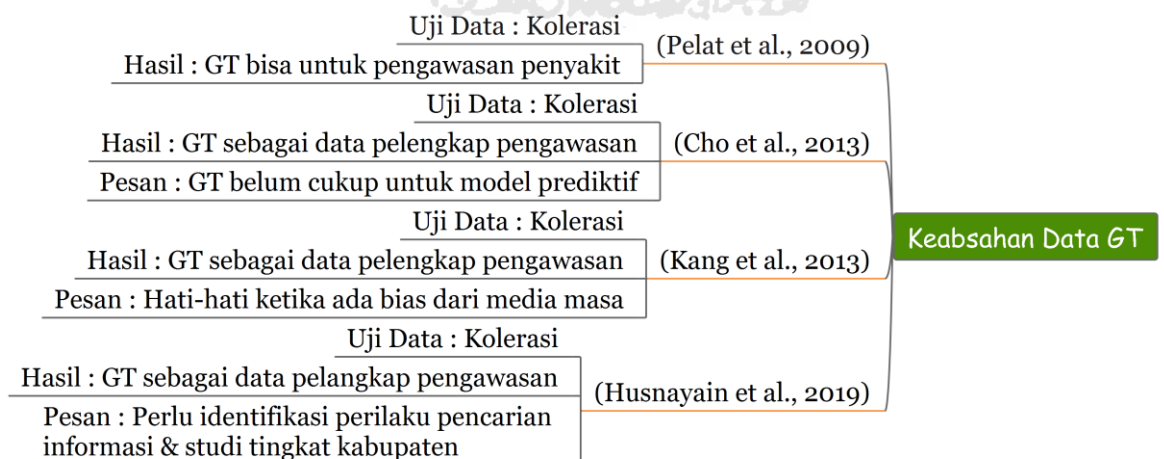
## 8. Main map pola penggunaan GT



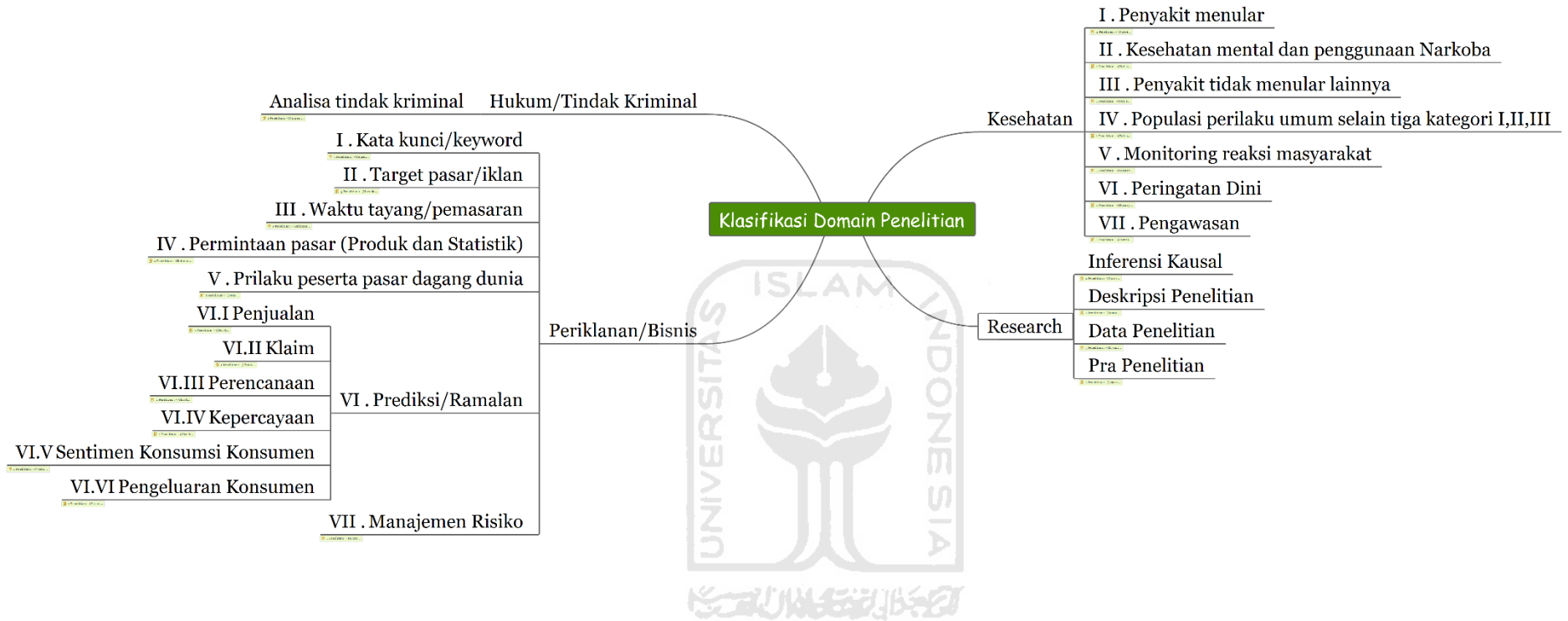
## 9. Main map standarisasi pengambilan data tren GT



## 10. Main map keabsahan data GT



11. Main map klasifikasi domain penelitian tentang GT



## 12. Pembahasan dan penggunaan GT sebelumnya

No	Peneliti	Tujuan	Model Penggunaan GT (✓) / Hasil Penelitian (✓)											Metode yang digunakan (✓)							Hasil/Evaluasi			
			Model/Strategi	Optimasi	Pelacakan	Pemantauan Reaksi Publik	Pembuktian	Pengawasan	Pengecekan	Peringatan Dini	Prediksi	Promosi & Pemasaran Produk	ReSearch	Moving Average	Action ReSearch	Correlation of Time Lag	Deskriptif	Komparatif	Kuantitatif	Kualitatif		Literatur Review	Pearson correlation	
1.	(Carneiro & Mylonakis, 2009b)	Mengusulkan tool pengawasan berbasis <i>web</i>						✓	✗															<ul style="list-style-type: none"> <li>- GT sebagai tool <i>web</i> pengawasan wabah penyakit.</li> <li>- Perlu nya standarisasi kata kunci penelusuran <i>web</i> yang cocok untuk pengawasan wabah penyakit tertentu.</li> </ul>
2.	(Olson et al., 2009)	Mencari pengawasan/ <i>surveillance</i> flu yang lebih baik						✓																<ul style="list-style-type: none"> <li>- GT berpeluang menjadi <i>surveillance</i></li> </ul>
3.	(Pelat et al., 2009)	Melacak lebih banyak penyakit dengan GT			✓			✓	✗															<ul style="list-style-type: none"> <li>- Data GT bisa digunakan untuk pengawasan penyakit di negara dengan bahasa selain bahasa Inggris.</li> <li>- GT berpeluang untuk pengawasan penyakit di negara lain.</li> </ul>
4.	(Seifter et al., 2010)	Menjelaskan pemanfaatan GT dalam bidang penelitian <i>epidemiologi</i>						✓	✗															<ul style="list-style-type: none"> <li>- 6 dari 10 kota yang termasuk tren pada GT tentang penyakit Lyme terdapat juga di tren CDC.</li> <li>- Bias GT : pengambilan sampel oleh sistem GT tidak representatif, dipengaruhi oleh asumsi matematika dan metode perkiraan.</li> </ul>



No	Peneliti	Tujuan	Model Penggunaan GT (✓) / Hasil Penelitian (✓)										Metode yang digunakan (✓)							Hasil/Evaluasi					
			Model/Strategi	Optimasi	Pelacakan	Pemantauan Reaksi Publik	Pembuktian	Pengawasan	Pengecekan	Peringatan Dini	Prediksi	Promosi & Pemasaran Produk	ReSearch	Moving Average	Action ReSearch	Correlation of Time Lag	Deskriptif	Komparatif	Kuantitatif		Kualitatif	Literatur Review	Pearson correlation		
10.	(Preis, Moat, & Eugene Stanley, 2013)	Menganalisis GT untuk keperluan keputusan jual beli saham di perdagangan pasar dunia.	✓																						- Menghasilkan strategi bisnis yaitu membeli saham ( <i>take a long position</i> ) jika <i>volume</i> pencarian GT menurun, dan menjual saham ( <i>take a short position</i> ) ketika <i>volume</i> pencarian GT naik.
11.	(Cho et al., 2013)	Mengecek korelasi data pengawasan <i>influenza</i> nasional dengan data GT di Korea Selatan.																							- GT untuk bahasa Korea dapat digunakan untuk sebagai data pelengkap <i>surveillance influenza</i> , tetapi tidak cukup untuk penggunaan model prediksi/ <i>predictive</i>
12.	(Kang et al., 2013)	Menggunakan GT untuk <i>influenza surveillance</i> di China Selatan.																							- GT bisa menjadi sumber data pelengkap untuk <i>influenza surveillance</i> di China Selatan.
13.	(Nuti et al., 2014b)	Menjelaskan pengetahuan tentang pemanfaatan GT	✓																						- Terdapat 4 model domain topik para penelitian tentang GT yaitu penyakit menular, kesehatan mental dan penggunaan narkoba, penyakit tidak menular lainnya, dan perilaku populasi umum. - GT digunakan untuk inferensi kasual, deskripsi, dan pengawasan.



No	Peneliti	Tujuan	Model Penggunaan GT (✓) / Hasil Penelitian (✓)										Metode yang digunakan (✓)							Hasil/Evaluasi				
			Model/Strategi	Optimasi	Pelacakan	Pemantauan Reaksi Publik	Pembuktian	Pengawasan	Pengecekan	Peringatan Dini	Prediksi	Promosi & Pemasaran Produk	ReSearch	Moving Average	Action ReSearch	Correlation of Time Lag	Deskriptif	Komparatif	Kuantitatif		Kualitatif	Literatur Review	Pearson correlation	
		Para diplomasi Kota Bandung.																						
19.	(Laila R Swari & Lakoro, 2016)	Merancang video promosi untuk membantu mempromosikan pantai di Kabupaten Tulungagung.																						
20.	(Siswanto & Fahriannur, 2016)	Menjelaskan fungsi GT untuk menganalisis pasar bisnis <i>online</i> dan pemilihan kata kunci pada <i>e-commerce website</i> .		✓																				
21.	(Satibi, Suharyono, & Abdillah, 2017)	Mengoptimasi SEO <i>website</i> dalam meningkatkan penjualan produk UKM		✓																				
22.	(Nuruddin & Udin, 2017)	Mensosialisasikan internet sehat kepada kelompok kerja sehingga lebih produktif.		✓																				

No	Peneliti	Tujuan	Model Penggunaan GT (✓) / Hasil Penelitian (✓)										Metode yang digunakan (✓)							Hasil/Evaluasi		
			Model/Strategi	Optimasi	Pelacakan	Pemantauan Reaksi Publik	Pembuktian	Pengawasan	Pengecekan	Peringatan Dini	Prediksi	Promosi & Pemasaran Produk	ReSearch	Moving Average	Action ReSearch	Correlation of Time Lag	Deskriptif	Komparatif	Kuantitatif		Kualitatif	Literatur Review
																						dan menemukan produk yang cocok untuk waktu-waktu tertentu.
23.	(Abiyyu, Andreswari, & Hasibuan, 2018)	Membuat aplikasi <i>mobile</i> untuk pemesanan tiket even di kota Bandung					✓		✗													<ul style="list-style-type: none"> <li>- GT digunakan untuk mengecek sekaligus membuktikan tren tentang kota Bandung sebagai kota kreatif dengan tersedianya banyak even-even.</li> <li>- Perlu nya identifikasi faktor-faktor yang mempengaruhi perilaku pencari informasi khususnya di Indonesia.</li> </ul>
24.	(Husnayain et al., 2019)	Menguji hubungan atau korelasi kasus Demam Berdarah <i>Dengue</i> (DBD) dengan tren DBD di GT khususnya di wilayah Indonesia.				✓	✗	✓	✗	✓			✓		✓	✓			✓		✓	<ul style="list-style-type: none"> <li>- Pola deret waktu data GT <i>linier</i> dengan laporan resmi kasus DBD (berdasarkan analisis <i>moving average</i>)</li> <li>- Uji korelasi <i>Pearson</i> nya menunjukkan korelasi yang tinggi pada tiga istilah/kata kunci yang diteliti (ditetapkan nilai rentang dari R 0,921 hingga 0,937 dengan <math>p \leq 0,05</math>, periode keseluruhan).</li> <li>- GT diindikasikan dapat digunakan untuk sistem peringatan dini.</li> <li>- GT juga dapat digunakan untuk melihat sekaligus</li> </ul>



No	Peneliti	Tujuan	Model Penggunaan GT (✓) / Hasil Penelitian (✓)										Metode yang digunakan (✓)							Hasil/Evaluasi		
			Model/Strategi	Optimasi	Pelacakan	Pemantauan Reaksi Publik	Pembuktian	Pengawasan	Pengecekan	Peringatan Dini	Prediksi	Promosi & Pemasaran Produk	ReSearch	Moving Average	Action ReSearch	Correlation of Time Lag	Deskriptif	Komparatif	Kuantitatif		Kualitatif	Literatur Review
																						memantau reaksi masyarakat tentang suatu kasus.
25.	Usulan	Membuat <i>template Search term</i> dalam melihat tren kasus penyakit di <i>Google Trends</i> untuk pengawasan wabah penyakit di Indonesia.			✓	✓	✓	✓	✗	✓	✓	✓	✓		✓	✓		✓	✓	✓		<u>Pembaharuan</u> - Tersedia <i>template list Search term</i> yang nanti nya menghasilkan kata kunci untuk melihat tren (penyakit) di GT untuk wilayah Indonesia.

Keterangan:

- Centang Warna Ungu (✓): Bermaksud bahwa *Google Trends* digunakan secara teknis pada awal dan saat penelitian.
  - Centang Warna Biru (✓): Bermaksud bawah *Google Trends* bisa digunakan atau dimanfaatkan untuk hal tersebut (sesuai kolom yang di centang) dan ini keluar atau didapat setelah penelitian selesai dilakukan, dalam arti lainnya yaitu hasil dari penelitian yang disebutkan oleh para peneliti di bab kesimpulan/hasil penelitian nya.
  - Centang Warna Hijau (✓): Bermaksud bahwa peneliti menggunakan/menyebutkan metode yang digunakan saat penelitian.
- Tanda silang warna merah (✗): Bermaksud bahwa peneliti dalam hasil penelitian nya menyebutkan bahwa GT belum cukup/tidak cocok untuk digunakan untuk hal itu (sesuai kolom yang dipilih).

### 13. Perbandingan penelitian sebelumnya dengan usulan peneliti

	(Carneiro & Mylonakis, 2009b)	(Olson et al., 2009)	(Pelat et al., 2009)	(Seifter et al., 2010)	(Cho et al., 2013)	(Kang et al., 2013)	(Nuti et al., 2014b)	(Husnayain et al., 2019)	Peneliti
<b>Domain</b>	P. Menular - <i>Influenza/Flu</i> P. Tidak Menular -	P. Menular - P. Tidak Menular -	P. Menular - <i>Influenza/Flu</i> - Gastroenteritis - Cacar Air P. Tidak Menular -	P. Menular - Lyme P. Tidak Menular -	P. Menular - <i>Influenza/Flu</i> P. Tidak Menular -	P. Menular - <i>Influenza/Flu</i> P. Tidak Menular -	P. Menular. P. Tidak Menular. Kesehatan Mental dan penggunaan Zat. Perilaku populasi Umum.	P. Menular - Demam Berdarah <i>Dengue</i> (DBD) P. Tidak Menular -	P. Menular - Tuberkulosis (TBC) P. Tidak Menular - Stroke
<b>Data Konvensional</b>	✗	✗	<i>Clinical surveillance data from the French Sentinel Network.</i>	Center for Disease Control (CDC) di US	CDC di Korea	- CDC provinsi - Sistem pengawasan Virologi Guandong.	-	Laporan resmi DBD dari Departemen Arbovirus, Kementerian Kesehatan Indonesia.	Laporan resmi kasus penyakit Tuberkulosis dan Stroke dari Ditjen P2P Kemkes RI.
<b>Data GT</b>	✗	✗	✔	✔	✔	✔	-	✔	✔
<b>Teknik</b>	- <i>Google Flu Trends.</i> - <i>Scaling the data.</i> - <i>Data normalization.</i>	-	- <i>Pearson correlation.</i> - <i>Correlation of Time Lag.</i>	-	- <i>Pearson correlation.</i> - <i>Correlation of Time Lag.</i>	- <i>Pearson correlation.</i> - <i>Scatter Plot.</i> - <i>Correlation of Time Lag.</i>	-	- <i>Pearson correlation.</i> - <i>Correlation of Time Lag.</i>	- <i>Pearson correlation.</i>
<b>Evaluasi</b>	<u>Search term</u> -	<u>Search term</u> -	<u>Search term</u> - <i>Brainstorming tim*</i> .	<u>Search term</u> - Memakai kata nama objek yaitu "lyme" dan ditambah "penyakit".	<u>Search term</u> - 100 pasien di survei**. - Definisi dari ILI. - Temuan peneliti.	<u>Search term</u> - Pengetahuan umum tentang <i>influenza.</i> - Definisi ILI.	<u>Search term</u> -	<u>Search term</u> - <i>Disease definition.</i> - <i>Symptom.</i> - <i>Treatment.</i> - <i>Vector of Disease.</i>	<u>Search term</u> - Literatur sebelumnya. - Buku Kapita Selekta Kedokteran.
	Lebih <i>review</i> untuk membuka wawasan tentang GT.	Berfokus membandingkan data GT dengan CDC dari US.		Condong ke <i>review</i> untuk membuka wawasan tentang GT. Kasus di US.	Kasus di Korea	Kasus Di Cina Selatan	Penelitian Literatur <i>Review.</i>	Kasus di Indonesia pada tingkat provinsi dan nasional.	Kasus di Indonesia pada tingkat provinsi dan nasional.

Keterangan:

\* : Tim memiliki beberapa peserta, setiap peserta memberikan istilah pencarian (pertanyaan) yang berkemungkinan akan digunakan dalam mencari informasi terkait penyakit yang sedang diteliti.

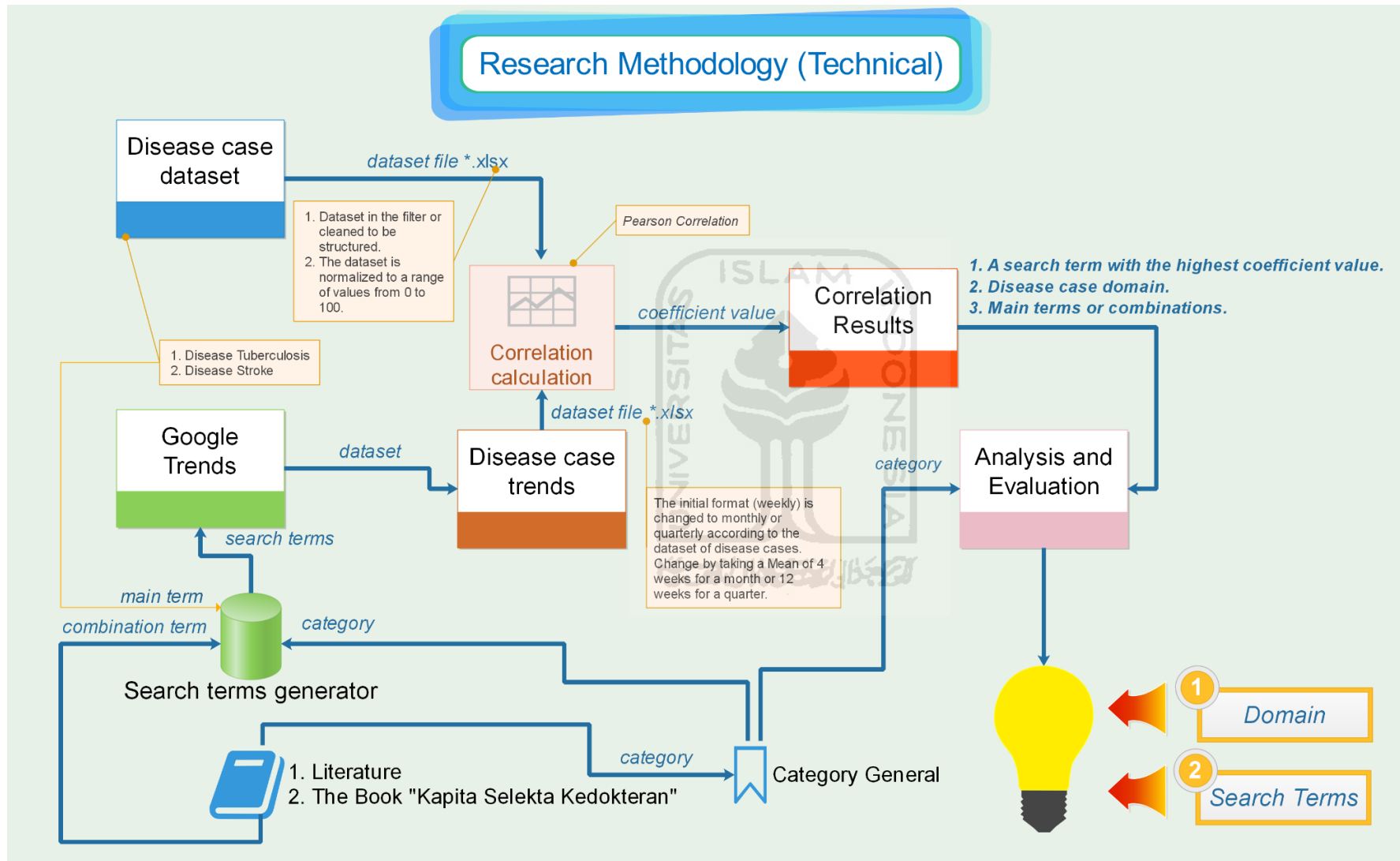
\*\* : Ada 100 pasien anonim (diketahui berturut-turut masuk ruang gawat darurat) di survei oleh peneliti dengan memberikan pertanyaan "Jika Anda mencari *Influenza*, pertanyaan atau istilah apa yang Anda gunakan?" .

## LAMPIRAN B

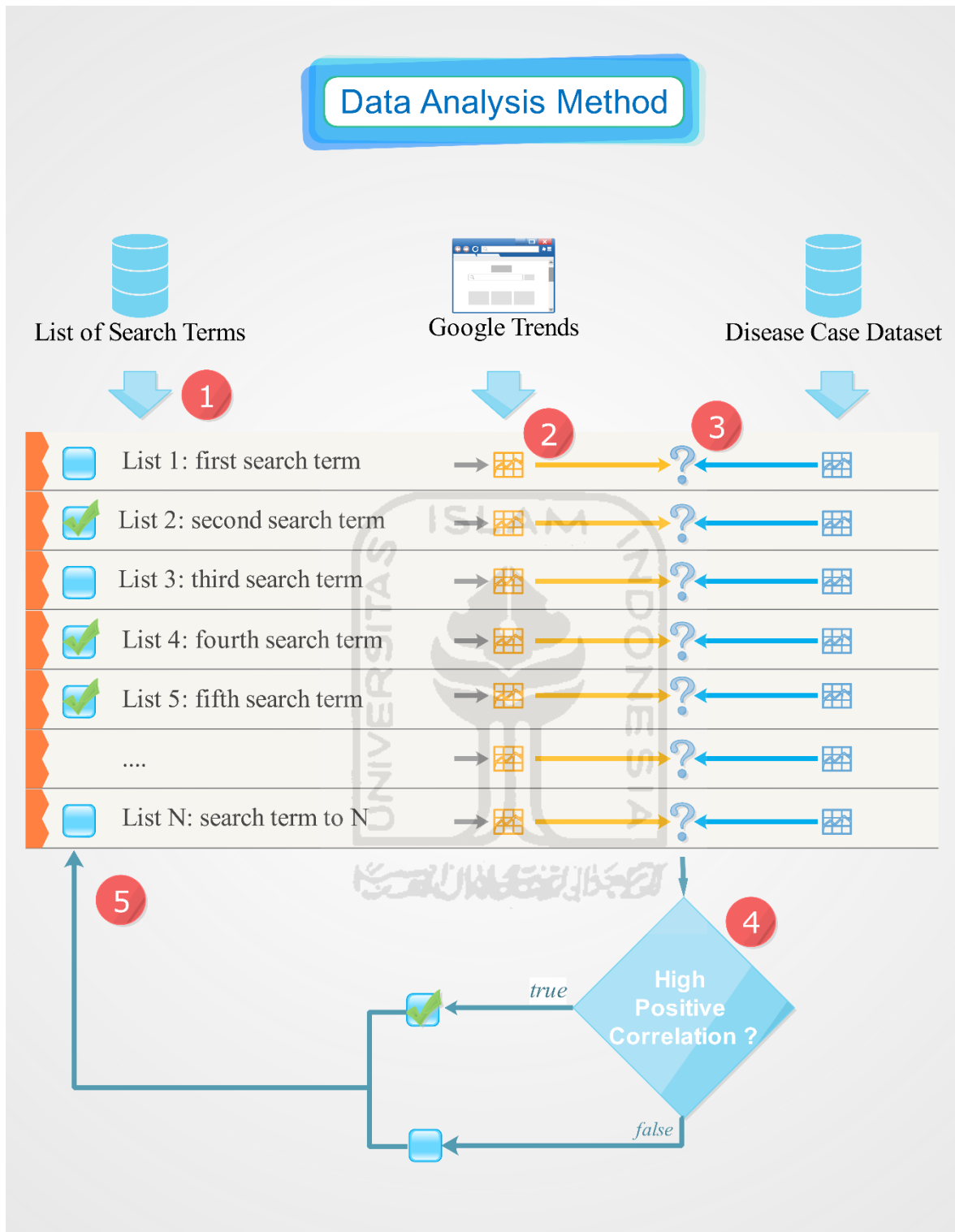
### 1. Jadwal pelaksanaan penelitian

No	Kegiatan	Waktu Kegiatan																			
		Feb 2019	Mar 2019	Apr 2019				Mei – Okt 2019	Nov 2019				Des 2019				Jan 2020	Feb 2020	Mar-Nov 2020	Des 2020	
		1-4	1-4	1	2	3	4	1-4	1	2	3	4	1	2	3	4	1-4	1-4	1-4	1-4	
1.	Studi literatur.																				
2.	Penetapan tema dan fokus penelitian.																				
3.	Mempelajari objek penelitian																				
4.	Pengumpulan data Kasus Penyakit dari lembaga pengawas.																				
5.	Perancangan Aturan dalam membuat <i>Search term</i> .																				
6.	Pengumpulan data tren GT.																				
7.	Perhitungan Korelasi.																				
8.	Analisis Hasil Korelasi dan Evaluasi.																				
9.	Penyusunan laporan.																				

## 2. Metodologi penelitian secara teknis



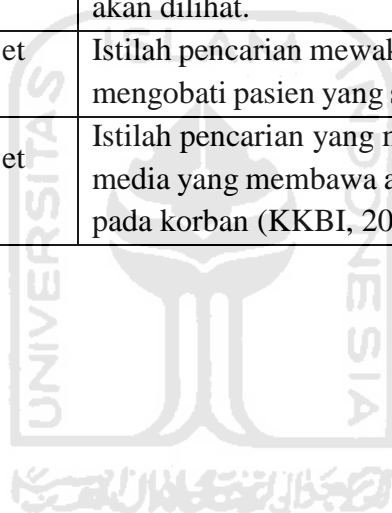
### 3. Metode analisis data



## LAMPIRAN C

### 1. Tabel definisi setiap kategori dari literatur

Kategori		Sumber	Pengertian Kategori / Keterangan
1.	Definisi	(Choi & Varian, 2012a; Husnayain et al., 2019; Kang et al., 2013)	Istilah pencarian yang dihasilkan mampu menggambarkan ciri utama, identitas utama, atau makna utama dari obyek yang sedang didefinisikan (KKBI, 2016a).
2.	Gejala	(Husnayain et al., 2019)	Istilah pencarian yang mewakili tanda-tanda dari suatu obyek (dalam kasus ini ialah penyakit) akan muncul atau terjadi (KKBI, 2016b). Kata “gejala” bisa juga digunakan sebagai istilah pencarian untuk penekanan makna yang dimaksud dari tren yang akan dilihat.
3.	Pengobatan	(Husnayain et al., 2019)	Istilah pencarian mewakili dari kegiatan atau proses mengobati pasien yang sedang sakit (KKBI, 2016f).
4.	Vektor Penyakit	(Husnayain et al., 2019)	Istilah pencarian yang menggambarkan hewan atau media yang membawa atau menularkan penyakit ke pada korban (KKBI, 2016g).





E = sub-bab pembahasan ‘Klasifikasi dan/atau Terminologi’  
F = sub-bab pembahasan ‘Epidemiologi’  
G = sub-bab pembahasan ‘Diagnosis dan/atau Banding’  
H = sub-bab pembahasan ‘Prinsip dan/atau Tata Laksana’  
I = sub-bab pembahasan ‘Pemeriksaan Penunjang dan/atau Lain’  
J = sub-bab pembahasan ‘Indikasi Rawat dan/atau Kriteria Inap’  
K = sub-bab pembahasan ‘Komplikasi’  
L = sub-bab pembahasan ‘Patofisiologi’  
M = sub-bab pembahasan ‘Etiologi’  
N = sub-bab pembahasan ‘Patogenesis’  
O = sub-bab pembahasan ‘Prognosis’  
P = sub-bab pembahasan ‘Pencegahan dan/atau Skrining’  
Q = sub-bab pembahasan ‘Komplikasi’  
R = sub-bab pembahasan ‘Profil Virus’  
S = sub-bab pembahasan ‘Diare Akut’  
T = sub-bab pembahasan ‘Diare Kronik’  
U = sub-bab pembahasan ‘Langkah Promotif/Preventif’  
V = sub-bab pembahasan ‘Pendekatan Klinis Diare Kronik’  
W = sub-bab pembahasan ‘Terapi Lainnya’  
X = sub-bab pembahasan ‘Strategi Nasional Pengendalian TB di Indonesia’  
Y = sub-bab pembahasan ‘TB Pada Dahak’  
Z = sub-bab pembahasan ‘Interpretasi Hasil Dahak’  
AA = sub-bab pembahasan ‘Pengobatan TB’  
AB = sub-bab pembahasan ‘Indikasi Rawat di Ruang Intensif’  
AC = sub-bab pembahasan ‘Kriteria Pulang’  
AD = sub-bab pembahasan ‘Patofisiologi’  
AE = sub-bab pembahasan ‘Penatalaksanaan Tekanan Darah’  
AF = sub-bab pembahasan ‘Penatalaksanaan Gula Darah’  
AG = sub-bab pembahasan ‘Rehabilitasi Stroke’  
AH = sub-bab pembahasan ‘Penilaian PPOK’  
AI = sub-bab pembahasan ‘Indikasi Rawat ICU’



### 3. Contoh pemilihan istilah kombinasi

3. For the sub-chapter of the "Epidemiologi" discussion:			
Number	Term	Count	
1	anak	11	
2	tahun	11	unit of time
3	hepatitis	8	disease name
4	kasus	7	
5	akut	6	
6	infeksi	6	
7	dunia	5	
8	indonesia	5	country name
9	juta	5	unit of numbers
10	negara	5	unit place
11	terjadi	5	
12		5	4 number
13	berusia	4	
14	kronis	4	
15	penyakit	4	
16	vhb	4	disease name
17	virus	4	
18		2	3 number
19	a	3	letters of the alphabet
20	angka	3	

Kotak *Cell* 'Number' yang berwarna Hijau merupakan nomor *term* yang akan dijadikan sebagai istilah kombinasi, sedangkan yang berwarna kuning merupakan nomor *term* yang tidak dijadikan sebagai istilah kombinasi. Alasan bagi *term* yang ditolak dapat dilihat pada samping *cell* 'Count' yang ditandai dengan tulisan berwarna merah.

### 4. Istilah kombinasi yang ditolak

Nomor	Nama	Keterangan
1.	Country name	Semua nama negara tidak dijadikan sebagai <i>combination term</i> .
2.	Disease name	Nama-nama penyakit juga ditolak karena <i>term</i> ini merupakan kombinasi untuk <i>main term</i> yang merupakan nama penyakit yang sedang diteliti.
3.	Letter of the alphabet	Huruf-huruf dari <i>alphabet</i> atau abjad juga ditolak.
4.	Name of the day	Nama hari seperti senin atau selasa tidak dijadikan <i>combination term</i> .
5.	Number	Angka seperti 1, 2, 3 dan seterusnya tidak dijadikan sebagai <i>combination term</i> .
6.	Number words	Kata bilangan atau kata-kata yang mewakili bilangan juga ditolak seperti sebagian, satu jam, setengah, seminggu, dan lainnya.
7.	Unit dose of the drug	Symbol atau satuan dari dosis obat tidak dijadikan sebagai <i>combination term</i> .
8.	Unit of numbers	Satuan angka seperti puluhan, seratus, atau juta, tidak dijadikan sebagai <i>combination term</i> .
9.	Unit of time	Kata-kata satuan waktu seperti tahun, bulan, minggu, dan hari.
10.	Unit place	Kata-kata yang termasuk satuan tempat atau memiliki makna tempat seperti negara, provinsi, kabupaten, kota, dan desa.
11.	Unit weight	Kata-kata yang merupakan satuan berat seperti kg, liter, atau mg.

Aturan jenis *Combination Term* yang ditolak ini digunakan untuk kedua obyek penelitian yaitu Tuberkulosis dan Stroke khusus pada tahap penelitian 'Generate New Term'.

5. Term baru untuk penyakit Tuberkulosis (hasil tahapan metodologi *Generate New Term*)

For the sub-chapter of the "Definisi" discussion		For the sub-chapter of the "Diagnosis" discussion		For the sub-chapter of the "Epidemiologi" discussion		For the sub-chapter of the "Faktor Risiko" discussion		For the sub-chapter of the "Tata Laksana" discussion	
Number	Term	Number	Term	Number	Term	Number	Term	Number	Term
1	primer tuberkulosis	1	pemeriksaan tuberkulosis	1	anak tuberkulosis	1	virus tuberkulosis	1	terapi tuberkulosis
2	akut tuberkulosis	2	dilakukan tuberkulosis	2	kasus tuberkulosis	2	anak tuberkulosis	2	untuk tuberkulosis
3	disertai tuberkulosis	3	infeksi tuberkulosis	3	akut tuberkulosis	3	cairan tuberkulosis	3	diberikan tuberkulosis
4	paru tuberkulosis	4	gejala tuberkulosis	4	infeksi tuberkulosis	4	faktor tuberkulosis	4	pasien tuberkulosis
5	sakit tuberkulosis	5	pasien tuberkulosis	5	dunia tuberkulosis	5	risiko tuberkulosis	5	pemberian tuberkulosis
6	afek tuberkulosis	6	ditemukan tuberkulosis	6	terjadi tuberkulosis	6	darah tuberkulosis	6	cairan tuberkulosis
7	akibat tuberkulosis	7	paru tuberkulosis	7	berusia tuberkulosis	7	individu tuberkulosis	7	terapi tb
8	hati tuberkulosis	8	bta tuberkulosis	8	kronis tuberkulosis	8	infeksi tuberkulosis	8	untuk tb
9	kronis tuberkulosis	9	darah tuberkulosis	9	penyakit tuberkulosis	9	lahir tuberkulosis	9	diberikan tb
10	penyakit tuberkulosis	10	foto tuberkulosis	10	virus tuberkulosis	10	penyebab tuberkulosis	10	pasien tb
11	virus tuberkulosis	11	normal tuberkulosis	11	angka tuberkulosis	11	usia tuberkulosis	11	pemberian tb
12	primer tb	12	tanda tuberkulosis	12	anak tb	12	ditemukan tuberkulosis	12	cairan tb
13	akut tb	13	untuk tuberkulosis	13	kasus tb	13	perinatal tuberkulosis	13	terapi tbc
14	disertai tb	14	berat tuberkulosis	14	akut tb	14	rendah tuberkulosis	14	untuk tbc
15	paru tb	15	pemeriksaan tb	15	infeksi tb	15	secara tuberkulosis	15	diberikan tbc
16	sakit tb	16	dilakukan tb	16	dunia tb	16	virus tb	16	pasien tbc
17	afek tb	17	infeksi tb	17	terjadi tb	17	anak tb	17	pemberian tbc
18	akibat tb	18	gejala tb	18	berusia tb	18	cairan tb	18	cairan tbc
19	hati tb	19	pasien tb	19	kronis tb	19	faktor tb		
20	kronis tb	20	ditemukan tb	20	penyakit tb	20	risiko tb		
21	penyakit tb	21	paru tb	21	virus tb	21	darah tb		
22	virus tb	22	bta tb	22	angka tb	22	individu tb		
23	primer tbc	23	darah tb	23	anak tbc	23	infeksi tb		
24	akut tbc	24	foto tb	24	kasus tbc	24	lahir tb		
25	disertai tbc	25	normal tb	25	akut tbc	25	penyebab tb		
26	paru tbc	26	tanda tb	26	infeksi tbc	26	usia tb		
27	sakit tbc	27	untuk tb	27	dunia tbc	27	ditemukan tb		
28	afek tbc	28	berat tb	28	terjadi tbc	28	perinatal tb		
29	akibat tbc	29	pemeriksaan tbc	29	berusia tbc	29	rendah tb		
30	hati tbc	30	dilakukan tbc	30	kronis tbc	30	secara tb		
31	kronis tbc	31	infeksi tbc	31	penyakit tbc	31	virus tbc		
32	penyakit tbc	32	gejala tbc	32	virus tbc	32	anak tbc		
33	virus tbc	33	pasien tbc	33	angka tbc	33	cairan tbc		
		34	ditemukan tbc			34	faktor tbc		
		35	paru tbc			35	risiko tbc		
		36	bta tbc			36	darah tbc		
		37	darah tbc			37	individu tbc		
		38	foto tbc			38	infeksi tbc		
		39	normal tbc			39	lahir tbc		

<i>For the sub-chapter of the "Definisi" discussion</i>		<i>For the sub-chapter of the "Diagnosis" discussion</i>		<i>For the sub-chapter of the "Epidemiologi" discussion</i>		<i>For the sub-chapter of the "Faktor Risiko" discussion</i>		<i>For the sub-chapter of the "Tata Laksana" discussion</i>	
<i>Number</i>	<i>Term</i>	<i>Number</i>	<i>Term</i>	<i>Number</i>	<i>Term</i>	<i>Number</i>	<i>Term</i>	<i>Number</i>	<i>Term</i>
		40	tanda tbc			40	penyebab tbc		
		41	untuk tbc			41	usia tbc		
		42	berat tbc			42	ditemukan tbc		
						43	perinatal tbc		
						44	rendah tbc		
						45	secara tbc		

## 6. Istilah pencarian untuk penyakit Tuberkulosis yang tidak memiliki tren pada GT

<i>Number</i>	<i>Term</i>
1.	afek tb
2.	afek tbc
3.	afek tuberkulosis
4.	akibat tuberkulosis
5.	akut tuberkulosis
6.	angka tuberkulosis
7.	berat tuberculosis
8.	berusia tb
9.	berusia tbc
10.	berusia tuberkulosis
11.	bta tuberkulosis
12.	cairan tuberkulosis
13.	diberikan tuberkulosis
14.	dilakukan tuberkulosis
15.	disertai tb
16.	disertai tuberkulosis
17.	ditemukan tb
18.	ditemukan tbc

<i>Number</i>	<i>Term</i>
23.	individu tb
24.	individu tbc
25.	individu tuberkulosis
26.	kasus tuberkulosis
27.	kronis tuberkulosis
28.	lahir tbc
29.	lahir tuberkulosis
30.	normal tuberkulosis
31.	pemberian tuberkulosis
32.	perinatal tb
33.	perinatal tbc
34.	perinatal tuberkulosis
35.	primer tbc
36.	primer tuberkulosis
37.	rendah tuberkulosis
38.	risiko tbc
39.	risiko tuberkulosis
40.	secara tuberkulosis

<i>Number</i>	<i>Term</i>
19.	ditemukan tuberkulosis
20.	dunia tuberkulosis
21.	foto tuberkulosis
22.	hati tuberkulosis

<i>Number</i>	<i>Term</i>
41.	tanda tuberkulosis
42.	terapi tuberkulosis
43.	terjadi tuberkulosis



## 7. Contoh perubahan dataset tren asli setelah di normalisasi

**a. Original results from Google Trends**

Minggu	pemeriksaan tuberkulosis: (Indonesia)
1/5/2014	54
1/12/2014	55
1/19/2014	0
1/26/2014	54
2/2/2014	0
2/9/2014	0
2/16/2014	0
2/23/2014	52
3/2/2014	0
3/9/2014	51
3/16/2014	50
3/23/2014	0
3/30/2014	51
4/6/2014	51
4/13/2014	100
4/20/2014	50
4/27/2014	0
5/4/2014	0
5/11/2014	0
5/18/2014	53
5/25/2014	56
6/1/2014	0
6/8/2014	0
6/15/2014	51

**b. After normalizing to quarterly format**

Period	pemeriksaan tuberkulosis	dilakukan tuberkulosis
January-March 2014	28.23076923	0
April-June 2014	31.92307692	0
July-September 2014	32.07692308	0
October-December 2014	59.58823529	0
January-March 2015	26.30769231	0
April-June 2015	24.61538462	0
July-September 2015	29.84615385	0
October-December 2015	30.92307692	0
January-March 2016	15.38461538	0
April-June 2016	25.92307692	0
July-September 2016	8.461538462	0
October-December 2016	26.84615385	0
January-March 2017	20.92307692	0
April-June 2017	18.46153846	0
July-September 2017	22	0
October-December 2017	16.5	0
January-March 2018	26.91666667	0
April-June 2018	20.84615385	0
July-September 2018	26.85714286	0
October-December 2018	16.23076923	0
Retrieval Time	12 November 2019 14:57 Wib	12 November 2019 14:03 Wib

Pada gambar ini, poin (a) menunjukkan kondisi tren dari istilah pencarian yang langsung di *download* dari GT, sedangkan poin (b) menunjukkan kondisi *dataset* tren istilah pencarian ‘pemeriksaan tuberkulosis’ telah dinormalisasi dari format mingguan ke triwulan.

## 8. Contoh perhitungan korelasi

**A. Correlation value for search terms**

Number	Category	Korelasi
11	4 Faktor Risiko	0,331612353
12	virus tuberkulosis	-0,392996971
13	cairan tuberkulosis	-0,29232888
14	faktor tuberkulosis	0
15	risiko tuberkulosis	0
16	darah tuberkulosis	0,396097571
17	individu tuberkulosis	0
18	infeksi tuberkulosis	-0,004099652
19	lahir tuberkulosis	0
20	penyebab tuberkulosis	0,209390783
21	usia tuberkulosis	0
22	ditemukan tuberkulosis	0
23	perinatal tuberkulosis	0
24	rendah tuberkulosis	0
25	secara tuberkulosis	0
26	virus tb	-0,11391277
27	anak tb	0,616291531
28	cairan tb	0,27787836
29	faktor tb	-0,24604347
30	risiko tb	-0,174112646
31	darah tb	0,467723607
32	individu tb	0
33	infeksi	0,66683394

**B. Disease Case Dataset**

Period	Number of Normalization Cases
January-March 2014	67.57526979
April-June 2014	66.60945122
July-September 2014	65.1410128
October-December 2014	66.84484486
January-March 2015	70.59755917
April-June 2015	68.20496589
July-September 2015	66.44355382
October-December 2015	66.61602142
January-March 2016	77.94550024
April-June 2016	76.50827023
July-September 2016	71.00408995
October-December 2016	70.70109973
January-March 2017	90.51017559
April-June 2017	81.17967839
July-September 2017	93.66551141
October-December 2017	86.53365948
January-March 2018	99.92741578
April-June 2018	90.05683218
July-September 2018	100
October-December 2018	94.13692285

**C. Search terms trends from Google trends**

Period	virus tuberkulosis	anak tuberkulosis
January-March 2014	3.615384615	45.6153846
April-June 2014	7.307692308	41.07692308
July-September 2014	3.461538462	33.46153846
October-December 2014	10.41176471	42.11764706
January-March 2015	3.307692308	33.46153846
April-June 2015	8.153846154	34.46153846
July-September 2015	7.307692308	28.76923077
October-December 2015	22.61538462	48.15384615
January-March 2016	11.46153846	36.76923077
April-June 2016	5.692307692	28.53846154
July-September 2016	10.84615385	30.76923077
October-December 2016	10.38461538	30.30769231
January-March 2017	0	32.61538462
April-June 2017	7.307692308	30.30769231
July-September 2017	17.38461538	24.76923077
October-December 2017	15.78571429	31.84285714
January-March 2018	20.58333333	24.41666667
April-June 2018	7	2
July-September 2018	20.57142857	30.14285714
October-December 2018	4.923076923	34.99441538
Retrieval Time	12 November 2019 17:45 Wib	12 November 2019 14:03 Wib

Perhitungan korelasi menggunakan fungsi ‘=PEARSON (array1, array2)’ yang tersedia di aplikasi *Microsoft Office Excel*. Sebagai contoh, untuk korelasi term ‘virus tuberkulosis’ maka akan menghitung kolom I dengan kolom L hingga menghasilkan nilai korelasi yang terdapat pada kolom E yaitu 0,331612353.

9. Daftar istilah pencarian dengan nilai koefisien korelasi besar sama dengan ( $\geq$ ) 0,444

Number	General Category	Search Term	Correlation Value
1	Definisi	paru tb	0,732352909
2		sakit tb	0,644330603
3		kronis tb	0,454969303
4		penyakit tb	0,728357716
5		paru tbc	0,735198943
6		sakit tbc	0,831721294
7	Diagnosis	dilakukan tb	0,531887366
8		infeksi tb	0,469892578
9		gejala tb	0,808484277
10		pasien tb	0,624515583
11		paru tb	0,732352909
12		darah tb	0,465430002
13		normal tb	0,671514872
14		untuk tb	0,776927681
15		berat tb	0,461684115
16		gejala tbc	0,474493561
17		paru tbc	0,735198943
18		untuk tbc	0,751299664
19	Epidemiologi	anak tb	0,627072174
20		infeksi tb	0,469892578
21		kronis tb	0,454969303
22		penyakit tb	0,728357716
23	Faktor Risiko	anak tb	0,627072174
24		darah tb	0,465430002
25		infeksi tb	0,469892578
26		usia tb	0,468592675
27	Tata Laksana	untuk tb	0,776927681
28		diberikan tb	0,467531671
29		pasien tb	0,624515583
30		untuk tbc	0,751299664

10. Daftar istilah pencarian dari kueri terkait pada GT dengan nilai koefisien korelasi besar sama dengan ( $\geq$ ) 0,8

Number	General Category	Search Term	Correlation Value
1	Other 3	bronkitis	0,820658044
2	Diagnosis	ciri ciri tbc	0,858677014
3	Diagnosis	ciri tb	0,832426032
4	Diagnosis	gejala tb	0,808484277
5	Other 4	google scholar	0,807526544
6	Diagnosis	icd 10 limfadenitis tb	0,885142371
7	Diagnosis	icd 10 tb	0,879452181
8	Diagnosis	icd 10 tb kelenjar	0,822974118
9	Diagnosis	icd tb paru	0,884785123
10	Definisi	kelenjar tb	0,814140951
11	Diagnosis	kode diagnosa tb paru	0,903883411
12	Diagnosis	kode icd 10 tb	0,866965476
13	Diagnosis	kode icd tb paru	0,841108996
14	Tata Laksana	obat tb	0,906818458
15	Tata Laksana	pro tb 4	0,899244973
16	Definisi	sakit tbc	0,831721294
17	Definisi	tb adalah	0,80776689
18	Definisi	tb mdr	0,813800313
19	Definisi	tbc adalah	0,878146589
20	Tata Laksana	tcm tb	0,865422029
21	Tata Laksana	toss tb	0,944734838

### 11. Istilah Kombinasi Utama dari kueri terkait pada GT

Number	Original Combination Term	General Combination Term
1	'adalah'	'adalah'
2	'gejala'	'gejala'
3	'sakit'	'sakit'
4	'ciri'	'ciri'
5	'ciri ciri'	'ciri ciri'
6	'kode diagnosa ... paru'	'kode diagnosa'
7	'obat'	'obat'
8	'icd 10 ... kelenjar', 'kode icd ... paru', 'kode icd 10 ...', 'icd 10 ...', 'icd ... paru', dan 'icd 10 limfadenitis'.	'kode icd'

### 12. Hasil uji konsistensi *template* istilah pencarian usulan sementara

Number	Term Position				Is Consistent? (yes or no)
	'Combinate Term' + 'Main Term'		'Main Term' + 'Combinate Term'		
	Search Term	Correlation Coefficient	Search Term	Correlation Coefficient	
1	sakit tb	0.644330603	tb sakit	0.771404096	no
2	adalah tb	0.8427403	tb adalah	0.80776689	yes
3	ciri ciri tb	0.872033114	tb ciri ciri	0.850012872	yes
4	ciri tb	0.832426032	tb ciri	0.838533619	yes
5	gejala tb	0.808484277	tb gejala	0.806129959	yes
6	kode icd tb	0.858101827	tb kode icd	0.858101827	yes
7	kode diagnosa tb	0.861335602	tb kode diagnosa	0.856012227	yes
8	obat tb	0.906818458	tb obat	0.939146312	yes
9	sakit tbc	0.831721294	tbc sakit	0.693854593	no
10	adalah tbc	0.880782022	tbc adalah	0.878146589	yes
11	ciri ciri tbc	0.858677014	tbc ciri ciri	0.853277529	yes
12	ciri tbc	0.792047624	tbc ciri	0.818081913	no
13	gejala tbc	0.474493561	tbc gejala	0.568754917	no
14	kode icd tbc	0.77120912	tbc kode icd	0.77120912	no
15	kode diagnosa tbc	0	tbc kode diagnosa	0	no
16	obat tbc	0.793847879	tbc obat	0.735676781	no

### 13. Tren kasus penyakit untuk setiap triwulan (level negara).

